PASHTO PHONOLOGY

An Evaluation of the Relationship between Syllable Structure and Word Order

MUHAMMAD KAMAL KHAN

...Publishing : eBook Collection (EBSCOhost) - printed on 2/10/2023 4:03 AM via
... 2...951 ; Muhammad Kamal Khan, Author.; Pashto Phonology: An Evaluation of the Relationship between Syllable Structure and Word Order
...ount: ns335141

Pashto Phonology

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Pashto Phonology:

An Evaluation of the Relationship between Syllable Structure and Word Order

By

Muhammad Kamal Khan

Cambridge
Scholars
Publishing

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Pashto Phonology: An Evaluation of the Relationship between Syllable Structure and Word Order

By Muhammad Kamal Khan

This book first published 2020

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Copyright © 2020 by Muhammad Kamal Khan

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

ISBN (10): 1-5275-4608-X
ISBN (13): 978-1-5275-4608-0

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

To my mother

- the most literary woman of the world, who was not formally educated!

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

TABLE OF CONTENTS

List of Figures ix

List of Tables x

Acknowledgements xii

Preface xiv

Abbreviations and Symbols xvi

Chapter 1 1
Introduction

Chapter 2 25
The Study of Typology and Language Universals

Chapter 3 60
Syllable Structure in Pashto

Chapter 4 121
Basic Word Order of Pashto

Chapter 5 146
Analysis of the Correlation between Syllable Structure and Word Order

Chapter 6 170
The Comparison between Pashto and English

Chapter 7 201
Global Analysis and Conclusions

Bibliography 207

Appendix A 219
Glossary of Concepts and Constraints

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Appendix B..... 221
Pashto Edge Phonotactics

Appendix C..... 224
Pashto Consonant Clusters

Appendix D 230
Pashto Words containing Consonant Clusters

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

List of Figures

1.1. The Graphic Representation of OT 10
1.2. Indo-European Language Family 16
1.3. The Classification of Pashto within the Indo-Iranian Family of Languages 16
1.4. The Geographical Area in which Pashto Is Spoken Natively 20
3.1. Universal Prosodic Hierarchy 67
3.2. The Structure of the Syllable 68
3.3. The Sonority Scale in Pashto 108
5.1. The Possible Order of Change in BWO 154
6.1. The Structure of MAXST in Pashto 173

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

List of Tables

1.1. The Five Main Dialects of Pashto 19
2.1. Implications/Correlations Identified by Lehmann (1973) 35
2.2. Comparison of Munda and Mon-Khmer Languages (Donegan & Stampe, 1983) 37
2.3. Polarizations in Munda vs. Mon-Khmer (Donegan & Stampe, 2004) 38
2.4. The Categories of Syllable Structure (Maddieson, 2005) 43
2.5. The Correlation of Word Order and Syllable Structure (Lehmann 2005) 44
2.6. Brief Comparison of Munda and Mon-Khmer (Donegan & Stampe, 1983) 45
2.7. Universal Number 196 of the Universal Archives 47
2.8. Universal Number 207 of the Universal Archives 47
3.1. Pashto Consonantal Phonemes 63
3.2. Pashto Vocalic Phonemes 64
3.3. Pashto Syllabic Patterns 70
3.4. Categories of Pashto Syllable Patterns 71
3.5. The Simple Syllables in Pashto 72
3.6. Bipartite Initial Clusters (Plosives + Liquids) 74
3.7. Bipartite Initial Clusters (Plosives + Glides) 75
3.8. Bipartite Initial Clusters (Fricatives + Liquids) 76
3.9. Bipartite Initial Clusters (Fricatives + Plosives) 77
3.10. Bipartite Initial Clusters (Fricatives + Glides) 77
3.11. Bipartite Initial Clusters (Fricatives + Fricatives) 78
3.12. Bipartite Initial Clusters (Fricatives + Nasals) 78
3.13. Bipartite Initial Clusters (Nasals + Others) 78
3.14. Bipartite Initial Clusters (Liquids + Others) 79
3.15. Bipartite Initial Clusters (Glides + Liquids) 79
3.16. Bipartite Initial Clusters (Affricates + Others) 80

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.17. Summary of Bipartite Consonant Clusters at Word-Initial Position in Pashto 81
3.18. Bipartite Final Clusters (Others + Plosives) 83
3.19. Bipartite Final Clusters (Others + Fricatives) 84
3.20. Bipartite Final Clusters (Others + Nasals or Affricates) 85
3.21. Summary of Bipartite Consonant Clusters at Word-Final Position in Pashto 86
3.22. Bipartite Medial Clusters 88
3.23. Summary of Bipartite Consonant Clusters at Word-Medial Position in Pashto 91
3.24. Summary of Bipartite Consonant Clusters in Pashto 92
3.25. Tripartite Consonant Clusters in Pashto 95
5.1. Six Types of Dominant BWO in Languages 151
5.2. The Change of BWO in Afro-Asiatic Languages 153
5.3. Different BWOs in the Niger-Kordofanian Macro-Family 156
6.1. The Case of ONSET in Pashto and English 180
6.2. The Case of NOCODA in Pashto and English 181
6.3. The Case of *COMPLEXONS in Pashto and English 182
6.4. The Case of *COMPLEXCOD in Pashto and English 182
6.5. Comparison between Pashto and English 196

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

ACKNOWLEDGEMENTS

This book is a slightly modified version of my PhD dissertation, which I submitted at the University of Azad Jammu and Kashmir (UAJK), Pakistan, in 2012. I completed this work as a visiting scholar at the School of English Literature, Language and Linguistics (SELLL), Newcastle University, Newcastle Upon Tyne, United Kingdom, in 2011. The basic ideas of the present study, however, go back even further, to my course on Experimental Phonology with Dr Sarmad Hussain at FAST University Lahore, during which I discovered the initial impetus for this work – the syllable structure of Pashto phonology.

A number of people played a role in the completion of this study. First, I owe a great debt of gratitude to my research supervisors Dr S. J. Hannahs (Newcastle University) and Dr S. N. H. Bukhari (UAJK), who were extremely kind to agree to be my advisers and in helping me formulate this work. Their invaluable guidance, feedback and, above all, very professional attitude steered me to come up with this study. Apart from them, I was very fortunate to have friends like Qadir, Mangrio, Rabia, Umar, Ghani and Riaz. Their comments helped me immensely through the early steps of this work. The administrative staff at the Department of English (UAJK) are also to be mentioned here, especially Javed and Ali – thank you Bhraoo!

My heartfelt thanks are also due to my external examiner, Professor Dr Ayaz Afsar (IIUI), and my committee members – Professor Dr Raja Naseem Akhtar and Dr Ayesha Sohail (both from UAJK), for reading the dissertation and asking very relevant questions, especially at the final defence. I also greatly benefitted from the reviews of Dr Henrik Liljerin (Stockholm University) and Dr Jalal Tamimi (Newcastle University) – both of them deserve a big THANKYOU!!

Special thanks are also due to the Higher Education Commission (HEC) of Pakistan for supporting my research financially during the course of research work at Newcastle University, through the International Research Scholar Initiative Plan (IRSIP). I am extremely fortunate to have received this support and, for sure, I do not take it for granted. Thank you, HEC!

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Worth mentioning are also my linguistics professors, whom I'll never be able to forget. They include Shahbaz Arif, Anjum Saleemi, Zafar Iqbal, Mubina Talaat, Zameer Naqvi, Mike Medley and Nazir Malik. Besides the study of linguistics, their inspiration gave me the required momentum to continue with my studies and enriched my life by bringing me into contact with so many like-minded people. My sincere prayers are with them, wherever they are!

Also my thanks are due to Cambridge Scholars Publications, United Kingdom, for their willingness to publish my work in the form of this book, for their suggestions on the material and for making it a comprehensive book.

Finally, and above all, my profoundest gratitude and love go to my family – my mother and my SAAZ and especially to my siblings, Naveed and Shafqa. My constant hard work and focused performance was only possible because of your care. We managed to weather the storm together and for me, at the moment, nothing could be a greater achievement than the completion of this work.

Lahore, March 2013
Muhammad Kamal Khan

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

PREFACE

A number of studies regarding dependencies between different levels of language have pointed out that languages with SOV word order tend to have a simple syllable structure (CV(C)) and vice versa (Lehman, 1973; Donegan & Stampe, 1983; Gil, 1986; Plank, 1998; Fenk-Oczlon & Fenk, 2004; Tokizaki, 2011; Tokizaki & Kuwana, 2012, among others). This correlation between syllable structure and word order has been typologically demonstrated by checking data from a number of languages and the phenomenon has been described as an implicational universal: an SOV word order implies a simple syllable structure.

However, the present study, by checking data from the Pashto language (an Eastern-Iranian language from the Iranian sub-group of the Indo-Iranian family), shows that the very phenomenon does not exist in the case of subject language. The study further explores the edge phonotactics and consonant clusters of the language for this purpose. The data was taken from three main dictionaries of Pashto. Besides, 5000 everyday words used by native speakers of Pashto were recorded for the study. Subsequently, the data was phonetically transcribed and the consonantal sequences were noted down. The findings showed that Pashto allows not only two-consonant clusters (CC) at all three positions (word-initially, -medially and -finally) of the word (various types of 108 consonant sequences) but also permits three-consonant clusters (CCC) at onset. The maximum syllabic template (MAXST) of Pashto which has been drawn from the current data is CCCVCC. Showing the rich bipartite consonant clusters at all three positions, Pashto syllabic structure questions the correlation between simple syllable structure and SOV word order. Thus, the accuracy of the implicational universal linking SOV word order with simple syllable structure needs to be re-examined.

The next important questions are: is there any correlation between basic word order and syllable structure? And, if yes, why does Pashto, being an SOV language, have complex syllable structure? These questions have been discussed in the light of the optimality-theoretic framework of the study. It is argued, on the basis of discussion of basic word order and syllabic structure of a number of language families from previous studies, that the correlation between syllable structure and word order does not in

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

reality exist and that the specific structure of a linguistic level is determined by the permutations of the constraints in CON. It is further suggested that Optimality Theory (OT) does provide the theoretical justifications for unmarked linguistic structures within various types of languages. Various interesting aspects of Pashto (e.g., reverse sonority clusters etc.) are also justified theoretically by using this framework of constraint interaction. Further, for broader theoretical impact and global analysis, the study compares Pashto (SOV) with English (SVO) language. This comparison reveals a number of interesting features, such as that syntactically distinct languages share a number of phonological characteristics. The frequency ratio of complex syllabic words (English 23% and Pashto 15% in content words) shows that syllable structure is not correlated to basic word order. Thus the fresh data and its novel analysis throw a new light on theoretical and typological assumptions based on the implicational correlations between different levels.

Finally, the present study also explores various features of Pashto phonology, such as syllabification rules, its sonority scale and reverse sonority clusters; points out a number of interesting aspects of Pakistani regional languages generally and Eastern-Iranian languages in particular; and recommends further detailed investigation on important topics concerned.

Keywords: Basic word order, correlation, East-Iranian, implicational universal, Optimality Theory, Pashto, syllable structure

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

ABBREVIATIONS AND SYMBOLS

ACC accusative
AUX auxiliary
BWO basic word order
C consonant
CC two-consonant cluster
CCC three-consonant cluster
CCCC four-consonant cluster
CL clause
DEM demonstrative
DO direct object
DIR direct case
EVAL Evaluator
FEM feminine
FUT future
FWO free word order
G glide
GEN Generator
IMPF imperfective
IO indirect object
IPA International Phonetic Alphabet
MASC masculine
MOP maximum onset principle
MSD minimum sonority distance
N noun
NSP natural serialization principle
NEG negative
NOM nominative
O object
OBJ objective
OBL oblique case
OT Optimality Theory
PART participle
PERF perfective
PL plural
PP past participle
PRES present tense
PROG progressive
PRT participle
Q question marker
R CL relative clause
REL relative
S subject
SSP sonority sequencing principle
SG singular
SOV subject + object + verb
SPE *Sound Pattern of English*
SSWL Syntactic Structures of World's Languages
TRANS transitive
UG universal grammar
V vowel (except in SOV)
VP verb phrase
WALS World Atlas of Language Structures
1P first person
2P second person
3P third person
σ syllable
(*) ungrammatical
1,2,3 first, second, third person
2p second-posit

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION TO THE STUDY

The relationship between syllable structure and word order has been one of the general issues in typological linguistics. In a number of studies on language structure (Lehmann 1973, 1978; Donegan & Stampe 1983, 2004; Donegan 1993; Gil 1986; Plank 1998; Tokizaki & Kuwana 2007, 2012), it has been stated that syllable structure has a strict correlation with the word order of a language. However, this correlation of syllable structure (the phonological component) with the word order (the syntactic component) of the grammar is poorly understood. In other words, the exact nature of the relationship between syllable structure and word order has not been well investigated. The aim of this book is to take a deeper view of the relationship between syllable structure and word order and, therefore, between phonology and syntax.

Typological linguistic theories have much to say about dependencies between different levels of natural languages. This idea has promoted the holistic approach towards the study of languages. Linguists (e.g. Gabelentz, 1901; Skalička, 1966; Greenburg, 1974; Klimov, 1974; Yartseva, 1979; Seiler, 1990) are of the view that human language has a definite tendency as a unit and that different levels of a language are related to each other. Among these levels, there is great potential for phonology to be related to syntax and morphology (Tokizaki & Kuwana, 2007, 2012). Linguists have typologically studied the nature of the relationship among various components of language and have demonstrated by checking data from a number of languages that languages with SOV (subject + object + verb) order tend to have simple syllable structure, and vice versa. They have defined implicational universals on the basis of their studies, providing a good deal of material for discussion regarding the correlation between syllable structure and word order. In the current study I argue and in subsequent chapters show that this phenomenon – and therefore the correlation – does not exist in the Pashto

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

language. On the basis of my data from Pashto, I claim that the authenticity of such implicational universals needs to be re-examined. The present study provides a detailed account of the phonotactics and syllabification of Pashto. It also describes the nature of the relationship between the syllable structure and word order of the language in the light of Optimality Theory (OT). Finally, the phonological and syntactic features of Pashto are compared with English and, on the basis of this comparison, the conclusion is drawn that syllable structure is not directly correlated with the word order of a language. Section 1.2 covers the primary goals, hypothesis and research questions of the study. Section 1.3 describes the significance of the study while Sections 1.4 and 1.5 provide background to the framework and justification for using OT as a methodology in the field of phonology. Section 1.6 gives a detailed introduction and rationale for the selection of the focus language. Finally, the organization of the book's contents and a summary of the present chapter are given.

1.2 GOALS, RESEARCH QUESTIONS AND HYPOTHESES

The primary goal of the current study is to offer a detailed account of the nature of the relationship between syllable structure and word order. As discussed above, Pashto has unexpected patterns of syllable structure and word order and provides evidence against the implicational universals and other previous works of various researchers. Thus the present work is a two-fold study consisting of a phonological and syntactic structuring of the language, on one the hand, and then the general application of the rules found to typological and theoretical linguistics, on the other.

Based on the previous studies mentioned in Section 1.1, the accepted nature of the relationship between syllable structure and word order can be summarized as follows:

- There is a correlation between syllable structure and word order.
- If the word order of a language is SOV (subject + object + verb), the syllable structure of that language will be simple, i.e. CV (consonant + vowel) or CV (C) (consonant + vowel + optional consonant).
- Similarly, if the word order of a language is SVO, the syllable structure of that language will be complex, e.g. CCCVCCCC (up to three consonants at onset level + vowel + up to four consonants at coda level).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Although these statements provide the starting points for an understanding of how word order is correlated with syllable structure, the case of Pashto is totally different. Pashto provides serious counter-examples to the above statements and generalizations. Although this language has SOV word order, it also has complex syllable structure (up to CCCVCC). So, the following is my key question:

Key Q: What is the nature of the relationship between syllable structure and word order?

This question is answered in the light of the study of the syllable structure and word order of the Pashto language. This language, as mentioned earlier, provides counter-examples to the generalizations of previous works. Therefore, the phonotactics and syllabification of Pashto are studied in detail. Further, to provide greater insight into the phenomenon, Pashto is also compared with English in Chapter 6. Based on the above key question, the following five sub-questions are formed:

Q1: What are the patterns of syllabification in Pashto language?
Q2: What is the basic word order of Pashto?
Q3: What are the similarities and differences between Pashto and English syllable structures? Q4: What are the similarities and differences between Pashto and English word order?
Q5: Is the nature of the syllabic pattern related to the word order of a language?

In order to further analyse the correlation, the current study also compares an SOV language (i.e. Pashto) with an SVO (i.e. English). Interestingly, both languages, despite their different word orders, allow complex syllable structures. The complex syllable structure of Pashto is strikingly similar to that of English. So questions 3 and 4 are focused on these points. Based on the above questions, my hypotheses for the current study are the following:

- The syllable structure of SOV languages is not always simple.
- The constraints on syllable structure are not directly related to the typology of word order in a particular language.

These hypotheses are subsequently investigated through the above questions. I will investigate them by checking data from Pashto and by showing consonant clusters at word boundaries and between words in the subject language. My arguments are based on a number of phonological and syntactic patterns of Pashto. Thus the authenticity of the implicational

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

universals regarding the syllable structure of SOV languages, especially in the case of Pashto, an Eastern-Iranian language, will be analysed and re-examined. These examples will be sufficient to demonstrate the viability of the research. From a conceptual and theoretical point of view, the possible reasons for complex syllable structure in Pashto will also be explored. The specific objectives are to:

1. Study in greater detail the relationship between syllable structure and word order,
2. Give a detail account of the phonotactics and syllabification in Pashto and figure out the overall syllable structure of the language,
3. Explore the possible word order of Pashto language, and
4. Compare the syllable structure and word order of Pashto with English

There are some general objectives of the study as well. These are to:

1. Analyse the authenticity of the implicational universals that SOV languages tend to have simple syllable structure,
2. Discuss the potential explanations for the possibility of complex syllable structure in an SOV language like Pashto,
3. Explore the relationship of phonology with syntax and morphology in the case of

Pashto, and finally,

4. Contribute to the existing theories of typological and theoretical linguistics.

The study will rely on the following sources for the data:

- Daryab Pashto dictionary (1994)
- The *Dictionary of Nine Languages* by Mengal (2009)
- The early dictionary of Pashto compiled by Captain Raverty (1859 [2001])
- A collection of 5000 words recorded from 532 native speakers from the rural area where the Yousafzai dialect is spoken natively
- Being a native speaker of the language, the observation and intuition of the researcher

For the sake of clarity and uniformity, the study will be delimited to the Yousafzai dialect of Pashto language.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

It is clear from the discussion so far that the current study is a constraint-based study which will examine different aspects of syllable structure of the Pashto language. In the framework of Optimality Theory, various constraints will be deployed to see the overall syllable structure of the language.

1.3 SIGNIFICANCE OF THE STUDY

As discussed earlier, very little work has been done so far in Pashto. Many of its phonological, morphological and syntactic aspects are yet to be studied. Typologically, it is the most interesting and historically important language of the area (Roberts, 2000). Many unique and interesting features of the language are yet to be explored. Similarly, the nature of the relationship of syllable structure with word order in the language is not in harmony with that of the majority of the languages of the world. Detailed discussion of this can be found in Chapter 2.

This research will provide basic description of phonological, syntactic and morphological features of the Pashto language. It may yield new information, as no research has so far been carried out on the relationship or dependencies of these levels in Pashto. As mentioned earlier, this study will add up to different strands of implicational universals which have been developed regarding the relationship (or covariation) between syllable structure and word order that have implications for theoretical and typological linguistics. Moreover, it is hoped that the information provided by this study of word order and syllable structure, while comparing Pashto and English, will also highlight interesting features as a contribution to the field of typological linguistics.

There are some further points which add to the significance of this study. For example, it will study phonotactics and syllabification in the Pashto language. It will also analyse some idiosyncratic features of Pashto, such as the reverse order of (or the violability of) the SSP (Sonority Sequencing Principle) in the language (Bell & Saka, 1982; Levi, 2004). It will also provide information about the nature of possible word order and possible complex syllable structure in Pashto. Another motivation for this project is the relative lack of precedents. There is no work completed in Pashto so far on the topic. Thus, this book will serve as a milestone in the field of theoretical linguistics and will also include Pashto and other Indo-Iranian languages in the discussion. The theoretical framework of the study is discussed in the following sections.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

1.4 METHODOLOGIES FOR PHONOLOGICAL STUDIES

Bofore moving on to discuss the theoretical framework for the current study, it seems appropriate to give here a brief account of the methodologies applied in the field of phonology. Subsequently, this comparison will also be helpful in justifying my selection of Optimality Theory (OT) as the framework for the present study.

According to *The Sound Pattern of English* (SPE) (Chomsky & Halle, 1968), phonology consists of sequentially ordered rules. This idea ultimately led to the strict application of rule-based analysis in phonological enquiry. In the line of generative grammar, rule-based phonology assumes the step-by-step derivation of rule, thus making it strictly derivational in its nature. The brief structure of rule-based phonology, based on Chomsky and Halle (1968), is given below.

In rule-based theoretic grammar, one needs:

a. A universal/finite set of features to describe elements, and
b. A restricted rule formalism to express generalizations.

It says that any phonology is a set of ordered rules. For example: Rule: [+ syl] → [+ nasal] / _[+ nasal]

As in: /hæt/ →/hæt/, /hænd/ →/hænd/

Rule-based phonology has several problems in terms of phonological analysis. The main problems are:

a. An infinite number of rule types are possible,
b. A phonology can contain an unbounded number of rules, and
c. Input restrictions are treated differently.

Secondly, in derivational phonology, the classical formal proposal is that grammars subject rules to rule-ordering constraints – "R1 is ordered before R2" – which regularize the sequence of application. Then, in principle, a rule may create the conditions for another rule to apply (called a feeding effect), although a rule may fail to apply when another rule creates the conditions for its application (a counter-feeding effect); one rule may wipe out the condition for another rule before the other can apply (a bleeding effect); or the rule whose conditions would be wiped out by another rule may be allowed to apply first (a counter-bleeding effect). Thus the explanation of the sequence among the rules has been a big issue in

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

phonological analysis. So, alongside rule ordering, or in place of it, other principles of application have also been tried, from which OT emerged as a theory for phonological analysis.

The idea of phonological conspiracies (the phenomenon that various rules aim at the representational goal) was also an issue to explain within rule-based phonology. In OT phonology, constraints on surface forms can easily express such conspiracies. In other words, within the OT paradigm, those phonological conspiracies can easily be justified by forming the specific hierarchy for a language.

Within rule-based phonology, constraints were inviolable. With the emergence of OT, constraints are no longer inviolable. Based on the nature of the violability of specific constraints within languages, low-level constraints can be violated in order to satisfy another higher-level constraint. For example, the NOCODA constraint is followed within some languages strictly by allowing only CV syllabification. But other languages may be more liberal about this constraint, thus allowing CVC syllable structure.

OT, as a major shift in the line of generative grammar, profoundly changed phonological enquiry from a rule-based to an output-based model. However, it may be kept in mind that theories have consequences only to the extent that they are formalized. All phonological theories have a common goal of describing the nature of human sound features as accurately as possible. The surface level differences just show the interaction of descriptive mechanisms. Moreover, formal comparison of the derivational and optimality-theoretic descriptions of the underlying-to-surface function of phonology compares alternative grammar designs. In a data-centred comparison, these theories may be compared by their ability to provide a natural analysis of some given data.

There are some common points in OT and derivational theory as well. For example, derivational and optimality-theoretic grammars for phonology are found to have a similar outline insofar as they associate surface representational forms with underlying representational forms. Mainly, what derivational theory and OT do is to provide two alternative descriptions of the function that maps underlying forms to surface forms. In a rule-based analysis, the rule configuration would be expressed twice – once as context of rule, then in a morphophonemic substitution (applied output) rule (as is clear from the above examples). In an optimality analysis, however, both processes follow from the single constraint

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

interacting in different ways within a single hierarchy of constraints. Derivational phonology and OT phonology are comparable on three fronts: rule operations and faithfulness constraint violations; serial rule interaction and evaluative constraint interaction; and derivational sequences and harmony scales. In each case, the correlation breaks down and pertinent data emerges. In Optimality Theory, a strong theory of substantive universals is attempted:

> U[niversal] G[rammar] provides a set of constraints that are universally present in all grammars … a grammar is a ranking of the constraint set. (McCarthy & Prince, 1994, p. 336)

Constraints in OT are not merely solutions to language-particular problems; they are claims about UG with rich typological consequences.

> Descriptive universals rarely make good constraints, but descriptive *tendencies* often do. Indeed, the success of OT in incorporating phonetic or functional generalizations is largely a consequence of its ability to give a fully formal status to the otherwise fuzzy notion of a cross-linguistic tendency. (McCarthy, 2002, p. 39–40)

OT provides a smart interaction of recognized constraints on a linguistic structure for analysis. In Optimality Theory, constraints must be well defined, so as to assign a particular number of violation marks to each candidate, and violated in just the way predicted by the theory – minimally, when in conflict with higher-ranked constraints (Prince & Smolensky, 1994). In some cases there is evidence that constraint interrelationships other than ranking are needed. Since the core of the theory employs markedness constraints and faithfulness constraints (as will be shown in the next section), any other constraint types would require a careful defence. Processes arise from adherence to markedness constraints at the expense of faithfulness to underlying forms. It has been argued that the formal universals of Optimality Theory enable a natural analysis of "conspiracies", whereby different processes achieve the same output generalization. Recent much broader application of OT in phonology has proved that OT is the best methodology for such analyses. The next section introduces OT formalism in detail.

1.4.1 Introduction to Optimality-Theoretic Framework

The current study will employ Optimality Theory, the most widely adopted constraint-based phonological methodology, as its theoretical framework. Among different approaches applied in the field of phonology, OT is considered the most important and powerful methodology by most

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

of the researchers (e.g. Kager, 1999; McCarthy 2008). The constraint-based approach of OT has had quite a strong impact on research work in linguistics. OT is now also applied in areas other than phonology. Recently, semantics and syntax have also benefited from the constraint-based approach of this theoretical framework.

OT proposes that grammars arise from the interaction of the violable conflicting constraints (Prince & Smolensky, 1993, 2004). It formalizes the idea of "conspiracies", or the triggering of one or more than one phonological rule by the evidence of a single phonological structure of a language. Under OT, briefly speaking, the phonological constraints are hierarchically ranked and are violable by the phonetic forms of their underlying representations. Such constraints are minimally violated by a set of potential candidates (possible forms of output) and the one which enacts the least serious violations wins. This candidate is called the optimal candidate. The violation of the constraints is defined in terms of their hierarchies, e.g. the violation of the higher-ranked constraints is the most serious one and is therefore avoided. There are two types of constraints within this framework, namely faithfulness constraints and markedness constraints. Faithfulness constraints require that the surface form (the output) must be identical to the underlying form (input) in some way or another, while markedness constraints demand that the structure of the surface form (output) should be well formed. The interaction of these two types of constraint decides the winning candidate from among the available choices and thus declares the optimal form for an input.

OT is basically a constraint-based competition mechanism among a number of possible candidates. According to the basic idea of OT, the generation of utterances involves two important functions, viz., GEN (or Generator) and EVAL (or Evaluator). GEN generates a set of unique output candidates for an input. Out of these candidates, one is identical to the input, while the rest are somewhat modified in their structure. Then EVAL evaluates the candidates and chooses the optimal one that best satisfies the set of ranked constraints depending on the nature of the violation. This means that the constraints in OT are violable. The most important process is, of course, the selection of the optimal candidate as the output from a list of the hierarchical choices. This is done by EVAL. Figure 1.1 shows the processes of GEN and EVAL and how they reach the output from the input (Davenport & Hannahs, 2005). This figure is only indicative for the sake of clarity: the actual tableaux are different from this representation.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

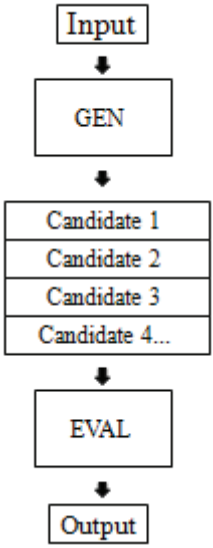


Figure 1.1. The Graphic Representation of OT

The two types of constraints discussed above, i.e. markedness and faithfulness constraints, function as part of EVAL. Markedness constraints make sure that the output candidate is well formed, prohibiting the forms which are difficult to produce or comprehend (Kager,

1999, p. 9). Such constraints ensure the utterance of certain segments and impose restrictions on others. The following are examples of markedness constraints:

(1) **Markedness Constraints**
a. Syllables must not have codas (NoCoda)
b. Syllables must have onsets (Onset)
c. Obstruents at coda position must not be voiced (*VDOBS (CODA))
d. Obstruents must be voiced (VDOBS)

On the other hand, faithfulness constraints ensure the similarity of the output to the input. For example, Kager (1999, p. 10) lists the following faithfulness constraints which are found in most languages of the world:

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(2) **Faithfulness Constraints**

a. The output must present all segments which are found in the input (DEP-IO)
b. Elements adjacent in the input must be adjacent in the output (CONTIGUITY)
c. Input segments must have counterparts in the output (MAX-IO)
d. The specification for the place of articulation of an input segment must be preserved in its output correspondent (IDENT-IO (PLACE))

The important point to remember is that the hierarchy of both markedness and faithfulness constraints is always language-specific. One constraint which is at the top of the ranking scale in a language may not be in the same position in another. In an analysis, various markedness and faithfulness constraints conflict; therefore, the ranking of the candidates is dependent on the violation pattern of the language. So, the ranking of the constraints is not universal; it is based on the violating nature of a language. The ranking hierarchy of one language may not be applicable to other languages. But every language has its own strict ranking for its constraints. A winning candidate must satisfy a high-ranking constraint for which that candidate may violate the lower-ranking constraints. In other words, the optimal candidate is not required to satisfy all constraints. It may violate a constraint and still be eligible to become the winner if it satisfies the top-ranking constraint.

I discussed the OT framework for the phonological analysis in the above section. The major portion of the current study is based on the phonological structure of the Pashto language, for which OT is considered the best methodology. But the current study, being typological in nature, also includes some portions related to syntax, as I am concerned with the basic word order of Pashto and English too. Therefore, for word-order analysis, I will apply the same methodology. It is also important to mention that, though initially meant for phonology, OT is now also benefiting other fields of linguistics such as syntax and morphology. Work in the field of syntax under the OT paradigm started quite late, but the appropriateness of the methodology is being proven here as well. Thus, OT is gaining ground in both morphological and syntactical analysis (Archangeli & Langendoen, 1997; McCarthy, 2002; Prince & Smolensky, 2004).

The organization of syntax in OT is not entirely different from that of phonology. According to Speas (1997), the particular group of words for a

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

sentence constitutes the input and the possible structures for that sentence are the candidates for the output. The process, from the generation of possible choices of sentence to the selection of the optimal candidate (the mapping from input to output), is mediated by the same options of GEN and EVAL. GEN provides the set of potential candidates to an input and EVAL selects the winning candidate from those options created by GEN. The following is the general organization of syntax in OT given by Speas (1997, p. 178).

(3) General Organization of Syntax in OT INPUT: group of words

GEN: creates candidate outputs for the input

EVAL: uses the constraint hierarchy to select the best candidate(s) for a given input from among the candidates produced by GEN; these constraints are ranked; lower-ranked constraints may be violated

This theoretical framework has not yet been used on the Pashto language. Since OT is the most innovative methodology in the field of linguistics in general and in phonology in particular, I will be analysing the data for Pashto within the theoretical framework of OT. I will give an account of how consonant clusters work at different positions in the syllable structure of Pashto. Similarly, I will relate the word order of the language to the syllable structures in order to investigate the hypotheses of the study. In the following sections, I justify the selected framework for the current study.

1.5 WHY AN OPTIMALITY-THEORETIC STUDY?

As discussed in the preceding section, OT is considered by many to be the best available methodology in the field of phonology. This constraint-based framework has influenced almost all fields of linguistic research in the last two decades. Thus the suitability of OT as the best framework qualifies it as a methodology for the current study. Though the main justification is, as mentioned above, the lack of any previous OT work in the language, there are some other potential justifications for the selection of OT as framework for the current study. For example, OT is inherently a typological framework. It is a constraint-based methodology and has provided answers to a number of long-standing questions in the field of linguistic research. For example, it has renewed the connections between language acquisition and typology in the field of phonology and thus it seems to capture some of the fundamental truth about human language (McCarthy, 2002, p. 237). Moreover, OT is an ideal framework as it spans

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

the traditional linguistic discipline, allowing for interface studies, and the nature of the current study is also of this type. It involves the relationship between syllable structure and word order and therefore between phonology and syntax. These points are further elaborated in the following section.

1.5.1 OT as Inherently Typological Framework

The first and foremost reason for the selection of OT as the research framework for the current study is that the nature of OT is inherently typological (McCarthy, 2002, p. 237). It classifies languages on the basis of their violation of various constraints. This aspect of OT is directly related to the current study. Here, I am concerned with typological issues. The basic point of the research is embedded in the classification of languages. The idea of the current study is taken from previous works which were carried out by typological and theoretical linguists. OT examines the restrictions available in the phonological patterns of a particular language and analyses relevant constraints to account for such idiosyncratic features. The main topic of the study, i.e. syllable structure, is one of the prominent topics researched in the OT framework today. Several issues of syllable structure have been explained by studying phonetic and phonological patterns typologically. So it is hoped that the current study, being typological in its scope, will benefit from the methodological force of OT.

1.5.2 A Constraint-Based Framework

The second most important reason is that OT is a constraint-based framework. As discussed in Section 1.4., the constraints of markedness and faithfulness conflict and interact with each other, and finally the optimal candidate is selected. This aspect of OT suits the current study well, because a number of violations of phonological constraints are found in this language. For example, the nature of the relationship between syllable structure and word order in this language is very interesting (SOV word order with complex syllable structure). Some patterns are not normally available and, therefore, are called “typologically unusual” (Roberts, 2000, p. 8). Thus this language violates the universal implications and provides interesting data for further research in this filed. Then, there are other features which are not easily found, such as the reverse sonority sequence of the language. Pashto violates the rules of Sonority Sequencing Principles (SSP) in a number of its syllable clusters (Bell & Saka, 1982; Levi, 2004). The following words are examples in this regard:

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(4)

a.	/wrəz/	day
b.	/wror/	brother
c.	/wla:ɽ/	upright
d.	/lmər/	sun

Such idiosyncratic features of Pashto can be best studied under the framework of OT, as it justifies these features simply in terms of different constraint hierarcy. Moreover, OT provides direct connection between acquisition and typology (McCarthy, 2002). Under this characteristic, one can answer the question of how these exotic patterns are acquired by the speakers of this language. Thus the advantage of the OT framework is that it provides the tools to go beyond the normal process of rule application or merely describing the process, by connecting specific phonological phenomena with the underlying tendencies of the phonology of one specific language as a whole.

1.5.3 An Ideal Methodology for Interface Studies

With the advent of OT, the focus of phonological approaches in relation to other fields of linguistics, such as morphology and syntax, has realigned in a useful direction. According to Ramasamy (2010), the flexibility of OT as a theoretical framework in tackling interface-related issues directly and competently has achieved significant attention in the field of linguistic research. He says “the approach showed a significant breakthrough in analysing the structural changes at interfaces, as opposed to previous approaches which appeared to have limited applicability” (p. 5). The markedness constraints versus faithfulness constraints of OT conflict with other forces (e.g., alignment constraints) at different intersections and thus provide enough evidence for the active role of interaction. The advancement of OT has given rise to a number of approaches within phonology, such as the connections between language acquisition and typology (McCarthy, 2002, p. 237). The current study, in one sense, is an interface study focusing on the relationship between syllable structure and word order. Therefore, OT is an ideal framework for it.

In this section, the selection of OT has been justified in the light of a number of points. The current study, being focused on the relationship

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

between syllable structure and word order (therefore, phonology and syntax), discusses the interface domain. The points raised in the above sections are evidence of the importance of OT in studying relationship issues between phonology and syntax. Through the application of OT as a methodology, it is hoped that the same benefits will be achieved for the current study of Pashto syllable structure and word order. The language is introduced below in Section 1.6.

1.6 THE FOCUS LANGUAGE: PASHTO

The focus of the current study is the exploration of the relationship between syllable structure and word order in the Pashto language. Pashto provides enough data as counter-examples to the generalizations of the World Atlas of Language Structures (WALS).[1] Details of such examples are given in Chapters 3 and 4. The language family of Pashto is detailed in Section 1.6.1 below.

1.6.1 Language Family

Pashto belongs to the Eastern-Iranian branch of the Indo-European family of languages (Morgenstierne, 1926; Payne, 1987). This group also includes languages like Ossete (North Ossetia, South Ossetia) and Yaghnobi (Tajikistan). Eastern-Iranian and Western-Iranian (which includes Persian) are major sub-groups of the Iranian group of languages. This group belongs to the Indo-Iranian branch of the Indo-European family of languages. Figure 1.2 shows the family tree of Iranian languages (adopted from Tegey and Robson, 1996, p. 5).

Figure 1.2 shows that Iranian languages belong to the Indo-European language family, and so is distantly related to English and other European languages. Indo-Iranian languages are spoken in a wide area stretching from portions of eastern Turkey and eastern Iraq to western India (see Crystal, 1987; Payne, 1987). The other main division of Indo-Iranian, in addition to Iranian, is the Indo-Aryan languages, a group comprising many languages of the Indian subcontinent including Sanskrit, Hindi, Urdu, Bengali, Gujarati, Punjabi and Sindhi. Figure 1.3 shows the classification of the Pashto language within the Indo-Iranian family of languages.

[1] http://wals.info (accessed on 20 August 2011)

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

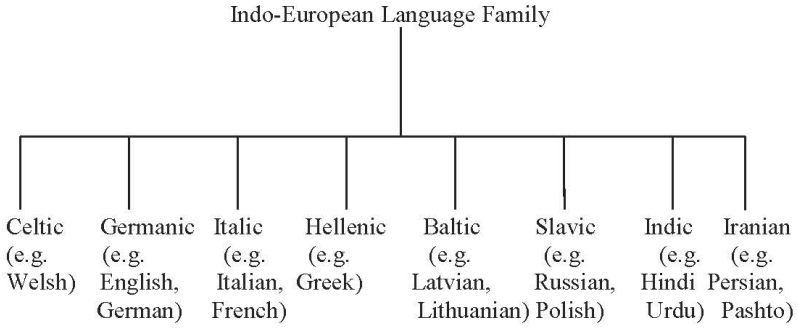


Figure 1.2: Indo-European language family

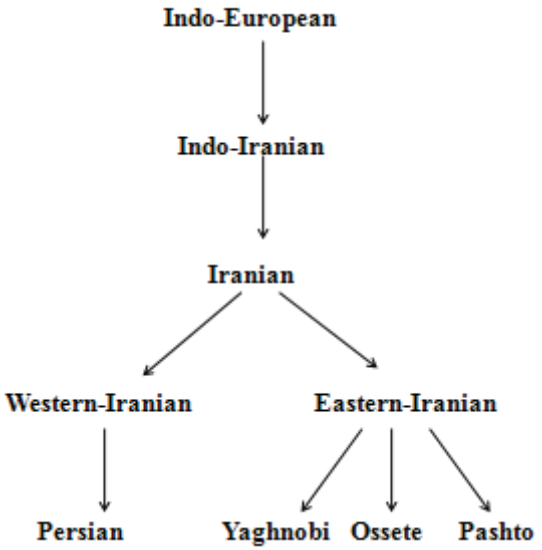


Figure 1.3: The Classification of Pashto within the Indo-Iranian Family of Languages

Figure 1.3 shows that Pashto is a sister language of Yaghnobi and Ossete,and a close cousin of Persian. In the following section, I discuss the historical background of the language.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

1.6.2 Historical Background

The traces of the history of Pashto are present in the vocabulary of the language. The majority of terms and expressions can be traced to the roots of Pashto roots as a member of the Eastern-Iranian language group. It has also borrowed words from other languages which co-existed at different times over the course of more than two thousand years in the region. The oldest borrowed words were taken from Greek, datinge back to the Greek occupation of Bactria in the third century BC. A few traces of contact with Zoroastrians and Buddhists are also found in the vocabulary of the Pashto language. Within the Islamic period, Pashto also borrowed many words and expressions from Arabic, as well as from Persian languages. On the other hand, due to its close geographic proximity to the languages of the Indian sub-continent, Pashto has also for centuries included a number of words borrowed from Indian language.

Pashto is an ancient language which is believed to have a history of thousands of years (Tair, 1989). According to Tegay and Robson (1996), it is the most conservative of the Iranian languages. The earliest written records of Pashto are believed to have been composed in the sixteenth century, in an account of Sheikh Mali's conquest of the state of Swat. Subsequently, in the seventeenth century, Khushhal Khan Khatak, considered the national poet of Pashto, wrote in the language. Since the twentieth century there has been a rapid increase Pashto writings within journalism and other modern genres. These new trends forced some innovations in the language and the creation of many new words and expressions in it.

Pashto has long been accepted as an important language in Afghanistan and Pakistan. In the nineteenth century, classical Pashto was the object of study and research by British soldiers and administrators in the area. Some of them also compiled classical grammars which are still used. The dictionary compiled by Captain Raverty (first compilation in 1859) is among the historical works on the language carried out by British officers. In 1936, Pashto was declared the national language of Afghanistan by royal decree. Today, Dari and Pashto are both officially declared national languages of Afghanistan. It is also taught in schools, colleges and universities in both Pakistan and Afghanistan. The next section discusses the various dialects of Pashto language.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

1.6.3 The Dialects of Pashto

According to the classification of Payne (1987), Pashto is an Eastern-Iranian language. It has many dialects. In one of the first studies on Pashto, Grierson (1921) recognized two dialects of Pashto, namely the North-East dialect and South-West dialect. Penzle (1955) argues for the division of Pashto into three main dialects, while Mackenzie (1959) talks about four dialectical areas of the language. Similarly, Tegey and Robson (1996), in their *Pashto Reference Grammar*, have divided Pashto into three major dialects. There are some other writers, such as Henderson (1983) and Skjaervo (1989), who insist on four dialectical varieties of Pashto. Most recently, Rehman (2009, p. 8) has recognized five dialects of the language. The main reason for these differences in opinion regarding the dialects of Pashto language is that very little variation is found among the various dialects.

There are a number of distinctions used by researchers for the division of Pashto into dialects. One of the main distinctions is the difference between *hard* and *soft* Pashto. According to this division, the speakers of the *hard* dialect can be recognized by their pronouncing the letter ښ as /x/, while speakers of the *soft* Pashto pronounce this letter as /ʃ/ (Hallberg, 1992, p. 9–10). On the other hand, Grierson writes that Pashto language is almost the same over the whole geographical area. But when one goes from the area of one tribe to another, one finds slight differences in their grammar and pronunciation. Grierson (1921) has given the example of differences in pronunciation between the North-East and the South-West dialects. According to him, speakers of the North-East dialect pronounce the letter ښ as /x/ and letter ږ as /g/, while those of the South-West pronounce them as /ʃ/ and /ʒ/ respectively (1921, p. 7).

The most recent work on Pashto language, by Rehman (2009), summarized the discussion by previous researchers on the dialects of Pashto. According to him, there are five main dialects within Pashto language. Table 1.1 shows the five main dialects with their geographical area given by Rehman (2009).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1.1: The Five Main Dialects of Pashto

Dialect	Areas where spoken
1. North-Eastern (Yusafzai) dialect	Most parts of KP (Khyber Pakhtunkhwa) province of Pakistan, e.g. Peshawar, Charsada, Swabi, Dir, Buner, Shangla, Swat, Bajaur and some parts of Hazara. This dialect is considered the "standard" dialect by many linguists.
2. North-Western (Central) dialect	Waziristan, Karak, Lakki Marwat and Bannu (southern part of KP). It is also called the Western dialect.
3. South-Eastern (Quetta) dialect	Balouchistan province of Pakistan, e.g. Quetta, Pishin, Chaman and Pishin Kakari.
4. South-Western (Kandahari) dialect	The main cities of Afghanistan, e.g. Jalalabad, Gardez, central parts of Kabul and the entire province of Kandahar.
5. The Middle Tribal Dialect	Hangu, Thal and Parachinar

The main distinction among the five dialects of Pashto is based on pronunciation. On the other hand, there is very little variation on the basis of lexical items. But the phonological variation of the language does not heavily affect mutual comprehensibility among the speakers of different dialects. For example, the second consonant of Pashto is a retroflex /ʃ/ in the Kandahari dialect and a palatal fricative in the Kabuli dialect. In general, however, a speaker from any dialect of Pashto readily understands another. The Central and Southern dialects are more divergent than the rest of the dialects.

The North-Eastern dialect of Pashto, which is also called Yousafzai dialect, is considered prestigious (Hallberg, 1992) and therefore the standard dialect of the language. The same Yousafzai dialect is reflected in the spelling system of the language. Moreover, Pashto literature is written in this dialect, as well as daily newspapers, dictionaries, etc. In the current study, the main data relates to this dialect, which is at the same time the mother tongue of the researcher. The geographical area in which Pashto is spoken is discussed in the next section.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

1.6.4 The Geographical Area of the Pashto Language

Pashto is one of the principal languages spoken in Pakistan and Afghanistan. It is spoken natively by over half of the population of Afghanistan. Pashto is also the native language of 90% of the population of Khyber Pakhtunkhwa province in Pakistan. It is also spoken in Baluchistan and by a community of 2.5 million in Karachi as their mother tongue. In some cities of Punjab, such as Attock, Mianwali and Rawalpindi, Pashto is spoken among other local languages. Pashto is also spoken in different parts of Iran, India, UAE, Canada, Australia and other countries. Figure 1.4 shows the geographical area of Pakistan and Afghanistan (adopted from McClean, 1990) in which the Pashto language is spoken natively.

Figure 1.4: The Geographical Area in which Pashto Is Spoken Natively

Pashto has been an official language of Afghanistan since 1936 (Rehman, 2009). Today its official status is shared by Dari. Section 1.6.5 provides details about the speakers of Pashto.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

1.6.5 Speakers of Pashto

According to Rehman (2009, p. 3) Pashto has approximately 50 million speakers. It is spoken in Pakistan, Afghanistan and Iran. After Punjabi, Pashto is the most widely spoken native language in Pakistan (Rehman, 2009). Rehman (2003) is of the view that Pashto is spoken by 15.42% of the total population of Pakistan. It is also spoken by over 50% of the population of Afghanistan. The picture given by Rehman (2009) about the population of Pashto speakers is realistic, as it not only considers Pashto as the second major language of Pakistan but also takes into account the number of Pashto speakers in Afghanistan, Iran, India and other countries of the world. There are some sources which give a figure half of that above for the number of Pashto speakers (e.g. Robson and Tegey, 2010, p. 721), but fail to give any solid grounds for this. Therefore, I consider the estimate of Rehman (2009) more accurate than others.

Major cities in Pakistan where Pashto is spoken as a native language include Peshawar, Mardan, Quetta and Karachi; in Afghanistan, they include Kabul and Kandahar. Today, Pashto speakers are found in almost every city of Pakistan. This might be due to the migration of Pashto speakers to these cities for business and other purposes. The recent decades of Afghan conflict have also brought Pashto speakers to far-flung areas of the world. A huge number of them have migrated to Australia, Canada, India and the United States. As a result of their ethnocentric social norms and particular lifestyle, they mostly move in groups and take their culture and language with them.

1.6.6 Previous Work on the Language

Despite the very rich history of its culture and folklore, very little academic work has been done so far on this language. Many of its phonological, morphological, semantic and syntactic aspects are yet to be examined. Many unique and interesting features (Morgenstierne, 1926, 1940) of the language are yet to be explored. Tegey and Robson (1996) are of the view that it is the most conservative language of the region, as it still contains a number of archaic sounds and patterns.

The grammatical tradition of Pashto is not very long-standing. It is not as standardized as other languages of the world such as English and other European languages (Rehman, 2009). As a result, there is no universally agreed variety of “standard” written or spoken Pashto among the speakers of this language. Some steps for the promotion of Pashto were taken at

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

governmental level in the past. For example, the Pashto Academy of the University of Peshawar was established in 1950s in order to promote Pashto. Along the same lines, the late king of Afghanistan, King Zahir Shah, established the Pashto Academy in Kabul in 1979. In spite of these efforts, the grammatical tradition of the language is still in infancy.

The initial works on Pashto, carried out mostly by British officers, were based on lexicography and the compilation of very simple grammars. In this regard, Raverty (1859), Khairullah (1905), Grierson (1921), Morgenstierne (1926, 1940), Mackenzie (1959, 1987), Penzle (1955), Shafeev (1964), Becka (1969), Khatak and Ishaq (1974), Henderson (1983), Skjaervo (1989), Tegey (1977, 1979), and Momand and Sehrai (1994) are worth mentioning. Some of these works were concerned with the alphabets, basic grammar and sounds of the language. Such grammars and dictionaries were written in order to fulfil teaching and learning needs. Among recent works, *Sociolinguistic Survey of Northern Pakistan, Volume 4: Pashto, Waneci, Ormuri* by Hallberg (1992) and *A Reference Grammar of Pashto* by Tegey and Robson (1996) are the important ones on the grammar of the language. There are some other studies which were carried out in order to explore the specific syntactic features of Pashto language. For example, Lorimer (1915), in his book *Pashtu Part 1: Syntax of Colloquial Pashtu,* discusses the syntax of the language. Jan Muhammad (1993), in his work *Second-Position Clitics in Pashto*, and Roberts (1997, 2000) in *Clitics and Agreement*, discuss the syntactic features of Pashto. Most recently, Rehman (2009), in his PhD dissertation, discusses the syntax of Pashto within the minimalist programme. These studies show the recent trend among researchers of Pashto towards the modern grammatical frameworks of pure linguistics.

1.7 THE ORGANIZATION OF THE BOOK

The present volume has seven chapters altogether. Chapter 2 reviews the related literature on the topic. This chapter starts from earlier theories. Typological approaches based on cross-linguistic data will also be discussed. The implicational universals given by Universal Archives[2] and the views of WALS are also discussed in connection with the correlation between syllable structure and word order. The progression of the background studies is set out chronologically. The chapter further goes on to discuss comparatively recent views on the topic given by linguists, and

[2] http://typo.uni-konstanz.de/archive/intro/index.php (accessed on 20 July 2010)

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

to highlight gaps in the approaches and implicational universals of various researchers.

Chapter 3 gives a complete description of Pashto phonotactics and syllabification rules. The relevant phonetic and phonological characteristics of Pashto vowels and consonants are described in this chapter. The chapter starts with a brief introduction to previous works in Pashto. It goes on to explore the phonotactic rules of Pashto syllables in OT format. Word-initial, word-medial and word-final consonant clusters will be analysed in detail. This portion of the thesis will be an overall analysis of the language and its complex syllable structure at every level.

Chapter 4 assesses the syntactic features of Pashto language by means of basic word order and other syntactic categories of the language. This chapter is fully devoted to proving the SOV structure of the language. Apart from the primary word order of the language, other possible word orders are also discussed here. Case system, simple and complex verbs, clitics, and agreement within the syntactic components of the language will be highlighted in this part of the book.

Chapter 5 discusses the nature of the possible relationship between syllable structure and word order within the Pashto language. The case of Pashto as a proof of negation to the implicational universals under study is outlined. This chapter further envisages the nature of the syllable structure of the language in the light of Pashto data. Various gaps in the subjectof correlation are analysed here.

Chapter 6 is devoted to comparison of Pashto with English in terms of syllabic complexity and word order possibility. This part of the study will single out the similarities and differences between the two languages. Being a typological study, the phonological and syntactic features of the Pashto language will be compared with English in this chapter. The basic idea of this comparison is that despite having different word orders, both of the languages have complex syllable structures. This part of the study will enhance the validity of the research and the inclusion of the English language will further increase the readership of the study.

Chapter 7, “Global Analysis and Conclusions”, is the last chapter of the study. This chapter closes the discussion raised in Chapters 3 and 4. It seeks to provide possible reasons for complex syllable structure in a natural language. By testing the nature of Pashto in connection with this, conclusions are drawn about the effect of differences in word order and

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

syllable structure. This chapter also points out areas for further detailed future research in the field.

1.8 SUMMARY OF THE CHAPTER

The present chapter covers three basic points. The introduction section describes the idea of the current study and its goals, with assumptions and hypotheses. It also elaborates on how previous research has tried to prove the correlation between syllable structure and word order in typological linguistics. The counter-features of Pashto language against such implicational universals were also introduced. The second issue covered by this chapter is the introduction to the framework of the study. The significance of OT as the most important methodology in the field of phonology was highlighted within this section. This chapter argues that OT is the most appropriate framework for the current study.

Next, this chapter provides an introduction to Pashto. It gives primary information on the background of the language and its dialects. Other closely related information, such as language family, historical background, geographical area and population of the language, was also discussed. The last segment of the chapter is the chapter overview of the thesis. Each part of the chapter is self-explanatory and provides the basic information required. Overall, this chapter illustrates the background of the topic and the language under study, and verifies the research interest and direction of the subject.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

CHAPTER 2

THE STUDY OF TYPOLOGY AND LANGUAGE UNIVERSALS

2.1. INTRODUCTION

In this chapter, I will discuss the background to the correlation of syllable structure with word order. Keeping in view the typological nature of the current study, the chapter begins by briefly introducing typology, and the scope and history of typological studies in the field of linguistics. It offers an introduction to some basic notions of typological linguistics and discusses various types of correlations and language universals. It also reviews the studies describing word order as the basis of typological analysis and universal implications. The main part of the chapter reviews previous scholarship on the correlation of word order with syllable structure. Along with discussing correlations and implicational universals, an analysis of the various aspects of the subject topic is also provided. Finally, the chapter summarizes the views of previous researchers on the correlations in hand. Through highlighting the gap in previous studies on the topic, the chapter closes by demonstrating the viability of the current study.

2.2 TYPOLOGY AND THE SCOP OF TYPOLOGICAL LINGUISTICS

By typology, one generally means classification, taxonomy or the study of types. In linguistics, it means the classification of structural types across languages. In other words, it is the study of linguistic patterns by cross-linguistic comparison. The term *typology* was used in linguistics for the first time by von der Gablentz in 1901 (1972 [1901], p. 481. It is now a subject in linguistics concerned with the study of cross-linguistic patterns of languages (Croft, 1990, p. 1).

The study of the classification of languages first emerged in the field of linguistics in the nineteenth century. Schlegel (1772–1829) was among the

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

first linguists who proposed the classification of languages on the basis of morphological characteristics. But typology, as a fully fledged sub-discipline studying cross-linguistic patterns, emerged in 1960s with the study of Joseph Greenberg's implicational universals on morphology and word order (Greenberg, 1963, 1966). With the spread of the work of Greenberg, a number of scholars started using typology as an approach in linguistics to comparing structural patterns across languages. Typology as an approach started influencing the study of languages in the second half of the twentieth century, and was developed as an effective methodology for cross-linguistic analysis comparable to American structuralism or formalism (Croft, 1990). As an approach, typology was close to functionalism and was referred to as the Greenbergian approach.

The current concept of linguistic typology is somewhat different. It is now taken up as a strategy in the field of linguistics (Croft, 1990). This strategy is used as a procedure for comparing and analysing the structures of languages cross-linguistically. Cross-linguistic comparisons are carried out in order to analyse particular grammatical phenomena across various languages and language families. Typological analysis is done on the basis of the survey of the structure of similar grammatical phenomena across languages. In short, typology is now a strategy which is used for the classification of languages into language types. In the current study, this concept of typology is used.

In typological study of languages it is assumed that, in spite of outwardly different structures, languages share a lot of structural similarities. According to Lehmann (1978, p. 1), "languages, in spite of their outward differences, are formed by identical principles".

This idea is supported for a number of reasons. First and foremost is the ability of human beings to master any natural language. A child learns any language which is spoken to him or her. There are other reasons as well: for example, the possibility of inter-linguistic translation and the formation of new languages, e.g. a pidgin. Similarly, the use of highly abbreviated language by logicians, mathematicians and philosophers is also an example of some commonness across languages. The identification of common patterns and principles, and the manifestation of this commonness, are the subject matter of typological linguistics. The subject of the typological linguistics is "what languages of the world are alike?" (Shopen, 2007, p. xiv). Within typological linguistics, languages are classified on the basis of the similarities and differences in their structures.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

There is has been extensive work done in typological linguistics since the 1960s. Among these works, the contribution of the Prague Schools, WALS and the Universal Archive of Konstanz University are worth mentioning. The extensive work of Russian linguists on language typology in the last decades has also been very popular. From Russia, the main contribution in this regard is made by St Petersberg or the Leningrad School of Grammatical Typology (Haspelmath et al., 2001). In the next section, I shall discuss the role of cross-linguistic comparison.

2.2.1 The Role of Cross-linguistic Comparison

The primary characteristic of typological study is cross-linguistic comparison. So the question that arises is: what is the role of cross-linguistic comparison? Being pursued rigorously, the typological study of languages has contributed a lot to the theory of general grammar and language universals. It has helped in a number of ways to analsze the deep structure of human language and is concerned with the study of universal grammar. Therefore, a comment on the role of cross-linguistic comparison is appropriate at this point.

Shibatani and Bynon (1995, p. 3) point out that there are two prevailing concerns of typological linguists:

1. The typological classification of the languages of the world, and
2. The identification of the correlative grammatical properties that define linguistic types.

So the main role of cross-linguistic comparison is the explanation of inter-linguistic phenomena from a number of different perspectives. Human languages have more than one construction for the representation of one linguistic structure. Cross-linguistic comparison enables us to observe any characteristics that are universal across languages. For example, the distribution of articles in English can be compared with that in the French or Japanese languages. Such analysis will reveal that the concept of articles in English is totally different from that in Japanese. On the basis of such analysis, languages can be classified according to their use of articles. One can also make possible generalizations about possible structures of languages. Another example is the difference in the word order patterns of different languages of the world. On the basis of word order pattern analysis, one can classify languages as, for example, SVO (subject + verb + object) or SOV (subject + object + verb).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Cross-linguistic comparison also helps in describing the universal aspects of grammar. With the help of comparison among languages of a specific grammatical phenomenon, one can predict the structure of human language. Cross-linguistic comparison also helps in understanding the nature of human language and thus contributes to the theoretical part of linguistics as well. It has also provided guidelines to assist the investigation of human languages on a broader level. Now linguists work with facts and figures, and typological linguistics has given new dimensions to the historical study of languages. In order to understand the nature of the current study, one must know what the bases of typological linguistics are, and this is discussed in the next section.

2.2.2 The Basis of Linguistic Typology

As discussed earlier, the typological study of languages is concerned with efforts to determine the common features of human languages. For such a study, a very large quantity of data is required. The assembling of materials from many languages of the world to determine the similarities and differences for analysis is the main characteristic of typological studies. Such a study can only be successful when it is known what the common patterns and principles are of human language. For this purpose, typologists first of all identify the common characteristics of human languages. In other words, successful typology requires an exact understanding of the nature of language and its common patterns. Linguistic typology is based on the analysis of such features.

According to Lehmann (1978), among linguistic components, the syntactic component is the most significant as a basis of linguistic typology. For this reason, syntax is used as the central component in typological analysis. Such analysis is based on the syntactic structure and its basic elements, such as subject, verb and object.

There are many reasons why syntax is taken up as the basis for linguistic typology. Firstly, syntax regulates and makes use of different orders of words within a sentence. Such regulating of words is more significant than the morphological and phonological characteristics of those words. Secondly, syntax contributes mainly to the fundamental meanings of the sentence. Thirdly, word order is the easiest way to observe the sequence of units within a sentence (Croft, 1990). Moreover, the *verb* is central in human languages (Lehmann, 1978) and the addition of an object to the verb decides the word order of a particular language. According to Lehmann, a verb has the ability to form sentence solely, and one cannot

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

form any sentence without verb. The simplest sentence is made of a verb or a verb and an object. Every language is classified as either VO or OV on the basis of the pattern of its verb use. Therefore, languages of the world are classified as SOV or SVO according to the patterns of their verb use in the basic syntactic structure.

Thus because of the central role of the syntactic component, it is expected that it will provide the basic criteria for the typological study of human languages. Many linguists have based their typological conclusions on the syntactic component. Major word orders are correlated with particular structural features of languages to enable the classification of languages. Word order is central to any correlation. The correlations between word order and other features of grammar are assumed on the basis of the syntactic component. Other features and constructions of a language are examined in relation to its fundamental syntactic order. Before discussing the subject universal based on word order, one must understand typological study and the study of universals. In the following section, various types of universals and correlations are discussed.

2.2.3 Typological Study and the Study of Universals

When languages are classified according to their structural types, the next step is to find restrictions on certain possible language types. According to Croft:

> The first step beyond typology is the classification of the types and towards the explanation of the cross-linguistic variation that classification describes is the discovering of restrictions on possible language types. (1990, p. 44)

The typological method applied for the discovery of restrictions or constraints on a typical language type is the strongly empiricist method of typological linguistics. When one particular type of language is attested by typologists, they next verify and analyse the grammatical features related to this specific type. This attestation and verification from data consisting of a number of languages helps them in deciding the possible restrictions on one particular type. For example, in studying the structure of relative clauses in different languages, a typologist first will take a large sample of languages. Then he will analyse what types of relative clause are present or absent in those languages. If there are any gaps, he will look for an explanation for such constraints, as per Greenberg's statement:

> Language universals are by their very nature summary statements about the characteristics or tendencies shared of all human speakers. (1966, p. xv)

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

On the basis of such analysis, unrestricted and implicational universals are decided for a particular grammatical feature. Since an implicational universal is the direct concern of the study, unrestricted and implicational universals are discussed with further explanation in the next section.

2.2.4 Unrestricted and Implicational Universals

When a large number of languages are distributed along a single grammatical parameter, such a grammatical feature is called an unrestricted universal as there is no restriction attested to for that particular universal. In other words, all the checked languages have the same grammatical feature and no restriction is found against such a parameter. All the languages studied are of the same parentage or class and all of them share the same grammatical feature. For example, it has been found that all languages have oral vowels (Croft, 1990):

(5)	**Oral Vowels**	**No Oral Vowels**
	X	−

In (5), (X) means that such languages exist and (−) means that no languages of this type exist.

This attestation shows that all languages are of the same class or type in relation to oral vowels and no other class or type was attested. Such features are called unrestricted universals. Greenburg's first universal on the order of subject and object within a sentence is an example of an unrestricted universal. Such a parameter actually allows for the logically possible other type, but only one type is attested. Therefore, unrestricted universals require uniformity among all tested languages: They do not display any variation, as is explicit from our example of oral vowels. In this unrestricted universal, it is found that oral vowels exist within almost all languages of the world. In linguistic theories, unrestricted universals require deeper explanations (Croft, 1990). There are relatively a very small number of unrestricted universals (Croft, 1990, p. 46).

On the other hand, implicational universals may have more than one type of representation, so there may be some restrictions on a specific type of structure. This means that implicational universals limit linguistic variation and do not preclude the existence of other types. For example, there is an implicational universal regarding the order of noun (N) before demonstrative (DEM) and relative clause (R CL).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

> If a language has N before Dem then it has N before R Cl. (Hawkins, 1983, p. 84) [universal no. XI]

In this example of an implicational universal, two independent parameters, N-DEM order and N-R CL order, have been correlated with each other. So, implicational universals state a dependency between two logically independent parameters.

Implicational universals capture a pattern in language variations and are not concerned with the uniformity of one grammatical or structural pattern (uniformity is the concern of unrestricted universals). They are concerned with the dependency or correlation between two different parameters and a large sample is required to determine the range of possible variations. They are paradigmatic examples of typological generalizations. Implicational universals attempt to assemble the independent parameters of grammar together into an integrated whole. They correlate one parametric variation with another. These variations are central to typological analysis, being the lowest-level link between grammatical parameters.

The discovery of implicational correlations was first made by Trubetzkoy (1958), who termed them the typology of vowel systems. He investigated and identified the correlation of "irreversible solidarity" between two grammatical elements. His correlation was:

> The presence of a certain phoneme, X implies the presence of another phoneme, Y in a given phonological system.

Jakobson (1962) later put forward the idea of "Implicational Laws" (*lois d'implication*), which were used by linguists in order to analyse the correlations between two objects in the study of typology. These rules were conspicuously applied by Greenberg in his typological work regarding the combinations of features in the structure of languages. His language universals are two types of cross-linguistic generalization:

1. Theoretically possible structures of languages, and
2. Implicational relations between different features of language. (Greenberg, 1966) Implicational universals and unrestricted universals are both universals, and they are not language-specific. They are, as is clear from their names, intended to be apply as a whole for all languages. An implicational universal is the application of propositional logic to the typological study of languages. This means that one has to apply the full process of

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

propositional logic in the process of defining an implicational universal. The current study is based on the analysis of implicational universals. The topic is chosen from the field of implicational universals based on the correlation between word order and syllable structure. The next section introduces word order typology and the contribution of well-known scholars to this field of study.

2.3 WORD ORDER TYPOLOGY

A number of scholars (e.g., von der Gabelentz, 1901; Skalička, 1935; Lehmann, 1973, 1978; Donegan & Stampe, 1983, 2004; Donegan, 1993; Gil, 1986; Plank, 1998; Fenk-Oczlon & Fenk, 2004; Haspilmath, 2005; Tokizaki & Kuwana, 2012, among others) have stated, described or assumed correlations among different levels of language, such as syntactic, morphological or phonological properties. The typological area which has attracted the most attention has been typological study on the basis of word order. As discussed in section 2.2.2, the basis of the typological study of languages is the syntactic component. Within syntactic patterns, word order is the most easily definable feature. It is related to the specific pattern of the basic syntactic elements in the simple sentences of a language. This pattern is called word order. For example, English and a number of languages are termed SVO (subject + verb + object) languages, as they have this typical order of their essential elements in a simple sentence. For example:

(6)	The child	saw	the train
	S	V	O

Other possible word orders are SOV, VSO, etc. This basic word order has fascinated a number of scholars in the field of typological linguistics so far. They have tried to predict the structure of grammatical constituents of language on the basis of its word order. Thus, the single property of word order has been used as a powerful tool to predict a whole set of type-specific characteristics. Therefore, the specific domain of word order typology attempts to make forecasts about other characteristics of a language.

Word order typology is based on the classification of languages according to their syntactic pattern of verb and object in simple sentences. As discussed earlier, the *verb* has the central role among other elements in the syntactic characterization of a language (e.g., Greenberg, 1963, 1966;

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Lehmann, 1971, 1978; Vennemann, 1972; Hawkins, 1983). Furthermore, the syntactic construction of the verb in relation to its *object* is considered the most fundamental aspect in the syntactic classification of languages. Some linguists believe that these two elements (verb and object) are enough to define the basic word order of a language. Some linguists also add the *subject* to this list and focus on the combination of three basic syntactic elements, namely subject, verb and object.

Word order typology, for the first time in the history of typological linguistics, discovered a new area for research with the work of Greenberg in 1963. According to him, the word order of a given language has the potential to predict a great deal about the structure of that language. Once the basic order of the major constituents is known, one can predict a great deal about the treatment of other features. For example, Greenberg (1966) identified a number of universals on the basis of his data. Some of these are mentioned below.

When the dominant word order of the language is VSO, the language has prepositions (Greenberg's universal no. 3) and the genitive follows the governing noun (Greenberg's universal no. 2). Similarly, if the question particles which are specified in position by reference to a particular word in the structure do not occur (Greenberg's universal no. 10), the language will always put interrogative words or phrases first in interrogative word questions (Greenberg's universal no. 12), and the inflected auxiliary will precede the main verb (Greenberg's universal no. 16). Another universal about such a language was that the adjective will come after the noun (Greenberg's universal no. 17).

Greenberg predicted some specific features of syntax on the basis of word order. But in 1973, Lehmann investigated word order typology with another dimension. He concluded both morphological and phonological implications on the basis of word order typology. Then he further investigated word order typology and contributed to the type characteristics of individual languages classified in terms of word order in 1978. The Cologne School and the St Petersberg/Leningrad School of typological linguistics also contributed to the study of language structures in this regard. They considered word order in typology as a highly sophisticated framework for capturing cross-linguistic generalizations about language types, essentially following the Gabelentzian ideal of being able to predict the complete structure of a language on the basis of a handful of typological properties, or even a single property (von der Gabelentz, 1901).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Greenberg's word order universals have had considerable impact on the typological study of languages. According to him, languages have the following three types of word order: verb + subject + object (VSO), subject + verb + object (SVO), or subject + object + verb (SOV). He studied a sample of 30 languages, and on the basis of his data he proposed 45 universal statements; 25 out of these 45 universals are related to word order. Vennemann (1972), in order to provide a theoretical perspective, reformulated and explained the Greenbergian universals. He redefined the basic word order on the basis of two basic verb positions. Thus, he collapsed VSO and SVO into one VO type. Vennemann (1974) also put forward the theoretical notion of the natural serialization principle (NSP) to provide the theoretical and historical background to Greenberg's word order universals. Lehmann (1971, 1973, 1978) also carried out research on word order typology and considered word order the area of greatest potential for typological work.

In his article on the structural principle of language and its implications, Lehmann (1973) discussed various morphological, syntactic and phonological implications on the basis of word order pattern. He concluded from his investigation that languages with VO word order have their sentence qualifier markers, such as interrogative or negative markers, before verbs, while in the case of OV languages these sentence markers are placed after verbs. In the same article, he identified that VO languages are inflectional, while the morphological structure of OV languages is agglutinative. Lehmann (1973) is of the view that word order typology is very important to the understanding of the linguistic structures of languages. His overall conclusion shows that word order typology is very helpful for the formation of correlations between the order of a language and other structures of that language. In Table 2.1 I have listed some of Lehmann's important correlations and their implications on the basis of word order typology:

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2.1: Implications/Correlations Identified by Lehmann (1973)

Structure	VO Languages	OV Languages
Syntactic	Syntactic quantifiers or markers are placed before verbs	Syntactic quantifiers or markers are placed after verbs
Expected String	Negative causative + V + O	O + V + causative negative
Morphological	Inflectional structure	Agglutinative structure
Examples	Classical Hebrew, Portuguese, Squamish	Japanese, Turkish, Quechua, Sanketi
Phonological	Complex syllable structure Closed, ends with consonants	Simple syllable structure Ends in vowels
Syllable Structure	CCVCC clustering at both sides	(C) CV

As Table 2.1 makes explicit, Lehmann was determined to generalize structural implications and correlations for almost every linguistic level based on the pattern of the basic syntactic element of word order. He gives a complete list of contrasting syntactic patterns and characteristic morphological structures which are correlated with the language type. For such generalizations, Lehmann (1978) has also tried to explain the morphological implications on the basis of a single syntactic rule. He tries to explore the phenomenology of language by looking into syntactic features of various types of languages on the basis of their word orders. The main types of languages discussed by him in 1978 were SVO (English), SOV (Japanese) and VSO (Eastern Island). He explored the typological characteristics of these languages for the following nine syntactic patterns with their sub-classes (Lehmann, 1978, p. 16):

I. Simple Clauses
II. Nominal Modifiers
III. Verbal Modifiers
IV. Sentence Adverbials
V. Compound and Complex Sentences
VI. Marking
VII. Grammatical Processes
VIII. Morphological Characteristics
IX. Phonological Characteristics

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

He says that the agglutinative structure of OV languages is significant in attempting to understand the nature of such languages. In the current study, I am concerned with the correlations between phonological patterns and syntactic features of language. The syllable structure of the Pashto language is not in accordance with Lehmannn's above-mentioned universals. The problem is that Pashto is SOV, but it has a number of complex syllable structures. I will consider this point in more detail in the following section, where I will analyse Lehmann's work on the correlation between word order and syllable structure.

In addition to linguists such as Greenberg (1963) and Lehmann (1973), there were a number of other scholars who used this single syntactic process (word order) as a typological or correlational source for a number of other features of language. Some even went so far as to say that the features of word order could be used to classify the linguistic characteristics associated with specific types of languages, while others studied word order diachronically and in connection with an analysis of linguistic change.

Donegan and Stampe (1983) and subsequently Donegan (1993) also contributed to the field of word order typology. They carried out their research on Munda and Mon-Khmer languages, demonstrating their oppositeness at every level on basis of word order. Their analysis is shown in Table 2.2.

It is evident from Table 2.2 that Donegan and Stampe pointed out many contrastive characteristics of two different language families on the basis of their classifications. In other words, they re-affirmed the already existent notion of typological studies that languages with SOV (also called dependent-head) order are prone to have agglutinative morphological structure and simple syllable structures, while SVO (head-dependent) languages have inflectional morphological structure and the possibility of complex syllable structure. What they showed was that the clear polar oppositeness of the languages extended to almost every level, from syntax to morphology and phonology. They included new aspects of linguistic level such as tone, register, vocalism, timing and even phrase accent in their discussion of language classification. Their idea of the correlation between word order and syllable structure will be further discussed in the following section, in the evaluation of the hypotheses of the current study. In subsequent works, Derbyshire and Pullum (1979) and Hawkins (1983) also contributed to the field of word order typology by discussing word order universals.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2.2: Comparison of Munda and Mon-Khmer Languages (Donegan & Stampe, 1983)

Level	Munda	Mon-Khmer
Phrase accent	Falling (initial)	Rising (final)
Word order	SOV dependent-head (AN, post-positions); variable	SVO head-dependent (NA, pre-positions); rigid
Syntax	Case-verb agreement	Analytic
Word canon	Trochaic, dactylic	Iambic, monosyllabic
Morphology	Agglutinative, suffixes, polysynthetic	Fusional, prefixing or isolating
Timing	Isomoric	Isoaccentual
Syllable canon	(C)V(C)	(C)V or (C)(C)V(C)(C)
Consonantism	Stable, geminate clusters	Shifting, tonogenetic
Tone/register	Level tone (Korku only)	Contour tones, register
Vocalism	Stable, monophthongal, harmonic	Shifting, diphthongal, reductive

Donegan and Stampe (2004) more recently calculated the diachronic change within the same language families. They compared the characteristics of South-Asian (Munda) languages with South-East Asian (Mon-Khmer) languages, both being branches of the Austroasiatic language family, to analyse their most divergent structural patterns. Table 2.3 shows the polarization in the languages discussed by them.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2.3: Polarizations in Munda vs. Mon-Khmer (Donegan and Stampe, 2004)

	Munda	**Mon-Khmer**
Grammar	Synthetic	Analytic
Word Order	Head-last: OV, post-pos.	Head-first: VO, pre-pos.
Phrases	Falling (initial)	Rising (final)
Words	Falling (trochaic)	Rising (iambic /monosyllabic)
Affixation	Pre-/infixing, suffixing	Pre-/infixing or isolating
Timing	Isosyllabic/isomoric	Isoaccentual
Fusion	Agglutinative	Fusional
Syllables	(C)V(C)	(C)V + (C)V(:/V)(C)
Consonants	Stable/assimilative	Shifting/dissimilative
Tonality	Level (rare)	Contour (common)
Vowels	Harmonizing/stable	Reducing/diphthongizing

Donegan and Stampe (2004) try to explain the structural opposition of languages from the same family, providing a focal point for discussion related to the historical development of languages. Thus they take up the issue raised by Grierson (1904, p. 2) in "Linguistic Survey of India" of whether such divergent languages could be historically related to each other. They raise a number of questions related to divergent typologies within language families and comment that such issues of contrastive features pervading every level of structure down to phonetics and prosody are the result of drift and shift among various features of languages.

More recently, Fenk-Oczlon and Fenk (2004) studied the cross-linguistic correlations between size of syllables, number of cases, adposition order and other features As OV word order was already correlated with other linguistic features, such as a very small number of phonemes per syllable and agglutinative morphological language structure, they analyse these interdependencies between different levels of language for further explanations. Thus the already existing correlation between OV order and simple syllable structure is re-investigated from a new perspective. They found that languages with OV order have simple syllable structures and

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

languages with VO arrangement, in contrast, have complex syllable structures. Moreover, the morphological structure of OV languages is, again, found to be agglutinative and associated with post-positions. Thus they also found other hypotheses of Greenberg (1966) to be correct, e.g. VO languages tend to have not only fusional or inflectional morphological structure but also complex syllable structure. They proved that it is the natural tendency of human language to keep abreast with the optimization of the interaction between articulatory and cognitive systems through natural self-regulatory processes. Thus all languages, irrespective of their typological character, are bound to adopt such constraints. Their explanation also proved the idea of "natural" tendency, or the "easy for the human being" scope of language.

Summing up the points discussed in the current chapter so far, one can conclude that:

- Typology is the strategy used for the classification of languages into language types.
- Cross-linguistic comparisons provide the universal characteristics of the grammar of human languages.
- Typological studies of human languages are mostly based on the syntactic component.
- The possible restrictions on the verified structure of grammatical features are realized through various types of universals.
- Unrestricted universals show the uniformity among all attested languages, while implicational universals, being central to typological analysis, capture structural variation by stating a dependency between two logically independent parameters.
- Word order plays a central role in typological study, and a number of structural features are correlated with the word order of languages.

In section 2.4, I take up the subject correlation between syllable structure and word order, which is the direct subject of the current study.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.4 THE CORRELATION BETWEEN SYLLABLE STRUCTURE AND WORD ORDER

In the previous section, it was made clear that almost all typological studies are based on the word order of languages. A number of studies were mentioned which have emphasized the correlations between syntax and other levels of language. The level at which such studies are most readily undertaken is the phonological level of language. In the light of this, the correlation between word order and syllable structure has been the area most researched by linguists and typologists.

Starting from von der Gabelentz (1901) and going through various stages and levels, Lehmann (1973, 1978), Donegan and Stampe (1983, 2004), Donegan (1993), Gil (1986), Plank (1998), Fenk-Oczlon and Fenk (2004), Haspilmath (2005) and Tokizaki and Kuwana (2012), among others, describe the correlation between word order and syllable structure. They have tried to show this correlation through a number of studies demonstrating that languages with SOV order tend to have simple syllable structure, while in the case of SVO order, syllable structure tends to be complex. The Universal Archives of the Universitate Constanz and the universals of WALS also demonstrate the same phenomena. In the current section, I plan to present a detailed overview of these studies. But before reviewing the previous work on the correlation between word order and syllable structure, it is appropriate to define the basic terms *word order* and *syllable structure* in order to ensure a precise understanding of the correlation under discussion.

2.4.1 Basic Word Order

The discussion of basic word order is an important point in the study of typological linguistics. Linguists define the basic word order (BWO), or simply the word order (WO), in consideration of the importance of the combination of basic constituents related to syntactic characterization. In simple and straightforward terms, the BWO of a language is the order of the most basic syntactic elements within a simple declarative sentence of that language. In the current section, I review the ideas of well-known linguists in relation to the notion of BWO and the possible types of BWO in languages of the world.

As established earlier, linguists (e.g. Greenberg, 1963, 1966; Lehmann, 1971, 1973, 1978; Vennemann, 1972; Hawkins, 1983; see Section 2.3 for further detail) agree that the *verb* has the central role among other

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

elements in the syntactic characterization of a language. Then, the syntactic construction of the verb in regard to its *object* is considered the most fundamental in the syntactic classification of languages. Some linguists believe that these two elements (verb and object) are enough to define the BWO. But there are others who think that in conjunction with these two elements, the *subject* is also important. Therefore, the BWO of a language is defined on the basis of the relationship between these three fundamental elements in a simple sentence of the language. Let me show the views of some experts on this point.

Starting from Greenberg (1966), whose WO niversals had considerable impact on the typological study of languages, languages have the following three types of BWO:

(7)
a. Verb + subject + object (VSO)
b. Subject + verb + object (SVO)
c. Subject + object + verb (SOV)

Hawkins (1983) calls this division of language types the "Greenbergian Trichotomy". According to this division, these three types of BWO are the most common orders. He based his linguistic universals on these types. Lehmann (1971) mentions two possible BWOs, the VO (verb + object) and OV (object + verb) and considers the order of these two fundamental elements (i.e., the verb in combination with its object). He divides languages on the basis of these two BWO types; VO languages (e.g., English, European and Semitic languages) and OV languages (e.g., Japanese, Turkish and the Dravidian languages of India). This division of Lehmann (1971) shows the typological significance of verb position in the study of languages. Based on the ideas of Lehmann, Vennemann (1972), in order to provide a theoretical perspective on Greenbergian universals, reformulates the "Greenbergian Trichotomy" into two possible order types. He describes two basic verb positions (as discussed above, the VO and OV). Thus collapsing the VSO and SVO word orders into one VO type, he gives his theoretical notion of the natural serialization principle (NSP). The NSP has contributed a great deal in defining the basic languages into two types and providing a logical explanation for BWO change. If S is added to V and O, it makes different combinations of S, V and O; subsequently, in his book on syntactic typology, Lehmann (1978) mentions three highly consistent prominent types of languages: SOV (Japanese), VSO (Eastern Island) and SVO (English). He also talks about the relatively infrequent VOS language type.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Apart from the above types, other logically possible verb positions have been discovered and studied by linguists. Derbyshire and Pullum (1979) reported that OVS and OSV languages also exist. Hawkins (1983) gives five possible BWO types. He considered three main types – SOV (e.g. Japanese), SVO (e.g. English) and VSO (e.g. Welsh) – and two relatively rare types: VOS (e.g. Malagasay) and OVS (e.g. Hixkaryana). Dryer (2005), in WALS (Chapter 83), divides languages into three possible orders: VO languages, OV languages and languages with no dominant order. The third type of language is those in which the order of V and O is determined by grammaticality, etc. More recently, Brody (2009) gives six logically possible orders of S, V and O (SVO, SOV, OVS, OSV, VOS and VSO).

In the light of the above discussion, it seems appropriate to adopt the definition of Lehmann (1971, 1978) and Vennemann (1972, 1974, 1975), among others, by considering the combination of verb and object as the fundamental order for the subject languages. The first reason for doing so is that the pattern of WO of all languages can easily be divided on the basis of these two basic syntactic elements and it is easy to relate all languages to these two formations. Secondly, the order of my subject language (Pashto) is OV, and the language which I will take for comparison in Chapter 6 (i.e., English) has VO order. So, I have the representation of both types of languages. Even if subject is included in our basic syntactic elements, the order of the languages will be SOV and SVO respectively. Finally, the correlations relevant to the current study are related to either VO or OV and this provides enough viability for my research. The section below discusses the categorization of syllable structure.

2.4.2 The Categorization of Syllable Structure

By syllable structure, in this study in the work of other linguists who have worked on the topic, I mean the possible patterns or categories of syllable structure found in various types of languages. This is basically related to the complexity or possible consonant clusters within the syllable structure of a language. According to Maddieson (2005), languages can be divided into three categories on the basis of syllable complexity, namely simple, moderately complex and complex. Further detail is given in Table 2.4 below:

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2.4 The Categories of Syllable Structure (Maddieson, 2005)

S. No.	Category	Structure	Examples
1	Simple	CV	Hawaiian and Mba languages (Adamawa-Ubangian, Niger-Congo, Democratic Republic of Congo)
		(C)V	Fijian, Igbo (Niger-Congo, Nigeria) and Yareba (Yareban, Papua New Guinea
2	Moderately Complex	CVC	
		CC2V	C2 = liquids (r/l) or glides (w/j)
		CC2VC	C2 = w in Darai (Indo-Aryan; Nepal)
3	Complex	(C)(C)(C)V(C)(C)(C)(C)	English

The division of syllable structure on the basis of complexity has been further modified by linguists such as Plank (2009) and Tokizaki and Kuwana (2012). According to them, the above division is problematic in determining the correlation between word order and syllable structure. Tokizaki and Kuwana (2012) suggest that CV(C) should also be considered simple rather than moderately complex. They take the example of Japanese, which has one variety of consonants at coda level, and call it a simple syllable structure. In the case of the subject language, Pashto, both these categorizations work. The syllable structure (for examples see Section 2.6.1) of Pashto is complex under both definitions.

2.5 CORRELATION BETWEEN SYLLABLE STRUCTURE AND WORD ORDER

A number of previous works (e.g. Lehmann, 1973, 1978; Donegan & Stampe, 1983, 2004; Donegan, 1993; Gil, 1986; Plank, 1998; Fenk-Oczlon & Fenk, 2004; Haspilmath, 2005; Tokizaki & Kuwana, 2012, among others) show the correlation between word order and syllable structure.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Plank (1998) and Tokizaki and Kuwana (2012) have presented an overview of them. They show that languages with OV (or SOV, more precisely) order tend to have simple ((C)V) syllable structure. In the current section, I am going to discuss the major contributions of linguists on the subject of the correlation or correlations between word order and syllable structure.

2.5.1 Lehmann (1973)

Among his other morphological, syntactic and phonological correlations and implications on the basis of word order, Lehmann (1973) also makes typological conclusions about the correlation of word order and syllable structure. I briefly analysed his arguments in the last section and listed a number of important points in Table 2.1. Here, I take up the question of correlation in further detail.

Lehmann (1973) says that VO languages have complex syllable structure, while in contrast languages with OV word order have simple syllable structure. Table 2.5 shows the syllabic implications of word order linearity.

Table 2.5: The Correlation of Word Order and Syllable Structure (Lehmann, 1973)

VO Languages	OV Languages
Complex Syllable Structure	Simple Syllable Structure
Syllables are closed, end in consonants	End in vowels, no consonant cluster possible at the end
CCCVCCC	CV
Consonants clusters possible at both side of the nucleus	Mostly, one consonant comes at the beginning

As can be seen from Table 2.5, Lehmann draws phonological implications for the syllable structure of a language type on the basis of word order. He gives the examples of Japanese, Turkish, Quechua and Sanketi for OV languages and Classical Hebrew, Portuguese and Squamish for VO languages.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.5.2 Donegan and Stampe (1983)

The comparison of Munda and Mon-Khmer languages by Donegan and Stampe (1983, 2004) and Donegan (1993) also contributes to the discussion of correlation between word order and syllable structure. The detail of their analysis was shown in Table 2.2 in the last section. They are of the view that the syllable structure of SOV (dependent-head) languages is likely to be simple, i.e. (C)V. In contrast, languages with SVO (head-dependent) order have complex syllable structure, i.e. (C)(C)V(G)(C). This conclusion reaffirms the conclusions of typological linguists already discussed.

Table 2.6: Brief Comparison of Munda and Mon-Khmer (Donegan & Stampe, 1983)

	MUNDA	**MON-KHMER**
Word order	SOV dependent-head	SVO head-dependent
Syllable canon	(C)V(C)	(C)(C)V(G)(C)

Table 2.6 illustrates that languages with SOV (dependent-head) order have simple syllable structures and languages with SVO (head-dependent) order maintain complex syllable structure.

The more recent analysis by Donegan and Stampe (2004) of languages from the same family, comparing South Asian (Munda) languages with South-East Asian (Mon-Khmer) languages, shows that languages from the same family can also display divergent features of structural patterns. Such a change in the structure of languages is also evidence of the diachronic or historical development of languages, reaffirming the notion of languages as living entities. These branches of the Austroasiatic language family were discussed in the previous section. The main point relevant to the current section is the relationship of syllable structure and word order. Once again in their 2004 paper, Donegan and Stampe maintain that languages with OV word order have simple syllable structure: (C)V or the possibility of (C) at coda level. Their results were based on the analysis of Austroasiatic languages and thus add another point to our understanding of correlation. Their comments are not different from those found in previous works discussed so far, except for the fact that they are interested in the

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

overall divergent structure of languages from the same family. Their main focus was on showing that their subject languages were altogether different despite being from the same family.

2.5.3 David Gil (1986)

Gil (1986) tests the same correlation between syllable structure and word order within a sample of 170 languages. The data for these 170 languages, taken from the Stanford Phonology Archives and the UCLA Phonological Segment Inventory Database, is used to show that the syllable structure of OV languages is simple. On the basis of his data, Gil concludes that the average number of segments in the syllable structure templates is lower for OV languages: (SOV) 4.04 < (SVO) 4.93. Although the study is based on very limited data, Gil claims that this tendency is found in SOV languages and suggests further detailed study with a greater quantity of data on the topic. However, Tokizaki and Kuwana (2012) say that the result is not very convincing because the difference was less than 0.9 (0.89). They added to the number of sample languages, providing a more authentic basis in the data for identifying the implicational universals shown below in Tables 2.7 and 2.8.

2.5.4 Data from the Universal Archives

The Universals Archives[3] lists 2029 *implicational universals* regarding various structures of languages. These implicational universals are correlations which are decided on the basis of the typology and cross-linguistic variations of the languages. Universals which are directly related to our study are correlations 196 and 207. The detail of these universals is given below as Tables 2.7 and 2.8, with comments from Plank as reported in Lehmann (1973).

[3] http://typo.uni-konstanz.de/archive/intro/index.php (accessed on 20 July 2010)

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2.7 Universal Number 196 of the Universal Archives

Original	OV languages tend to have simple syllable structure
Standardized	IF basic order is OV, THEN syllable structure is simple (tending towards CV)
Formula	OV ⇒ simple syllable structure
Comments	Languages with flexive morphology (which tend to be OV) tend to have the ends of syllables closed, with consonant clusters occurring in this position as freely as in initial position (Lehmann, 1973, p. 61).

This implicational relation is the case in some of the SOV languages of the world, for example Turkish (Turkic, Altaic), Japanese (Japanese-Ryukyuan), Ijo (Niger-Congo), Yareba (Papua New Guinea) and Warao (Venezuela). The syllable structure of these languages is simple (CV).

Table 2.8 Universal Number 207 of the Universal Archives

Original	VO languages tend to have complex syllable structure
Standardized	IF basic order is VO, THEN syllable structure is complex (permitting initial and final consonant clusters).
Formula	VO ⇒structure complex syllable
Comments	-

In the above two universals, it is implied by the experts of language typology that there will be considerable difference between SOV and SVO languages based on their syllable complexity. Both of these universals serve as the basis for further studies on the correlations between syllable structure and word order (e.g. Tokizaki, 2011).

2.5.5 Evidence from WALS

A number of researchers discuss the correlation between syllable structure and SVO order using the data from WALS, which includes 2561 languages. They use the categories of languages according to the definition of syllable structures given by Maddieson (2005) in WALS (Chapter 12). Plank (1998) and Tokizaki and Kuwana (2012) concluded

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

that such correlations, in which the simple syllable structure of SOV word order can be attested, are certainly possible. They conclude from their studies that there exists a possibility of such correlations because they believe that phonology and morphology tend to be correlated with syntax.

The same correlation is tested by Tokizaki and Kuwana (2012) using data from WALS. They studied a sample of 359 languages with both OV and VO orders. They also used tMaddieson's definition (2005) for syllable structure as given in Section 4.2.

Tokizaki and Kuwana (2012) are of the view that the three types (simple, moderately complex and complex) are quite effective in classifying languages on the basis of syllabic complexity. However, this categorization is not fine enough to enable the correlation of the subject, as also pointed out by Plank (2009). Tokizaki and Kuwana (2012) combined the classification of Maddieson (2005, chapter 12) for syllable structure and the definition of Dryer (2005, chapter 83) for word order, and applied them to the data available from WALS. The results showed that only 18 out of 171 SOV languages had simple syllable structure, while 47 out of 165 SVO languages had complex syllable structure.

Following the remarks of Dryer (2005) that typological work should be based on the number of genera rather than languages, Tokizaki and Kuwana (2012) ultimately divide languages into their generic groups. By genera, Dryer meant the grouping of languages on the basis of similarity and uncontroversial genetic relatedness (Dryer, 1992, p. 84). According to Toziaki and Kuwana (2012), in order to control the most severe genetic bias, typologists should count genera rather than languages. Thus, on the basis of counting the genera instead of the number of languages, they prove with the help of data from WALS that OV languages have more simple consonant clusters than those of VO languages. This is further discussed separately at the end of this section.

2.5.6 Fenk-Oczlon and Fenk (2004)

Recently, Fenk-Oczlon and Fenk (2004) have analysed the correlation between syllable structure and word order. Their research concerns the cross-linguistic correlations between size of syllables, number of cases, adposition order, etc. They prove that OV word order is rightly correlated with other linguistic features, such as a very small number of phonemes per syllable and agglutinative morphological language structure. They analyse these interdependencies between different levels of language for further explanations. Thus the already existing correlation between OV

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

order and simple syllable structure was re- investigated from a new perspective.

Their results (2005, p. 4) show that:

a. The more syllables per clause, the fewer phonemes per syllable,
b. The more syllables per word, the fewer phonemes per syllable,
c. The more syllables per clause, the more syllables per word, and
d. The more words per clause, the fewer syllables per word.

They further state:

> OV order is not only associated with less complex syllables, but also with a tendency to postpositions (e.g. Greenberg 1966 and our sample, where 72% of the postpositional languages showed OV order and 90% of the prepositional languages VO order.) (2005, p. 4).

They calculate from their results that languages with OV order have simple syllable structure and vice versa. The view that OV languages tend to be agglutinative and associated with post-positions is also found to be correct by them. They also found support for the hypothesis (Greenberg, 1966) that VO languages tend not to have only fusional or inflectional morphological structure but also complex syllable structure.

2.5.7 Tokizaki and Kuwana (2012)

Tokizaki and Kuwana (2012) discuss the same correlation with the help of data from WALS and universals from the Universals Archives. They suggest a number of orientations for the reconsideration of syllable structure:

- Geographical gradation of the variety of word-final consonants,
- The simplification of syllable structure within words and between words,
- Head-complement orders,
- The gradation of coda inventory, and
- The fine classification of syllable complexity.

On the basis of their suggestions, Tokizaki and Kuwana (2012) analyse the correlation between syllable structure in OV languages and find that OV languages do have simple syllable structure (p. 6).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Because the three-way distinction given by Maddieson is unable to prove the correlation between syllable structure and word order, as suggested by various scholars as mentioned above, Tokizaki and Kuwana reconsider the syllable structure. They suggest that syllabic complexity should be decided on the basis of the number and variety of coda consonants. They follow Hashimoto's (1978) argument that both coda and tone are simpler in the OV languages of North Asia (e.g. Manchu) than in the VO languages of South Asia (e.g. Thai, having the most complex syllables in the region). Thus, Tokizaki and Kuwana redefined the complexity level of syllable structure. The main problem for them is with the moderately complex category of Maddieson's classification of the syllable structure. According to the new parameter of syllable structure put forward by Tokizaki and Kuwana (2012), if the syllable structure of one language is (C)V(n), it should be considered quite close to simple syllable structure, i.e. (C)V. In such a language, only one variety of consonant at coda level is possible.

Following the arguments made by Dryer (1992, 2005), Tokizaki and Kuwana (2012) count the genera rather than the languages. They count the number of segments in a syllable on the basis of language grouping and calculat the average value of head- complement orders. Thus, on the basis of their results, they show that there is a tendency for the value of the head-complement to increase with an increase in the number of segments. Finally, they conclude that there is a fine correlation between syllable structure and head- complement order on the basis of the structures of VO/OV.

Tokizaki and Kuwana (2012) also hypothesize the same correlation on the basis of the calculation of coda inventory in OV languages. They list a number of languages with constant OV order and count the possible coda consonants. Their results show that OV languages have more limited variety of consonants in coda than VO languages. For example, they argue that SOV languages (e.g., Japanese) tend to have nasals, liquids and some voiceless obstruents at coda. On the other hand, VO languages may have a full variety of obstruents and glides at this position. From this, they suggest that OV languages may have simpler syllable structures.

Tokizaki and Kuwana (2012) prove their hypothesis by exploring the nature of syllable structure alongside some other phonological characteristics. They say that languages with SOV order have more phonological changes, such as epenthesis and deletion. According to them, epenthesis and deletion simplify the structure of their syllables by reducing them. For example, consonant clusters are avoided through

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

processes such as epenthesis, deletion and coalescence, such as in the following cases (Tokizaki & Kuwana, 2012):

1. Epenthesis changes CC cluster into CVC
2. Deletion reduces CC to C
3. Coalescence also shapes CC into C.

They give examples from a number of languages, such as Persian, Japanese, Basque, Korean and Kannada, some of which are shown below:

Epenthesis:

i. Nambiqra: *w"aklsú→w'akəlisú* alligator
ii. Persian: *drožki* (Russian)→*doroške* droshky
iii. Basque: *libru* (Latin)→*liburu* book
iv. Kannada: *magal* (Old)→*magalu* daughter
v. Japanese: *drink* (English)→*dorinku*
vi. Korean: *text* (English)→ *teyksuthu*

In these examples, the consonant clusters are reduced to one cluster by the phonological process of epenthesis of vowels.

Deletion/assimilation:

i. Basque: *gloria→loria* glory
ii. Basque: *ecclesia →eliza* church

The consonants are deleted at word-initial and word-medial positions.

Coalescence:

i. Korean: *stʌr→ttal* daughter
ii. Korean: *skur→kkul* honey

The process of coalescence in the Korean language, in which the consonant clusters change into reinforced consonants, is also shown.

According to Tokizaki and Kuwana (2012), such processes are very rare in the case of VO languages. VO languages do not simplify their consonant clusters with epentheses and the processes of deletion and coalescence. To support this argument they give the example of English. On the basis of their analyses, Tokizaki and Kuwana predict that OV languages will have simple consonant clusters and VO languages complex clusters.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Finally, Tokizaki and Kuwana try to prove that OV languages have simple consonant clusters through the analysis of limited consonants between words. They argue that consonant clusters between as well as within words are simpler in OV languages than in VO languages. They give the example of Korean and Moghol (Mongolic) languages having a number of nouns with coda consonants. Such particles are attached to nouns in order to show their cases. Their examples are given below.

Korean:

a. Nominative: *-i/ka*

b. Accusative: *-ul/lul*

c. Instrumental: *-ulo/lo*

d. Comitative: *-kwa/wa*

e. Vocative: *-a/ya*

f. Topic: *-un/nun*

Every case from the above list has two forms, one starting with a vowel and another with a consonant. Thus this language avoids consonant clusters by adding a consonant to a noun ending with a vowel, and adding a vowel starting form to a noun ending with a consonant. In both cases, the syllable template will remain the same (CV.CV). On the basis of this morph- phonology examination of Korean language, they claim that SOV languages have simple syllable structure.

From such examples, Tokizaki and Kuwana prove that OV languages do have simple syllable structure if one analyses the data by considering the geographical gradation of the variety of word-final consonants, and the nature of the classification of syllable complexity; by examining the variety of coda inventory; and by simplifying the syllable structure within words and between words. Next, they look for an explanation for why OV languages have simple syllable structure and fewer consonant clusters between and within words. They argue that left-branching structures have weaker juncture between its constituents than the structures of right-branching languages. This juncture is asymmetrical between the left and right branching of the structure. Taking data from Japanese, Korean, Dutch and German lnguages, they argue that this asymmetry is the result

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

of phonological and morpho-branching of the structure. From these examples of inter-fixation and translating the left-/right-branching asymmetry into prosodic phrases, Tokizaki and Kuwana prove that the juncture in right-branching (VO) languages is stronger than that in left-branching (OV) languages.

By applying the asymmetry found in prosodic phrases to word order and syllable structure, they show a marked difference in the structure of languages. They say that OV languages permit CV, as it fits the weaker juncture found in such languages and thus makes no consonant clusters. On the other hand, VO is permitted to have complex syllable structure because the boundaries between words (verbs and nouns) are strongly demarcated. These three points are interesting and provide a potential basis for discussion for analysis of the correlation which is the focus of this study. This discussion, with proper evidence from the subject language, is taken up in Chapter 5.

The final analysis of Tokizaki and Kuwana is that such points limit consonant clusters in OV languages at the within-words and between-words levels. All this is predicted by the notion that left-branching languages have weaker juncture than right-branching languages. In a more recent work, Tokizaki (2011) has tried to prove the correlation between word order and phonology in dialects and mutually related languages.

2.6 OVERVIEW OF PREVIOUS STUDIES

I started by reviewing previous scholarship on the subject of the correlation between syllable structure and word order, from the word order universals of Greenberg (1963, 1966) to the recent contribution of linguists like Tokizaki (2011). All of them are of the view that syllable structure is directly or indirectly related to the word order of a language. Most of them see this correlation as a tiny realization of the huge bulk of the strong relationship or interface between phonology and syntax (particularly Plank, 1998; Tokizaki & Kuwana, 2012). Summing up, with respect to the correlation between syllable structure and word order according to the studies discussed in the present section, I conclude with the following points:

- The syllable structure of a language is determined by its word order.
- If the word order of a language is SOV, the syllable structure of the language will be simple, i.e. (C)V.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

- The same correlation also implies that languages with SVO will have complex syllable structures, i.e. (C)(C)(C)V(C)(C)(C).

Thus, researchers such as Plan (1998) and Tokizaki and Kuwana (2012) observe that the possibility of correlations and interfaces between two or more linguistic levels may be a potential area for exploration and for proving the truth of "Gabelentz's ideal".

But the study of grammar may not be so simple. This point is proved by the violation of the above concluding points in our subject language. Pashto violates both the above correlations and provides counter-examples to them, and thus, in the language of logic, provides the "proof of negation" on this topic. In the following section, I will analyse the case of Pashto in this regard. Having discussed the issue, the feasibility of the current study will be considered.

2.7 THE CASE OF PASHTO AND THE VIABILITY OF THE CURRENT STUDY

As discussed earlier and realized at every relevant point, I argue that the Pashto language violates the discussed correlations and provides enough data to prompt the re-examination of their reliability. This language has SOV order, but it does have a number of complex syllabic templates. In order to address the exact problem, the case of Pashto must be introduced. In this section, I will discuss how the case of Pashto is different from other languages, creating a very solid basis for the current study. Let me first provide more detail about the case of Pashto

2.7.1 The Case of Pashto

The basic word order of the Pashto language is SOV, as illustrated in the following section:[4]:

(8) ba:ʈʃa kita:b wa:i:
Bacha book read-do
Bacha reads a book

[4] The sentences within this study are taken from Yousufzai, the standard dialect of Pashto. It is the native dialect of the researcher and is considered the standard dialect of the language (see Section1 1.6.3 for detail).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(9) mor maʃu:m xkolkʈo
Mother baby kiss-did
Mother kissed the baby

(10) gwəl pətəŋ əluzəwi:
Gwal kite fly-do
Gwal is flying the kite

Roberts (2000) reports Pashto to be a rigidly head-final language across its lexical categories, and the same is clear from the above examples. He has also discussed the possibility of OSV order in some contexts and if prosodic conditions are fulfilled. He further argues that the SOV order is rarely changed into OSV, and only in such cases where it is demanded by the context and there is a heavy pause after the first word, i.e. the object. Such order is subject to cliticization and topicalization. Furthermore, this freedom is only possible in the past tense, and is not possible in other tenses.

From the above examples, Pashto can easily be concluded to be an SOV language. Now, according to the implicational universals provided by the studies discussed in the previous section, SOV languages should have simple syllable structure. Since Pashto has SOV word order, it must have simple syllable structure. But this is not the case. Now let me show the possible syllable structure of the language, as proved by its syllabic templates.

The syllable structure in the Pashto language can easily be reported to be a complex one. Based on our definitions of the three categories of syllable structures in Section 2.4.2, one can argue that Pashto syllable structure is complex. Pashto is very rich in its syllable structure. There are various types of consonant clusters which can form individual words in the language. According to Tegay and Robson (1996), Pashto is remarkable for the possibility of different consonant clusters within its syllables (see also Morgenstierne, 1926, 1940, among others). For example, English has about 35 possible combinations of consonant clusters for two-consonant clusters at onset level, while in Pashto this number goes up to 100 combinations (Tegay & Robson, 1996). They also maintain that Pashto syllables may consist of at least one vowel and as many as three consonants at onset and two consonants at coda level. Levi (2004) shows two aspects of Pashto phonology, namely the presence of many consonant clusters (including reverence sonority) and the features of glides (palatal /j/ and labial /w/). She is of the view that “Not only does Pashto contain

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

clusters … but also several reverse sonority clusters" (Levi, 2004, p. 203). Similarly, Bell and Saka (1982) show a number of initial reverse sonority clusters in Pashto. The point about reverse sonority is covered in detail in Section 3.5. Here, I show only the possible syllabic templates of Pashto which suggest that the syllable structure of the language is not simple. Tegey and Robson (1996) suggest the possible syllabic patterns for Pashto given below:

(11)

V	VC	VCC
CV	CVC	CVCC
CCV	CCVC	CCVCC
CCCV	CCCVC	CCCVCC

(Tegey & Robson, 1996, p. 19)

To reinforce the view of Tegay and Robson, examples for all the possible clusters in the form of words or at word boundaries have been provided in (12). English translation for them has also been given.

(12)

V	/o/	sirrah
	/a:o/	yes
	/ə/ as in /ə.na:r/	pomegranate
VC	/əm/	mango /or/ fire
	/əs/	mare
VCC	/ərʃ/	heaven
	/əks/	jealousy
	/Iʃq/	love
CV	/xa:/	okay
	/nə/	no
	/ta:/	you
CVC	/xog/	sweet
	/dəm/	moment
	/ɖer/	very/more
CVCC	/lu:nd/	wet
	/ʃənɖ/	barren/impotent
	/ɖənɖ/	pool

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

CCV	/srə/	red
	/mla:/	waist
	/ɣla:/	theft
CCVC	/kləк/	hard
	/stər/	big
	/gra:n/	difficult/expensive
CCVCC	/drəst/	right
	/ʃna:xt/	identity/known
	/dʒwənd/	life
CCCV	/xwlə/	mouth
	/xwdai/	God
	/xwrə/	eat
CCCVC	/ndror/	sister-in-law
	/xkwəl/ as in /xkwəl.kro/	kiss
	/ʃxwən/ as in /sxwən.dər/	bull
CCCVCC	/ʃxwand/	chewing

The possible syllable structure of the Pashto language as derived from the above discussion is the following:

(C)(C)(C)V(C)(C)

The above words are everyday words, used frequently by Pashto speakers. This list also shows that a variety of syllable structures is possible in Pashto, and that most of the consonant sounds are acceptable at both onset and coda positions of the syllables (Tegey & Robson (1996). The detailed features of Pashto syllabification and phonotactics of Pashto phonology are discussed in Chapter 3. I will discuss a number of word-initial, word-medial and word-final syllables in greater detail there. This shows that the syllable structure of Pashto, as based on our earlier categorization of syllable structure in Section 2.2.1, is complex.

The current section shows that the word order of Pashto is SOV, and now it is clear that it has complex syllable structure. This phenomenon provides a clear-cut contradiction to the implicational correlation assumed, stated and proved by a number of studies discussed earlier. In the next section, I take this issue as a problem statement for the current research.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.7.2 The Problem Statement

The case of Pashto in the preceding section raises the core research question of the current study as discussed in Section 1.2. According to the correlation, the syllable structure of Pashto, being an SOV language, must be simple, i.e. (C)V. But the problem is that this is not the case: it has complex syllable structure. Pashto presents a very special case in the shape of a challenge to the correlation of syllable structure and word order assumed by a number of researchers mentioned in the current chapter. In order to formally address this problem, I shall explore the syllable structure of Pashto in the following chapter. Word order is the focus of our analysis in Chapter 4.

This characteristic of Pashto seriously questions the validity of the current correlation. The features of the Pashto language show that there is something more involved in defining the relationship between these two structures. A detailed study of Pashto syllable structure and word order will provide more data on the study of implicational universals and correlations based on them.

2.7.3 The Viability of the Current Research

It is necessary to state the viability of the current research from the perspective of the formal issue just discussed. Focusing on this problem, the study provides a comprehensive picture of the case of Pashto. Exploring the formal issue under study, it exhibits a number of phonological and syntactic features of the language within the Optimality Theory framework, with the intention of making a detailed analysis of relevance to readers with a background in Eastern-Iranian as well as English languages, especially the comparative study in Chapter 6. Pashto is typologically a very important language in this respect. So, while it will be of great interest for general linguists, the g25study also serves as an important document for comparative linguists interested in contrasting related phenomena across languages (e.g. word order variations and syntactic structures) or studying the Pashto language community, and, of course, for specialists in Pashto language.

2.8 SUMMARY OF THE CHAPTER

The chapter has briefly reviewed the main field of current study and the subject to be studied in greater depth in the following chapters. Having provided a short history of typological linguistics and its basic notions, the chapter considers well-known studies on the basis of word order typology.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Specific explanation of the subjectof correlation, in the light of Pashto data, is provided in detail. In the later part of the chapter, the basis for this correlation is brought into question with the help of the syllable structure and word order of Pashto. Thus, a straightforward motivation is provided for discussion of the problem and for requiring independent analysis of the phonotactics of Pashto phonology as a whole. Lastly, the existence of the problem suggests the potential viability of the study.

The next chapter analyses the syllabification of the Pashto language, addressing the so-far unexplained exception of the language to the correlation under discussion.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

CHAPTER 3

SYLLABLE STRUCTURE IN PASHTO

3.1 INTRODUCTION

The current chapter has three main aims: a detailed analysis of Pashto syllable structure, as well as the representation of Pashto syllabic complexity and the sonority-related specifications of the language. It starts by introducing Pashto phonology, goes on to cover various types of syllabic patterns in the language, and develops the discussion on syllabic complexity by demonstrating consonant clusters at word-initial, -medial and -final positions in Pashto. Along with a description of syllabification rules and the phonotactics of the language, the chapter examines the typology of Pashto syllables. OT will be used as a framework to provide the theoretical basis for the analysis of Pashto syllabic patterns. The combined information on the above points will pave the way to better understand and acknowledge the capacity of Pashto syllable structure, before the discussion moves on to the discussion of the word order of the language in the next chapter. All this will ultimately lead to our discussion of the methodology required to deduce the nature of the implicational correlation discussed in greater detail in Chapter 2.

3.2 PASHTO PHONEMIC INVENTORY

Previous studies of Pashto show some differences of opinion regarding the quantity[5] and quality[6] of the phonemes (e.g. regarding the presence/absence of consonants like /ts, dz, ʂ, ʐ, f,ŋ, j, q/ and vowels /e:, ʌ, æ/) in the

[5] For example, Penzl (1955) and Levi (2004) include /ʂ, ʐ/ and exclude /f, ŋ, j, q/. Tegey and Robson (1996)
include /ts, dz, f/ and exclude /ʂ, ʐ/. Ijaz (2003) includes /ɪ, e, ʊ/ too.

[6] For example, Ijaz (2003) shows /ʈ, ɖ, n, s, z/ as alveolar but Penzl (1955) describes /s, z/ as dental and /ʃ, ʒ/ as alveolar, while Bell and Saka (1982) consider /s, z, ʃ, ʒ/ all to be alveolar.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

segmental phonemic inventory of Pashto. Rehman (2009) details this lack of unified description in the phonetics and phonology of the Pashto language. According to him, the absence of a unified phonemic inventory is somehow the result of dialectical variations in the language. Despite these differences in opinion about segments and place of articulation, as mentioned in Chapter 1, a good thing about the language is that a speaker of any dialect readily understands the speakers of other dialects. What follows is a brief overview of the variations in the previous studies on this issue.

Penzl (1955), Roberts (2000) and Levi (2004) show /ʂ, ʐ, ʒ/ sounds, but according to my data these sounds do not exist today, specifically in Yousafzai dialect. Similarly, they do not show sounds /ŋ, f, q, j, ħ, ʕ/, which exist in Pashto, especially in a number of borrowed words from Arabic, Persian and Urdu. Penzl (1955) does not even consider the affricates /ts, dz, tʃ, dʒ/ as single segments, although these sounds are mentioned by Elfenbein (1997) as single segments. Likewise, Roberts (2000) does not even include /ts, dz, tʃ, dʒ/ in his inventory for study on Pashto clitics. Bukhari (1960), Ijaz (2003) and Rehman (2009) have based their studies on the Yousafzai dialect. They also do not show all the sounds of the language. For example, Bukhari (1960) does not show /f, q/, Ijaz (2003) does not include /ŋ, q/ and Rehman (2009), along with both of the others, does not add /ħ, ʕ, ts, dz/ in his studies. It is true that the sounds /f, q, ħ, ʕ/ are not native to the language, but they are used very frequently by native speakers of Pashto today.

The variation in the descriptions of Pashto phonemes is explicit in the differences among the scholars discussed above. In some cases, variation is found even within studies by the same scholars. For example, Tegey and Robson (1996), which includes /ŋ/, the velar nasal, is different from Robson and Tegey (2009), which excludes /ŋ/. In order to solve this unsettled dispute and to develop consensus among scholars, the above points suggest the need for one unified phonemic inventory for the language. Since such a task may be out of the scope of the current study, it is appropriate to mention that such a contribution might be taken as an important future study for the phonetics and phonology of the language. Phonetic study will solve the dispute of the quality of Pashto sounds and phonemic research will settle the quantity issue in this regard.[7]

[7] Phonetics is the study of the human sounds in general and Phonology is the study of sounds of a specific language. I used them both here because I mean that

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

In the current study, I try to remain as consistent as possible for the sake of clarity and uniformity. For this purpose, I propose the Pashto segmental phonemic inventory based on the Yousafzai dialect of Pashto which is shown in Tables 3.1 and 3.2 below. The Yousafzai dialect is considered the standard dialect of Pashto by a number of studies.[8] This phonemic inventory is closely related to Tegey and Robson's (1996) for consonants. In addition to some variation in the place of articulation of some sounds, my difference with them regarding consonants is in sounds /ʔ, ʒ/ (the glottal stop and the voiced post-alveolar fricative), on which, being a native speaker of the Yousafzai dialect, my intuition and the data do not match. Symbols are, of course, changed as I have applied IPA. For vowels and diphthongs, my study is based on the same source (Tegey & Robson, 1996). I have also given a brief view of some other sources as well.

3.2.1 Consonants

Table 3.1 shows the segmental consonants of Pashto. According to this description, the language has 30 consonants. They are subdivided into stops, nasals, fricatives, affricates, liquids and glides. Oral stops and fricatives constitute the major groups of the sounds, having nine sounds each. The majority of them are voiceless (11, including /q, f, ʃ, h/ single sounds, in addition to one in each pair). Among four nasals, /ɳ/ is a peculiar sound. Similarly, /ts/ among three fricatives is also very interesting among Pashto sounds. The three liquids are lateral /l/, trill /r/ and flap /ɽ/, one each. Pashto shares the flap /ɽ/ sound with other languages of the area (e.g., Urdu). The remaining are two glides, /w, j/. In terms of the place of articulation, Pashto has four bilabials, one labiodental and two dentals which are all common sounds. Four retroflex sounds are quite interesting for linguists. According to Tegey and Robson (1996, p. 16), Pashto is the only Iranian language to have these retroflex sounds.[9] The rest of the sounds are six alveolar, four palatal, five velar, one uvular, two pharyngeal and one glottal, all of which are quite common in South Asian

phonetically unified inventory will solve the disputes in the place and manner of articulation of Pashto aforementioned and phonological study will settle down the issue of the number of phonemes that how many basic sounds are there in Pashto.

[8] See Penzl (1955, p. 8), MacKanzie (1959, p. 233) and more recent work, Hallberg (1992, p. 44), for further detail on the topic. After Hallberg (1992), no study has queried this point as it is acceptable to most scholars.

[9] Some linguists (e.g., Penzl, 1955; Levi, 2004) show six retroflex sounds in Pashto by adding /ʂ, ʐ/ to the list.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

and Arabic languages.

Pashto retroflex sounds are unfamiliar to non-natives who have no retroflex sounds in their languages. According to Ray et al. (1996), retroflex sounds are made by the underside of the tongue tip flapping down only once against the alveolar ridge while not completely blocking the airflow down through the middle. The case of Pashto retroflex sounds is somewhat different to this description, except for /ɽ/. Pashto /ɽ/ is somewhat similar to the American pronunciation of /rd/ in the word *hardy*. These sounds are found in abundance in the Pashto vocabulary.

Table 3.1: Pashto Consonantal Phonemes

	Bilabial	Dental	Alveolar	Retroflex	Palatal	Velar	Glottal
Stops	p b	t̪ d̪		ʈ ɖ		k g	
Nasals	m		n	ɳ		ŋ	
Fricatives			s z		ʃ	x ɣ	h
Affricates			ts		t͡ʃ d͡ʒ		
Liquids			l r	ɽ			
Glides	w				j		

Table 3.1 illustrates of the place and manner of articulation of the consonants of Pashto. This inventory is mainly based on Tegey and Robson (1996), as discussed earlier. The vowels of the language are discussed in the following section.

3.2.2 Vowels

The case of Pashto vowels, in terms of overall agreement, is not different from that of its consonants. The vocalic segments are mentioned differently in quality as well as quantity by different sources. Penzl (1955) and Bell and Saka (1982) mention seven vowels but with a difference of quality in two low vowels, /a/ and /ɑ/. Tegey and Robson (1996) show nine vowels. Studies based on the Yousafzai dialect show varied views among the scholars.[10] The inventory of Pashto vowels shown in Table 3.2

[10] Ijaz (2003) shows 10 vowels and Rehman (2009) shows 13 vowels.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

is based on Tegey and Robson (1996).

Table 3.2: Pashto Vocalic Phonemes

	Front	Central	Back
High	i:		u:
	I		u
Mid	e:	ə	o
Low		ʌ	a:

There are a number of sounds in Pashto which demand the inclusion of /ʌ, e:/, which are missing in Tegey and Robson (1996), in the inventory. The minimal pairs in (13) show these sounds and the difference among their counter parts.

(13)	/ʌ/	/ʃʌl/ (cripple)	/ʃəl/ (twenty)
		/kʌr/ (to sow)	/kər/ as in /xkər/(horn)
	/e:/	/ku:ze:/ (ewers)	/ku:ze/ (down)
		/wale:/ (carriageway)	/wale/ (why)

Rehman (2009) and Ijaz (2003) include them in their studies. Since the scope of the current study is the complexity of the consonantal clusters in the syllable structure of the language, the question related to quality and quantity will not change the data because the current study is focused on the analysis of consonant clusters only. Therefore, for the sake of consistency, I base my study mainly on Tegey and Robson (1996).

3.2.3 Diphthongs

Tegey and Robson (1996) describe six diphthongs in their *Reference Grammar of Pashto*. They do not use IPA symbols for the representation of them. Similarly to my approach to consonants and vowels, I take their diphthongs as the basis of the current study. They show the symbols given in (14) for their diphthongs.

(14) /ej, əj, aj, a;j, uj, aw/

According to Jones (1964), "[Diphthongs] … must necessarily consist of one semi- vowel". Based on this principle Tegey and Robson's diphthongs might be recognized as such, but the issue about the quality and quantity of diphthongs is taken up by Ijaz (2003) and Rehman (2009) in detail in

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

their studies. According to them, there are four diphthongs in Pashto. Rehman (2009, p. 145) proposes the following diphthongs, with their minimal pairs (English translation is my own).

(15)

/ei /	/buɖei/	(old lady)	/buɖɑ/	(old man)
/ɑi/	/ɣwɑi/	(ox)	/ɣwɑ:/	(cow)
/əi/	/rɑzəi/	(come) *2p. imp*	/rɑzi/	(come) *3p. dec*
/iɑ/	/ziɑt/	(much)	/zɑ:t/	(caste)

The basic description of phonemes introduced in the above sections is used throughout this study. Having discussed the phonemic distribution of Pashto, I will now discuss how these sound segments are brought together to form minimal sound blocks and syllables, and their prosodic features. I start from the definition of *syllable* and will move on to highlight the nature of the study of syllables in OT. The analysis of Pashto syllables under OT will enable the extraction of syllabification rules which will ultimately be used for the representation of Pashto syllables.

3.3 THE SYLLABLE

The main focus of this study is the exploration of Pashto syllabic complexity. To demonstrate the complex syllable structure of the language, one must first analyse the nature of its syllables. A syllable is a unit of pronunciation which can easily be described by native speakers of a language, as the concept is more intuitive than linguistic to native speakers, but at the same time it is very hard for either a native speaker or a linguist to formally define what a syllable is. There are two levels of representation of syllables: the phonetic level and the phonological level. Phonetically, it has been claimed that syllables are accompanied by a chest pulse (which is “an individual burst of action by the expiratory muscles”; Devine & Stephens, 1994, pp. 9–10). This idea has been rejected by a number of scholars, e.g. Ladefoged (1967). Ladefoged and Maddieson (1996) are of the view that the syllable should be viewed strictly as a phonological unit. Actually, all the evidence mentioned in defining a syllable as a true constituent is in favour of its phonological nature. The present study is concerned with the phonological level of the syllable.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Blevins (1995, p. 207) mentions four reasons to regard the syllable as a true constituent of a phonological nature. Firstly, it functions as a domain for phonological rules and constraints. It is a domain which is larger than a phonemic segment, but smaller than a word, and contains only one sonority peak. This is clear from the assignment of stress in various languages on the basis of the syllable. Kenstowicz and Kisseberth (1979, p. 260) give the example of Classical Latin, in which stress is assigned on the basis of the number of syllables in the word and the weight of the penultimate syllable.

Secondly, some phonological rules may target the margins (edges) of syllables, as I will show the case of Pashto syllable clusters in this chapter in detail. Another common example which is given by experts (e.g., Kahn, 1980, p. 73) is the process of aspiration in English, as aspiration targets only initial voiceless stops in syllables. Thirdly, syllables are the target of certain morphological rules. Reduplication and language games are common examples. Finally, native speakers have clear intuitions about syllables in their language. The existence of syllabaries in many languages of the world is a demonstration of this fact. Another example is the observance of poetic metre in the language of poetry. Pashto provides a very handy example in this regard. Within Pashto folklore, the two-line *tappah* is composed of 22 syllables, out of which the first line always consists of nine syllables and the second of 13 syllables.[11].

The above discussion shows that the syllable is very important in the phonology of a language and that it is easier to define the syllable phonologically than phonetically. The syllable is part of the fixed prosodic hierarchy, and the syllable node dominates the onset and rhyme in universal prosodic hierarchy. Figure 3.1 shows this hierarchy, based on Blevins (1995, p. 210).

[11] Like other genres of folklore, most *tappah* are composed and transmitted by Pashtun women. The prosodic analysis was provided initially by Mackanzie (1958).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

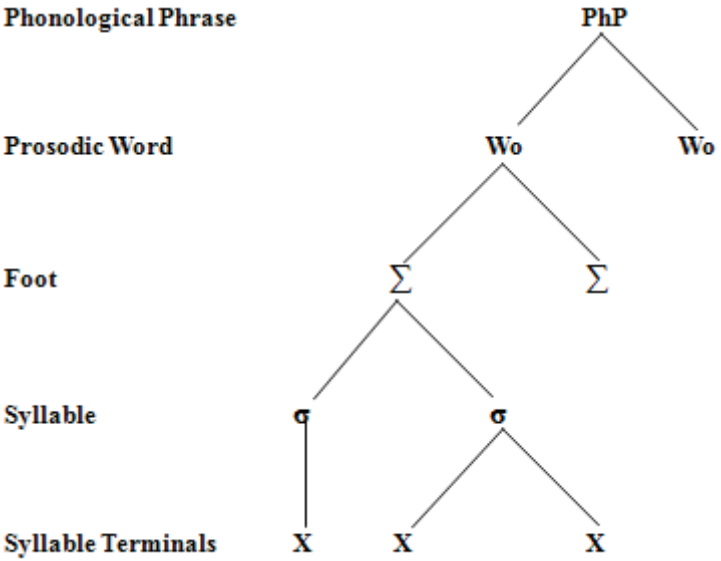


Figure 3.1: Universal Prosodic Hierarchy

Figure 3.1 shows that the syllable is the part of a fixed prosodic hierarchy. Another special characteristic of the syllable is that it is frequently organized in regard to the sonority of the segments within it. This aspect of Pashto is also part of the current study and will be taken further in Section 3.5.

A syllable comprises two basic components: onset and rhyme. The two elements are obligatory in the syllable, the onset and the first component of the rhyme[12] (peak, or in some languages it can be a syllabic consonant). In languages such as English, nasals and liquids are also allowed to take the position of the nucleus (Davenport & Hannahs, 2005, p. 15). A syllable ending with a vowel (CV) is universally considered harmonic and, therefore, some languages avoid syllables ending with a consonant. Figure 3.2 shows the structure of the syllable.

[12] There are languages in which the onset in a syllable is optional. Pashto is one of them.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

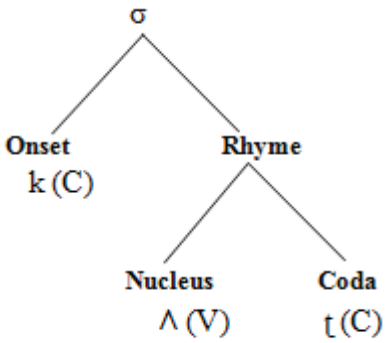


Figure 3.2: The Structure of the Syllable

The three components of syllable structure are the obligatory onset and nucleus, and the non-obligatory coda. Figure 3.2 shows the structure of Pashto word /kʌt̪/ (wooden bed), which has a consonant both at onset and coda. Universally, languages prefer to start with a consonant as onset but end without a coda consonant (Kager, 1999). But this universal syllabic process is challenged by language-specific syllabic preferences within various languages. In other words, syllabification is certainly a language-specific enterprise.

OT assumes that syllables are generated by the function GENerator from the base. According to Prince and Smolensky (2004), syllables are part of the base. OT considers syllables as larger components which are constructed out of the process of incorporation of the segments into higher prosodic constituents. Kager (1999, p. 91) points out that syllables are capable of providing proper phonological generalizations based on phonological patterns and verifying the well-formedness of the sequence of segments. He is of the view that the syllable plays an important role in the study of the well-formedness of structural patterns, the prosodic nature of a structure and the demarcation of the edges of morphemes. He also believes that processes such as epenthesis, deletion and alternation are triggered in order to avoid syllable ill-formedness.

A number of studies assume that syllables are organized in the form of minimal sound blocks. Rhyme Theory (Selkirk, 1984), Timing Slot Organization and Cs and Vs theory (Bell & Bybee, 1978; Clements & Keyser, 1983) are some of the examples from phonological literature

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

which have shown the sound organization within the syllable. The present study aims to analyse the syllabic structure of Pashto and to identify its syllabic-specific constraints.

3.4 SYLLABLES IN PASHTO

Pashto syllables have been the subject only of introductory-level studies so far. Bell and Saka (1982), Tegay and Robson (1996), Zaman (2002) and Levi (2004) are a few studies which have given some importance to the role of syllable structure in the phonology of the language. I recapitulate these studies in the current section. The current study is the only contribution to the constraint-based study of the language associating syllables and sonority of the language in it.

The source of the Pashto syllables discussed in this section has already been shown in Section 1.2. Apart from the three dictionaries, the syllable templates mentioned here are taken from the collection of 5000 real words (collected from 532 native speakers of the Yousafzai dialect of Pashto from a rural area) from everyday life.

Pashto neither allows nor abandons marginal segments. It has both types of syllables, open syllables (CV) (C for Consonant and V for Vowel), and closed syllables (CVC). The language allows nothing other than a vowel as the nucleus of its syllables. This means that Pashto contains at least a vowel as a peak, which is preceded or followed by consonant sounds. Although the language has the CV pattern, it also allows other more complex patterns. The maximum syllabic pattern is CCCVCC, as reported by Tegey and Robson (1996). In other words, the vowel in a Pashto syllable may be preceded by three consonants and it may be followed by up to two consonants. Because of the consonant clusters and reverse sonority,[13] Pashto syllables are considered one of the most interesting aspects of the language (Levi, 2004). The organization of sounds within a syllable and the syllabic templates are given importance by some of the previous studies.

Till now, Pashto syllabification and syllabic structure have not received appropriate attention. The literature available in the field of Pashto

[13] Reverse sonority means the violation of SSP – the sonority sequencing principle. This is an interesting feature of
Pashto phonology and is considered in greater depth in Section 3.8.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

phonology is developed on the basis of syllables based mostly on the intuition of the scholars, or it is only of introductory level. Hence, the core study of syllabification has been more or less neglected and is yet to be undertaken. Only two important studies, those of Bell and Saka (1982) and Rehman (2009), have tried to explore the articulatory and acoustic aspects of Pashto phonemes with limited orientation. The focus of Bell and Saka (1982) is on reverse sonority, while Rehman (2009) is concerned with the learning and teaching of English consonantal phonemes by L1 Pashto speakers. However, an amount of emphasis on the syllabic structure of the language and related issues is still required. In the following paragraphs, I will try to recapitulate (then critically evaluate) the previous studies on Pashto syllable structure.

Bell and Saka (1982) show a number of initial consonant clusters in Pashto. They believe that a number of Pashto initial clusters have reverse sonority hierarchy. They specifically show two consonant clusters (/wr-/ and /wl-/) and many other combinations which violate the sonority sequencing principle. Mainly, they address the violation of sonority sequencing in the initial clusters of the language, but they also demonstrate that /w/, the bilabial glide, is actually a glide and not /β/, a bilabial fricative, as previously reported by Shafeev (1964). Detailed reference to Bell and Saka (1982) is made in Section 3.5.

Tegey and Robson (1996, p. 19) give some detail about Pashto syllables and syllabic structures in their reference grammar. They suggest that the language has a maximum of three consonants t the beginning and two at the end of the syllable in root words. Giving the evidence of a number of consonant clusters, they also claim that the language is remarkable for its abundance of two-consonant clusters at the beginning. They show that it has 12 core syllabic patterns, as shown in Table 3.3.

Table 3.3: Pashto Syllabic Patterns

V	VC	VCC
CV	CVC	CVCC
CCV CCCV	CCVC CCCVC	CCVCC CCCVCC

According to the above patterns the maximum syllabic pattern is CCCVCC, but Zaman (2002) believes that maximum syllable structure of Pashto only goes up to CCCVC. His appears to be another version of the

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

syllable patterns of Pashto, showing nine syllabic patterns (V, CV, VC, CVC, CCV, CVCC, CCVC, CCCV, CCCVC). Thus he disregards three types of pattern (i.e., VCC, CCVCC, CCCVCC) from those shown by Tegey and Robson (1996). The three patterns which are shown by Tegey and Robson (1996) but not by Zaman (2002) are found in Pashto. In Pashto, words with a VCC pattern such as /əks/ (jealousy), /Iʃq/ (love), /Ilm/ (knowledge) and /r**I**zq/ (food) are found in abundance.

Similarly, CCVCC is also a very common pattern in the language. For example, words such as /drəst/ (right/total), /ʃna:xt/ (identity/known), /dʒwənd/ (life) are used in daily language. The last pattern (CCCVCC) is also found but in very few words. I will discuss them in detail in the section on syllabification rules of the language.

Levi (2004) highlights two very important phenomena of the language, namely the features of glides (labial glide /w/ and palatal glide /j/) and the presence of many consonant clusters in the language. She is of the view not only that Pashto contains a number of clusters but that several of them are reverse sonority clusters (Levi, 2004, p. 203).

The version of syllabic patterns submitted by Tegey and Robson (1996) is advantageous for the present study because it represents every possible syllable structure of the language. In Table 3.4, I group all the possible syllabic patterns into three main categories on the basis of consonant complexity, which I will further use for the analysis of consonant clusters in the following sections.

Table 3.4: Categories of Pashto Syllable Patterns

S. No:	Category	Syllabic Patterns
1	Simple syllables	(V. CV. VC. CVC)
2	Bipartite consonant clusters	(CCV. VCC. CVCC. CCVC. CCVCC)
3	Tripartite consonant clusters	(CCCV. CCCVC. CCCVCC)

Table 3.4 shows syllabic patterns which are based on the recorded collection of 5000 words spoken by Pashto native speakers. The recording was analysed for every possible syllable pattern. It is clear from the above

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

syllabic patterns that the language has a number of onset and coda clusters. I will show, in following sections, all of the above three categories of consonant clusters along with their sub-patterns of syllable structures in root words at different positions: word-initially, word-medially and word-finally.

3.5 SIMPLE SYLLABLES IN PASHTO

As per Table 3.4, the simple syllable patterns are (V.CV.VC.CVC[14]). Overall, there is no problem with all patterns shown in this category in Pashto syllables. They can occur easily anywhere in the world languges. For clarity, Table 3.5 briefly shows examples of them (with English translation in round brackets).

Table 3.5: The Simple Syllables in Pashto

Pattern	Examples
V	/o/ (sirrah), /a:o/ (yes), /ə/ as in /ə.na:r/ (pomegranate)
VC	/əm/ (mango), /or/ (fire), /əs/ (mare)
CV	/xa:/ (okay), /nə/ (no), /ta:/ (you)
CVC	/xog/ (sweet), /dəm/ (moment), /ɖer/ (very/more).

Since the scope of the current study is to show complex clusters in the syllables of Pashto, I restrict the analysis of simple syllables to the Table 3.5. However, more data related to the above patterns is given in the appendices. The possibility for complex consonant clusters is discussed in greater detail in the following sections. Section 3.6 shows the two-consonant clusters which are possible in the language.

3.6 BIPARTITE CONSONANT CLUSTERS IN PASHTO

In this section, I will analyse the bipartite consonant clusters at all three positions (i.e., word- initially, word-medially and word-finally) in Pashto. Pashto is remarkable for allowing various two-consonant clusters in the onset of the syllable. According to Tegey and Robson (1996), English has

[14] CVC pattern is not a simple pattern in the true sense as it violates the NOCODA constraint. It is included in this category only because in this study I am concerned with clusters, and within CVC no cluster is found.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

around 35 combinations of two-consonant clusters at onset level, but in Pashto this number is close to 100.

Tegey and Robson (1996) give some detail about initial clusters in their *Reference*

Grammar of Pashto, with 52 possibilities given below in (16).

(16)
/br/, /bʈ/, /bj/, /dr/, /dw/, /dzɣ/, /dzm/, /dzw/, /gr/, /gʈ/, /gw/, /ɣl/, /kl/, /kr/, /kʈ/, /kw/, /xp/, /xw/

/lm/, /lw/, /mj/, /mr/, /mʈ/, /ng/, /nɣ/, /nm/, /nj/, /pl/, /pr/, /pʈ/, /pj/, /tl/, /tr/, /sk/, /sx/, /sp/, /sr/

/st/, /sw/, /ʃk/, /ʃp/, /xk/, /wl/, /wr/, /wʈ/, /zg/, /zɣ/, /zm/, /zʈ/, /zj/, /zhm/, /zhw/

According to Tegey and Robson (1996), the 17 consonant clusters given below in (17) are also found in Pashto, but in few words.

(17)
/bl/, /dj/, /xj/, /ml/, /nj/, /rɣ/, /ʃx/, /ʃm/, /tw/, /tsk/, /tsw/, /zb/, /zd/, /zr/, /zw/, /xʈ/, /tʈ/

Thus, they submit the above sound combinations as possible complex consonant clusters in the language.

As discussed earlier, Levi (2004) points out that Pashto is remarkable for two very interesting aspects: first, the presence of consonant clusters, and second, the reverse sonority clusters. Her point regarding reverse sonority is discussed in detail in Section 3.8. In (18), her proposed 87 clusters are given.

(18)
/wr/, /wʈ/, /wl/, /rɣ/, /lw/, /lm/, /lɣ/, /mr/, /mʈ/, /mn/, /ml/, /nʂ/, /nʐ/, /nɣ/, /sw/, /sr/, /sʈ/, /sm/,

/sx/, /sp/, /st/, /sʈ/, /sk/, /ʃw/, /ʃr/, /ʃʈ/, /ʃl/, /ʃm/,/ʃn/, /ʃx/, /ʃp/, /ʃʈ/, /ʃk/, /ʂw/, /ʂk/, /xw/, /xr/, /xʈ/,

/xl/, /xp/, /zw/, /zʈ/, /zm/, /zɣ/, /zb/, /zd/, /zɖ/, /ʒw/, /ʒr/, /ʒɣ/, /zm/, /zb/, /zd/, /ɣw/, /ɣr/, /ɣʈ/,

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

/ɣl/, /pr/, /pɽ/, /pl/, /ps/, /pʃ/, /pʂ/, /tw/, /tr/, /tɽ/,/tl/, /ts/, /tʃ/, /ʈw/, /kw/, /kr/, /kɽ/, /kl/, /kʂ/, /br/,

/bɽ/, /bl/, /dw/, /dr/, /dɽ/, /dz/, /dʒ/, /gw/, /gr/, /gɽ/, /gz/

Since there are some differences in the phonemes between the data provided by Tegey and Robson (1996) and that of Levi (2004), I will give my own data from the Yousafzai dialect of Pashto. First, I will give detail of bipartite consonant clusters possible at all three positions (word-initially, -medially and -finally) one by one, starting with word-initial position in the following section.

3.6.. Bipartite Consonant Clusters at Word-Initial Position

In the current section, I will show the clusters, with their class combinations, which are possible at word-initial position. By word-initial position, I mean the onset of the root words (morphologically simple words). Table 3.6 shows the combinations of plosives and liquids as word-initial two-consonant clusters.

Table 3.6: Bipartite Initial Clusters (Plosives + Liquids)

Type	Cluster	Examples	English Translation
	pr	prəde	Stranger
Plosives	**pɽ**	pɽək	Moment/glitter
+	**pl**	pla:r	Father
Liquids	**br**	brəg	Spotted/piebald
	bɽ	bɽa:s	Heat/exhalation/steam
	bl	blarba	Pregnant
	tr	tror	Aunt
	tɽ	tɽək	Breaking noise
	tl	tləl	Going
	dr	drənə	Heavy
	ɖr	ɖrənka	Harsh, braying noise
	kr	krəka	Disgusting
	kɽ	kɽəp	Bang/slam
	kl	klək	Tight
	gr	gra:n	Precious
	gɽ	gɽəng	Bang/snotty-nosed

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3.6 shows the bipartite consonant clusters at word-initial level which have a plosive as the first and a liquid as the second consonant of the cluster. There are a number of such combinations in Pashto. In this type of combination, some clusters are not possible. They are given below:

/*dɽ,*dl,*dʐ, *ɖl, *gl,*qr,*qɽ/

Plosive /ʈ/, in combination with a liquid, is possible in words like /ʈrək/ (truck). But such words are foreign in origin, therefore I do not add such a combination to my list because the focus of the study is to analyse basic Pashto words.

In Table 3.7, various combinations of plosives with glides are shown in the initial clusters.

Table 3.7: Bipartite Initial Clusters (Plosives + Glides)

Type	Cluster	Examples	English Translation
	pj	pja:z	Onion
Plosives	**bj**	bja:	Then/again
+ Glides	**dw**	dwə	Two
	dj	dja:rləs	Thirteen
	gw	gwəʈa	Brown sugar
	tw	twa:n	Power/ability
	kw	kwənɖə	Widow
	qj	qja:mət	Day of judgment

Apart from the eight combinations in this type shown in Table 3.7, other combinations are not possible in basic Pashto words, although the combination /qj/ is possible in words borrowed from Arabic, and such words are now in daily use in the Pashto vocabulary. Table 3.8 demonstrates the consonant clusters combining fricatives and glides.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3.8: Bipartite Initial Clusters (Fricatives + Liquids)

Type	Cluster	Examples	English Translation
	sr	srə	Gold
Fricatives	**zɽ**	zɽə	Heart
+ Liquids	**ʃr**	ʃrəng	Clink/jingling
	ʃɽ	ʃɽək	Thunder/swift
	ʃl	ʃla:nʤə	Frog/heavy
	xr	xrə	Donkey
	xɽ	xɽi:s	large and bulky
	ɣr	ɣrət	A sudden blow
	ɣɽ	ɣɽap	Quick gulp
	ɣl	ɣla:	Theft

Other combinations, such as those shown below, are not possible in this type of cluster:

/*sɽ, *sl, *zr, *xl/.

Fricatives /f, ħ, ʕ/ are not possible in combination with liquids in Pashto. Some linguists (Bukhari, 1960; Levi, 2003; Rehman, 2009) are of the view that these are not Pashto sounds by their origin. This may be the reason that these phonemes are not found in combination with liquids. Similarly, fricative /h/ never takes the position of first consonant in the onset cluster in a syllable.

It is important to note a point raised by previous studies (e.g., Bell & Saka, 1982; Levi, 2004): that clusters starting with fricatives (Tables 3.8–3.12) have no epenthetic vowel in them. Table 3.9 shows the combinations of fricatives with plosives.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3.9: Bipartite Initial Clusters (Fricatives + Plosives)

Type	Cluster	Examples	English Translation
	sk	skor	Charcoal
Fricatives	**sp**	spa:i	Soldier
+ Plosives	**st**	stən	Wooden pillar
	ʃk	ʃku:ṇ	Porcupine
	ʃp	ʃpəg	Six
	xk	xka:r	Hunting
	xp	xpəl	Relative
	zg	zgerəl	Moan/sew
	zb	zba:dəl	To prove
	zd	zdakol	To learn

Among fricatives, /f, ɣ, ħ, ʕ, h/ have no combinations with plosives while others have limited combinations with them. Table 3.10 exhibits the initial clusters of two consonants which are made of fricatives and glides.

Table 3.10: Bipartite Initial Clusters (Fricatives + Glides)

Type	Cluster	Examples	English Translation
	sj	sja:l	Equal
Fricatives	**sw**	swəzedəl	To burn
+ Glides	**zj**	zja:t	Much
	zw	zwa:n	Young
	xj	xja:l	Care
	xw	xwa:	Side
	ɣw	ɣwa:	Cow

Other fricatives, such as /f, ʃ, ħ, ʕ, h/, are not found in combination with glides in onset clusters. Table 3.11 illustrates the pairing of fricatives with fricatives.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3.11: Bipartite Initial Clusters (Fricatives + Fricatives)

Type	Cluster	Examples	English Translation
Fricatives + Fricatives	**zɣ**	zɣaməl	To tolerate
	ʃx	ʃxəʈa:	Dispute
	sx	sxər	Father-in-law

Only the three cluster combinations shown in Table 3.11 are possible between fricatives and fricatives at onset level. Table 3.12 shows the combination of fricatives with nasals.

Table 3.12: Bipartite Initial Clusters (Fricatives + Nasals)

Type	Cluster	Examples	English Translation
Fricatives + Nasals	**zm**	zməka	Land/earth
	ʃm	ʃmer	Counting
	ʃn	ʃna	Greenery

Nasals are possible not only as the second consonant in combinations at onset level but also as the first consonant. Nasals are found in combination with all other sound classes except affricates, but they are very few in number. Table 3.13 gives a picture of nasals in their various combinations.

Table 3.13: Bipartite Initial Clusters (Nasals + Others)

Type	Cluster	Examples	English Translation
Nasal + Plosive	**ng**	ngor	Daughter-in-law
Nasal + Fricative	**nɣ**	nɣəre	Hearth/stove
Nasal + Liquids	**ml**	mla	Waist
	mr	mrəsta:	Help
	mj	mja:ʃt	Month
Nasal + s Glides	**mj**	mja:ʃt	Month
	nw	nwəri:	Clothes
	nj	nja:	Grandmother
Nasal + Nasal	**nm**	nməsi:	Grandchildren

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3.13 shows the possible combinations of nasal sounds with others as first consonant at onset level. Combinations of two nasals are found in very few words, such as /nməsi:/ (grandchildren) and /nmər/ (sun), but nonetheless they are used. Similarly to nasals, liquids are also found at both positions (first or second consonant of the cluster at onset level). The possibility of syllabic nasals is also evident from some of the above examples; this point needs further detailed work. Table 3.14 illustrates the possibility of liquids as first consonant of the onset cluster, in various combinations.

Table 3.14: Bipartite Initial Clusters (Liquids + Others)

Type	Cluster	Examples	English Translation
Liquid + Fricative	**rɣ**	rɣəʈi:	Rolls
Liquid + Nasal	**lm**	lmər	Sun
Liquid + Glide	**lw**	lweʃt	Span

In Table 3.14, liquid sounds are the first sound of the onset cluster and the second combinatory sound varies. Liquids can also be the second sound in combination with glides in some onset clusters. Table 3.15 demonstrates such combinations.

Table 3.15: Bipartite Initial Clusters (Glides + Liquids)

Type	Cluster	Examples	English Translation
Glide + Liquid	**wr**	wra:rə	Nephew
	wʈ	wʈəl	Carry
	wl	wluna	Cleaning

It is clear from Table 3.15 that such combinations violate the sonority sequencing principle (SSP). Pashto is remarkable in this sense, in that it has such attraction towards bipartite consonant clusters that it even allows SSP-violating clusters. I will discuss this point in greater detail in Section 3.8. For the time being, I note that such a possibility is a peculiarity of the language, allowing for combinations which are rare in other languages.

Affricates have not been touched upon so far in this section on consonant clusters. The story of Pashto affricates is also very interesting. They have not been considered even as a unit of sound by some scholars. Penzl (1955), for example, does not consider them a single unit. Even Roberts

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(2000) does not show them in his inventory of Pashto. But there are some linguists who do see them as a unit of sound. Elfenbein (1997) argues that Pashto affricates are single segments, and he proves that they behave as single units in metathesis. This point is very important because if I consider them as single segments, which they are, they form combinations as part of bipartite clusters. According to Elfenbein, if they were not single unit, they would form the only clusters of four consonants in a syllable. He gives the word /tsʈməx/ (lining) as an example in this regard (Elfenbein, 1997, p. 741).

In Pashto syllables, affricates always take the first consonant cluster position in clusters at onset level. In Table 3.16, I confirm the possible clusters which combine affricates and other sounds.

Table 3.16: Bipartite Initial Clusters (Affricates + Others)

Type	Cluster	Examples	English Translation
Affricate + Liquid	**dʒr**	dʒrənda	Watermill
Affricate + Glides	**dʒw**	dʒwənd	Life
	tsw	tswa:rləs	Fourteen
Affricate + Fricative	**dʒɣ**	dʒɣorəl	To preserve/ keep
Affricate + Plosive	**tsk**	tska:k	Drinking

It is clear from Table 3.16 that affricates can take a liquid, a glide, a fricative or a plosive as their partner in onset clusters. Though such combinations are not very great in number, these words are very common in everyday Pashto language. The following section summarizes all the possible word-initial clusters in the language.

3.6.2 Summary of Bipartite Consonant Clusters at Word-Initial Position

In section 3.6.1, I have shown a number of onset clusters which are found in the language. As shown in Tables 3.6 to 3.16, a total of 77 combinations are found which take place at the position of onset. This shows that the language has a very rich structural pattern from the point of view of syllable complexity. Table 3.17 gives evidence of all these available clusters, showing first consonant in the left-hand column and second along the top row.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3.17: Summary of Bipartite Consonant Clusters at Word-Initial Position in Pashto

	p	b	t	d	ɖ	k	g	q	s	z	ʃ	x	ɣ	m	n	r	ɽ	l	w	j	ts	dʒ
p	-															√	√	√		√		
b		-														√	√	√		√		
t			-													√	√	√	√			
d				-												√			√	√		
ɖ					-											√						
k						-										√	√	√	√			
g							-									√	√		√			
q								-												√		
s	√		√			√			-			√				√			√	√		
z		√		√			√			-			√	√			√		√	√		
ʃ	√					√					-	√		√	√	√	√	√				
x	√					√						-				√	√		√	√		
ɣ													-			√	√	√	√			
m														-		√	√	√	√			
n							√						√	√	-				√	√		
r													√			-						
ɽ																	-					
l														√				-	√			
w																√	√	√	-			
j																				-		
ts						√													√		-	
dʒ													√			√			√			-

It is clear from Table 3.17 that Pashto is quite rich in terms of consonantal clusters possible at the onset level of words. The first consonants in such clusters are mostly plosives, fricatives or nasals, while affricates, glides and liquids rarely take this position. In the case of second consonant, this is mostly a liquid or a glide, while fricatives and nasals hardly appear in this position. The largest number of possible clusters are found in combinations between plosives and liquids, of which there are 16. Plosives as second consonant in the cluster are found in combination with fricatives, which violates the standard SSP. Such violation by consonant clusters is discussed later in the chapter. In the next section, I will show the syllable clusters at coda position (word-finally) of the syllable in root words of the language.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.6.3 Bipartite Consonant Clusters at Word-Final Position

In the previous section, I discussed the consonant clusters which are possible in word-initial position. Now, I will analyse the bipartite consonant clusters which are possible in word-final position in Pashto. In this position, Pashto also has various types of possible clusters. Tegey and Robson (1996) show a number of clusters which could possibly take up the position of coda clusters. For example, they show 30 combinations at word-final position, which are given here as (19). They are of the view that word-final clusters contain a maximum of two consonants (bipartite).

(19)
/kht?/, /nd/, /nɖ/, /rg/, /ʈx/, /sk/, /st/, /xt/, /bz/, /fz/, /ks/, /xs/, /mp/, /mz/, /ndz/, /nj?/, /nk/, /ns/,

/nz/, /qʃ/, /rd/, /rɣ/, /rm/, /rn/, /rq/, /rs/, /rt/, /rz/, /ʃq/, /ʃt/

However, they have missed some clusters which are very common at the end of words in the Yousafzai dialect: for example, /rʃ/ in words such as /derʃ/ (thirty) and /nɳ/ in /tərkaːnɳ/ (carpenter). Such clusters are very common at word-final position and are frequently used. Similarly, their clusters which contain the /?/ mark are not tested in my data.

They have also shown /kht/ and /xt/ as two distinct clusters, but in my dialect they are the same. Table 3.18 demonstrates the possible clusters which end with plosives at word-final position in Pashto.

Table 3.18 shows that a number of word-final cluster combinations are possible which end with plosives. Fricatives, nasals and liquids are the possible pairs which are placed near to peak and thus have no problem from a sonority point of view. In Table 3.19, I show word-final clusters which end with fricative sounds.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3.18: Bipartite Final Clusters (Others + Plosives)

Type	Cluster	Examples	English Translation
Fricatives	**xt**	ba:ləxt	Pillow
+	**ʃt**	səlweʃt	Forty
Plosives	**ʃq**	Iʃq	Love
	ʃk	əʃk	Tear
	sk	həsk	Cloth
	st	drəst	Right
Nasals	**nd**	dʒwənd	Life
+	**nɖ**	ɖənɖ	Pool
Plosives	**mp**	pəmp	Pump
	nk	Benk	Bank
Liquids	**rg**	tʃərg	Chicken
+	**rq**	ərq	Essence
Plosives	**rd**	dərd	Pain
	rt	pərt	Property record letter
	rp	tərp	Radish
	ɽd	ʃa:gəɽd	Student

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3.19: Bipartite Final Clusters (Others + Fricatives)

Type	Cluster	Examples	English Translation
Plosives + Fricatives	**bs**	həbs	Humidity
	ks	əks	Contrary/reverse
	bz	sərsəbz	Lush green
	qʃ	nəqʃ	Copy/drawing
Glide + Fricative	**wz**	ləwz	Promise
Nasals + Fricatives	**mz**	rəmz	Secrecy
	nz	munz	Prayer
	ns	həns	bird
	nʃ	benʃ	beam
Liquids + Fricatives	**rs**	tərs	Pity
	rz	tərz	Form
	rk	tərk	Cleft
	ɽx	əɽx	Aspect
Fricative + Fricative	**xs**	ʃəxs	Person

In clusters ending with a fricative, they are found in combination with plosives, nasals, liquids and even another fricative as the pairing consonant.

There are some clusters which end with nasal or affricate sounds, but such clusters are very few in the language. Table 3.20 exhibits examples.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3.20: Bipartite Final Clusters (Others + Nasals or Affricates)

Type	Cluster	Examples	English Translation
Liquids + Nasals	**rm**	fa:rm	Form
	rn	qərn	Century
Nasals + Affricate	**ndʒ**	torəndʒ	Citron tree

It is clear from Table 3.20 that nasals and affricates can also take the position of the ending consonant in clusters at coda level.

3.6.4 Summary of Bipartite Consonant Clusters at Word-Final Position

Section 3.6.3 exhibits a number of coda clusters which are found in the language. In Tables 3.18 to 3.20, a total of 33 combinations are shown, with the help of examples from the language, at coda level. It is evident from this data that the language has a very rich pattern of syllable complexity at word-final position as well. Table 3.21 summarizes all possible clusters found at word-final position. The left-hand column shows the first consonant and the top row the second consonant in the cluster.

As well as being rich in word-initial consonantal clusters, Table 3.21 shows that Pashto also has a number (33 in total) of possible bipartite consonantal clusters at word-final position. The last consonant in such cases is either a fricative or plosive (in 30 clusters out of 33) or, rarely, an affricate or nasal (only 3 out of 33). The first consonant can be any sound except a glide.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3.21: Summary of Bipartite Consonant Clusters at Word-Final Position in Pashto

	p	b	t	d	ɖ	k	g	q	s	z	ʃ	x	m	n	r	ɽ	l	w	ʤ
p	-																		
b		-							√	√									
t			-																
d				-															
ɖ					-														
k						-			√										
g							-												
q								-			√								
s			√			√			-										
z										-									
ʃ			√			√		√			-								
x			√						√			-							
m	√									√			-						
n				√	√	√			√	√	√			-					√
r	√		√	√		√	√	√	√	√			√	√	-				
ɽ				√								√				-			
l																	-		
w										√								-	
ʤ																			-

Words having two-consonant clusters at word-final position (coda) are very interesting from a morphological point of view. For example, the pluralizing process breaks the clusters of such sounds when they are made plural. I show some examples in this regard below. Syllable boundaries are also shown.[15]

[15] It is interesting to look at this phonology–morphology interface. The discussion on morpheme and syllable boundaries may be very interesting with regard to alignment constraints in OT. However, here I just point out these features of Pashto, because the scope of the current study is to look at consonant clusters and at word order in light of the subject correlation which is introduced in Chapter 2 in depth.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(20)

/to.rəndʒ/	(citron tree *singular*)	››	/to.rən.dʒa:n/ (citron trees *plural*)
/əks/	(contrary act *singular*)	››	/ək.su:.nə/ (contrary acts *plural*)
/mu:nz/	(prayer *singular*)	››	/mu:n.zu:.nə/ (prayers *plural*)
/ʃənɖ/	(sterile *singular*)	››	/ʃən.ɖa:n/ (sterile *plural*)

Since morphological analysis is not the focus of the current study I do not go into the detail of such processes, but it is the case that the final clusters are broken as a result of such morphological operation in the language. The words selected for analysis in this section are morphologically simple, so the clusters present there are by default available and they have no concern with morphological operations.

Having shown bipartite consonant clusters at both word-initial and word-final position, I demonstrate various possible consonant clusters at word-medial position in the following section.

3.6.5 Bipartite Consonant Clusters at Word-Medial Position

In the previous two sections, consonant clusters at word-initial and word-final positions – which are in fact onset and coda positions, respectively – were shown. In the section, to complete the analysis of consonant clusters, I show consonant clusters at word-medial level also. There has been no previous study carried out on Pashto word-medial clusters: the studies which have discussed word-initial or word-final clusters have said nothing about word- medial clusters.

Pashto morphological processes break most clusters at word-medial level as just discussed. For example, the consonant clusters of the word /to.rəndʒ/ (CV.CVCC) (citron tree *singular*) are broken through the process of pluralization, such that it becomes /to.rən.dʒa:n/ (CV.CVC.CVC) (citron trees *plural*). Many such changes take place at the word level; more examples can be seen from the word list given in the appendices. Despite such breaking of consonant clusters, a great number of clusters are found at word-medial position. The analysis in the current section boosts the thesis of the study that consonant clusters are possible at every level of the Pashto lexicon.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3.22: Bipartite Medial Clusters

Cluster	Examples	Pattern	English Translation
sw	zʈə.swun.ke	CCV.CCVC.CV	Sympathetic
xw	məl.xwə.ze	CVC.CCVC	Grass-like vegetable/fenugreek seed
ɣw	nar.ɣwəʈ	CVC.CCVC	He-cat (tom)
sj	təŋ.sja:	CVC.CCV	Distress
gw	məŋ.gwəl	CVC.CCVC	Hand's width
kw	mə.ʈə.kwər	CV.CV.CCVC	Having hard skin or nature
tj	ʃin.tjə	CVC.CCV.CV	Greenery
dr	tən.drəst	CVC.CCVCC	Healthy
pʈ	pI.ʃo.pʈa:ŋ	CV.CV.CCVC	Wild cat
kʈ	wər.kʈə.ra.kʈə	CVC.CCV.CV.CCV	Give and take
tl	ra.tləl	CV.CCVC	Coming (N)
zʈ	ra:n.zʈə	CVC.CCV	Box for collyrium
nd	drund.wale	CCVCC.CV.CV	Soberness
ʃm	kəm.ʃme.rə	CVC.CCV.CV	Minority
ʃp	jəu.ʃpe.tə	CV.CCV.CV	Sixty-one
st	stə.ʈe.sto.ma.na	CVC.CV.CCV.CV.CV	Extremely tired

Table 3.22 exhibits the possible two-consonant clusters which are found at word-medial position in the language. It is clear from the combinations that the first consonant is most frequently either a fricative or a plosive (15 out of 16) and the second consonant is generally either a glide or a liquid (12 out of 16). Plosive and nasal sounds can take both positions. Another interesting thing is that almost all of the possible word-medial clusters are found at the onset position of the internal syllable.

One important point which was identified during this analysis is that Pashto root words mostly have one syllable, as shown in the data, in word-final and word-initial positions. Although words with more than five or six syllables (e.g. /bər.gʈən.di:.ke.dəl/, to overtake someone/to embrace, CVC.CCVC.CV.CV.CVC) are also found in the language, in case of longer words morphological processes affect the structure of the word. As already discussed, my focus throughout this study remains on morphologically simple (root) words and I have given 16 possible combinations from within word clusters in Table 3.22. Interestingly, the morphologically complex words (words with one or more affixes based on grammatical meaning) also contain many clusters within them. Examples of polymorphemic words are given in (21) .

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(21)

/Il.tʃja:n./	Ambassadors	VC.CCVC
/bər.gɽən.di:.ke.dəl/	to embrace	CVC.CCVC.CV.CV.CVC
/be.stər.go/	blind/shameless	CV.CCVC.CV

In Pashto, /be/ is actually a negative morpheme in Pashto. But words having clusters do not break their clusters even when /be/ is added to them. There are some other examples as well.

(22)

/be.dre.ɣə/	pitiless	CV.CCV.CV
/be.dja:.nə.tə/	untrustworthy	CV.CCV.CV.CV
/be.ʃme.rə/	countless	CV.CCV.CV

Similar to the negative morpheme /be/, there are others, such as verbal clitics /wo/, /wər/ and /ra/, which do not disturb the word-medial consonant clusters of the words. Examples are given as (23).

(23)

/wo.prən.dʒe.dəl/	to sneeze	CV.CCVC.CV.CVC
/ra.pre.wə.tal/	to lay down	CV.CCV.CV.CVC
/ra.ɣwən.ɖə.wəl/	to collect	CV.CCVC.CV.CVC
/wər.xwə.re.dəl/	to spread	CVC.CCV.CV.CVC
/wər.pra.nəs.təl/	to free	CVC.CCV.CVC.CVC
/wər.tləl./	to go	CVC.CCVC

Another interesting morphological characteristic of the language is the adding of the suffix /tja/ to make a noun from any verb. Such words are in abundance in Pashto, and they also contain consonant clusters (the combination between plosive + glide /t/ + /j/) within them. Note the following examples.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(24)

/ʃɨn.tjə/	greenery
/roɣ.tjə/	healthiness
/kə.mə.qəl.tjə/	foolishness

Since our concern in the current study is to analyse the nature of morphologically simple words, I do not include these polymorphemic words in the study or go into detail about any morphological process related to consonant clusters. The only fact to be pointed out here is the possibility of clusters at word-medial position, which is clear from Table 3.22. The section below summarizes these word-medial clusters in Pashto.

3.6.6 Summary of Bipartite Consonant Clusters at Word-Medial Position

In the previous section, two-consonant clusters are demonstrated in detail. A total of 16 combinations are found in morphologically simple words. It is evident from this data that Pashto has a number of clusters at every position, and that it has syllable complexity at word-medial position as well. Table 3.23 summarizes all clusters at this position. The left-hand column shows the consonants which are first in the cluster and the top row shows those which are second.

Table 3.23 summarizes the possible clusters at word-medial position in the language. It is evident from the chart that most of the second consonants are glides or liquids. In the following section, a summary of bipartite consonant clusters at all three positions is given.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3.23: Summary of Bipartite Consonant Clusters at Word-Medial Position in Pashto

	p	t	d	k	g	s	z	ʃ	x	ɣ	m	n	r	ɽ	l	w	j	ʤ
p	-													√				
t		-													√		√	
d			-										√					
k				-										√		√		
g					-											√		
s		√				-										√	√	
z							-							√				
ʃ	√							-			√							
x									-							√		
ɣ										-						√		
m											-							
n			√									-						
r													-					
ɽ														-				
l															-			
w																-		
j																	-	
ʤ																		-

3.6.7 Summary of Bipartite Consonant Clusters in Pashto

In the previous three sections, bipartite consonant clusters in the Pashto language were shown with the help of data from the Yousafzai dialect. I have analysed the various combinations of consonant clusters possible at three different positions. It is explicit from the data that the maximum number of combinations – 77 – are possible at word initial (onset) position. The second highest number is the 33 bipartite consonant combinations possible at word-final position. The number of possible word-medial consonant combinations is only 16. Adding together all possible combinations at all positions gives a total of 126 two-consonant clusters in the language. But there is one problem in this calculation: there is a risk of double-counting some clusters, as some are possible at more than one position. For example, the combination of fricative /s/ with plosive /t/ is possible at all three positions of the word (although this is the only combination which is found at all three positions), as shown in (25).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(25)

1. Word-initial position /stən/ Needle/pillar
2. Word-medial position /stə.ɽe.sto.ma.na/ Extremely tired
3. Word-final position drəst/ Right/total

It is clear from these examples that one combination may take place at more than one position. While it is only one combination, it may be counted three times. Therefore, in order to avoid such risk of double counting, Table 3.24 provides an accumulative account of all two-consonant clusters in Pashto. The initial consonant of the cluster is shown in the left-hand column and the second along the top row of the table.

Table 3.24: Summary of Bipartite Consonant Clusters in Pashto

	p	b	t	d	ɖ	k	g	q	s	z	ʃ	x	ɣ	m	n	r	ɽ	l	w	j	ts	dʒ
p	-															√	√	√		√		
b		-							√	√						√	√	√		√		
t			-													√	√	√	√	√		
d				-												√			√	√		
ɖ					-											√						
k						-			√							√	√	√	√			
g							-									√	√		√			
q								-			√									√		
s	√		√			√			-			√				√			√	√		
z		√		√			√			-			√	√			√		√	√		
ʃ	√		√			√		√			-	√		√	√	√	√	√				
x	√		√			√			√			-				√	√		√	√		
ɣ													-			√	√	√	√			
m	√									√				-		√	√	√	√			
n				√	√	√	√		√	√	√		√	√	-				√	√		√
r	√		√	√		√	√	√	√	√			√	√	√	-						
ɽ				√								√					-					
l														√				-	√			
w										√						√	√	√	-			
j																				-		
ts						√													√		-	
dʒ													√			√			√			-

Table 3.24 merges all clusters found at any of the three positions of words in Pashto and gives an accumulative picture of bipartite consonant clusters. It is clear from the table that, in reality, 108 distinct two-consonant clusters are found in the language. Out of the higher figure of 126 given above, 16 combinations were repeated once while one

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

combination was possible at all three positions of the word. So, the picture given by Table 3.24 is the real one.

With regard to bipartite consonant clusters, Pashto is among the richest languages in the world (Tegey & Robson, 1996; Levi, 2004). Now in the following section, I will discuss the case of tripartite clusters in the language.

3.6.8 Tripartite Consonant Clusters in Pashto

The existence of tripartite clusters in Pashto is reported by a number of previous studies. According to Tegey and Robson (1996, p. 19), tripartite clusters have three types of patterns in Pashto (CCCV, CCCVC and CCCVCC). The last one is the maximum syllabic pattern and, at the same time, the least common one. Tegey and Robson (1996, p. 23) are of the view that such combinations are possible only within a handful of words in Pashto.

According to them, these combinations include /xwl/, /xwd/, /ndr/, /skw/, /sxw/, /ʃxw/, and /xkw/. Elfenbein (1997) and Levi (2004) also identify words such as /tsʈməx/ (lining), but such a word, with a CCCVC pattern, does not exist in the Yousafzai dialect of Pashto today. Another word with similar sounds exists, /tsəʈ.məx/ (having an ugly face), with a CVC.CVC syllabic pattern. These scholars do not include the combination /xwʈ/, which is found in a number of words in Pashto. Table 3.25 illustrates the claimed tripartite consonant combinations.

The case of tripartite consonant clusters in Pashto is very interesting for two reasons. Firstly, almost all these three-consonant combinations have at least one glide (in seven out of eight combinations) or nasal (in one combination). Secondly, many of these combinations have reverse sonority sequence (e.g. /xwl/, /xwd/, /ndr/). Both these reasons slightly throw doubt on the existence of tripartite consonant clusters in the language. The point of reverse sonority was taken up by Bell and Saka (1982) and Levi (2004), and they say that the language is quite rich in terms of reverse sonority clusters. Reverse sonority is discussed in greater detail in Section 3.11. In this section, I discuss the point of glide /w/ in such combinations.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3.25: Tripartite Consonant Clusters in Pashto

Type	Cluster	Examples	English Translation
Fricative + Glide +	**xwl**	xwlə	Mouth
Liquid	**xwɽ**	xwɽi:n	Softened
Fricative + Glide +	**xwd**	xwda:i	God
Plosive Fricative +			
Fricative + Glide	**sxw**	sxwəndər	Bull
Fricative + Plosive	**ʃxw**	ʃxwənd	Chewing
+ Glide	**skw**	skwərəʈə	Ember
Nasal + Plosive +	**xkw**	xkwəle	Beautiful
Liquid	**ndr**	ndror	Sister-in-law

Previously it was made clear by a number of studies[16] that /w/ is a bilabial glide and not a /β/ (bilabial fricative) in such combinations, as had been claimed by Shafeev (1964). Bell and Saka (1982) experimentally proved that it produced no frication and its patterns were like those of a sonorant. Levi (2004) says that in Pashto, only a sonorant can combine with voiced or voiceless obstruents. It is also clear from the above clusters that the glide /w/ can occur with both voiced and voiceless obstruents, which in turn must agree in voicing among themselves[17] in the shape of clusters. Another point which can be submitted in favour of such tripartite consonant clusters is the length of duration of glide. When such combinations (in words like /sxwəndər/, /ʃxwənd/, /skwərəʈə/, /xkwəle/) were analysed through PRAAT, the duration of the glide was around 180 ms. Since the average duration of schwa /ə/ is around 60–70 ms,[18] the duration of the glide suggests that it is confirmedly a glide, /w/, thus proving the existence of the tripartite clusters with experimental satisfaction.

[16] For example Penzl (1955) and Bell and Saka (1982).

[17] See the /sxw/, /ʃxw/ and /skw/ combinations in the above examples of three-consonant clusters.

[18] The length of duration of schwa varies from language to language. For example, it is around 60–70 ms in English. Its length in Pashto is 50 ms (Fougeron, Gendrot & Burki, 2007).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Another point to observe about these tripartite clusters is the presence of nasal /n/ in one out of eight clusters of tripartite combinations. In some dialects of Pashto, this word, /ndror/ (sister-in-law), has two different pronunciations given here as (26).

(26)

1. /ndror/
2. /In.dror/

Yet it is true that it is mainly pronounced as /ndror/ in Yousafzai dialect. Once the so-called reverse sonority clusters are accepted in the language, there is no problem in accepting that such combinations exist.

It is very true that the number of tripartite consonant clusters is quite few compared to the number of bipartite clusters. In the case of bipartite clusters the language is extremely fertile, but in the case of tripartite consonant clusters it is less so. This is often the case with languages having complex clusters, because the more complex the clusters, the less they are found. Importantly, such syllabic-shaped words are in abundance in Pashto and are frequently used by speakers. Fifteen per cent of the 5000 content words collected have complex syllabic patterns and are in frequent use by the speakers. These three-consonant clusters are found at word-initial position only.

3.6.9 Interim Conclusion

An extensive illustration has been given on the syllable structure of Pashto. It has been established that the language has 108 possible bipartite consonant clusters which are possible at every position of the word, while eight tripartite consonant clusters are also exhibited with the help of data from the language. What I have found so far is that the syllable structure of the language is undoubtedly complex, providing enough data for the hypothesis of the study regarding the complexity of the syllable structure of Pashto.

Now, in order to theoretically justify the possibility of such consonant clusters and to answer a number of questions regarding the sonority-related specification of the language, I analyse the syllable structure of Pashto in an OT framework in the following section. I have shown in the preceding sections that there are a number of syllabic clusters which violate the usual pattern of clusters. I have also pointed out that the language has some sonority-violating clusters at more than one position. Therefore, it is appropriate to address this aspect of the language before

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

moving into the detail of the syllabification rules. The next section deals with such characteristics of the language.

3.7 AN OT ANALYSIS OF PASHTO SYLLABLES

As introduced in Section 1.4, OT is an ideal framework for the current study because it provides the best possible constraint-based analysis for syllable structure. An OT methodology assumes that syllables are not part of algorithmic structure-building (Prince

& Smolensky, 2004); rather they are part of the base, a point which is justified by the concept of Richness of the Base (ROB) in OT. OT assumes that syllables are generated by the component GENerator which is constructed as a process of integrating the segments into higher prosodic blocks. I have also established, in Section 3.3, that the functional benefits of syllables are well known in the study of phonological features. Spencer (2002) mentions ways in which syllables play an important role in the organization of phonological processes of a language. Kager (1999) observes that syllables provide proper phonological generalizations for the verification of well-formedness of the sequence of the segments in a language.

I have also discussed, in Section 3.6, the fact that Pashto syllables constantly challenge the universally preferred syllable pattern because of the language's specific syllable preferences. For example, a number of Pashto syllables are complex; the onset cluster goes up to CCC- and the coda up to -CC. It is also shown in the same section that the language has onset as well as onset-less syllables. In the current section, I will provide a theoretical justification for such a tendency in the language, using the OT framework.

3.7.1 ONSET and NOCODE in Pashto Syllables

Beside the nucleus, onsets and codas are both very important in syllable structure. In this section, the constraints relating to onset and coda and their rankings in Pashto are discussed in detail. According to typological studies on syllable structure (e.g., Jakobson, 1962;

Greenberg, 1972; Ito, 1986; Blevins, 1995), the presence of an onset in a syllable is an unmarked situation. Kager (1999) is of the view that no language can entirely exclude onset even if some of its syllables lack onsets. This idea is expressed in the structural well-formedness constraint ONSET.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(27)
ONSET
*[σ V Syllables must have an onset (Ito, 1989; Prince & Smolensky, 1993)

This is a powerful constraint, and by itself it can generate all ill-formed syllable structures regardless of their context. It requires that all syllables within a language must begin with an onset but not with a vowel (nucleus). Syllables such as CV and CVC satisfy this constraint, but syllables such as V or VC do not. Pashto is not an ONSET-dominant language, as it allows many V and VC or even VCC structures at word-initial position. The following examples are given in this regard:

(28)
Satisfying the constraint Violating the constraint
/xa:/ CV (Okay) /a.ba.si:n/ V.CV.CVC (Indus River) /na:/ CV (No) /It.bar/ VC.CVC (Sunday)
/ra:.za:/ CV.CV (Come) /a:w/ VC (Yes)

It is clear from the above examples that the constraint is not at the top of the hierarchy of constraints in the language. Pashto, unlike Arabic, is not very strict about onset in every syllable, and both types of syllable are evident from the above examples. As a result, it is required that the constraint change its position to avoid malfunction. This shows that the ranking of this constraint is motivated by some additional constraints.

Onsets are of two types, simple and complex. Simple onsets have one consonant, while complex onsets have more than one. ONSET itself does not say anything about the number of consonants at this position. In Pashto, mixed types of onsets are possible and the maximum number of consonants in an onset is three. Tables 3.6 to 3.16 show various types of onset combinations possible in two-consonant clusters in the language. Similarly, Table

3.25 shows the three-consonant clusters at this position in the language. Now I will show the position of the NOCODA constraint in the language.

(29)
NOCODA
*C] σ Syllables are open (Kager, 1999, p. 94)

The coda is a very controversial component in the structure of the syllable. The presence of consonants at coda level is considered marked universally. When there are two syllables creating an intervocalic

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

consonant (CVCV), languages prefer to syllabify them according to a universally expected syllabification trend (CV.CV). This principle is called MOP (maximum onset principle), which states that the onset of the syllable must be maximized by putting the consonant in it wherever it is possible. Pashto likewise prefers onset syllabification to coda syllabification, but at the same time it allows syllable patterns which end with consonants.

Most languages (e.g., Boumaa Fijian; Kager, 1999) prefer to avoid syllables ending with codas because they are treated as marked cases. As I have shown, Pashto also has syllables with codas, so it is important for the constraint to be ranked in order to avoid an ungrammatical outcome. Preference between syllables with and without codas is not fixed. If it is required by the well-formedness of the structure, the coda is provided; otherwise, it maintains NoCoda. The following examples show the violation of the constraint in Pashto.

(30)
/ʃi:n/ CVC (Blue) /sel/ CVC (Visit) /ɣa:x/ CVC (Tooth)
So the ranking of the constraint in Pashto can be read as:

(31)
NoCoda›› Coda (the coda is not allowed)

In the above lines, I have shown that the language does not give top priority to structural well-formedness constraints, such as Onset and Nocoda. At this stage, the question arises of why the language allows syllables of "imperfect" types, such as CVC, V, VC, etc. OT has a very simple answer to this question. It says that in such a situation, the language simply gives priority to faithfulness constraints. Faithfulness constraints may prevent the segmental changes which are necessary in order to create perfect syllable structure (Kager, 1999, p. 97). In other words, the faithfulness forces the markedness violation.

There are other phonological phenomena in Pashto which are relevant here, e.g. complex clusters at both edges of the syllable. The language not only allows the "imperfect" types of syllables mentioned above but also permits consonant clusters at both edges of the syllable. This pattern is examined in the following section.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.7.2 Complex Onsets and Codas in Pashto Syllables

According to syllable typology, languages differ in terms of the complexity of syllable margins. Complex onset and complex coda are indeed marked cases and, therefore, avoided by some languages. But there are some languages which prefer both of them in a limited way. Pashto is a type of language which allows up to three consonants at onset position and two at coda, as seen in Section 3.6. Thus it is evident that the language is violating the constraints of *COMPLEXONS and *COMPLEXCOD. Before giving examples from the language in this regard, let me define both constraints here.

(32)
*COMPLEXONS
*[σCC Onsets are simple (Kager, 1999, p. 97)

This constraint is based on the idea that onset of a syllable is simple and cluster in onset is something marked. This constraint requires onsets to be simple, and *COMPLEXONS (No Complex Onset) would punish a candidate with a complex onset. It belongs to the category of well-formedness constraints. In Pashto, the constraint *COMPLEXONS is dominated by other faithfulness constraints and, therefore, complex clusters in the onset are possible, as shown in (33).

(33)
/tri:x/ CCVC (Bitter)
/sxər/ CCVC (Father-in-law)
/brex.na:/ CCVC.CV (Lustre)
/ʃxwənd/ CCCVCC (Chewing)

The above examples show that Pashto does not give preference to the constraint

*COMPLEXONS. This markedness constraint is dominated by faithfulness constraints because Pashto allows not only two-consonant clusters but also three-consonant clusters in the onset. The tables in Section 3.6 show a number of two- and three-consonant clusters in the language.

Like onsets, there are some clusters possible in the coda position as well. Let me discuss the position of codas in this regard.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(34)
*COMPLEXCOD
*CC]σ Codas are simple (Kager, 1999, p. 97)

Complex codas in syllables also have universally marked positions. Like onsets, codas are also complex in Pashto. Therefore, I can say that the constraint *COMPLEXCOD is also dominated by other faithfulness constraints in the language. Pashto allows the possibility of complex clusters in this position. Note the following examples in (35):

(35)
/benʃ/ CVCC (Beam) /tərs/ CVCC (Pity) /tərz/ CVCC (Form)

It is clear from the discussion so far that the faithfulness constraints dominate the markedness constraints in Pashto. The faithfulness constraints may prevent the segmental changes which are necessary in order to create perfect syllable structure in the language (Kager, 1999, p. 97). DEP-IO and MAX-IO are the correspondence (faithfulness) constraints over ONSET and NoCoda (the well-formedness constraints) and now also over *COMPLEXONS and *COMPLEXCOD. Now before drawing a table, it is appropriate to discuss these faithfulness constraints.

3.7.3 DEP-IO and MAX-IO in Pashto Syllables

Correspondence constraints DEP-IO and MAX-IO are the faithfulness constraints dominating the markedness constraints in Pashto.

(36)
DEP-IO
Output segments must have input correspondents (“No epenthesis”; Kager, 1999, p. 101)

DEP-IO is basically an “anti-epenthesis” constraint which is ranked highly in Pashto, which does not allow any epenthesis into onset or coda clusters. As a result, not only are clusters found but no additional segment is added to the structure. McCarthy and Prince (1995) argue that this Dependency-IO constraint says that output segments should correspond to the input segments and, therefore, nothing additional should be inserted into the output. This constraint governs the faithfulness between input and output in the syllables of Pashto.

On the other hand, MAX-IO is an anti-deletion constraint which states that the segments which are present in the input should be preserved in the

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

output. So, no deletion of the segments is allowed in the output. This constraint punishes any deletion of segments which is part of the input. The following is the constraint.

(37)
MAX-IO
Input segments must have output correspondents ("No deletion"; Kager, 1999, p. 102)

Similarly to DEP-IO, MAX-IO is also a faithfulness constraint which makes sure that the input is identical to the output. In Pashto syllables, epenthesis of a segment is prohibited by DEP-IO, while the deletion of a segment is restricted by the constraint MAX-IO. All this shows that these correspondence constraints in the language are dominant over these particular well-formedness constraints. The tableau below shows this tendency.

(38) /tri:x/

Input: /tri:x/	DEP-IO	MAX-IO	*COMPLEX^{ONS}	NOCODA
a. ☞/tri:x/			*	*
b. /tə.ri:x/	*!			*
c. /ti:x/		*!		*

In the above tableau, it is shown that the language gives priority to faithfulness constraints DEP-IO, MAX-IO over markedness constraints, such as *COMPLEXONS and NOCODA. Candidate (a) has consonant cluster at the onset, while candidate (b) has no cluster because one segment is added (the epenthesised /ə/). Candidate (c) has no cluster because one segment from the input is dropped. So, all the candidates have the violation of at least one constraint. The optimal output (a) violates the lowest-ranking markedness constraints *COMPLEXONS and NOCODA, which are not further ranked because they are equally ranked in the hierarchy of the language. The violation of these two constraints is the least expensive in the language. The second candidate violates the faithfulness constraint DEP-IO, which is ranked above *COMPLEXONS and NOCODA, and therefore its violation is fatal and hence ruled out. Similarly, the third candidate violates another faithfulness constraint, MAX-IO, by deleting one segment in the input. This violation is again fatal because it is also ranked above both *COMPLEXONS and NOCODA.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

This ranking of the constraints allows complex onsets in the language along with codas. This listing can generate the factorial typology given in (39):

(39)
DEP-IO, -IO ›› *COMPLEXONS, NOCODA.
"Codas and complex onsets are allowed"

Codas are not the essential part of the syllables, but still languages allow them, and some such as Pashto allow clusters at coda position as well. Section 3.6.4 shows 33 combinations of coda clusters possible in Pashto. The following tableau shows the position of constraint related to coda cluster in the language.

(40) /drəst/

Input: /drəst/	DEP-IO	MAX-IO	*COMPLEX^{ONS}	*COMPLEX^{COD}
a. ☞/drəst/			*	*
b. /drəs/		*!	*	
c. /drə.si:t/	*!		*	

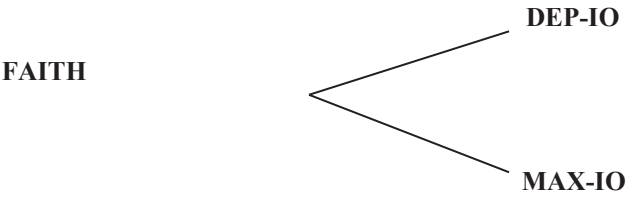
Tableau (40) shows that faithfulness constraints DEP-IO, MAX-IO dominate both the markedness constraints *COMPLEXONS and *COMPLEXCOD. Candidate (a) has consonant cluster at onset and coda position, while candidate (b) has no coda cluster because one segment is deleted. Candidate (c), again, has no coda cluster because one vowel is added, which creates another syllable in the word. So, all the candidates have violation of some constraints. The optimal output (a) violates the lowest-ranking markedness constraints *COMPLEXONS and *COMPLEXCOD, which are the equally lowest-ranking constraints related to the syllable structure of the language. The violation of these two constraints is the least expensive in the language. The second and third candidates violate the faithfulness constraints MAX-IO and DEP-IO, which are ranked above *COMPLEXONS and *COMPLEXCOD. Since the violation of both of them is fatal, therefore, they are ruled out. I can now add *COMPLEXCOD to replace NOCODA from our ranking of the constraints in the factorial typology of the syllable structure of the language:

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(41)
DEP-IO, MAX-IO ›› *COMPLEXONS, *COMPLEXCOD .
“Both complex onsets and codas are allowed”

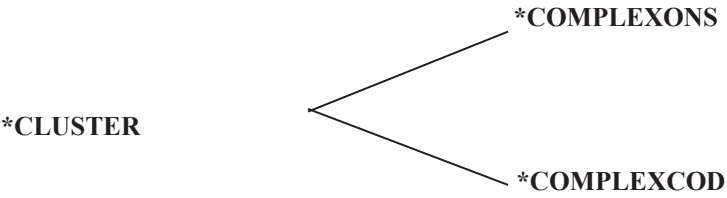
Now, in order to simplify our ranking of the constraints, I include both DEP-IO and -IO within one collective constraint with the name of FAITH. (42) shows both these constraints collectively within FAITH.

(42)
DEP-IO and MAX-IO are faithfulness (FAITH) constraints (Kager, 1999).



Similarly, I also include both *COMPLEXONS, *COMPLEXCOD within one single constraint with the name of *CLUSTER.

(43)
*CLUSTER No complex syllable codas and onsets.



Now, in order to include the words of three-consonant clusters at onset and two at coda, the following tableau is given.

(44) /ʃxwənd/

PLEASE REINSERT AS NOT FORMATTED CORRECTLY IN WORD OR PDF FORMAT

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Tableau (44) shows that faithfulness constraints are strictly dominant in Pashto. *CLUSTER constraints (*COMPLEXONS and *COMPLEXCOD) are not that important in this ranking and, therefore, they are easily violated by syllables in Pashto. The mutual ranking between DEP-IO and MAX-IO, and between *COMPLEXONS and *COMPLEXCOD, is irrelevant here because in this study I am only concerned with the possibilities of the consonant clusters in the language. In three-consonant clusters, such as above, another constraint is quite relevant which is related to the sonority among the segments in Pashto syllables and is discussed in detail in a following section. The following section provides a summary of the OT analysis carried out in the current section.

3.7.4 Summary

In order to provide theoretical justification for complex clusters in Pashto, I applied the theoretical framework of OT in the last section. To avoid unnecessary discussion on other issues, I focused exclusively on consonant clusters in the syllable structure of the language. I have shown with the help of tableaux that the faithfulness constraints (DEP-IO and MAX-IO) are dominant in the language. Other constraints, such as ONSET, NOCODA, *COMPLEXONS and *COMPLEXCOD, are easily violated in the syllable structure of the language. The final factorial typology of the syllable structure of Pashto is the following:

(45)

FAITH (DEP-IO, MAX-IO) ›› *CLUSTER (*COMPLEXONS, *COMPLEXCOD).
“Both complex onsets and codas are allowed in Pashto syllables”.

(46) shows the hierarchy of these constraints in the language. This hierarchy covers all the syllabic templates of Pashto shown in Section 3.4.

(46)

FAITH (DEP-IO, MAX-IO)

*CLUSTER (*COMPLEXONS, *COMPLEXCOD)

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

The basic setting of the Pashto syllabification system has been described so far. However, this analysis does not elaborate on the morphological or syntactic intensity of syllable alteration. The main focus of the current study is to analyse the complex syllable structure of the language. Therefore, I have restricted the analysis to the required level of the study. The next section covers one of the important phenomena related to consonant clusters: their sonority specification.

3.8 SONORITY-RELATED SPECIFICATION OF PASHTO

The current section describes one of the most interesting aspects of Pashto phonology. The sonority-related specifications of the language have fascinated a number of scholars in the past. Bell and Saka (1982) were the first to study sonority-related idiosyncrasies of the language. They describe the language as having two very peculiar reverse sonority clusters (i.e. /wr-/ and /wl-/) at word-initial position. According to Levi (2004), the language not only has consonant clusters "but also several reverse sonority clusters" (p. 203). Previous studies have only analysed the initial reverse sonority clusters of the language. In this section, I present the reverse sonority clusters found at all positions based on the data presented in Section 3.6. But before considering in detail the sonority-related specifications of the language, some theoretical background is necessary.

3.8.1 Sonority Sequencing Principle (SSP)

As early as Sievers (1881), linguists identified the phenomenon of sonority decreasing when moving away from a syllable nucleus cross-linguistically, which is typically termed the sonority sequencing principle (SSP). According to this principle, the sonority of the syllable must increase in the onset towards the peak and decrease towards the coda. It says:

> Between any member X of a syllable and the syllable peak P, only sounds of higher sonority rank than X are permitted. (Clements, 1990; Dost, 2004)

SSP is a presumed cross-linguistic universal, which governs the permissible sequences of consonant clusters within syllables. For example, in a regular syllable pattern (CV), a consonant forms the onset and is followed by a vowel, which is higher than the onset in terms of sonority (the peak of the sonority). Then, in some cases, it may be followed by another consonant which is lower than the vowel at nucleus (peak) and which forms the coda of the syllable (CVC). In general, the sonority sequencing principle is related to the existence of consonant clusters under a specific order in a position (onset or coda) in the syllable. There are

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

some languages which do not allow the coda in a syllable, while the onset is a universal occurrence. SSP is relevant to languages such as Pashto, which not only allows a coda but also permits clusters of consonants at both edges of the syllable.

Following SSP, in the case of clusters at onset position, the first consonant must not be higher in sonority than the second consonant and so on. Rather, it must have lower sonority value than the following consonant. Likewise, the final consonant at coda cluster position must have lower sonority than the previous one. This rising and falling of sonority is ranked on a sonority scale. The sonority hierarchy of Jesperson (1904) gives basic information regarding the sonority ranking of onset and coda. It points out the harmonic contact between the peripheral components of the syllable. His hierarchy claims that the segments with higher sonority values towards peak and those of lower values towards coda would form a better combination at syllable margins because this mutual sonority arrangement makes for an easy articulation. Example (47) below shows the patterns of such harmonic combinations at coda and onset levels:

(47)
a. Onset:
t > s > d > z > n > l > r > w
b. Coda:
w > r > l > n > z > d > s > t (Based on Jesperson, 1904)

The above patterns show that the sonority value in consonants is highest in glides and lowest in the voiceless stops, as shown in (48):

(48)
Glides › rhotics › laterals › nasals › voiced fricatives › voiced stops › voiceless fricatives › voiceless stops (abbreviated as: w › r › l › n › z › d › s › t)

Although this sonority hierarchy is basic in its scope, it has influenced most of the follow-up theories (e.g. Clements, 1990; Dost, 2004) related to sonority sequencing of consonantal clusters at the syllable level. The influence of such sonority sequencing has affected the phonological resolutions in the literature. Such theories have addressed a number of complexities of phonological gaps and their relevance to the sonority hierarchy. Since I am concerned here with SSP and its hierarchy, it is appropriate to mention the sonority scale and discuss the clusters of Pashto in the light of that scale.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.8.2 Sonority Scale

Before discussing the consonant clusters allowed in Pashto, I will discuss the scale of sonority in greater detail.[19] Though there has been some controversy relating to the particular detail of the formulation of the sonority scale, the basic idea is uncontroversial. For example, very little agreement is found on the point of the relative sonority of laterals and rhotics, and the voiced and voiceless segments. Similarly, Clements (1990) is of the view that all vowels have the same sonority, while Kenstowicz (1994) and Crosswhite (1999) state that they differ in terms of sonority. Apart from such relative issues, most of the researchers agree on something like the following:

(49) Vowels › glides › liquids › nasals › obstruents
(Bell & Hooper, 1978; Clements, 1990; Smolensky, 1995)

Clements (1990) argues that this single sonority scale portrays the sonority distinctions available in all languages and that it does not require any further divisions on any bases. The discussion on the controversial part of SSP is neither relevant nor possible within the scope of current study. Therefore, I follow the most widely agreed view of the phenomenon.

The sequence of rising and falling sonority is marked on a sonority scale, which is a measurable index. Within this index, the plosives are the least sonorous segments and the vowels are the most sonorous ones. Plosives are taken together with fricatives within one category, i.e. obstruents. After obstruents, the sonority rises through nasals and liquids, then glides, and finally vowels. The following sonority scale, based on Burquest and Payne (1993), is further modified for the purpose of current study. It shows the sonority hierarchy of

segments in Pashto. This is the five-point sonority scale and shows that vowels are the most sonorous and obstruents are the least sonorous segments in the language, while glides, liquids and nasals are ranked in order of most to least sonorous in the middle.

[19] Sonority scales have been discussed by a number of scholars in greater detail in literature on phonology. See Sievers, 1881; Jesperson, 1904; Hooper, 1976; Hankamer & Aissen, 1974; Murray & Vennemann, 1983; Clements, 1990a. Clements 1990a offers an excellent overview of them.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

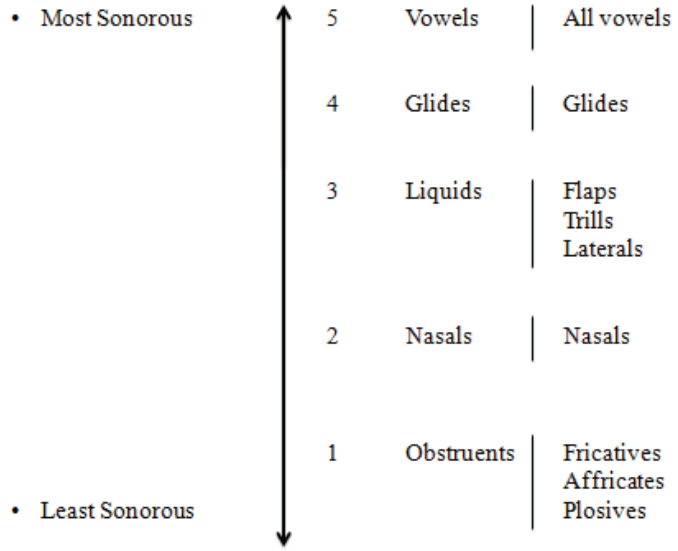


Figure 3.3: The Sonority Scale in Pashto

The above scale is important, as most of the combinations follow it. However, Pashto gives least importance to the constraint of sonority sequence and does not even maintain minimal sonority distance (MSD). This is discussed in the following section in detail. In short, the following two points (a) and (b) precisely show the basic crux of SSP.

a. Every syllable peak must be the peak of sonority in the syllable.
b. A syllable must not contain more than one sonority peak.

The above points collectively mean that segments progressively decrease in sonority from the peak to the margins in terms of the segments. Now having the above scale, I analyse the consonant clusters of Pashto in terms of sonority.

3.8.3 Sonority Hierarchy of Segments in Pashto

SSP is a strong cross-linguistic tendency relied on by many theories of syllable structure. As discussed earlier under this theory, a syllable has exactly one sonority peak, and sonority decreases through segments to the margins of the syllable. Some of the early works on Pashto clusters show

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

that Pashto violates this principle in some of its syllable patterns (Bell & Saka,

1982; Levi, 2004). With an overview of the past works, I will discuss the topic in the light of my current data from the Yousafzai dialect. In the end, I will demonstrate theoretical justification for such behaviour by the language.

One of the early works on Pashto phonetics, a study by Bell and Saka (1982), shows a number of consonant clusters. The authors opine that some of the Pashto initial clusters have reverse sonority sequence. Apart from the clusters, they reinforce the idea of Penzl (1955) that /w/, the bilabial glide, is actually a glide and not /β/, a bilabial fricative, as had previously been reported by Shafeev (1964). On the basis of segment /w/ as a glide, they confirm the idea of reverse sonority in the initial clusters. They acoustically examine the nature of two very different reverse sonority clusters /wr-/ and /wl-/ and show that the output segment of such clusters is a glide, thus providing enough evidence for the existence of reverse sonority sequence in the language. They compare such clusters with prevocalic /wV-/ and then with some unstressed short vowels /u-/ and find that the glides, in reverse sonority clusters, have the same spectral shape as in singleton onsets. They also prove the presence of glides by comparing it with short vowels and through stretching.[20]

Apart from Bell and Saka (1982), Levi (2004) also discusses the possibility of reverse sonority. She bases her data on Penzl (1955). According to her:

> Not only does Pashto contain a wide variety of normal clusters, but also many reverse sonority clusters. (Levi, 2004, p. 202)

She gives the following 23 bipartite consonant clusters as being reversed in sonority (2004, p. 204):

(50)
/wr, wɽ, wl, rɣ, lm, lɣ, nʂ, nz̧, nɣ, sp, st, sʈ, sk, ʃp, ʃt, ʃk, ʂk, xp, zb, zd, zḍ, ẓb, ẓḍ/

[20] This is a method used to increase the overall duration of an utterance (Levi, 2004).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

She states that Pashto allows reverse sonority in such clusters, a very interesting phonological characteristic of the language. Keeping in view the argument made by Levi (2004), I show similar clusters from my data. As many segments[21] of the current study are different from Levi's, I also present the violation of sonority. Most of the consonant clusters show the phenomenon of SSP. The combinations in (51) show that the initial bipartite consonant clusters observe SSP:

(51)

1. Plosives + liquids	(detail given in Table 3.6)
2. Plosives + glides	(detail given in Table 3.7)
3. Fricatives + liquids	(detail given in Table 3.8)
4. Fricatives + glides	(detail given in Table 3.10)
5. Fricative + nasals	(detail given in Table 3.12)
6. Nasals + liquids	(detail given in Table 3.13)
7. Nasals + glides	(detail given in Table 3.13)
8. Liquids + glides	(detail given in Table 3.14)
9. Affricates + liquids	(detail given in Table 3.16)
10. Affricates + glides	(detail given in Table 3.16)
11. Affricates + fricatives	(detail given in Table 3.16)

However, at the same time, some of the combinations that may violate as reverse clusters are found in the data. Combinations in (52) show the bipartite consonant clusters which violate the standard universal of SSP in Pashto.

(52)

1.	Fricatives + plosives /sk, sp, st, ʃk, ʃp, xk, xp,zg, zb, zd/	(detail given in Table 3.9)
2.	Fricatives + fricatives /zɣ, ʃx, sx/	(detail given in Table 3.11)
3.	Nasals + plosives /ng/	(detail given in Table 3.13)
4.	Nasals + fricatives /nɣ/	(detail given in Table 3.13)
5.	Nasals + nasals /nm/	(detail given in Table3.13)

[21] I do not include /ʐ, ʂ, ʒ /, which are included by her.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

6. Liquids + fricatives /rɣ/ (detail given in Table 3.14)
7. Liquids + nasals /lm/ (detail given in Table 3.14)
8. Glides + liquids /wr, wɽ, wl/ (detail given in Table 3.15)
9. Affricates + plosives /tsk/ (detail given in Table 3.16)

It is clear that there are 22 types of consonant combination which violate the cross-linguistic phenomenon of sonority sequencing in Pashto. Some combinations are simple violations of SSP (e.g. no. 2, fricatives + fricatives, or no. 5, nasals + nasals) through equal sonority: SSP demands that every following segment must rise in sonority towards peak; this tendency is violated by such combinations. But at the same time there are combinations which are severely reverse in nature (e.g. no. 3, nasals + plosives, or no. 8, glides + liquids). Having demonstrated reverse sonority in initial consonant clusters, I will show the position of final and medial clusters not discussed by any researcher before.

Combinations in (52) also show that the language does not strictly follow the MSD (minimal sonority distance) principle. Some languages maintain MSD between two segments in a strict manner.[22] Some combinations of the same segments, such as fricatives + fricatives or fricatives + plosives, have the same value, 1 (given in Figure 3.3), as shown in the sonority scale. So the MSD is 0 in such clusters, and in reverse sonority sequences it is in minus. Under SSP, the coda is required to be more sonorant than the onset. However, I have just shown the falling sequence of sonority in coda clusters. The data shown in Tables 3.18 to 3.20 is simplified in the following lines. It is evident that word-final clusters show SSP. But in a few combinations it is also violated (e.g., plosives + fricatives /-bs, -ks, -bz, -qʃ/ and fricatives + fricatives /-xs/). Out of these five combinations, two words are originally Pashto (/əks/ *contrary* and /həbs/ *humidity*). The other three combinations are from words that are not of Pashto origin (/sərsəbz/ *lush green*, /nəqʃ/ *copy*, /ʃəxs/ *person who is Urdu by origin*), but these words are commonly used today. I can thus conclude that coda clusters exhibit SSP, but not strictly. Medial clusters fully show SSP, as is

[22] Spanish is a classic example of such languages which follows strict MSD. This value is 2 in Spanish (Harris, 1983).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

clear from Table 3.22 in the above section. Now, I turn to tripartite consonant clusters.

Table 3.25 shows that some of the tripartite initial consonant clusters do not abide by SSP. These violating combinations are given below.

(53)

1. /sxw-/
2. /ʃxw-/
3. /skw-/
4. /xkw-/
5. /ndr-/

These combinations show that some of the tripartite consonant clusters also violate the sonority hierarchy of segments, which is to be expected in the light of so many reverse clusters featuring in bipartite initial clusters in the language. These clusters are from words of Pashto origin, as in the case of the two-consonant clusters. The theoretical justification for such a pattern under OT theory will be presented in the following section.

3.8.4 Theoretical Analysis of Reverse Sonority Clusters in Pashto

As discussed in Section 3.7, OT is a linguistic framework which observes the continuous interaction between conflicting markedness and faithfulness constraints within a language. It has also been discussed that CV is the most common and universal syllable structure. Nonetheless, a language may violate this pattern due to parametric variations based on the hierarchy of universal components. So, there are languages which have a (C)V(C) pattern. In such a language, the faithfulness constraints dominate the markedness constraints, in the case that underlying material surfaces in coda position. Similarly, there are languages which not only allow onsets but also allow clusters at onset and coda positions. Like English, Pashto belongs to the latter class.

It was established above that Pashto consonant clusters violate sonority sequencing in a number of consonant combinations. For theoretical justification, let me show the constraint first.

(54)
SON-SEQ
Complex onsets rise in sonority and complex codas fall in sonority (Kager, 1999, p. 267).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kager (1999) shows this as a prosodic markedness constraint. It is a restriction on languages that words should not end with segments of rising sonority. Inputs which end with segments of rising sonority are repaired in a number of ways. The most common is the process of epenthesis (addition of a vowel). Pashto is very strict against such addition and gives priority to the anti-epenthesis constraint, DEP-IO. Another means of sonority repair is the deletion of the segment. Such deletion is punished by the constraint of MAX-IO, which is another high-ranking constraint in Pashto. No epenthesis and no deletion take place. As a result, the reverse sonority clusters are allowed in the language. Tableau (55) shows this interaction of three candidates for the input.

(55) /wɽəl/

Input: /wɽəl/	DEP-IO	MAX-IO	SON-SEQ
a.☞/wɽəl/			*
b. /wa.ɽəl/	*!		
c. /ɽəl/		*!	

Faithfulness constraints DEP-IO and MAX-IO take precedence over markedness constraints SON-SEQ. Candidate (a) has a consonant cluster of reverse sonority at the onset, while candidate (b) has no cluster because one segment is added (the epenthetic ə) and, as a result, two syllables are found. No cluster is present in candidate (c) because one segment from the input is dropped. So, the optimal output (a) violates the lowest-ranking markedness constraints SON-SEQ. The violation of constraint SON-SEQ is least expensive in this interaction process.

It is clear from the above tableau that the avoidance of epenthetic vowels and deletion of consonants take priority over avoidance of syllabic ill-formedness. In terms of constraint ranking, it is:

(56)
DEP-IO, MAX-IO ›› SON-SEQ

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

The case is the same with three-consonant clusters in the onset position of Pashto syllables. (44) shows the tableau for three-consonant clusters. In Section 3.7, I discussed in detail the violation of complex onset in the language. Here I include *COMPLEXONS and its equally ranked constraint SON-SEQ in the cover constraint *CLUSTER.

(57) /ʃxwənd/

Input: /ʃxwənd/	FAITH		*CLUSTER	
	DEP-IO	MAX-IO	*COMPLEX^{ONS}	SON-SEQ
a.☞/ʃxwənd/			*	*
b. /ʃənd/		***!		
c. /ʃI.xə.wənd/	**!			

Tableau (57) shows, again, that the faithfulness constraints take top position in the hierarchy. *CLUSTER constraints (*COMPLEXONS and SON-SEQ) are least important in this ranking and are violated by the syllables shown. The mutual ranking between *COMPLEXONS and SON-SEQ, and between DEP-IO and MAX-IO, is irrelevant in this context, but the constraint ranking given in (46) is relevant here too. Candidate (a) has both a cluster and reverse sonority in the onset. Candidate (b) has no onset cluster and therefore no reverse sonority in the onset because three segments from the input are dropped. Candidate (c) has epenthetic vowels and creates two more syllables in the word. So, the optimal output (a) violates the lowest-ranking markedness constraints in *CLUSTER (*COMPLEXONS and SON-SEQ). Their violation is least expensive in this interaction.

Reverse sonority is not only found in onset clusters but also in some of the coda clusters. In coda position, some of the consonants do not exhibit SSP, as shown below:

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(58) /əks/

Input: /əks/	FAITH		*CLUSTER	
	DEP-IO	MAX-IO	*COMPLEX^{COD}	SON-SEQ
a.☞/əks/			*	*
b. /ək/		*!		
c. /ək.i:s/	*!			

In (57), just as in (58), the faithfulness constraints are dominant over well-formedness *CLUSTER constraints (*COMPLEXONS and SON-SEQ). Candidate (a) has a coda cluster with reverse sonority. Candidate (b) has neither coda cluster nor reverse sonority in it. Candidate (c) has an epenthetic vowel, thus adding one more syllable. This hierarchy of the constraints allows not only complex codas but also the violation of SSP. All the tableaux discussed in this section generate the following factorial typology:

(59)
DEP-IO, MAX-IO ›› SON-SEQ, *COMPLEXONS, *COMPLEXCOD .
"Complex onsets and codas are allowed with reverse sonority"

In Pashto, the "anti-epenthesis" constraint DEP-IO dominates SON-SEQ because no epenthesis takes place. Similarly, the choice of (consonant) deletion also dominates SON-SEQ. The violation of SSP is also recorded in other languages. For example, Malayalam shows the violation of sonority sequencing (Mohanan, 1989). The language violates SSP in its onset clusters. A summary of the current section is given in the section below.

3.8.5 Summary

I have noted multiple violations of SSP in both bipartite and tripartite consonant clusters in Pashto. The language not only allows complex clusters but also violates the sonority sequencing principle. Faithfulness constraints are ranked higher and markedness constraints are less important. Similarly, SSP is not only violated in initial clusters but also in some of the coda clusters. Abandoning SSP (as a markedness constraint) and creating complex onsets and codas which violate it happens frequently in Pashto consonant clusters at word-initial and word-final edges.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Although SSP is a robust cross-linguistic property widely respected by most languages, it is nonetheless a language-specific parametric choice which can be violated, like other constraints. Pashto allows clusters at both edges of a word, onset and coda, and both types, observing and violating SSP.

I can confirm, now, that the maximum syllable cluster of the language goes up to (CCC), three consonants, at onset and (CC), two consonants, at coda. This confirmation is helpful for deciding other patterns of the syllable structure. Now it is imperative to set out the syllabification rules of the language with the help of the syllables discussed in the above sections. Once syllabification rules are framed, syllable constraints will be easily noted. Syllable constraints enable one to decide about the representation of the syllable structure in terms of theoretical representation. The next section discusses the syllabification rules.

3.9 PASHTO SYLLABIFICATION RULES

Syllabification plays a significant role in the phonology of a language. The rules of Pashto syllabification have not yet received proper attention. Not only the phonetic study but also the representative work in phonology is developed on the basis of so-called intuitive native-speaker knowledge. The main work on Pashto has been more or less neglected. The first and the only work, so far, on Pashto syllable structure is a short article by Zaman (2002) which addresses the issues at the surface level. One can find only very little emphasis on the syllabification rules based on the synchronic evidence. The article is just a brief discussion with fairly narrow scope of investigation. Apart from Zaman (2002), Tegey and Robson (1996) also point out a few syllable constraints which are discussed in this section.

The data shown in Section 3.6 and elsewhere in the current study was taken from the Pashto dictionaries and my collection of 5000 words. This provides enough information regarding the syllabification rules of the language. Therefore, I try to pinpoint Pashto syllabification rules which are based on the phonotactic rules mainly found in my data. Moreover, previous studies are also referred to where required. These rules cover the points relating to consonant clusters only:

- Based on the analysis of data in the current study, it is argued that syllabic consonants are not allowed in Pashto. This constraint is shown as (60).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(60)
Syllabic Consonant Constraint
* V [+ consonantal] No syllabic consonant allowed.

This means that nothing other than a vowel can bear the feature of syllabic [+syll] i.e., the ability to be the peak of a syllable.

- Pashto syllables allow codas, which is a significant structural property of the language.

All consonants and consonant sequences form part of the onset in the syllabic templates of the language. Detail of possible clusters is given in Section 6.3.

- The most preferred syllable pattern in Pashto is CV, which is described as an unmarked syllabic pattern by Jakobson (1972). This pattern is not only the basic one but also the most natural syllable pattern in Pashto.
- Any unitary V+V sequence (i.e., the true diphthong) never consists of a first member which is higher than the second and is interpretable as being [-syll].
- Normally, a semi vowel (/j, w/ or /i, u/) can occur as the second member of a diphthong.

Very rarely can another vowel become the second member of it.

- Glides can also occur as word-initial consonants.
- The word-initial consonant clusters become the onset of the following syllable. For example, in words like /zməka/ (earth/land) and /ʃmer/ (counting) the clusters of /zm/ and /ʃm/ are the onsets to the peak in their respective syllables.
- Similarly, the word-final consonant clusters become the coda of the preceding syllable.

For example, in words like /tərs/ (pity) and /tərz/ (form), the clusters of /rs/ and /rz/ are the codas to the peak in their respective syllables.

- The language follows some of the rules related to syllable structure. It follows the principle of maximum onset (MOP).
- It is also clear from the data that in the case of bipartite consonant clusters at word-initial level, the language shows some acceptance of the sonority sequencing principle (SSP). But multiple violations

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

of SSP are also recorded at word-initial onsets (also reported by Bell & Saka, 1982; Levi, 2004; Rehman, 2009). The current study also confirms the SSP violation in the onset and finds some violations in the coda clusters as well.

- In majority of the cases of word-final position, SSP is observed.
- At word-medial position, some violations of SSP are also found.
- It is clear from the tables on bipartite consonant clusters that the second consonant in such a combination must be a glide.
- The maximum cluster possible at onset is three consonants, and at coda two consonants (CCCVCC), as reported by Tegey and Robson (1996).
- In all three-consonant syllables, one out of three is a /w/ glide. There is only one word, /ndror/ (sister-in-law), in which this /w/ glide is not found and that has /n/ as the first sound in it.
- Based on the observations from the data, the following 12 syllabic patterns can be established as the canonical syllable patterns in Pashto in terms of consonant complexity, arranged in descending order of frequency (also reported by Tegey & Robson, 1996): V, CV, VC, CVC, VCC, CCV, CVCC, CCVC, CCVCC, CCCV, CCCVC, CCCVCC. The last three patterns (i.e., CCCV, CCCVC, CCCVCC) are not very frequent.
- These 12 possible shapes range from a single vowel (V, e.g. /o/ *sirrah*) to a maximum of three consonants in the onset and two in the coda (CCCVCC, e.g. /ʃxwənd/ *chewing*) (also reported by Tegey & Robson, 1996).
- With respect to the syllabification of medial consonants, a single intervocalic consonant always syllabifies with the following vowel both word-internally and across a word boundary.
- Thus a variety of possible syllable structures are possible and most consonant sounds are allowed at onset position.
- There is one crucial fact about Pashto syllabification: that intervocalic singletons are always syllabified in the onset together with the following vowel.
- *Geminate is ranked higher in the language. Therefore, real geminate representation is not available to surface form in the language.
- The Pashto syllable prefers an obstruent to be the first consonant and an approximant as the second consonant in a two-consonant cluster, but other combinations are also possible in this position. The data shows that in sonority-observing bipartite consonant clusters, the first consonant is an obstruent. This rule has only four exceptions, in which the sound could be /r, l, m, n/. Similarly, the

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

second consonant in such clusters is an approximant, with only three exceptions, / m, n, ɣ /. (51) shows the detail of such combinations.

- In reverse sonority clusters, any segment can take the position of first consonant in a two-consonant cluster (obstruent + nasals + liquids + glides). The second consonant is an obstruent in such clusters with the exception of nasals and liquids. In the case of nasals or liquids, the first consonant must be either a nasal or a glide. See (52) for details.
- In two-consonant SSP clusters, the consonant towards peak is always nasal or liquid. The second consonant must be an obstruent in such arrangement. There is only one exception to this rule, in which the second consonant is nasal. But in this case, the consonant adjacent to peak must be a liquid, which is higher than a nasal in sonority.
- In coda clusters violating SSP, both consonants must be obstruent.
- Single vowel syllables can also start words. For example, /a.na:r/ *pomegranate* or /aw.la:d/ *offspring*.
- Phoneme /h/ can only come at the start of a syllable, e.g. /həlwa/ *sweets* or /wa.hal/ *to beat.* The same was also reported by Tegey and Robson (1996).
- Glide /w/ can also take the position of the first consonant in an onset cluster which violates SSP in such combinations. See (52) for details.
- Previously Tegey and Robson (1996, p. 19) reported that sounds /ɳ/ and /g/ can only occur at the end of syllables, as in /ʃku:ɳ/ *porcupine* or /jəg/ *beer*, but the current study does not attest such constraints in the language. Both these phonemes can be found at the start of syllables, as in words like /gi:rə/ *beard* or /rə.ɳa:/ *light*.
- In the above rules, the focus of the analysis is consonant clusters in Pashto.

Therefore, the study is restricted to them and the rules related to vowels are not discussed in detail, simply for the reason that such analysis is outside the scope of the study. In respect of consonant clusters, the following section describes the maximum syllable template.

3.10 THE MAXIMUM SYLLABLE TEMPLATE IN PASHTO

Through the examination of data presented in Section 3.6 and the subsequent discussion on consonant clusters, I find that the maximum

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

word-initial consonant cluster in Pashto consists of three consonants (CCC-). In the word-final position, the maximum is two (-CC). These combinations indicate that the maximum syllable template is CCCVCC.

(61)
Maximum syllabic template (MAXST) in Pashto: CCCVCC

The maximum syllable consists of three consonant in the onset and two in the coda. The onset and coda may both violate the SSP. Given the MAXST of the language, I will conclude the chapter in the next section.

3.11 SUMMARY AND CONCLUSIONS

The current chapter started from the discussion of Pashto phonemes, based on previous literature. This was followed by a detailed analysis of syllable structure. I discussed both bipartite and tripartite consonant clusters in the language. From this discussion, I conclude that some 108 combinations of two-consonant clusters at all three positions of the word are found. The largest number of clusters were found at initial position. This is followed by final position, and the smallest number of clusters are found in word-medial position. Similarly, eight types of combinations are found in tripartite clusters in Pashto, all of which are at initial position.

The conclusion from the discussion on consonant clusters is that Pashto not only permits various types of complex consonant clusters but it also tolerates a number of sonority-violating combinations. The following sections provide the theoretical justification for such complex clusters allowing even reverse sonority under the OT framework. It is found here that Pashto gives priority to faithfulness constraints (such as DEP-IO and MAX-IO) over well- formedness constraints (such as *COMPLEXONS and *COMPLEXCOD). Complex syllable structure and the violation of SSP are justifiable in such a hierarchy of constraints. Based on the data, the syllabification rules are given at the end.

Relating to the representation of Pashto syllabic complexity, it is established from the data that the maximum range of the syllabic template (MAXST) is CCCVCC. This syllable structure provides a very interesting point for the discussion on the main subject of the study, i.e. the correlation between syllable structure and word order. The word order of the language is the topic of the next chapter.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

CHAPTER 4

BASIC WORD ORDER OF PASHTO

4.1 INTRODUCTION AND OVERVIEW

In the previous chapter, I established with the help of my data the maximum syllable template (MAXST) for Pashto. The MAXST makes possible predictions about the consonant clusters at Pashto word edges. The present chapter describes the basic word order of Pashto. The main focus of the chapter is to illustrate the word order patterns of the language in the light of the theoretical assumptions of OT (Grimshaw & Samek-Lodovici, 1998; Costa, 1997c, 1998, 2001). It also discusses the constraints used in determining the word order pattern and their ranking hierarchy for Pashto. In order to determine the basic word order, the initial focus is on the structure of simple sentences in the language.

The chapter starts by introducing the main contribution of the position of the verb in establishing the word order of a language. It provides an overview of past scholarship on the topic of syntax in Pashto and highlights important contributions in this regard. Then it goes on to analyse the basic word order of Pashto using the OT framework as a methodology. In Chapter 5, I will apply the findings of the present chapter in order to test the validity of the implicational correlation under discussion, that is, the correlation between word order and syllable complexity.

It should be kept in mind that this chapter is by no means a complete analysis of Pashto syntax. It only covers the portion of basic word order which is relevant to the subject correlation. The main goal is to provide a better understanding of basic word order patterns of the language in light of the implicational correlation and highlight the most crucial aspects of it.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.2 BASIC WORD ORDER (BWO)

Chapter 2 provides a discussion regarding basic word order (BWO) in languages. Just a short recap is given here. BWO is, in simple words, the order of basic constituents in a simple sentence. Thisa means the arrangement of the most basic syntactic elements within a simple declarative sentence (having a single verb) in a language. According to some linguists (among others, see Li, 1976; Dik, 1978; Costa, 2001), BWO means the order which is taken by sentences in out-of-the-blue contexts or within a specific contextual sentences in which the question "what happened?" is answered. In such an arrangement, the role (place) of the *verb* is particularly important: BWO is determined on the basis of its position.

Linguists (e.g., Greenberg, 1963, 1966; Lehmann, 1971, 1978; Vennemann, 1972; Hawkins, 1983) agree upon the point that it is the verb which has the most basic role among the syntactic elements. Its relationship with its *object* is considered the most fundamental thing in the syntactic classification of a language. Lehmann (1971) talks about the following two BWOs:

(62)

a. VO (verb + object) Examples: English, European and Semitic languages
b. OV (object + verb) Examples: Japanese, Turkish and the Dravidian languages of India

But there are some linguists who opine that there is another important element in the syntactic constituents: the *subject*. Such linguists think that these three constituents together define the BWO of a language. So, if I put all these three elements together, the following are the two most frequent orders found within languages of the world (Gell-Mann & Ruhlen, 2011). Traditional typology mainly treats these two patterns of basic word order. They are given in (63).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(63)

d. Subject + verb + object (SVO) Example: English
e. Subject + object + verb (SOV) Example: Japanese

Linguists such as Derbyshire and Pullen (1979), Hawkins (1983), Brody (2009), and Gell-Mann and Ruhlen (2011) give six logically possible orders of these three elements in different language families. (64) lists all of these logically possible BWOs.

(64)

a. Subject + verb + object (SVO)
b. Subject + object + verb (SOV)
c. Verb + subject + object (VSO)
d. Verb + object + subject (VOS)
e. Object + verb + subject (OVS)
f. Object + subject + verb (OSV)

As far as the frequency of these orders is concerned, the first two, as discussed earlier, are the most common, followed by the third one. The remaining three orders are rare, but since they are logically possible they are given here. The current study is related to two types of BWO, i.e. SOV and SVO which are given in (63). This categorization suits the study because the languages (Pashto and English) included in the study instantiate these word orders.

The variations in the BWO of languages are a reflection of the fact that languages vary at surface level. BWO actually reveals the discourse-configurational character of a language. It shows language-specific tendencies and, therefore, BWO is considered very important by language experts. In typological linguistics, many things are decided on the basis of it. In Chapter 2, I discussed at greater length the fact that BWO is the main source of language classification. The implicational correlation which I focus on in this study is also based on the word order of languages. Chapter 3 shows that Pashto has complex clusters in its syllable structure. According to our implicational universal, it is therefore predicted to have an SVO word pattern. But the data in the current study shows something different. It shows that the BWO of Pashto is contrary to the expectations of the implicational universal. What is shown by the data collected in previous works is the subject matter of the next section.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.3 PREVIOUS WORKS ON PASHTO WORD ORDER

The syntax of Pashto has fascinated a number of scholars in the past. The current section gives an overview of past works specifically concerned with the BWO of Pashto. A number of studies have been carried out, ranging from a simple reference grammar on the language (e.g. Tegey & Robson, 1996) to a specialized analysis of Pashto syntactic patterns (e.g. Roberts, 2000, on Pashto clitics). Here I will discuss the important works on Pashto grammar in general and its word order in particular.

Penzl (1955) has written a short grammar of Pashto. Though mainly exploring phonological aspects of the language, it also covers the syntactic detail of Pashto. In addition, it discusses various grammatical properties of the language, such as case-marking and the system of gender assignment. Shafeev (1964) is another work which is among the first completed on Pashto syntax. It discusses various grammatical properties of the language, such as the order of syntactic elements in sentences. This study is the data source for a number of entries in WALS. It discusses various word orders in the language: for example, the order of genitive and noun, the order of subject and verb with object, the order of numeral or relative clause and noun, etc. Shafeev (1964) concludes that Pashto is an SOV language, as demonstrated by the order of the syntactic constituents in his examples.

Tegey and Robson (1996) assign Chapter 11 of their reference grammar to the sentence structure of Pashto. They describe the structure of simple sentences in this chapter and discuss various conditions related to them. They highlight the structure of syntactic elements such as noun phrases, pronouns, adjectives and prepositional phrases along with their modifiers in the language. In discussing elements of verb phrases, they talk about negative and positive verb phrases, verb particles and clitics. They also give a summary of agreement rules in their discussion on Pashto ergative construction and their presentation of some unusual constructions.

According to them, the basic word order of Pashto is SOV. To demonstrate this order of the syntactic elements, they give a number of examples. Some of them are given in the following section.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(65) (Present progressive)

a. asad	li:k	legi
S	O	V
Asad	*letter is*	*sending* (3P)
Asad is sending a letter.		

b. pla:r me	jaw loe ba:ɣ	axli:
S	O	V
Father my	*a big orchard*	*is buying* (3P)
My father is buying a big orchard.		

(66) (Interrogative)

tə	tsə ka:r	kawe
S	O	V
You	*what work*	*do* (2PQ)
What work do you do?		(Tegey & Robson, 1996, p. 165)

They argue on the basis of their analysis that in Pashto, contrary to English, all modifiers precede the verb in a sentence (p. 166). They give examples such as the following in the favour of their claim.

(67)

de aman korəne	pə ka:bəl ke	wosegi
S	modifier	V
Of Aman family	*in Kabul in*	*lives* (3P)
Aman's family lives in Kabul. (p. 166)		

They also discuss the possibility of the addition or deletion of subject and object (or weak pronouns) in certain kinds of sentences (such as imperatives) and submit that the structure is the same for both statements and questions in Pashto. All this discussion leads to the conclusion that Pashto is an SOV language and that its verb occurs as the final element, at the end of the sentence. Next, I discuss the work of Roberts (2000) on Pashto syntax.

Roberts (2000) focuses mainly on Pashto clitics, but he also discusses word order at the start of his dissertation. According to him, the basic word order of the language is SOV. He claims that "Pashto is a fairly rigidly head-final language" (p. 11). He also gives some examples to support his claim, some of which are given in (68).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(68)

a. surgwəl	kita:b	nəwiʃtakəi
S	O	V
Sur Gwel	*book*	*write do*
Sur Gwel wrote a book (Kandahar)		

b. spi	pi:ʃo	xu:gkɽa
S	O	V
Dog	*cat*	*hurt did*
The dog hurt the cat.		(Roberts, 2000, p. 11)

Roberts mentions other possible patterns of word order in sentences (OSV), and declares them not only unnatural but also subject to a heavy pause (topicalization) after the direct object and occurring only in certain specific contexts. OSV is possible with the help of cliticization, but only in the past. For the present and other tenses only SOV order works, and this is therefore the basic word order of the language. Like Babrakzai (1999, p. 14) and Tegey and Robson (1996), Roberts also confirms that the verb always occurs at the end of the sentence. As well as sentences, he also analyses lexical categories such as the noun phrase and finds that these are also head-final.

WALS maintains a large database on the structural properties of world languages, gathered using descriptive materials. It currently holds 76,492 data points for 2678 languages. On the order of verb and object, it provides data for 1519 languages. Reporting on the syntactic structure of Pashto, WALS provides more than 144 entries. These entries on the syntactic properties of the language are based mainly on the work of Shafeev (1964). The main points relating to Pashto BWO are summarized below.

(69)

Order of subject, object and verb:	SOV
Order of subject and verb:	SV
Order of object and verb:	OV

Another website based on data of the syntactic structure of languages is the Syntactic Structures of the World's Languages (hereafter SSWL).[23]

[23] http://sswl.railsplayground.net/browse/languages/Pashto (accessed on 26 September 2011)

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

This is a searchable database which allows researchers to discover the syntactic and other properties of languages and how these properties work across languages. For Pashto, the property of subject + object + verb has the value of "yes" because this order is used in a neutral context. The data is added by Roberts, whose work was discussed above. The clause which is used to demonstrate this property is an active declarative clause.

Most of the above-mentioned works are based on different varieties of Pashto, such as the Kandahari and Kabuli dialects of the language. This discussion shows that Pashto belongs to the SOV group of languages. On the point of word order, there is no variation among the dialects of Pashto: it is same in all varieties of the language. In the following lines, I provide my own data on the Yousafzai dialect and discuss what this suggests about the BWO of the language.

4.4 AN OPTIMALITY-THEORETIC ANALYSIS OF PASHTO WORD ORDER

The above discussion has made it clear that the word order of Pashto is SOV. Due to the variation of postulating norms regarding the word order of a language, it is difficult for one to suggest that there is only one possible word order. In reality, one can only speak of the word order which is strongly dominant and is most frequently observable in the usage of native speakers. Therefore, what is presented in this study regarding the basic word order of Pashto is based primarily on the native speaker's intuition, and past scholarship is given in evidence of the phenomenon. Recorded sentences are used which were spoken by native speakers in different contexts. Two main principles were applied in seeking to identify BWO patterns. These are:

(70) a. Sentences in out-of-the-blue contexts
b. Specific contextual sentences answering the question "What happened?"

Sentences based on the above criteria are explored in the current chapter. Before going into the detailed analysis of BWO in Pashto, it is appropriate to introduce the functions and mechanisms of OT framework which are applied in this study for the analysis of word order.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.4.1 Word Order in OT

I discussed in Chapter 1 that the organization of word order (syntax) in OT is not different from that of phonology. OT as a theory of language variation has been increasingly applied in the field of syntax in recent times (e.g. Grimshaw & Samek-Lodovici, 1998; Costa 1997c, 1998, 2001; Bresnan, 2000; Choi, 1999; Grimshaw, 2001; Sells, 2001, among others), and it has been shown that word order variations are not optional but are the result of various functional specifications in the input. The application of OT in this chapter has two main advantages. Firstly, it provides a possible architectural explanation for the type of BWO employed in Pashto by ranking the set of concerned constraints of the OT framework in syntax. Secondly, OT answers a number of important questions relating to the analysis of implicational universals on the subject (i.e. the relationship between word order and syllable structure), on the basis the findings of the current chapter. Moreover, according to OT, the variation among languages is actually the sample of the typological space of possible grammars. Therefore, I first give a brief introduction to the architecture of OT in syntax.

Here, I restrict the application of OT to determining the word order of the language. Irrelevant detail is avoided simply in order to save time and space. This is followed by presentation of the BWO found in Pashto, along with other variations, and is justified under the theoretic framework of grammar in OT. Finally, the conclusion of the chapter summarizes the findings about the language.

According to the commonly held view of OT (Prince & Smolensky, 1993), the grammars of individual languages consist of a categorically set ranking of constraints which dominate one another, and this hierarchy of constraints is based on language-specific tendencies. According to Grimshaw (1997), the primary idea of OT lies in the points given as (71):

(71)

1. Constraints are universal.
2. Constraints can be violated.
3. Grammars are rankings of constraints.
4. The optimal form is grammatical; all non-optimal candidates are ungrammatical. (p. 373)

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

This shows that while one constraint may be given priority in one specific language, the same constraint may be easily violated in another language. In relation to syntactic order, OT suggests that in the input, the particular group of words in a sentence provide for their possible structures as output candidates. The GENerator provides a handful of potential candidates to an input which are then evaluated by EVAL. The winning candidate is chosen by the EVALuation procedure (Speas, 1997, p. 178).

The ordering of the syntactic elements which target a specific position in a clause has been modelled by a number of recent studies in syntax. The positioning of elements such as subjects, verbs, adverbs, topics and operators has been successfully represented by extending the mechanism of Generalized Alignment (McCarthy & Prince, 1993) to the domain of syntax (e.g. Sells, 1999, 2001; Morimoto, 2002). The main idea of Generalized Alignment is applied to the placement of syntactic elements within a clause. The idea is based on restricting the alignment to a specific edge of the clause within the constraint-based, output-oriented model of grammar. Within head-initial (VO) languages, the structural position of the elements (object and focus, etc.) is after their head, as they follow the head, while in head-final languages (OV), the head occurs at the end of the clause, with the other syntactic elements preceding it. These positions are based on the hierarchy of constraints among the syntactic elements within a language.

Bearing in mind the above points, the OT framework is used in the current section to analyse the word order of the focus language. However, it must be kept in mind that the current study is concerned only with the primary word order of Pashto or with that portion of the syntactic patterns which is related to the subject matter of the correlation under discussion.

4.4.2 Constraints on Basic Word Order

In order to determine the BWO of Pashto under OT, I have adopted the constraints relating to clause structure which are proposed by Grimshaw (2001) and others, as mentioned already. As discussed at the start of the current section, I am concerned only with BWO. This means that I only include sentences which consist of simple main clauses, are of declarative nature and have all the basic three elements (subject + verb + object). Therefore, I only include the specific OT constraints which are relevant in such an analysis for deriving the typology of BWO in Pashto. It is also made explicit that, for the current purpose, the sentences included are declaratives. All the possible patterns of such sentences are covered within

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

this discussion. The interaction of, and hierarchy among, the following constraints is analysed in order to justify theoretically the SOV order of Pashto.

(72) Constraints on clause structure:
a. HD-RT: The head is rightmost in its projections (Grimshaw, 1997).
b. HD-LT: The head is leftmost in its projections (Grimshaw, 1997).
c. SUBJ-LT: A subject aligns left in the clause (Morimoto, 2001).
d. SO: Subject proceeds object (Choi, 1999).
e. STAY: Trace is not allowed (Economy of Movement) (Grimshaw, 1997).

Before exploring the BWO of Pashto, the above-mentioned constraints are introduced in greater detail. These constraints illustrate some of the fundamental variations in OT. HD-RT and HD-LT are in continuous competition with each other. Two rankings are possible between them. If HD-RT outranks HD-LT, the head of a projection (e.g. V within a VP) is in the rightmost position (at the end of the projection). If the case is reversed, then it is in the leftmost position (at the start of the projection). SUBJ-LT is relevant for the placement of a subject within a clause. It punishes the positioning of a subject in anything other than the left position of the clause. This is relevant here because I have included the element of subject within the basic three syntactic elements.

The constraint SO makes it possible for the subject to precede the object. It is included here in this list of constraints because the interaction between subject and object is also important in the study of the BWO pattern. Finally, STAY is added to the current list. It is a constraint of minimal projection. It is violated when any of the syntactic elements is moved from one position to another. STAY is relevant to the VP-internal subject hypothesis highlighted by a number of scholars (Zagona, 1982; Kitagawa, 1986; Koopman & Sportiche, 1991; McNally, 1992; Burton & Grimshaw, 1992). According to this hypothesis, the subject always originates within the VP. Then, in both declarative and interrogative sentences, it is raised from Spec-VP to Spec-IP. In such cases, the constraint STAY is violated.

Now, to illustrate how these constraints interact with one another in the hierarchic grammar of Pashto, let me consider the interaction of the above-mentioned constraints in the next section.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.4.3 The Analysis

The discussion in Section 4.3 shows that Pashto exemplifies the SOV structure as its BWO. Now, before the OT analysis, I give examples which are then further analysed in the light of OT. The current work is based on the Yousafzai dialect; therefore, I add sentences spoken by speakers of this variety of the language. This makes the claim more patent in its nature. Now, as discussed in (70 b), I give the first answer to "What happened?" in the following example.

(73)

a.	maʃu:m	pəlet	ma:tkɽo
	S	O	V
	Child	*plate break-did*	
	(MASC SG)	(PAST PERF)	(M3SG)
	The child broke the plate		

The verb in Pashto is always at the final position. Even sentences with complements also have S(ubject) preceeding V(erb). Consider example

(73 b):

b.	brəg	stəɽe stomana	raɣe
	S	C	V
	Brag	*very tired*	*come-did*
		(PAST PART)	(M3SG)
	Brag became very tired		

Examples (73 c–e) show the present progressive form:

c.	pla:r	srəd	r**I**bi:
	S	O	V
	Father	*grass cut*	
	(MASC SG)	(PRESENT PROG)	(M3SG)
	The father is cutting the grass		

d.	/zwana:n/	/xka:r/	/la rəwa:n di:/
	S	O	V
	Young men	*hunting for going are*	(M3PL)
	(MASC PL)	(PRESENT PROG)	
	The young men are going hunting		

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

e. wra:rə spe lobawi:
S O V
Nephew *dog play-do* (M3SG)
(MASC SG) (PRESENT PROG)
The nephew is playing with the dog

Let us show some other examples of simple declarative sentences.

(74) ba:ʧa ki:ta:b wa:i
S O V
Bacha *book* *read-do* (TRANS IMPERF) (3SG)
Bacha reads a book

(75) mor maʃu:m xkol kʈo
S O V
Mother *baby* *kiss-did* (PAST PERF) (M3SG)
Mother kissed the baby

(76) gwəl pətəŋ əluzawi:
S O V
Gwal *kite* *fly- do* (TRANS IMPERF) (3SG)
Gwal is flying the kite

Roberts (2000) also discusses the possibility of the order transforming into OSV in some contexts and if prosodic conditions are fulfilled. Sometimes the SOV order is changed into OSV, but only in cases where it is demanded by the context and there is a heavy pause given after the first word, i.e. the object. Such order is subject to cliticization and topicalization. Secondly, this freedom is only possible in past tense and is not possible in present tense. The following examples will clarify the possible use of OSV order:

(77)
ki:ta:b, ba:ʧa waele-de
Book Bacha read (PAST PERF)
Bacha read the book

Topicalization and the heavy pause shown here by the comma are mandatory. Without these, the meaning would be changed.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(78)

ki:ta:b ba:tʃa yee waee

Book Bacha (3SG) *read* (PRESENT IMPERF) (3SG)

Bacha reads the book

Example (78) shows that the clitic /yee/ co-refers to the left-dislocated nominal /kitab/. From the above examples, I can conclude that Pashto language has SOV word order as its basic order.

In all the above examples, the idea of Pashto being an SOV language is made explicit and reconfirmed. These sentences show SOV as the canonical word order of the language. Now, before exploring some other possible patterns of words within clauses in Pashto, I will analyse the above structure within the OT theoretic framework. For this purpose, I illustrate the following tableau which shows how the constraints discussed in 4.4.2 interact with each other.

(79) *maʃu:m pəleţ ma:t kţo* The child broke the plate

pəleţ maʃu:m ma:t kţo	HD-RT	SO	SUBJ-LT	STAY	HD-LT
a. *ma:t kţo maʃu:m pəleţ*	*!		*		
b. ☞ *maʃu:m pəleţ ma:t kţo*				*	*
c. *pəleţ ma:t kţo maʃu:m*		*!			*

The above tableau illustrates the basic rules of OT. Each of the candidates, including the optimal one, violates some of the constraints. The optimal one, which is marked with the pointing finger, also violates two of the constraints – STAY and HD-LT – but is still the ideal one. The other candidates violate other constraints which are higher-ranked than STAY and HD-LT. The tableau also shows the highest ranking of the constraint HD-RT. Any candidate which violates this constraint will fatally lose the position of optimal candidate. Some other rankings are also crucial here. SUBJ-LT dominates STAY; otherwise, it would not be possible for the subject to be in the leftmost position of the clause. The constraint SO is also very important, making it possible for the subject to precede the object. In a declarative sentence like the above, the constraint SO is always satisfied.

So, of these three possibilities – candidates (a), (b) and (c) – candidate (b) is the best one. Though it violates the STAY and HD-LT constraints it is

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

still the optimal choice, as it satisfies the top-ranking constraints HD-RT, SO and SUBJ-LT. The violation of STAY has the benefit of allowing SUBJ-LT to be satisfied. As discussed earlier, under the VP-internal subject hypothesis, the subject is raised from the Spec-VP position to Spec-IP position in a declarative sentence, satisfying SUBJ-LT but violating STAY. These three elements (subject, object and verb) are analysed in the above-mentioned candidates and their possible three patterns are given (SOV, VSO and OVS), out of which SOV is taken as the optimal one.

The tableau (79) also shows that constraint HD-RT is the most important constraint and that it is at the top of this hierarchy of syntactic constraints in Pashto. In the following lines, I will represent an analysis of some relevant sentences to confirm the top position of this constraint. Note the following example.

(80)

maʃu:m	mor ta	pəleʈ	war kṛo
S	IO	DO	V
Child	*mother to*	*plate give-do*	(PAST PERF)

The child gave the plate to its mother

The word order is considered by Roberts (2000) to be somewhat flexible in ditransitive sentences such as the above. For such sentences, six types of ordering (given as 81) are possible.

(81)

a. *S* maʃu:m	*IO* mor ta	*DO* pəleʈ	*V* war kṛo
b. *S* maʃu:m	*DO* pəleʈ	*IO* mor ta	*V* war kṛo
c. *IO* mor ta	*S* maʃu:m	*DO* pəleʈ	*V* war kṛo
d. *IO* mor ta	*DO* pəleʈ	*S* maʃu:m	*V* war kṛo
e. *DO* pəleʈ	*IO* mor ta	*S* maʃu:m	*V* war kṛo
f. *DO* pəleʈ	*S* maʃu:m	*IO* mor ta	*V* war kṛo

All the above six options are possible for the sentence shown. The only similarity among all of them is the occurrence of the verb in the final position. It is evident from these examples that the constraint HD-RT is at the top position of the hierarchy of these constraints. Apart from this final verb position within such examples, there are some conditions which make possible the above choices. Firstly, it is only possible within ditransitive

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

verbs. Secondly, the grammatical function of the arguments must be clear from the context. Thirdly, the relevant case marking is also very important, as the particle ta, along with the IO shows the position of the IO, wherever it is. Therefore, on the basis of such conditions, one cannot declare this relaxation a canonically possible one. The canonical order is the one which is shown as candidate (b) in tableau (79) and (81 a) above (SOV). Roberts (2000) also talks about the possibility of relaxation of word order in sentences like the above. To such comments, I would say that it is only possible due to the presence of case marking in the IO ta and in a specific and pragmatically possible context. On one hand the case marking shows the position of the IO and, on the other, the position of the DO makes it clear in the context that it is the object which is handed over.

Another point arises from the above discussion about the mutual ranking of constraints SO and SUBJ-LT. The analysis of sentences in (81 a–f) challenges the necessity of their mutual ranking in the hierarchy of these syntactic constraints in Pashto. The relaxation of the pattern of subject and object (as reflected in these examples) throws doubt on the mutual ranking of SO and SUBJ-LT. So, the question arises: does one really need their mutual ranking in the hierarchy, or are they equally ranked? The simple answer to this question is yes. Their mutual ranking is required as this relaxation is possible only in the past tense. In other tenses, this relaxation is not possible. Note the following examples given by Roberts (2000, p. 12).

(82)

a: spi	piʃo	xu:g	kɽa
S	O	V	
Dog	*cat* (DIR FEM)	*hurt-do* (PAST PERF)	(F3SG)
(OBL MASC)			
The dog hurt the cat			

b: piʃo	spi	xu:g	kɽa
O	C	V	
Cat	*dog*	*hurt-do* (PAST PERF)	(F3SG)
(DIR FEM)	(OBL MASC)		
The dog hurt the cat			

Roberts says that both the above orders (82 a and b) are possible in past tense. My analysis is that this possibility is based on two things: first, the case marking between the noun spi and verb kɽa and, second, the

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

contextual meaning of the sentence. Both these points make the sentence possible. From these two positions, it is clear that the doer of the action is a male (dog) as its gender is reflected in the name of the subject spi. Similarly, the case marking of the verb kɽa shows the female gender of the object (patient) piʃo. If the action was reversed, then the sentences would have been as follows.

(83)

a: spe	piʃo	xu:g	kɽo
S	O	V	
Dog	*cat* (OBL FEM)	*hurt-do* (PAST PERF)	(M3SG) (DIR MASC)
The cat hurt the dog			

b: piʃo	spe	xu:g	kɽo
O	S	V	
Cat	*dog* (DIR MASC)	*hurt-do* (PAST PERF)	(M3SG) (OBL FEM)
The cat hurt the dog			

It is clear from (83 a and b) that in such sentences, the case marking is important and the order of subject and object is not really important. The difference between (82) and (83) is only that of the case marking (noun spi››spe and verb kɽa››kɽo.[24] But as Roberts (2000) himself points out, such a possibility is not available in present tense. The same sentence does not have this OSV order in present tense. Note the example given by the same source:

(84)

spe	piʃo	xu:gawi:
S	O	V
Dog (DIR)	*cat* (DIR)	*hurt* (TRANS IMPERF) (3SG)
The dog is hurting the cat		

[24] Since I am not concerned with such semantic interpretation or cliticization of the language in the current study, I restrict my analysis to the basic word order pattern of Pashto. For Pashto clitics, Roberts (2000) and Rehman (2012) provide very interesting and detailed insight.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

In the present form, this sentence has this (SOV) order and cannot be given the OSV order. With OSV pattern, the meaning would be totally changed: the sentence would become: *"The cat is hurting the dog". The reason is that both of these words (spe and piʃo) take the same case marking in the present tense, wi:. Let me show the same in a sentence in a tableau.

(85) *spe piʃo xu:gawi:* The dog hurts the cat.

piʃo spe xu:gawi:	HD-RT	SO	SUBJ-LT	STAY	HD-LT
a. *piʃo spe xu:gawi:*		*!	*		*
b. *xu:gawi: spe piʃo*	*!		*		
c. *xu:gawi: piʃo spe*	*!	*	*		
d.☞ *spe piʃo xu:gawi:*				*	*

Tableau (85) now clarifies the whole situation. It illustrates that both SO and SUBJ-LT are needed. Consider candidate (a), piʃo spe xu:gawi: (in fact, OSV pattern), which means "the cat is hurting the dog". Now it is totally contrary to the input because it fatally violates the constraints SO and SUBJ-LT. Both of these constraints are ranked at a higher position than STAY and HD-LT. This candidate does not violate STAY but it cannot be the optimal one because it violates the other higher constraints. This candidate is representative of OSV order. It is clear from the violations committed by such a pattern that it is unacceptable and meanings are totally changed in such a sentence.

Candidate (b) is unnatural as it also violates the top-ranking constraint HD-RT, which is a fatal violation. There is another crucial point regarding this candidate: it makes the ranking of candidate SUBJ-LT very clear. Although this candidate satisfies the constraint of SO, it does not satisfy SUBJ-LT, which is also fatal. So it is clear that both of the constraints (SO and SUBJ-LT) must be satisfied in order for either of them to maintain their higher-ranking positions within the hierarchy. Though this candidate satisfies both of the lower-ranking constraints, STAY and HD-LT, it cannot be the optimal candidate as it does not satisfy the other important constraints. Candidate (b) is representative of VSO order, and it is clear from its violations that it is also unacceptable in Pashto.

Now turning to candidate (c), it violates all the fatal constraints and, therefore, is the most unnatural one. Though it satisfies STAY and HD-LT,

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

these are low-level constraints and are not important in the hierarchy of constraints. This candidate is representative of a VOS pattern, and it shows that there is no place for such a pattern. The optimal candidate is candidate (d). It violates two of the constraints, STAY and HD-LT, but it is the optimal candidate because it satisfies the high-ranking constraints HD-RT, SO and SUBJ-LT. The violation of constraint HD-LT makes it possible for constraint HD-RT to be in the top position, SO makes it possible for subject to precede object, and SUBJ-LT places the subject in the leftmost position of the projection. Similarly, the violation of STAY benefits the leftmost position of SUBJ-LT. This candidate is representative of the SOV pattern, which fully fits the hierarchy of constraints in Pashto.

In tableaux (79) and (85), the optimal candidates for present and past tenses are discussed. Both of these tableaux show that SOV is the canonical word order of the sentences. As far as future is concerned in Pashto tenses, it is realized by putting second position clitic ba (Roberts, 2000, p. 42) next to the subject in a sentence. Look at the following example.

(86) *maʃu:ma:n ba ki:tabuna wa:i* The children will read the books.

ki:ʈabuna maʃu:ma:n ba wa:i	HD-RT	SO	SUBJ-LT	STAY	HD-LT
a. *wa:i maʃu:ma:n ba ki:ʈabuna*	*!		*		
b. *ki:ʈabuna maʃu:ma:n ba wa:i*		*	*		*
c. *maʃu:ma:n ba wa:i ki:ʈabuna*	*!			*	*
d.☞*maʃu:ma:n ba ki:ʈabuna wa:i*				*	*

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Within the above tableau, again, the role of constraint HD-RT is crucial in deciding the optimal among the candidates. Candidate (a) satisfies one of the high-ranking constraints (SO), but violates two of the main constraints (HD-RT and SUBJ-LT), so it is excluded. Similarly, candidate (b) also satisfies one of the high-ranking constraints but still is excluded, as at the same time it fatally violates two of the main constraints. Now only candidates (c) and (d) remain; the latter performs better at keeping the head of the projection at the right-hand edge of the clause by satisfying HD-RT. In candidate (c), the head is in the centre and HD-RT and HD-LT are both violated. Since the constraint HD-LT is low in the hierarchy, its violation is not fatal. But in order to be the optimal candidate, the highest constraint of HD-RT is required to be satisfied. This is done by candidate (d), which is therefore chosen as the optimal candidate. In candidate (d), the head is positioned at the right-hand edge of clause, hence this candidate is the grammatical one. Note also that the optimal candidate (d) incurs two violations, of constraints STAY and HD-LT, but these are irrelevant since they are low in the ranking.

The above tableau illustrates the future tense of Pashto. Interestingly, the crucial role is played by the clitic ba. Roberts (2000) calls it a second-place clitic in Pashto, as it is placed second to the noun in the sentence. The ranking of the syntactic constraints remains the same. This means that the ranking hierarchy of tableaux (79) and (85) does not change here, and all three tenses can be best explained under the same hierarchy of the constraints.

It is also clear from the above analysis that the future tense is marked with the help of split ergativity and that the word order remains the same in this tense. The same idea is put forward by Roberts (2000). He says that future tense is created with the help of split agreement of the compound verb in Pashto. According to him, the two parts of compound verbs may agree with different syntactic constituents in sentences of future tense. This agreement mainly occurs between second-place future clitic ba and a past perfective auxiliary verb in the sentence. He has given various examples in evidence of this suggestion. These are important here because they maintain the same SOV order; therefore, they are given in the following lines.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(87) The future tense in Pashto (split agreement)[25]

a. sangi:n ba kəɽke ma:ta kɽi:
Sangin (MASC) (FUT) *window* (FEMSG) *broken* (FEM SG) *do* (PAST PERF) (MASC 3SG)
Sangin will break the window

b. sangi:n ba kəɽke ma:te kɽi:
Sangin (MASC) (FUT) *window* (FEM PL) *broken* (FEM PL) *do* (PAST PERF) (MASC 3SG)
Sangin will break the windows

c. sangi:n ba war ma:t kɽi:
Sangin (MASC) (FUT) *door* (MASC SG*)* *broken* (MASC SG) *do* (PAST PERF) (MASC 3SG)
Sangin will break the door

d. sangi:n ba waru:na ma:t kɽi:
Sangin (MASC) (FUT) *doors* (MASC PL) *broken* (MASC PL) *do* (PAST PERF) (MASC 3SG)
Sangin will break the doors

e. ta:so ba waru:na ma:t kɽəi
(2P) (FUT) *doors* (MASC PL) *broken* (MASC PL) *do* (PAST PERF) (MASC 2PL)
You (PL) will break the doors

(Roberts, 2000, p. 43) These five examples given by Roberts are very interesting in relation to the current discussion. Roberts used these examples to show split agreement. The subject in all of the above sentences agrees with the auxiliary, while the object agrees with the main verb. To me, these examples are interesting because all of them show the SOV word order pattern, and no other pattern would be grammatically correct here in the future tense. Only pragmatically variant order can be applied within these sentences, which would show emphatic stress. Note the following example:

[25] Since the current study only covers the word order patterns of Pashto and is not concerned with the nature of clitics, I only discuss matters which are related to word order.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(88)
kəɽke ba sangi:n ma:ta kɽi:
window (FEM SG) (FUT) *Sangin* (MASC) *broken* (FEM SG) *do* (PAST PERF) (MASC 3SG)
OSV: Sangin will break the window

The sentence in (88) shows that there is emphatic stress on "Sangin". There are many reasons why this order (OSV) in such a context has this pragmatic meaning. Firstly, ba is a modal verb and in such a sentence it shows that only Sangin has the ability to break the window. Secondly, the meaning of such a sentence can be changed if both roles of subject and object are played by animate things. So, one can say that the only canonical order is SOV in future as well. Since the current study only covers he word order pattern, I am not concerned with the study of cliticisation in Pashto. This are dealt with in depth by the same source (Roberts, 2000). The above analysis of future tense reconfirms the canonical SOV pattern for the language. An additional tableau is unnecessary here because the hierarchy of constraints decided within last three tableaux suits this type of sentence equally well.

It is evident from both these tableaux that the constraints STAY and HD-LT are equally ranked in the hierarchy of constraints. They are positioned as the lowest constraints in the list. The same is the case with SO and SUBJ-LT. They have a similar position in the hierarchy of the constraints. But they are among the high-ranking constraints of BWO in Pashto. Now the only thing remaining to be explained is the equal position of these constraints, which is shown in the following tableau.

(89) *mende maʃu:ma:n səmbalawi:* The mothers are taking care of the children.

maʃu:ma:n mende səmbalawi:	HD-RT	SO	SUBJ-LT	STAY	HD-LT
a. *səmbalawi*: *mende maʃu:ma:n*	*!		*		
b. *maʃu:ma:n mende səmbalawi:*		*!	*		*
c. *mende səmbalawi: maʃu:ma:n*	*!			*	*
d.☞ *mende maʃu:ma:n səmbalawi:*				*	*

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

The above tableau shows that the top-ranking head constraint is HD-RT, which makes it possible for the verb to be on the rightmost edge of the clause. This ensures that the optimal candidate (i.e., d) will be head-final. When this constraint is violated the meanings are lost, especially in sentences like the one above. This constraint is at the top in the hierarchy. The ranking of SO and SUBJ-LT is equal, and both are crucial. Similarly, the two lower-ranking constraints are ranked equally, and both are violable because their violability justifies the satisfaction of the top-ranking constraints. In other words, the satisfaction of constraints HD-RT, SO and SUBJ-LT is based on the violation of STAY and HD-LT. Revising the ranking, the other word patterns, such as SVO or VOS, can be given. So, the present ranking is crucial here: the head-final structure in languages such as Pashto is derived from the ranking of HD-RT at the top of the hierarchy. This constraint exhibits an interesting interaction with other constraints in the hierarchy such as SO and SUBJ-LT, particularly with those which are related to head positioning (such as STAY and HD-LT).

On the basis of the above analysis, the ranking of these constraints is given in the following section.

4.4.4 The Ranking of Constraints in Pashto

As a theoretical justification for the BWO of Pashto, only one unified hierarchy of syntactic constraints under OT is needed that can explain all grammatically correct patterns of syntactic elements for all tenses and contexts. The hierarchy shown below best suits this objective for the current analysis. Focusing on the constraints relating to BWO in Pashto discussed in the above analysis, the final hierarchy of these constraints is considered in (90).

(90) The ranking of constraints related to BWO in Pashto: HD-RT»SO, SUBJ-LT»STAY, HD-LT

This ranking shows that HD-RT (the right positioning of the head) is at the the top of the hierarchy. SO and SUBJ-LT are equally crucial: recall the sentences discussed in the past tense in the analysis portion. The remaining two constraints, STAY and HD-LT, are both presumed to be low-ranked constraints. Keeping in mind the above ranking, I shall finalize the discussion on the patterns of BWO in Pashto in the next section.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.5 THE PATTERNS OF BWO IN PASHTO

On the basis of the optimality-theoretic analysis in the above discussion, the pattern of BWO in Pashto can be explained. OT is very well suited to explaining and justifying every aspect of BWO patterns in Pashto. Examples given in the analysis section of the study confirm the findings of previous studies by linguists such as Roberts (2000) and Tegey and Robson (1996) that the canonical word order of Pashto is SOV. Roberts (2000) reports Pashto to be a "rigidly head-final" language across its lexical categories. Here one can strongly concur with him because the same is clear from the above examples. Overall, I argue that SOV (subject + object + verb) is the only order which is available in unmarked and syntactically simple sentences in all three tenses of the language. The detailed discussion on the basis of the OT analysis in the above lines proves SOV to be the unmarked BWO of the Pashto language.

The above discussion also shows the power of OT as a theory of grammar which covers the interesting aspects of the language as discussed in last two chapters and, with regard to BWO, the current chapter. It shows that the data of a specific language may be accounted for simply by evaluating the constraint hierarchy and interaction of a small set of constraints as they apply to that language.

There are some languages which allow some word orders other than the canonical ones in some contexts. Pashto is not among such languages, because it is fairly rigidly SOV. Only in sentences of ditransitive verbs does Pashto allow OSV, but in such cases some other conditions must also be fulfilled by the context. I discussed this point in detail in Section 4.3, and I still conclude that the canonical order of Pashto is SOV, because only within this word order are all tenses of the language possible. Finally, I am in a position to determine the typology of BWO in Pashto. This is discussed in the following section.

4.6 THE TYPOLOGY OF BWO IN PASHTO

I proposed in (72) that the BWO variation of Pashto and its application to all tenses of the language may be captured by assuming the constraints given in (91).

(91)
a. HD-RT b. HD-LT c. SUBJ-LT
d. SO e. STAY

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

I considered these constraints and their interaction in a number of sentences of pragmatically neutral and descriptively simple nature. On the basis of an OT analysis, I showed the BWO of Pashto to be SOV. Then I proposed the constraint ranking which is specified in (92).

(92) HD-RT » SO, SUBJ-LT » STAY, HD-LT

This ranking characterizes the behaviour of Pashto with regard to its BWO. It determines the following points:

1. Subject moves to Spec IP, which is the leftmost position of the projection.
2. Objects always follow their subjects.
3. The top ranking of HD-RT reflects the rigid behaviour of the language with regard to the head-final position in a clause.

Points 1 and 2 are based on the behaviour of arguments in the Pashto clause. All these points suggest the classification of Pashto as an SOV language, along with many other major languages such as Japanese, Dutch, Urdu, etc. The following section summarizes and concludes the current chapter.

4.7 SUMMARY AND CONCLUSIONS

In the current chapter, I analysed the BWO of Pashto in the light of previous scholarship. For this purpose, I kept my focus on three basic syntactic elements (S + O + V) of Pashto. Further, the optimality-theoretic analysis of the BWO of the language supports the previous findings of various researchers, including Roberts (2000) and Tegey and Robson (1996), that the language is fairly rigidly head-final in its syntactic typology. Having discussed the other possible patterns of word order in the language, I finally give the typology of the BWO of Pashto as an SOV order and, in the light of this hierarchy, I conclude that Pashto exemplifies a language the BWO of which is strictly SOV.

This conclusion has broader implications for a general theory of typological linguistics which is based on the correlation between word order and syllable structure. According to the theory of implicational universals and the records in WALS discussed in Chapter 2, an SOV language should have simple syllable structure. Since Pashto has SOV word order, it must have simple syllable structure. But this is not the case. The next chapter discusses the possible relationship between syllable

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

structure and word order in broader prospective.

The current chapter provides additional evidence for the discussion on our subject of implicational correlations. This analysis observes that the BWO of the Pashto language is not compatible with the implicational universal discussed in the current study. This chapter provides a counter-example to the generalization of the so-called implicational universal. Thus, it seriously challenges the validity of said universal. It observes that such universals may not necessarily be in accordance with the data found in languages such as Pashto. This topic is considered in greater detail in the next chapter.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

CHAPTER 5

ANALYSIS OF THE CORRELATION BETWEEN SYLLABLE STRUCTURE AND WORD ORDER

5.1 INTRODUCTION AND OVERVIEW

The present chapter examines the correlation between syllable structure and word order. In Chapter 3 I showed that the maximum syllable template (MAXST) of Pashto is CCCVCC, and the analysis in Chapter 4 showed that the language exemplifies the SOV class of word order. These two tendencies of Pashto seriously challenge the validity of the implicational correlation which says that the syllable structure of an SOV language tends to be a simple one (e.g. CV). So, within this chapter of the study, I examine the application of this correlation to Pashto and try to figure out the theoretical perspective under the paradigm of OT.

The chapter starts by recapping the important points of the correlation between syllable structure and word order suggested by a number of scholars. Then it goes on to explore said correlation in the light of the data found in Pashto. A detailed analysis demonstrates that there exist a number of problems with the subject correlation and implicational universals based on it. The focus in this section is to figure out issues relating to the subject correlation. Finally, the OT perspective of typological study is discussed in greater detail to cover the theoretic contribution of the study. The chapter ends by pinpointing the need to add languages in addition to Pashto to such an analysis to increase the validity of the study. Thus, the concluding remarks set the ground for the comparison of Pashto and English in Chapter 6. I recap the correlation first.

5.2 THE CORRELATION BETWEEN SYLLABLE STRUCTURE AND WORD ORDER

Chapter 2 describes a number of fields in which there is clear-cut scope for studies based on cross-linguistic comparisons. A number of important

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

works were also discussed there as a background to the current study on typological linguistics. I discussed in greater detail studies which suggest that there is a correlation between syllable structure and word order, or vice versa. Here in this section, I give a recap of the important points suggested by such studies.

Typological studies have tried to explore the of correlation between syllable structure and word order. The important ones are discussed here. Many studies maintain that there is a strict correlation between syllable structure and word order, starting with von der Gabelentz, 1901, followed by Lehmann, 1973, 1978; Donegan and Stampe, 1983, 2004; Donegan, 1993; Gil, 1986; Plank, 1998; Fenk-Oczlon and Fenk, 2005; Haspelmath, 2005; and Tokizaki and Kuwana, 2012. They claim that the syllable structure of an SOV language is always simple (e.g. CV). The same idea is put forward in the implicational universals of the Universitate Constanz and the universals of WALS. The detail is presented in Chapter 2: only the main points are provided here.

Lehmann (1973) concludes from the study of languages such as Japanese, Turkish, Quechua and Sanketi (OV languages), and Classical Hebrew, Portuguese and Squamish (VO languages), that languages with OV word order have simple (e.g. CV) structure of syllables. On the other hand, VO languages have complex syllables (e.g. CCCVCCC). Other important studies on this correlation include Donegan and Stampe (1983, 2004) and Donegan (1993). They also explore the correlation between word order and syllable structure. They conclude from data on the Munda and Mon-Khmir languages that the correlation does exist. They maintain that the syllable structure of dependent-head (SOV) languages is supposedly (C)V while in contrast, head-dependent (SVO) languages have complex syllable structure (e.g. (C)(C)V(G)(C)). Gil (1986) also investigates the subjectof correlation in data from 170 languages and concludes that the syllable structure of OV languages tends to be simple.

There are two implicational universals in the Universals Archives[26] which are directly related to the question of correlation. Universal 196 says that "OV languages tend to have simple syllable structure". In contrast, Universal 207 suggests that "VO languages tend to have complex syllable structure". The same universals have also been discussed by Tokizaki and Kuwana (2012) as evidence in support of their own data. Conducting

[26] http://typo.uni-konstanz.de/archive/intro/index.php (accessed on 20 July 2010)

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

research using the data in WALS, Maddieson (2005), Plank (1998) and Tokizaki and Kuwana (2012) suggested that the correlation does exist between simple syllable structure and SOV word order.

Similarly, Fenk-Oczlon and Fenk (2005) analyse the same correlation with the help of their own data and conclude on the basis of their findings that OV word order is rightly correlated with simple syllable structure. Recently, Tokizaki and Kuwana (2012) completed their study on the topic. They add data from a number of sources such as WALS and the Universal Archives. They use some other techniques such as geographical gradation and the gradation of consonants found in the coda and again conclude in favour of this correlation between syllable structure and word order.

To conclude regarding the views of the linguists mentioned above, (93) sets out the implicational correlation:

(93)
a. A language with SOV word order has simple syllable structure, and
b. A language with SVO word order has complex syllable structure.

The classification of syllable structure is discussed in Section 2.4.2. According to this classification, a simple syllable structure is supposed to be CV, and complex structure is the possibility of the presence of more than one C on one or both edges of the syllable.

Relating this correlation to the Basic Word Order (BWO) and syllable structure of Pashto, I showed that Pashto violates this correlation significantly and provides enough evidence against it. Let me show the case of Pashto first.

5.3 THE CASE OF PASHTO: A PROOF OF NEGATION

The data in the previous two chapters shows that Pashto allows complex syllable structure and that its BWO is SOV. This means that not only is the language an SOV language but also it allows consonant clusters at both edges of the syllable. Section 3.6.7 provides a summary of bipartite consonant clusters which shows that there are 108 consonant clusters which can take place in word-initial, -medial and -final positions in Pashto. Tripartite clusters are also presented. The examples demonstrate that Pashto allows not only a number of consonant clusters but also SSP-violating clusters at both onset and coda positions. In addition, the MAXST given as (60) in Chapter 3 shows that the language allows a maximum of three consonants at onset level and two at coda, and the

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

MAXST of the language is therefore CCCVCC. I applied OT as a methodology for the analysis of Pashto syllable structure and found that the approach provides a theoretical justification for the existence of such SSP-violating clusters.

As far as the BWO of the language is concerned, I analysed it in Chapter 4 and argued that the language is fairly rigidly head-final in its clauses of simple and declarative nature, and that this is the canonical BWO which is workable in all tenses of the language. In the analysis of BWO in Pashto, I also applied the same optimality-theoretic approach and showed that the syntactic constraints hierarchy gives top preference to the constraint HD-RT, thus making it possible for sentences to be of SOV order. The typology of Pashto exemplifies the SOV class of languages.

Having clarified both the syllabic (phonological) and word-order (syntactic) characteristics of the language through the analysis in the previous two chapters, I showed that Pashto violates the correlation robustly. Although it has SOV order, it does not have simple syllable structure. Similarly, it has complex syllable structure but, at the same time, it does not have the SVO order. It is evident in the light of Pashto data that this language does not observe this correlation. This point gives rise to two questions:

(94)

a. Why does Pashto fail to observe the expected correlation?
b. Is there any correlation between Pashto syllable structure and word order?

The point behind the questions in (94) is the core question within the current study. Pashto phonological and syntactic characteristics negate the idea that there is any correlation or relationship between these two linguistic structures – at least, not of the type one would expect from the previous literature. Pashto maintains its stance and exhibits a tendency which is against every aspect of this correlation. So, the next point which arises is that of where the problem lies. Is there any quandary or weakness within the language, or any gap in the correlation? This violation of correlation demands a thorough investigation and analysis of the correlation and challenges its validity as the basis for implicational universals.

An answer to the questions above is that Pashto linguistic patterns of syllable and word order are as normal as any other linguistic tendency within languages. The main problem lies within the correlation, and the

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

implicational universals which are based on this correlation. Such gaps and shortcomings are discussed below in Section 5.4. Pashto appears to be a language which violates said correlation and provides "the proof of negation" of it. Even though it is only a single language, this suggests that the correlation has not enough force to be observed by all languages. From such a perspective, OT very strongly provides theoretical justifications for the nature of Pashto syllable structure and BWO because OT, as described by McCarthy (2002), is typological in nature and can give answers to a number of such questions. Before discussing the application of OT, let me consider what weaknesses, gaps or shortcomings there are in the correlation in question, in light of the fact that it is not observed in Pashto.

5.4 GAPS IN THE CORRELATION

Having made it clear that the correlation between syllable structure and word order does not appear to be valid in the case of Pashto, it is appropriate to consider the correlation in further detail here. Pashto serves as an ideal case to challenge the subject of this correlation and to make room for other languages of its class which are not yet documented by researchers. In the following lines, I discuss the gaps in the correlation which can be pointed out on the basis of its violation.

5.4.1 Word Order: A Weak Base

The first important point, which is also the source of some subsequent issues, is the weak base of word order for such correlations. The term *BWO* is itself quite tricky in its approach. If the classical idea of BWO (which I have applied in the current thesis, answering the question "what happened?" or considering simple declarative sentences having one verb) is considered correct, then the question arises: what about the rest of the language? And even if one unified definition of BWO is agreed upon, then other problems arise. The most important one is: what is the base word order? This is a debated issue in syntax. Kayne (1994) says that the base word order is universally SVO, and that other orders are derived from this one base form. On the other hand, there are linguists such as Gell-Mann and Ruhlen (2011) who think that SOV is the base order from which other orders are derived. Something which is itself still under debate may not serve ideally as the basis for drawing conclusions about such a correlational phenomenon.

There are six logically possible patterns for the ordering of subject, object and verb, as discussed earlier. On the other hand, only three syllable

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

structures (simple, moderately complex and complex) are possible. Now, previous studies mention two word orders (SOV and SVO) and their two syllable patterns (simple and complex). But they do not say anything about the rest of the patterns of BWO and syllable structure. According to WALS, Chapter 81, on order of subject, object and verb (submitted by Dryer[27]) there are six patterns logically possible for BWO. The data entered for this input value consisted of 1377 languages. The allocation of language types based on the BWO of these langauges is given in Table 5.1.

Table 5.1: Six Types of Dominant BWO in Languages

Basic Word Order		**No. of Languages**
Subject + object + verb	(SOV)	565
Subject + verb + object	(SVO)	488
Verb + subject + object	(VSO)	95
Verb + object + subject	(VOS)	25
Object + verb + subject	(OVS)	11
Object + subject + verb	(OSV)	4
Languages lacking dominant word order		189
Total:		**1377**

It is clear from the above table that much research has been carried out on the first two types of languages but far less has been said about the other four types. Additionally, there are languages lacking any dominant order. So, this is a weak base for the correlation and its ability to demonstrate the correlation is not patent. Even if languages are divided on the basis of the ideas of Vennemann (1972), there are some problems with it. For example, his natural serialization principle (NSP) approves only two types of BWO (VO and OV). In such a division, the inclusion of SVO, VSO, VOS into VO and that of SOV, OSV and OVS into OV again raises the question of

[27] Accessed on 25 October 2011

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

the validity of the correlation. It is impossible to prove it in the form of Pashto, as it turns out to be an OV language with complex syllable structure. This topic is the subject of the next chapter, where Pashto and English are compared in order to further discuss the validity of the correlation. The other point which is noticeable in such a division of BWO of languages is the inclusion of languages of free word order which have no dominant BWO. Dryer (2005) shows that there are 189 languages of no dominant word order (free word order) out of the total of 1377 languages tested on word order. In other words, 13.73% of languages have no place in the correlation under discussion. Another weakness related to this implicational correlation is the changing nature of BWO; this is discussed in the following section.

5.4.2 The Changing Nature of BWO

In many languages and language families of the world, the changing nature of BWO is reported. Such change seriously challenges the validity of a correlation which is based on BWO, because it is not necessarily the case that with a change to BWO, everything will be changed. In particular, the changing of syllable structure of a language is not reported so far in the literature. Syllable structure is directly the representation of the pronunciation of a language; it is not necessarily correlated with a change of BWO.

There are many languages which allow various types of BWO at the same time or that, with the passage of time, change their BWO. Another type of change in BWO takes place when pattern of BWO in a mother language is changed in the daughter languages. Gell-Mann and Ruhlen (2011) show a number of languages within the same language family in which the BWO changes with the passage of time. The following table shows the detail of this phenomenon found by the same source.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5.2: The Change of BWO in Afro-Asiatic Languages

Language or Language Family	Basic Word Order
Afro-Asiatic	SOV
Omotic	SOV
Erythraic	SOV
Cushitic	SOV
Chado-Afro-Asiatic	SVO
Chadic	SVO
North-Afro-Asiatic	VSO
Ancient Egyptian	VSO
Semito-Berber	VSO
Semitic	VSO
Berber	VSO

It is clear from the above table that within this one language family, all three major BWOs are well attested. The original order in the mother language was SOV but this changed with the passage of time, and as more and more languages evolved within the same family they took various orders as their BWO. Gell-Mann and Ruhlen (2011) discuss this calculation in order to show the effect of various processes, such as evolution, diffusion and reversion of word orders, in the origin of languages. They try to prove that the base word order for languages was originally SOV, but that this has further evolved, diffused and reversed within different language families. Here my point is that when such processes are involved within the BWO of languages, it may seriously affect the correlation of word order with syllable structure because syllable structure may not necessarily change at the same time as word order.

Other examples of changes in BWO are found in the literature. Vennemann (1973, p. 40) talks about possible word order change in languages. According to him, this process of change follows an exact order. For example, an SOV language can only change to SVO. Similarly, this SVO order either changes to a VSO pattern or becomes a free word order (FWO) language. A VSO language may also revert back to SVO. He maintains that an FWO language could finally become an SOV again with the passage of time. None of these researchers talk about changes of syllable structure, and logically it is not possible for the syllable structure to change exactly along with word order in terms of pattern, space and time. So, this is another significant issue with word order being regarded as the basis for the correlation under discussion here. Figure 5.1 shows the changing process proposed by Vennemann (quoted in Gell-Mann &

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ruhlen, 2011, p. 2).

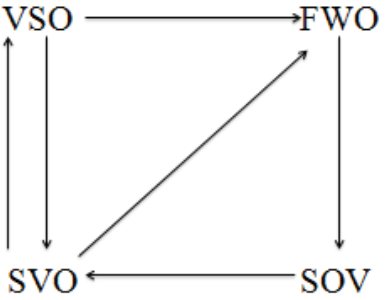


Figure 5.1: The Possible Order of Change in BWO

Figure 5.1 shows the circularity of the changing of BWO within languages. Vennemann (1973) does not mention any reason for such a change. He suggests that during this process of change an FWO language would change into an SOV language, and this suggestion seems to be true because Table 5.1 shows the maximum number of languages with this type of order. It shows that a total of 565 languages out of 1377 (41%) are SOV. Similarly, Givon (1977) talks about the diffusion of language orders and maintains that as a result of this process, the BWO of languages changes to SOV. Discussing the same point, Harris and Campbell (1995) conclude that most languages have an SOV order, and that this is the result of their contact with other SOV languages. An interesting hypothesis is put forward by Gell-Mann and Ruhlen (2011), who maintain from their observations that SOV was the order of the initial language from which other languages evolved with the passage of time. Even if the other version of the theory of BWO (that other word orders evolved from SVO; (Kayne, 1994) is accepted, the change of BWO is witnessed. So in both cases the order is subject to change, but there is no evidence in the literature of a concomitant change in syllable structure.

The example of English is also very interesting in this regard. Fries (1940, p. 199–208) and Lehmann (1978) report that English had the characteristics of an SOV language in earlier periods. They suggest from their analysis of old literature in English (e.g. *Beowulf*) that genitives used to precede nouns in 90% of the language of that time – a strong SOV pattern. They maintain that this order gradually changed with the passage of time. More interestingly, this direction of syntactic change concedes to the order of change suggested by Vennemann (1973), as shown in Figure 5.1.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Keeping in mind the above discussion on BWO change, I want to repeat that the validity of the supposed correlation between syllable structure and word order is seriously under question. When it is not sure that a newly evolved language will take the BWO of the mother language, how can it be guaranteed that the syllable structure will automatically be either the same or different? It is therefore hard to believe the validity of a robust implicational correlation of syllable structure with word order. There is no evidence found in the literature to show that the syllable structure of a language also changes exactly with the ratio and pattern of its word order. Any languages with any word order pattern can have any type of syllable structure.

It is pertinent to note here that syllable structure also changes from mother to daughter languages. This point is further elaborated in Section 5.4.4, where examples of CV patterns are given from pidgin and creole languages. Before taking up this point for detailed analysis, I show that, apart from changing nature of BWO, there are cases in which one language family has various BWOs in the daughter languages. This point is raised in the next section.

5.4.3 One Family, Different BWOs

There are cases in the history of languages which show that one family of languages can have more than one type of BWO. The example of Niger-Kordofanian is quite interesting in this regard. Table 5.3 shows the case of the Niger-Kordofanian macro language family, further divided into minor families, and their various types of BWO. The figures emphasize the above point (Gell-Mann & Ruhlen, 2011): that languages in the same family do not necessarily share a BWO.

Table 5.3 shows that a large number of languages with different BWOs are found within a single language family. So, one cannot exactly count on a pattern which is so variable even within a single family. The same is the case with the Eastern-Iranian language family. Pashto and Persian are cousin languages and both have SOV as their BWO, but there is a significant difference within their syllable structures. Pashto has complex syllable structure (Bell & Saka, 1982; Levi, 2004), while Persian has simple structure. So there is a problem with this correlation; it is true in the case of Persian but it is violated in the case of Pashto. The problem lies within the basis of the correlation, which has BWO as the deciding factor even though BWO is highly changeable across languages and across time. Here arises the question of the nature of phonological and lexical features

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(structural patterns) in a new language. This point is taken up in detail in Section 5.4.4.

Table 5.3: Different BWOs in the Niger-Kordofanian Macro-Family

Language or Sub-Family	Basic Word Order		
	SOV	SVO	VSO
Niger-Kordofanian	39	279	1
Kordofanian	4	15	1
Niger-Congo	35	264	0
Mande	22	0	0
Niger-Congo Proper	13	264	0
Atlantic	0	16	0
Kru	1	3	0
Dogon	1	0	0
Gur	8	22	0
Adamawa	0	16	0
Ubangian	0	21	0
South Central	2	52	0
Broad Bantu	0	16	0
Bantu	1	118	0

5.4.4 Language Evolution and the Correlation

In the last three subsections, a number of points relating to the changing nature of BWO are discussed. All of them have common grounds that show the same phenomenon, so it is appropriate to discuss logically the process of new language formation, i.e. pidginization or creolization, and the change of BWO. Within the process of pidginization a new language evolves, and according to linguists the grammatical system (which includes BWO) is supposedly taken from the superstratum language.

The topic of language change and creolization has been discussed by a number of scholars within the fields of general and historical linguistics. The structure of a superstratum is subject to change within a creole. But there is no consensus among grammarians about the nature of change which takes place within creole languages: this process of change in creoles is a debatable issue. They debate about whether the features of substrate or superstratum languages are more likely to be dominant in the structure of the new language, i.e. the creole. What is agreed by all of them is that drastic changes take place in the process of the formation of new

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

languages such as pidgins and creoles. Meillet (1912) and Schuchardt [(1842–1927) in Alvar 1967] report that the nature of language change in creoles is mixed. They invalidate the genetic model of languages which says that creoles are the daughter languages of their superstratum languages. They say that there is no direct relation between a creole and its superstratum language. On the other hand, there are linguists who say that the grammar of a new language (creole) is based on the superstratum and its phonology is based on the substrate language. Wardhaugh (2002, p. 60) calls it the re-expansion of syntax in the process of creolization. These points show that there is a significant chance for languages to change their word order in the process of new language formation. The data in Table 5.2 also shows that a greater number of languages change their word order in the process of new language formation. But even if one admits that they do not change their BWO and maintain the characteristics of the parent language, there is a clear-cut chance that the syllable structure of the new language will be different: it tends to be simple in a new language.

The process of pidginization or creolization always gives rise to the formation of simple syllable structure (i.e. CV) and the breaking of consonant clusters (Romaine, 1988). It is mentioned by a number of scholars that there is a general tendency in pidgin and creole languages not to have consonant clusters. This can be seen in Tok Pisin, West African Pidgin and Jamaican Creole. Note the following examples, which are given by Romaine (1994, p. 588) in order to show how the clusters are broken up.

(95)

1. Tok Pisin /gara:n/ ground
2. West African Pidgin /sakan/ skin
3. Jamaican Ceole /taka/ talk

The process of epenthesis is visible here, showing the tendency of these new languages to avoid complex consonant clusters. In creoles, consonant clusters are found but in a much later stage, when the language has grown up more strongly towards the superstratum language. To show how these creole languages exhibit a tendency towards maintaining open syllables, Romaine (1988) has given a word list for the English word *worm*. Her analysis shows that how these languages avoid consonant clusters. The following is the list given by her.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(96)
Word List for English word *worm* given by Romaine (1988)

1. Krio *worom*
2. Sranan *woron*
3. Saramaccan *wanu*
4. Cameroons *wom*
5. Guyana *worm*
6. Jamaican *worom*
7. Gullah *Burum*

The avoidance of the consonant cluster of /r + m/ in the above word list shows that there is a clear-cut tendency among newly formed languages, including creoles and pidgins, to not tolerate clusters and ultimately maintain no consonant cluster.

This shows that there is no chance for a new language to maintain the same syllable structure of the parent or superstratum language. Similarly, in Table 5.3 it is shown that the BWO of the superstratum language is not always followed. All this shows that there is no possibility of the correlation between syllable structure and word order in newly formed languages. The syllable structure of a pidgin language is expected to be simple, but no specific word order is predictable for it. Commonly they take simple syllable structure with any form of BWO.

Since languages are considered as living entities, they never complete their process of formation, diffusion and evolution. These are continuously in progress throughout the history of a language. The example of classical languages such as Arabic, which is now rare in its classical form, is also important. The case is similar with Latin, which has given rise to many languages such as French and Italian but itself is no more in use. So, no time-frame can be given for any language as to when it will flourish and when it will adopt the position of a fully fledged language. It is a continuum which is in progress and continues till a language completely dies.

Another question which is related to the discussion of new language formation, then, is: how many languages form through the process of pidginization and creolization and how many of them directly evolve? To answer this question, one has to undertake a deep diachronic analysis of a language. This can reveal that the reason for having complex clusters in one language is contact with other languages of the same nature. There is s

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

valid point in this idea, because many words are borrowed by languages from one another. In the case of Pashto, it is among the oldest surviving languages of South Asia and has a history of some 3000 years (Raverty, 1859). The abundance of consonant clusters shows that they are not the result of such contact with other languages.

The examples of languages given by Lehmann (1973) show that Mon-Khmir (SVO with complex syllable structure) and Munda (SOV with simple syllable structure) are from the same language family. But they have adopted different syllable structures and BWOs. He suggests on the basis of his analysis of these two languages that word order is correlated with syllable structure. I say that this correlation does not exist in reality and cannot be suggested on the basis of data from one language family. The case of the Eastern-Iranian family is important here. Our subject language, Pashto, is an SOV language with complex syllable structure, while its cousin language Persian has the same BWO but a simple syllable structure. Similarly, the case of English (SVO) and Pashto (SOV) can also be submitted as evidence against this correlation. Both of them have complex syllable structures, but they have different BWOs. Naturally, to generalize from this would need a larger set of data. English and Pashto are compared in detail in the next chapter, in which Pashto is discussed as the "proof of negation" against the correlation. Another challenging point is the existence of languages with no fixed syntactic patterns . This is taken up in the next section.

5.4.5 Languages with no Fixed Syntactic Patterns

Table 5.1 has shown that there are a large number of languages in the data of WALS which have free word order: 189 out of a total of 1377 languages. Vennemann (1973) maintains that free word order is very natural and is part of the change of order which takes place with the passage of time. So, in connection with the subjectof correlation, this point of FWO in languages must also be addressed. If the points of the above two sources are valid, then there is no place for the correlation between two totally different language structures, those of phonology and syntax.

This is a very strong point against the correlation of syllable structure and word order, and the implicational universals which are based on this correlation. It says indirectly that FWO is as much possible as any other BWO. One can infer from this that syllable structure is not bound in the correlation under discussion. It can also take place as freely as possible. The case of Spanish is very important in this regard. Two different BWOs

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

are reported for this language. Ordonez and Trevino (1995) and Zubizarreta (1995) report it as VSO, while according to Hernanz and Brucart (1987) it has SVO order. This point of difference is due to the flexible nature of BWO in languages where there are discrepancies among its speakers over whether SVO and VSO word patterns are used. In the previous few sections, I discussed the inability of BWO to stand as a criterion for the correlation or a basis for implicational universals based on the correlation. The next section highlights the positions of BWO and syllable structure as independent domains of linguistic investigation.

5.4.6 No Correlation between BWO and Syllable Structure

The analysis of the correlation between word order and syllable structure in this chapter suggests that, logically, such a correlation is not applicable within the study of linguistic structures. From this analysis, I can assert that BWO and syllable structure, and therefore syntax and phonology, are different levels of linguistic analysis. They exist within their circle of influence, creating and forming interfaces just like morphology and phonology, but they are not correlated. The categories of these levels can be used for the classification of languages and language families independent of their mutual compatibility or correlation. For example, languages can be classified according to either their syntactic structure (SOV, SVO, etc.) or their phonological characteristics (simple or complex syllables). The important thing is that syllable structure cannot be correlated with the word order of a language family or even an individual language. The data from Pashto clearly supports this argument and suggests that any type of syllable structure for an SOV language is possible. And similarly, a language with any word order pattern may have complex syllable structure.

Now, before moving on to provide a theoretical justification for the "proof of negation" against the correlation between word order and syllable structure, I give the interim conclusion for the present chapter.

5.4.7 Interim Conclusion

Sections 5.1 to 5.4 introduce and analyse the correlation between syllable structure and word order in the light of data from Pashto as the primary source and previous research as supporting data. I conclude these points in the following section.

- A number of previous studies regarding the correlation between syllable structure and word order and the implicational universals

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

based on such a correlation say that the syllable structure of an SOV language tends to be simple (e.g. CV).

- Similarly, the syllable structure of languages with SVO word order tends to be a complex one (e.g. CCCVCC).
- Pashto clearly negates such a correlation. The language is SOV, but contrary to said correlation it has complex syllable structure. Thus our data from Pashto robustly refutes such a correlation and provides enough evidence against it.

This conclusion raises a number of points which throw doubt on the logical existence of the correlation based on a weak source like BWO. Further, it shows that the implicational universals based on the same correlation are quite questionable. Though syllable structure is a comparatively stable entity, the problems with BWO make it unreliable. The main problems with the correlation are the following:

- BWO is a weak basis for this correlation and for the implicational universals constructed on it. There are six logically possible patterns for BWO, but the correlation is only concerned with of them (i.e. SVO and SOV).
- Both BWO and syllable structure have a changing nature. Recent literature reports a number of changes in the BWO and syllable patterns of languages and language families. Especially, pidgins and creoles prefer to break consonant clusters by default.
- This change in BWO at times follows a fixed pattern, but the same pattern is not reported in the case of syllable structure.
- There are language families which have various BWOs within the same family and continuous change is witnessed in the BWO of a language.
- In the case of pidginization and creolization, simple syllable structure is likely, but the BWO is not predictable in such a process of language formation.
- There are a number of languages which are reported to have no fixed BWO pattern.

Such languages are not covered by this correlation.

These conclusive remarks show that the logical existence of the correlation is quite questionable when it is built on the weak basis of BWO. Syllable structure is a comparatively important and stable entity but the problems with BWO make it doubtful. It is very hard to base typological

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

correlations on the correlation of different levels of linguistic structures. In other words, phonological and syntactic properties are different linguistic structures. They cannot be correlated with each other in the sense suggested by the correlation in question here. They have separate hierarchies of constraints, independent of the influence of each other. It is concluded on the basis of the gaps pointed out in the above discussion and proved by the data from Pashto that such a correlation is not logically possible.

This interim conclusion leads to a number of questions related to the "proof of negation" against the correlation. The main question is given as (97):

(97)

Why is the correlation of syllable structure and word order not possible?

(97) leads to a number of related sub-questions, given as (98):

(98)

a. How can one theoretically explain the variation in syllable structure and basic word order?
b. Why does a language choose one out of various patterns of BWO?
c. Is it optional for a language to adopt one specific word order?

These questions are addressed in the next section, which also covers the OT perspective on typological studies. Further suggestions are also given for cross-linguistic studies using OT as the methodology for dealing with such linguistic variations. It suggests a very solid approach to such typological studies and has a very interesting solution within this constraint-based paradigm.

5.5 OPTIMALITY-THEORETIC ANALYSIS OF THE CORRELATION

The power of OT as a theory of language variation provides a theoretical explanation in responses to the questions given as (97) and (98). As discussed in detail in Chapter 1, and motivating the selection of OT as the methodology for the current study, OT provides support for the investigation of variation in linguistic structures. Before giving the optimality-theoretic analysis of the subjectof correlation and trying to answer the questions given above, I briefly recap the main approach of OT as a theory of variation among languages. This approach fits as a

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

methodology to resolve a number of important issues regarding the subject correlation. Then I will look into the OT perspective on typological studies.

The OT idea of grammars as an interaction among violable constraints suggests that constraints are universal, while the ranking of them is the only difference among languages (Prince & Smolensky, 1993, 2004). OT further maintains that the ranking of constraints changes the grammar of a specific language. For example, in Chapters 3 and 4 the hierarchy of universal constraints for Pashto syllable structure and word order was explored, respectively. The effects of different ranking of universal constraints for syllable structure, such as faithfulness constraints (such as DEP-IO, MAX-IO) and markedness constraints (such as SON–SEQ, *COMPLEXONS, *COMPLEXCOD and NOCODA) were explored and it was concluded that Pashto ranks the faithfulness constraints highly, while the markedness constraints are easily violated in the hierarchy because of their low importance. Resultantly, Pashto not only allows complex consonant clusters but also permits reverse sonority clusters at both edges. The data presented in Chapter 3 supports such analysis of constraints for Pashto syllable structure and exhibits the extent of Pashto complex syllable structure. Similarly, the analysis of Pashto word order in Chapter 4 suggests that Pashto is a rigidly head-final language and exemplifies a strict SOV type. I based this classification on the permutation of important syntactic constraints for word order, such as HD-RT, SO, SUBJ-LT, STAY, HD-LT, in the hierarchy of Pashto. All this shows that language universals are held by all languages and the permutations of these constraints cover the complete range of permitted variations in the grammars of languages.

Keeping in mind the above-discussed approach of OT, I take up the questions given as (96) and (97) in the last section. Focusing on these questions, an OT analysis of the correlation between syllable structure and word order is provided. I take the sub-questions first.

a. How can one theoretically explain the variation in syllable structure and basic word order?

OT, being a typological methodology on language variation, provides an answer to the question above. The possible answer is that *the variation in any linguistic structure is due to the differences in the constraint hierarchy of languages for that specific structure*. For example, six logically possible word order patterns are available in languages. Section 4.2 shows these

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

variations in detail. Now it is up to the constraint hierarchy of the syntactic structure of a language to adopt one specific order out of them. Costa (2001) suggests that the observed variation in BWO is due to the effects of constraints which are not normally active. This variation in the constraint hierarchy is the result of various functional specifications in the input of a language. The selection among the constraints available in a language hierarchy has previously been suggested as "The Emergence of the Unmarked"[28] by McCarthy and Prince (1994) and McCarthy (2008). For example, the syllable structure of Pashto is complex (It allows more than one consonant at either edge), while the case of Japanese is different: it does not allow consonant clusters, and the syllables found in it are normally simple ones. The important point to be noted is that both of these languages are SOV, but their syllable structures are different.

In case of syntactic structure, Pashto is SOV. This is clear from the analysis in Chapter 4. On the other hand, English is an SVO language. Note the word order for the following sentence, which is given in both languages.

(99)

a. /ba:tʃa/ /jao ki:ta:b/ /li:ki/ (Pashto)
S O V
Bacha *a book* *write-do (TRANS IMPERF) 3SG*
Bacha writes a book

b. He writes a book
S V O
He writes a book

It is clear from the examples (a) and (b) in (98) that English is an SVO and Pashto an SOV language. OT justifies this difference of syntactic order as the result of permutations of universal constraints. Since Pashto is a head-final language, it keeps the head of the IP at the end of the sentence, while English is head-initial and therefore V comes before O. So, one can theoretically explain that this is the result of variation in the hierarchy of constraints for structures like syllable or word order. OT offers theoretic

[28] "The Emergence of the Unmarked" (TETU) is a key point in OT. According to it, "preference for some universally unmarked structure, such as syllables with onsets, can emerge under the right circumstances even if the language as a whole permits the corresponding marked structure" (McCarthy, 2008, pp. 24–25).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

explanation for this variation as the arranging and rearranging of the universal constraints in the hierarchy of different grammars. The above examples provide solid reasoning for the permutation of the relevant constraints in the example languages. Now take our remaining two sub-questions, which are inter-related.

b. Why does a specific language choose one out of various patterns of BWO?

c. Is it optional for a language to adopt one specific word order?

Previous studies on OT syntax (e.g., Grimshaw & Samek-Lodovici (1998); Samek-Lodovici (1996); Choi (1996); Costa, 1997c, 1998, 2001) suggest that variation in BWO is based on various functional specifications of a language and that it is not optional or randomly occuring. A language chooses one particular word order structure out of various available patterns. This formalism is already discussed as the permutation of the constraint hierarchy in the discussion on the first question. It has been discussed that this selection is not optional to a language; rather it is based on the functional specifications of a language. In Section 4.4.1, It is discussed under the heading of "word order in OT". It was suggested that OT has its own idea about BWO pattern. The permutation of CON is the same point as suggested by linguists for OT phonology.

As discussed above, the selection of one specific word order pattern is not optional for a language. Rather it is the result of the functional specifications of different types within a language. For the same reason, the hierarchy of constraints is different in various languages. The hierarchy of various syntactic constraints in Pashto, as discussed in Chapter 4, determines the SOV order of the language. So, it can be concluded that the selection of a specific word order is not optional to a language. Moreover, in the selection of one specific pattern as the BWO for a language, one may have to refer to the philosophical orientation behind such a selection by the speakers of a language. This would ultimately lead one to include the long-existing debate in linguistics around the Linguistic Relativity Hypothesis suggested by Sapir and Whorf (in Columas, 1999) early in the last century. The idea that world-view is conditioned by the grammar of a language was first initiated by German linguist Wilhelm Von Humboldt and was further refined by American linguists Sapir and his student Whorf (in Columas, 1999). They concluded from their research that languages are conceptual filters for their speakers

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(Columas, 1999). Trudgill (1987) also contributed to the soft version of the hypothesis: that languages are influenced by the social and physical environment of the society. Whatever may be the reason for language variations, OT quite strongly explains the differences in linguistic structures. The OT mechanism says that they are the result of the hierarchy of constraints selected by the grammar of a language, which in turn is based on the functional inclination of that specific language.

Having discussed the sub-questions, the main question is addressed here. It is given below.

(97) Why is the correlation of syllable structure and word order not possible?

OT suggests that all the available options for a linguistic structure are covered within the permutations of CON (McCarthy, 2008). CON is universal, while the constraint hierarchy is language-specific. Now, if a deeper analysis of the constraints discovered for syllable structure and word order pattern is given, one can easily judge that syllable structure and word order are both separate domains of constraint hierarchies. For syllable structure, the constraints such as DEP-IO, MAX-IO, NOCODA and *CLUSTER have been given. Similarly, the constraints of syntactic patterning (BWO) of a language (e.g., HD-RT, HD-LT SO and STAY) are defined in Chapter 4.

This listing of the syllabic and word arrangement shows that both levels independently choose their specific hierarchy for constraints within a language. Therefore, our OT analysis does not support a correlation between the two. The following example clarifies the issue.

Language A has a specific constraint (say constraint z) at the top of the hierarchy of its syntactic structure and another (constraint y) topmost in the hierarchy of syllable structure. This may not necessarily be the case with language B because it is not bound to take the same hierarchies for both the phonological and the syntactic structures. When it is said that SOV pattern is correlated with simple syllable structure, it is implied that the hierarchical structure of syntactic constraints determines the phonological constraints for a language, which is not true. In other words, I submit that the ranking of one structure (say syllable structure) may not be correlated with another (word order), and that this absence of correlation is predicted by the theory.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

The justification against such a correlation is provided by the idea that syllable structure (i.e., phonology) and word order (i.e., syntax) are two different domains of linguistic investigation. The possible hierarchical structure of one domain may not emerge in another. In this scenario, it is important to note that one structural domain is not correlated with another unrelated one, and that the effects of different re-ranking of a constraint may not necessarily influence the ranking constraints on another structure. OT countenances this typological approach and it confirms most strongly the idea of the functional inclination of a language being exhibited in its hierarchy. The same idea is exhibited in our data from Pashto. The main problem lies in the approach towards the typological studies. The previous studies which advocated the correlation between syllable structure and word order were based on the holistic approach to language typology.

Our analysis in the current chapter shows that the Gabelentzian ideal – that is, to construct the whole structure of a language on the basis of a single linguistic property – is quite impossible to attain. There are many problems relating to such an approach, and it is very difficult to classify the whole language on the basis of any a single tendency. Secondly, there are certain domains of grammar and these grammatical properties are the objects and may serve as the targets of classification and characterization of languages, but they cannot serve as signals to the hidden bulk of language as the result of correlations between two unrelated linguistic structures. Syntactic process may not necessarily be connected with phonological process by means of such a correlation. The validity of the correlation between the two may not be given any importance in the typological study of languages. OT, thus, exhibits itself as the most suitable contemporary method for typological investigation in linguistics. The OT paradigm has changed the approach of typological studies and has shown that the most strongly acceptable justification for the variation in grammars is the different ranking of the related constraints.

Having pointed out a number of problems in the subjectof correlation, it is now necessary to resolve such issues and shortcomings as have been found in the current chapter. Linguistic typology is of course one of the most important areas of study in linguistics. As discussed in Chapter 1, there are a number of benefits to typological studies and cross-linguistic comparisons. OT provides an ideal formalism for typological investigation which covers the basic objectives of cross-linguistic studies. There are two goals behind such an approach: to ascertain the whole range of variations for one specific structure; and to understand the characteristics of the members of each typological group. These objectives are easily achievable

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

under the paradigm of OT. OT maintains that all constraints are part of universal grammar and, therefore, available in the grammar of and common to all languages. Further, every language has its own typical ranking for these constraints. There is strong reason to consider that this suggestion is true, because the same is the case for our subject correlation.

The above discussion provides enough evidence for the stance taken in the above lines. It is also shown by the data that the constraint of SSP, which is not obeyed in Pashto, is followed in other languages. That this constraint is not active in the case of Pashto shows that this is a language-specific constraint. OT also gives a diachronic explanation of the change in the constraint hierarchy of language for a specific linguistic structure. Section 5.4 shows various aspects of change in the BWO of languages that can be explained using the mechanism of OT. It says that the rearranging of the constraint sequence is the main factor in bringing about change in the hierarchy of a linguistic structure.

The analysis of the (absence of) correlation in this chapter is supported by the notion of language type given by the OT mechanism. This leads to the view that language types are more abstract objects than they were assumed to be in the past. This revolutionizing of typological studies may ultimately lead to a point where no place is found for the Greenbergian trichotomy of VSO/SVO/SOV. Language classification within the OT paradigm is considered as a family of variant subtypes. These subtypes obey certain linguistic regularities and each of them shares at least one of the typological indicators with every other subtype. These variations may be witnessed throughout the grammars of the languages in very radical ways. If one is to define the full picture of possible human language, one must survey the complete range of variation patterns for a specific language structure across languages. Then one can also observe the principled limitations on every variation. Such regularities and their explanations for attested and non-attested variants can give the full picture of possible human language. Needless to say, OT, as constraint-based methodology, justifies the change in grammars of languages as the result of the permutations in the mechanism of constraint. OT suggests comparative typology based on the variation in the ranking of constraints.

5.6 CONCLUSION

The examination of the purported correlation between syllable structure and word order in this chapter shows that this correlation does not exist. The typological features of Pashto are very interesting and make it distinct

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

from other languages of the region. Roberts (2000, p. 8) calls Pashto "typologically unusual" because of its second-place clitics. It does provide enough evidence against the correlation under discussion here. I disagree with previous works such as those of Lehmann (1973, 1978), Donegan and Stampe (1983, 2004), Donegan (1993), Gil (1986) anda Plank (1998) that the syllable structure of an SOV language tends to be a simple one. In short, it is clear from the foregoing discussion that syllable and word orders are separate domains of investigation and they are not correlated.

The chapter also attempts to formalize OT methodology as an ideal framework for cross-linguistic studies. OT as a theory of language variation and constraint interaction demonstrates that the minimum differences between languages are due to the minimum re-ranking of the relevant constraints. The power of the constraint-based network is applied further in the next chapter in comparing Pashto (SOV language) with English (SVO), both of which have complex syllable structures. This comparison will show how important the OT mechanism is for the explanation of such a typological relationship.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

CHAPTER 6

THE COMPARISON BETWEEN PASHTO AND ENGLISH

6.1 OVERVIEW AND INTRODUCTION

The Pashto data in Chapters 3 and 4 provides robust evidence against the so-called correlation between syllable structure and word order. Similarly, the analysis of the same correlation in the light of the data from various language families in the preceding chapter supports my conclusion that the syllable structure of an SOV language may not necessarily be simple. This therefore suggests explicitly that simple syllable structure is not correlated with SOV word order. As a theoretical justification for this point provided by OT, it is suggested at the end of the previous chapter that these two separate linguistic domains have their own hierarchies for active domain-specific constraints which determine the relevant structure within a specific language. The current chapter continues the same discussion that linguistic structures are determined by the permutation of active constraints for a particular structure by comparing Pashto and English. Irrespective of their distinct word orders (SOV and SVO respectively) these languages have similar syllable structures (complex syllables). Accordingly, they provide very interesting comparison in support of my suggestions regarding typological studies.

The present chapter compares the characteristics of Pashto and English, focusing on the features that are relevant to the current study. The phonological and syntactic patterns of both languages are described, concentrating on the correlation between syllable structure and word order. This chapter also examines the usefulness of OT as the constraint-based framework to be used for typological studies in light of the evidence from the subject languages. In support of OT as an ideal framework for cross-linguistic studies, analysis of the relevant structures of the languages is presented. Syllable structure, syllabic templates and the strategies of syllabification in the subject languages are also presented. This is followed

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

by a description of the basic word order of Pashto and English. The chapter also summarizes the comparison of the relevant features of the two languages.[29] Finally, it supports and makes recommendations for typological study under the OT paradigm.

6.2 COMPARISON OF SYLLABLE STRUCTURES

This section describes the characteristics of Pashto and English phonology which are relevant to the current study. It is mainly concerned with the description of syllable structure in terms of consonant clusters, as complete phonological examination is beyond the scope of the current study. So, for example, the purpose of this comparison is to investigate consonant clusters and it does not cover the vowel systems of these languages.

The comparison here presents the similarities in the syllable structure of the languages under study. Since data from Pashto has seriously challenged the validity of the subject correlation in the previous parts of this study, here I test my hypothesis by comparing these two languages. The main objective of this comparison is to verify the point that the correlation between syllable structure and word order does not exist, particularly in the case of Pashto. I would also challenge the nature of the implicational universals which are based on the correlation which claims that the syllable structure of an SOV language tends to be simple (open, ending with vowel, (C)V).

The MAXST of Pashto (CCCVCC) outlined in Chapter 3 demonstrates the characteristic syllable structure of Pashto and shows a large number of bipartite consonant clusters at all positions of words. So, among others, one of the objectives is to compare Pashto syllable structure with that of English, which is considered a highly consistent SVO language (Lehmann, 1978, among others). The results of this comparative investigation are not only relevant to the current study but also useful for syllable theory, typological linguistics and applied linguistics.

Section 3.3 defines a syllable as the smallest possible unit of speech (Ladefoged, 2001). According to Blevins (1995), it is a phonological unit

[29] It should be borne in mind that the comparison in this chapter is not exhaustive in terms of phonological and syntactic characteristics. It focuses only on the comparison of syllabic and word order patterns of Pashto and English which are directly relevant to the study.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

which organizes segmental melodies in terms of sonority. As a structural property, a syllable has at least the nucleus (which is a vowel in most languages), while consonants are optional. More than one consonant at one edge of the syllable is called a consonant cluster. According to Jones (1976), a consonant cluster is the sequence of consonants that appear in a syllable without a vowel between them and which can be studied in terms of graphemes, phones and phonemes. So, what is meant by a consonant cluster is a blend of consonants with no intervening vowels in between them and which occur in a single syllable. Based on the nature of the consonant(s), Section 2.3 gives three possible structures for syllable categorization (i.e. simple, moderate and complex). Consonant clusters or sequences which are found in a single syllable are called intra-syllabic clusters, while consonant clusters which belong to two different syllables in a word are called inter-syllabic. The current study is concerned with intra-syllabic consonant clusters which are found at word-initial (onset) and word-final (coda) positions. Word-medial clusters are not taken into account as they are mostly the repetition of the same combinations found in these two positions. First, I take the case of Pashto.

6.2.1 Pashto Syllabic Structure

Pashto syllables range from a single vowel to (V, e.g., /ao/ yes) to the maximum of three consonants in the onset and two in the coda position (CCCVCC, e.g., /ʃxwɑnd/ chew).

There is the possibility of a total of 12 syllabic patterns in this language. Section 3.3 shows all of these, with detailed discussion and examples. Since I have discussed the syllable structure and consonant clusters in detail in Chapter 3, I will only recap the important points here.

A Pashto syllable contains at least one vowel as the peak in the nucleus, which may be preceded or followed by consonant sounds. In Pashto, the nucleus is always filled by a vowel because the language does not have any syllabic consonant. The nucleic vowel may be preceded by a maximum of three and followed by a maximum of two consonants. The maximum syllabic template, MAXST, is as shown in Figure 6.1.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

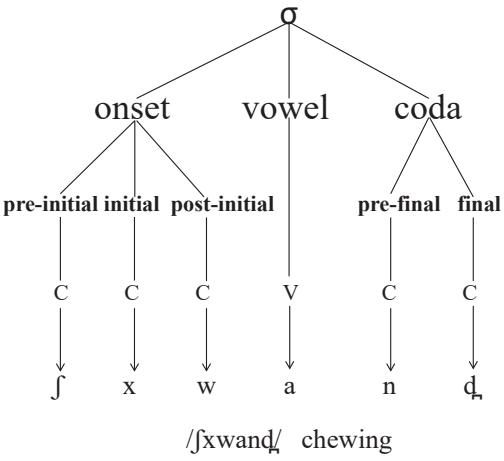


Figure 6.1: The Structure of MAXST in Pashto

It is clear from the MAXST shown in Figure 6.1 that the possible consonant cluster at onset can consist of up to three consonants (labeled pre-initial, initial and post-initial, respectively), while this goes up to two consonants at coda (pre-final and final). This division of the syllable is based on the picture put forward by Roach (2002). Thus, a variety of possible syllable structures is available in Pashto and most of the consonant sounds are allowed at onset position. The crucial point about Pashto syllabification and, as discussed below, the main characteristic that distinguishes it from English, is the abundance of reverse sonority clusters. The sonority sequencing principle (SSP) is generally satisfied in Pashto, but there are also some combinations in which this rule is violated. This point is discussed and theoretically justified in Chapter 3. Similarly, the rule of the MOP (maximum onset principle) is satisfied in Pashto.

Consonant clusters are possible at all three positions (word-initial, -medial and -final positions) but the maximal consonant combinations as clusters are possible at word-initial position. There is no restriction for onset consonants, and a large number of consonants are possible at coda level also. A total of 108 combinations of consonant clusters are possible in the base word forms in Pashto.[30] The next section explores consonant clusters in Pashto.

[30] For detail of these clusters see Chapter 3, Sections 3.5 and 3.6.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

6.2.2 Consonant Clusters in Pashto

Chapter 3 gives a clear-cut picture of Pashto consonant clusters. The first part of the chapter is dedicated to the analysis of the permissible strings in Pashto, while the second half is allocated to the OT analysis of the syllabic structure of the language. The maximal size of onset is CCC. Maximum combinations are possible at the initial position of Pashto words (syllables), while other combinations are also possible at word-medial and -final positions. Here I briefly summarize the main characteristics of and statistics relating to consonant clusters in Pashto.

6.2.2.1 Two-Consonant Clusters (CC)

Pashto syllables are quite rich in terms of two-consonant clusters. Maximum combinations are possible at onset level of syllables. As stated earlier, the language has an abundance of bipartite consonant clusters. It is remarkable for the initial permissible consonant combinations. Pashto not only allows a large number of two-consonant clusters but it also permits clusters of reverse sonority. The positions, combinations and number of these clusters are given below.

(100)

Position: word-initial
Combinations: /pr, pɽ, pl, br, bɽ, bl, tr, tɽ, tl, dr, dɽ, kr, kɽ, kl, gr, gɽ, pj, bj,
dw, dj, gw, tw, kw, qj, sr, zɽ, ʃr, ʃɽ, ʃl, xr, xɽ, ɣr, ɣɽ, ɣl, sk, sp, st, ʃk, ʃp, xk, xp, zg, zb, zd, sj, sw, zj, zw, xj, xw, ɣw, zɣ, ʃx, sx, zm, ʃm, ʃn, ng, nɣ, ml, mr, mɽ, mj, nw, nj, nm, rɣ, lm, lw, wr, wɽ, wl, dʒr, dʒw, tsw, dʒɣ, tsk/
Total: 77

Apart from word-initial position, a number of two-consonant clusters are possible at the word-final (coda) position. Such clusters are 33 in number and their detail is given below.

(101)
Position: word-final
Combinations: /xt, ʃt, ʃq, ʃk, sk, st, nd, nɖ, mp, nk, rg, rk, rd, rt, rp, ɽd, bs, ks, bz, qʃ, wz, mz, nz, ns, nʃ, rs, rz, rk, ɽx, xs, rm, rn, ndʒ/
Total: 33

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Among these two types of consonant clusters, there are several combinations which violate the sonority sequence. I have highlighted them in Section 3.8.3. The 22 combinations of reverse sonority are witnessed among all the two-consonant clusters. The behaviour of reverse sonority has also been examined by a number of studies in the past (e.g. Levi, 2004, and Bell & Saka, 1982, among others). This point is a very interesting aspect of Pashto phonology. The next section covers tripartite consonant clusters in Pashto.

6.2.2.2 Three-Consonant Clusters (CCC)

Apart from the two-consonant clusters, Pashto syllables also have the possibility of three consonant clusters. These three-consonant clusters are found at onset level only, and also include the reverse sonority combinations. They are shown as (102). (For further detail, see Appendix C.)

(102)
Position: word-initial
Combinations: /xwl, xwʈ, xwd, sxw, ʃxw, skw, xkw, ndr/
Total: 8

The above break-down of Pashto clusters shows that a number of consonant clusters are possible even in an SOV language such as Pashto. Pashto is especially rich in the case of initial consonant clusters. I calculated a total of 77 + 33 + 16 = 126 bi-consonant clusters in all three positions. It was also pointed out in Chapter 3 that some of the word-final consonant combinations were the same as clusters found in the initial position in the word. Therefore I recalculated those combinations and concluded that the true number of bi-consonant combinations in Pashto is 108. Even after excluding the repeated clusters, Pashto still enjoys a richness of consonant clusters. This clearly shows that the correlation between syllable structure and word order is not absolute. The large number of consonant clusters at both edges of Pashto words negates this correlation robustly. Let me now consider the case of English in the next section.

6.2.3 English Syllabic Structure

The English phonological inventory has a moderate number of vowel and consonant phonemes. Twenty-four are consonants, while 20 vowels are identified as phonemes. In English, there are some consonant sounds, such as /l, r, n, m/, which may function as nucleus in a syllable. These sounds

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

are called syllabic consonants when they do so. For examples, in words like “bottle” (where /l/ is syllabic) or ‘student” (where schwa is optional), the last syllables have no vowel but the function is performed by consonants. From a syllabification point of view, most of the previous studies assume that English satisfies the MOP and SSP (Giegerich, 1992; Duanmu, 2009, among others). Since much has been done in the field of syllables in English, the data presented here is based on other well-known studies, such as Jones (1976), Giegerich (1992), Roach (2002) and Duanmu (2009).

A vowel can also constitute a minimal syllable in English (e.g. I /ai/). As in Pashto, any English vowel can begin a syllable (word), and may be preceded by one, two or in some words three consonants. Similarly, a word can end with a vowel or it may be followed by one, two, three or in very few words four consonants. English syllables have a variety of possible structures. Syllabic patterns ranging from a single vowel (as in the words a or I: V) to those open and closed by several consonants (as in sprints: CCCVCCC) are available. The maximal syllabic template (MAXST) of English is CCCVCCCC. The detail of consonant clusters is given in the following section.

6.2.4 Consonant Clusters in English

In English, just as in Pashto, consonant clusters are possible at all three positions (initial, medial and final) of a word. So, both onset and coda are possible in English. The number of consonants possible in onset (word-initial consonant cluster) is up to three (Roach, 2002, p. 71), while a vowel can also begin a syllable. In the majority of the cases, the first consonantal sound is /s/, while the second is an approximant /l, w, r, j/ (Duanmu, 2009, p. 160). In coda position, the possible number goes up to four consonants. Note the following examples:

(103)

CC word-initial:	shrine, snake, sphere, etc.
CCC word-initial:	spring, strong, screw, etc.
CC word-final:	depth, books, graphs, etc.
CCCword-final:	films, holds, links, etc.
CCCC word-final:	prompts, sixths, texts, etc.

Apart from initial and final clusters, English also has some word-medial consonant clusters. Both intra-syllabic and inter-syllabic clusters are possible at word-medial position. Word-medial intra-syllabic clusters are

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

found in words like extra /ekstra/, and inter-syllabic clusters in words like description /diskripʃn/. Since the current study, as mentioned before, is concerned with intra-syllabic word-initial and -final consonant clusters, I will not mention inter-syllabic clusters here. A break-down of the possible consonant clusters is given below.

6.2.4.1 Two-Consonant Clusters (CC)

(104) shows the possible combinations of two-consonant clusters in English.

(104)
Position: word-initial
Combinations: /pr, pl, pj, pw, pf, ps, pʃ, br, bl, bj, tr, tw, tj, dr, dj, dw, kr, kl, kw, kj, km, kn, kv, gr, gl, gw, ŋj, mj, mw, fl, fr, fj, vj, vw, θr, θw, θj, st, sp, sk, sl, sw, sn, sm, sf, sj, sr, sv, zl, ʃr, ʃm, ʃn, ʃp, ʃw, hj/
Total: 55

The number of bipartite consonant combinations possible at word-final position is also 55. They are given below.

(105)
Position: word-final
Combinations: /pθ, pt, ps, bd, bz, tθ, ts, dz, kθ, kt, ks, gd, gz, mp, mf, mt, md, mz, nθ, nt, nd, ns, nz, nʃ, nʃ, ŋθ, ŋt, ŋd, ŋz, ŋk, ŋg, fθ, ft, fs, vd, vz, θs, ðd, ðz, sp, st, sk, zd, ʃt, ʃd, lp, lf, lθ, lt, ld, ls, lz, lk, ʃt, ʃd/
Total: 55

As mentioned earlier, English also allows three- and four-consonant clusters. Below is the case of three-consonant clusters.

6.2.4.2 Three-Consonant Clusters (CCC)

The possibility for three-consonant clusters in word-initial position is quite limited. Nine types of such combinations are available in English. All of them start with /s/, as given shown under (106).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(106)
Position: word-initial
Combinations: /skw, skj, skr, skl, stj, str, spj, spr, spl/
Total: 9

At word-final position, there are as many as 40 tripartite consonant clusters are many. Detail of these combinations is given below.

(107)
Position: word-final
Combinations: /pθs, pts, pst, tθs, tst, kθs, kts, kst, mps, mft, mfs, mts, mst, nts, ndz, nʃt, nʃd, ŋts, ŋst, fθs, fθl, fts, sts, lmd, lmz, lpt, lps, lbd, lbz, lft, lvd, lθs, lnd, lnz, ldz, lʃt, lʃd, lʃt, lks, lkt/
Total: 40

As discussed, there are also some four-consonant clusters in English; these are given below.

6.2.4.3 Four-Consonant Clusters (CCCC)

In English, at word-final position the syllable can end with a maximum of four consonants (Roach, 2002, p. 73). Such clusters are not allowed at word-initial position. The number of such cluster combinations is seven. Detail is given below.

(108)
Position: word-final
Combinations: /ksθs, ksts, mpts, ntst, lfθs, ltst, lkts/
Total: 7

The above survey on English consonant clusters shows that English is quite rich in the case of both bipartite and tripartite consonant clusters. Additionally, the combinations in (108) show that English tends to obey SSP, as all of the final consonants are voiceless. As far as the categorization of the syllables for both languages is concerned, it can be claimed at this point that Pashto and English both have complex syllable structures. The next section presents a brief OT analysis of them.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

6.3 AN OT ANALYSIS OF PASHTO AND ENGLISH SYLLABLE STRUCTURES

At every relevant point throughout this study, I have highlighted the power of OT as an ideal framework for typological studies. As introduced in Chapter 1 and subsequently applied in Chapter 3, OT exhibits a theoretically justified analysis for syllable structure. Section 3.7 provides a detailed analysis of Pashto syllable structure. The OT approach towards syllable structure is given there. Here I give an examination of the relevant features in Pashto and English for the purpose of comparison. OT assumes that there are a number of universally preferred syllabic patterns available to all languages. Such universal patterns are reflected in specific constraints, which may be violated variously by different languages. I first mention the important constraints below.

6.3.1 Relevant Syllabic Constraints

The constraints shown as (109) and (110) are relevant for the ongoing comparison of Pashto and English consonant clusters. These constraints are based on Itô (1989), Prince and Smolensky (1993) and Kager (1999).

(109) **Markedness Constraints**

a. ONSET Syllables must have an onset
b. NOCODA Syllables are open
c. *COMPLEXONS Onsets are simple
d. *COMPLEXCOD Codas are simple

(110) **Faithfulness Constraints**

e. DEP-IO Output segments must have input correspondents
"No epenthesis"

f. MAX-IO Input segments must have output correspondents
"No deletion"

Based on the above two types of constraint, an analysis of Pashto and English syllable structure is given below.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

6.3.2 Analysis

Starting from constraint (109 a), ONSET, I show that both Pashto and English have onset in their syllables. Both languages have CV and CVC types of syllable structures. These templates satisfy this powerful constraint. But on the other hand, there are syllables such as V, VC and VCC which do not satisfy this constraint, as shown by the examples in Table 6.1.

Table 6.1: The Case of ONSET in Pashto and English

	Pashto	English
Syllables satisfying the constraint	/xa:/ CV okay	/səʊ/ CV so
	/ra:.za:/ CV.CV come	/plei/ CCV play
Syllables violating the constraint	/am/ VC mango	/ɪz/ VC is
	/a.ba.si:n/ V.CV.CVC the Indus River	/ʌn.laɪk/ VC.CVC unlike

Table 6.1 shows that the constraint is not at the top of the hierarchy of the constraints in either Pashto or English. Both types of syllable are possible: those that satisfy and those that violate the constraint. Next, I consider the status of the NOCODA constraint in the languages.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6.2: The Case of NOCODA in Pashto and English

	Pashto	English
Syllables satisfying the constraint	/ta:/ CV you	/wi:/ CV we
	/plə/ CCV stick	/ʃəʊ/ CV show
Syllables violating the constraint	/ʃi:n/ CVC green	/nekst/ CVCCC next
	/əks/ VCC contrary	/weɪst/ CVCC waist

Consonants at coda level are considered to be marked universally. Intervocalic consonants (CVCV) in languages are preferred to be syllabified in onset position. This process is called MOP (maximum onset principle). Similarly, Pashto and English both prefer onset to coda but they also allow syllable patterns which end with coda consonants. Table 6.2 demonstrates these trends in the languages. It is clear from these examples that this constraint is ranked low and has no impact in the hierarchy of either language. Next I focus on the constraint *COMPLEXONS.

It is clear from Table 6.3 that neither Pashto nor English gives a high ranking to *COMPLEXONS. This means that this markedness constraint is dominated by faithfulness constraints, because both languages allow not only two-consonant clusters but also three- consonant clusters in the onset (and even four in case of the English coda). The discussion of consonant combinations in Section 6.2 has more examples as witness against this constraint. The same is the case with the next constraint, which is *COMPLEXCOD.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6.3: The Case of *COMPLEX in Pashto and English

	Pashto	English
Syllables satisfying the constraint	/na:r/ CVC hungry	/sʌm/ CVC some
	/ra:z/ CVC secret	/ʃeɪm/ CVC shame
Syllables violating the constraint	/tri:x/ CCVC bitter	/tri:/ CCV tree
	/sxər/ CCVC father-in-law	/streɪt/ CCCVC straight

Table 6.4: The Case of *COMPLEXCOD in Pashto and English

	Pashto	English
Syllables satisfying the constraint	/sa:n/ CVC whet	/fon/ CVC phone
	/ba:ɣ/ CVC garden	/ʃeɪd/ CVC shade
Syllables violating the constraint	/benʃ CVCC beam	/rest/ CVCC rest
	/tərz/ CVCC form	/ænd/ VCC and

As discussed earlier, since codas are not allowed, both codas and complex codas in syllables are universally considered marked cases among languages. But as in Pashto, English codas are also complex. Therefore, I can say that the constraint *COMPLEXCOD is also dominated by other

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

faithfulness constraints in both languages.

It is shown in the above examples that the constraints discussed so far are not the dominant ones, as their violation is possible in English and Pashto equally. It is also clear from the discussion that the faithfulness constraints are more important than the markedness constraints. In other words, DEP-IO and MAX-IO are dominant over ONSET, NOCODA, *COMPLEXONS and *COMPLEXCOD. Now I consider these faithfulness constraints.

I have already concluded from my analysis in the previous chapter that the correspondence constraints DEP-IO and MAX-IO are dominant over the markedness constraints in Pashto. Here, it will become clear that the same is the case in English. First, I consider an example from Pashto in tableau 111.

(111) /dʒwənd/ Life

Input:/ dʒwənd̪/	DEP-IO	MAX-IO	*COMPLEX^{ONS}	*COMPLEX^{COD}
a.dʒəwənd̪	*!			*
b.☞dʒwənd̪			*	*
c. dʒwən		*!	*	

Tableau (111) shows that DEP-IO and MAX-IO (faithfulness constraints) dominate *COMPLEXONS and *COMPLEXCOD (markedness constraints). Candidate (a) has no cluster at onset position, but it is not allowed as it does not satisfy the top-ranking constraint DEP-IO. Similarly, candidate (c) fatally violates MAX-IO as it deletes one segment from the input. So candidate (b) is the winning candidate. Although it violates both *COMPLEXONS and *COMPLEXCOD, it is still the optimal form as it satisfies the high-ranking DEP-IO and MAX- IO. Now I consider an example from English, given as tableau (112).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

(112) /prompt/ Prompt

Input:/ prompt/	DEP-IO	MAX-IO	*COMPLEX^{ONS}	*COMPLEX^{COD}
a. pərompt	*!			*
b. rom		***!		
c.☞ prompt			*	**

Tableau (112) shows that, as in Pashto, the English syllabic constraint hierarchy is dominated by faithfulness constraints DEP-IO and MAX-IO over *COMPLEXONS and *COMPLEXCOD. Candidate (a) has no consonant cluster at the onset as /ə/ is added in order to satisfy *COMPLEXONS. But it is not the optimal form as it violates the more powerful constraint of DEP-IO. Candidate (b) has neither onset nor coda cluster because segments are deleted and constraint MAX-IO is violated. Candidate (c), on the other hand, has onset as well as coda clusters but it is still the optimal candidate, as it fulfils the top-ranking faithfulness constraints. The optimal output (c) violates the lowest-ranking markedness constraints, which are equally low-ranking constraints relative to the syllable structure of the language. The violation of these two constraints is the least expensive in the language. Candidates (a) and (b) violate the faithfulness constraints MAX-IO and DEP-IO, which are ranked at the top, and therefore they are ruled out.

It is clear from the above points that the faithfulness constraints MAX-IO and DEP-IO are dominant over the others. Since I am not further concerned with the mutual ranking of MAX- IO and DEP-IO or *COMPLEXONS and *COMPLEXCOD, I delimit my analysis to the points explored so far. Finally, I can now consider the hierarchy of the constraints for syllable structure in Pashto and English. This is given as (113).

(113)
MAX-IO, DEP-IO›› *COMPLEXONS, *COMPLEXCOD.
“Both complex onsets and codas are allowed”.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

6.3.3 Summary

The OT analysis in the previous section provides enough evidence in favour of our hypothesis that both Pashto and English have *complex* syllable structures. Such sequences are further justified theoretically as they are allowed by the hierarchy of their relevant constraints. Though this was, as pointed out in the beginning, a brief analysis of the consonant clusters of the two languages, it demonstrates the point of similarity between the two. This does not mean that the languages are identical in every aspect of consonant clusters: there are a number of differences as well. For example, English allows four–consonant clusters, while Pashto does not. Similarly, English permits three-consonant clusters at both edges, but this is not permissible in Pashto. On the other hand, Pashto allows a large number of reverse sonority clusters which are not allowed in English. Since the focus of the current study is the consonant sequences, other points are left open for future studies. The next section takes up the comparison of the word order of the languages.

6.4 SYNTACTIC CHARACTERISTICS OF PASHTO AND ENGLISH

I have discussed above the syllable structures of both languages. The OT analysis further suggests that syllable structure formation is the result of the permutation of the relevant constraints in CON. In this section, I compare the word orders of Pashto and English. As with my analysis of syllable structure, here too I am concerned only with the main features of word order, as further explanation is out of the scope of the current study.

6.4.1 Pashto Word Order

In Chapter 4, I discussed in detail the fact that Pashto is a characteristic SOV language because this is the only structure which fits all three tenses of the language. This section gives a typological sketch of the syntactic structure of modern standard Pashto, i.e. Yousafzai dialect (Hallberg, 1992). Discussion of these features is necessarily brief and to the point, as I am concerned only with the most relevant features of Pashto word order. There are many important syntactic features to which this study gives only a passing mention or excludes as irrelevant to the present discussion. As discussed in Chapter 4, this study is only concerned with features which are important in the study of word order correlation.

As per the basic word order of the language in simple sentences, Pashto is a verb-final language. Word order in the sentence is relatively free as long as the verb-finality condition is fulfilled. Section 4.4.3 provides enough

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

material in this regard. Note that the following example, "Jan introduced Pari to Speen", can have the following variations in Pashto:

(114)

a.	dʒa:n	spi:n ta pəri milao kɽa
	Jan	Speen *(Acc.)* Pari introduce *(Past Dat.)*
b.	dʒa:n	pəri spi:n ta milao kɽa
c.	pəri	dʒa:n spi:n ta milao kɽa
d.	pəri	spi:n ta dʒa:n milao kɽa
e.	spi:n	ta dʒa:n pəri milao kɽa
f.	spi:n	ta pəri dʒa:n milao kɽa

In the above sentences, the only common thing is the verb-final position. Otherwise, the rest of the sentence differs in terms of word order patterns. As discussed in Section 4.4, this freedom is a result of the post-positional particles ta and kɽa, which represent accusative and dative case marking. These sentences are identical in terms of logical content but are distinct in discourse presupposition in a very subtle way. In the same section, I also pointed that this freedom only exists in specific tenses and contexts. Similarly, the subject-initial pattern is also the most common among the permissible word orders of the language in all tenses. SOV is the most frequent and most possible pattern in Pashto. It is the most dominant order, and is almost always obeyed in declarative sentences with nominal object and subject. Pashto, as a constant SOV language, shows all the features that Greenberg (1963) attributes to SOV languages. One of these characteristics relates to adpositions. Greenberg's language Universal 4 suggests that languages with normal SOV pattern are post-positional. Pashto has a number of post-positional adpositions. A representative sample of some adpositions is given, classified according to their functions. Quantifier-like particles are given below.

(115)

a. spi:n ham qabɪl de
Speen also genius is
Speen, too, is a genius.

b. spi:n jawa:ze ra:ɣlo
Speen only came
Only Speen came.

There are some other particles, such as noun-coordinating particles, case-marking particles and sentence-final particles, which show the same

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

behaviour. Some examples are given below.

(116) **Noun-Coordinating Particles**
a. spi:n ao dʒa:n ra:ɣlo
Speen and Jan came

b. spi:n ao dʒa:n de raɣlo
Speen and Jan (and others) came

c. spi:n ja dʒa:n ba ra:zi:
Speen or Jan will come

(117) **Case-Marking Particles**

spi:n ao ba:tʃa de ba motər ke pexəwər na mərdan ta zi:

Speen and Bacha (and others) will motor by Pashawar from Mardan up-to go

Speen and Bacha (and others) will go by car from Pashawar to Mardan

(118) **Sentence-Final Particles:**
ta da kɪtab wele de
You this book read is (Q)
Did you read this book?

Note that Pashto uses the sentence-final question particles which are characteristic of post-positional languages. Also note another example of the type:

(119)
Isla:məba:d ta dʒa:n təle de
Islamabad to Jan went
Jan went to Islamabad

b. səudi: na xət ra:ɣəle de
Saudia from letter came
A letter came from Saudia

The above examples show that these particles are post-positions rather than suffixes. All of them highlight that Pashto is a strictly verb-final language.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

In Pashto, interrogative sentences are formed by using the question particles /sa/ or /ka/ at sentence-final position, regardless of whether it is a "yes or no" question or another type of interrogative sentence. Examples are given below in (120).

(120)
a. xaːn raːɣe sə
Khan come Q
Did Khan come?

b. xaːn kaːr xlaːs kɽo kə
Khan work finish did Q
Did Khan finish his work?

From the above examples, I reconfirm the suggestion already made by a number of studies and theoretically justified by our OT analysis in Chapter 3 that Pashto is a characteristic SOV language. To continue my comparison, I now turn to the word order of English.

6.4.2 English Word Order

English is among the most highly consistent of SVO languages. It exemplifies the characteristics of SVO languages, such as the grammatical processes used to highlight the elements of sentences. Other SVO features include the heavy use of auxiliaries as substitutes in negative and interrogative sentences, high usage of clefting and passives, and the presence of prepositions, etc. (Lehmann, 1978). What follows is a brief examination of the characteristic typological patterns of English which are directly relevant to the comparison of the language with Pashto and required for the ongoing discussion on the subject correlation. So, needless to say, I am mainly concerned with the basic word order structure of English.

In English, as discussed earlier, simple unmarked clauses agree with the SVO pattern and require the three constituents of subject, verb and object. Note the following example given by Lehmann (1978).

(121)
Alice folded her hands

Lehmann says that none of the three constituents is omissible from (121). Similarly, the order of them is also very important, as English does not permit any order other than SVO in unmarked sentences occurring as a

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

single utterance. So, all the sentences in (122) are impermissible in an isolated context:

(122)
a. *Folded Alice her hands
b. *Folded her Alice hands
c. *Folded her hands Alice
d. *Her hands folded Alice
e. *Alice her hands folded

The above mandatory syntactic pattern requires the classification of English as an SVO language. So the basic pattern of modern English is quite consistent.

In sentence (122) the agent of the action is also the subject, and it occupies the initial position of the structure. Similarly, the object and verb have their exact positions in order to fulfil related grammatical, logical or psychological functions. Theoretically, all sentences of an SVO language should have this canonical pattern. Lehmann (1978) shows a number of related features of SVO languages. The abundance of pronouns which are used for placing subjects in their specific positions is one of them. Another characteristic is the common feature of passivization in English. All these patterns show the constant type of the language. According to the same source, passivization allows an SV sequence when the agent is unknown and unimportant. Such examples are very common in English. Some are given below.

(123)
William the Conqueror, whose cause was favoured by the Pope, was soon submitted to by the English. (Lehmann, 1978, p. 208)

Lehmann is of the view that avoidance of mentioning an agent may lead to emphasis on the verbal phrase as in the above sentence: the action is highlighted more than the agent. Secondly, the passivization (or even deletion, another characteristic feature of SVO languages) provides foregrounding for the predicate, whether its object or its verb, or both.

There are some morphological features which are associated with specific types. Affixation is one of them. Lehmann (1978) finds that in VO languages prefixes are more common, while OV languages have suffixes. In the case of English, both types of affixation are used. As far as the inflections are concerned, previous studies (e.g. Greenberg, 1966) show

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

that such affixes in VO languages are pre-posed, while in OV languages they are post-posed.

6.5 AN OT ANALYSIS OF PASHTO AND ENGLISH BWO

Chapter 4 analyses the BWO of Pashto in depth. Here I provide an OT analysis of the BWO of both Pashto and English. The main aim of this comparison is to show the different word orders of the languages. The point I wish to make is that syntactically distinct languages can have identical syllable structure. This point ultimately leads me to highlight the power of OT to justify such linguistic structures within languages. I will finally conclude that OT is the ideal framework for cross-linguistic studies. Needless to say, the present analysis is to the point and strictly relevant to the basic word order of Pashto and English. It is also pertinent to say that the sentences for both languages have been taken from everyday language. For Pashto, recordings of native speakers were analysed, while for English, as mentioned before, the sentences were taken from real-life contexts. The relevant syntactic constraints are introduced first.

6.5.1 The Relevant Constraints

A detailed OT analysis of Pashto sentence structure is given in Chapter 4.[31] Similarly, English word order is an area frequently explored using this framework (Grimshaw & Samek- Lodovici, 1995, 1998; Costa, 1997, 1998, 2001, among others). This section is dedicated to the comparison of Pashto and English, therefore I use the same constraints which I applied to Pashto BWO analysis earlier in the study. The relevant five constraints are given here briefly. They are all previously introduced constraints, based on Grimshaw (1997), Choi (1999) and Morimoto (2001).

(124)

a. HD-LT: The head is leftmost in its projections
b. HD-RT: The head is rightmost in its projections
c. SUBJ-LT: A subject aligns left in the clause
d. SO: Subject precedes object
e. STAY: Trace is not allowed (Economy of Movement)

[31] Chapter 4 also provides in-depth insight on previous works and an introduction to OT as a framework for syntactic analysis. To avoid repetition, the analysis provided here is necessarily brief.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

6.5.2 Analysis

Having introduced the relevant constraints, let me consider how these constraints interact with each other in the hierarchy. I begin by exploring the word order of Pashto observed previously in Chapter 4. I consider the following tableau here as the final evidence for Pashto being an SOV-type language.

(125) dza:n pətəŋga:n xərtsəwi: Jan sells kites.

pəʈəŋga:n dʒa:n xərtsəwi:	HD-RT	SO	SUBJ-LT	STAY	HD-LT
a. xərtsəwi: dʒa:n pəʈəŋga:n	*!		*		
b. pəʈəŋga:n dʒa:n xərtsəwi:		*!	*		*
c. dʒa:n xərtsəwi: pəʈəŋga:n	*!			*	*
d.☞ dʒa:n pəʈəŋga:n xərtsəwi:				*	*

Tableau (125) shows that the top-ranking head constraint is HD-RT. This constraint makes it possible for the head (verb) to be on the rightmost edge of the sentence. Candidate (d) is optimal due to this constraint. If this constraint is violated, the meanings are changed, especially in sentences as given in (125). Constraint HD-RT is at the top in the hierarchy. SO and SUBJ-LT are ranked equally and both are crucially important as well. Likewise, both of the lower-ranking constraints are equally violable, as their violability makes possible the satisfaction of the top-ranking constraints. Thus, the satisfaction of top-ranking constraints HD-RT, SO and SUBJ-LT is based on the violation of low-ranking STAY and HD-LT. When I revise the ranking, the word pattern will be SVO or VOS. Therefore, the present ranking is crucial. This type of structure in head-final languages such as Pashto is possible due to the top ranking of HD-RT in the hierarchy. HD-RT performs an interesting interaction with other constraints in the hierarchy, such as SUBJ-LT and SO and especially those on head positioning (HD-LT and STAY).

It is clear from the above analysis why sentences (126 b–d) do not convey identical meanings.

(126)

a. dza:n pətəŋga:n xərtsəwi: Jan sells kites

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

b. *xərtsəwi: dʒa:n pətəŋga:n sells Jan kites
c. *pətəŋga:n dʒa:n xərtsəwi: kites Jan sells
d. *pətəŋga:n xərtsəwi: dʒa:n kites sells Jan

The following ranking for Pashto BWO is evident from the analysis above.

(127)
HD-RT» SO, SUBJ-LT» STAY, HD-LT

This shows that SOV (subject + object + verb) is the only order which is available in unmarked and syntactically simple sentences in Pashto. Let me now show the case of English.

Section 6.4.2 gives a number of sentences demonstrating English word order pattern. I consider the same constraints and their interaction in (121), a sentence of pragmatically neutral and descriptively simple nature: "Alice folded her hands".

(128) Alice folded her hands.

Alice folded her hands	HD- LT	SO	SUBJ-LT	STAY	HD- RT
a. Alice her hands folded	*!			*	
b. her hands Alice folded	*!	*		*	
c.☞Alice folded her hands				*	*
d. folded her hands Alice		*!	*		*

Tableau (128) demonstrates the top-ranking constraint HD-LT. This makes it possible for the verb to be on the left-hand side of its projection and ensures that the optimal candidate (i.e. c) is head-initial. Violation of this constraint results in the loss of meaning, as is evident in candidates other than the optimal one in the above tableau. So, HD-LT is at the top in the hierarchy. The ranking of SO and SUBJ-LT is equally crucial here. The two low-ranking constraints are equal and both are violable. Their violability justifies the satisfaction of the top-ranking constraints (i.e., HD-LT). Revising the ranking of HD-LT with HD-RT gives patterns such as SOV. It is clear from the above analysis that sentences (a), (b) and (d) are not possible in a pragmatically natural context. Similarly, it is also clear that none of the three constituents is omissible from the sentence. Additionally, their order is crucial, as English does not permit any order other than SVO.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

The above mandatory syntactic pattern requires the classification of English as an SVO language because this order is quite consistent in modern English. I suggest the ranking given in (129) on the basis of my analysis above.

(129) HD-LT» SO, SUBJ-LT» STAY, HD-RT

This ranking of English BWO characterizes the behaviour of the language. It demonstrates the following points:

(130)

4. Subject moves to Spec IP, which is the leftmost position of the projection.
5. Objects always follow their subjects.
6. The top ranking of HD-LT reflects the head-initial position in a clause.

These points suggest the classification of English as an SVO (subject + verb + object) language. In the following sections, I summarize the observed similarities and differences between Pashto and English.

6.6 OBSERVED SIMILARITY BETWEEN PASHTO AND ENGLISH

This section summarizes the observed similarity between the syllable structures of the two languages. This similarity is measured in terms of the nature and statistics of possible consonant clusters found in these languages.

- Pashto and English both allow consonant clusters in onset position in all types of words (function and content words).
- They also allow consonant combinations at coda level in all types of words.
- Such consonantal clusters are possible at all three positions (word-initial, -medial and -final).
- The category of syllable structure as per our categorization in Section 2.3, in both languages, is *complex*. Such words are frequently used in everyday language.
- Both languages have syllables ranging from a single segment (a vowel) to a maximum of three consonants in one string (although English further allows four-consonant strings in some words and positions: see below).

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

The above points show that most of the features of the syllable structures of Pashto and English are similar, though some differences are also found, for example:

- English additionally allows up to four consonants at coda position; in Pashto, this limit is two.
- Pashto permits a larger number of word-initial clusters (77) than English does (55).
- Pashto has a large number of SSP-violating combinations which are not permitted in English.

Despite these minor differences, the syllable structures of the subject languages can be summed up as similar to each other. According to the categorization of Dyer (2005), they can be classified as *complex* syllable structures. The theoretical justification for identifying these similarities is given using OT analysis in the next section.

6.7 OBSERVED DIFFERENCES BETWEEN PASHTO AND ENGLISH

The previous section highlights a number of points of similarity in terms of the syllable structure of the two languages. Now, having gone through the syntactic orders of both Pashto and English, I can summarize points of difference in the following lines.

- The main difference between Pashto and English is in the word order. Pashto is SOV, while English is SVO.
- Pashto is mainly post-positional, while English is pre-positional.
- Pashto is dependent-head, while English is head-dependent.

As shown in the above two sections, I conclude that Pashto and English are phonologically similar (in terms of complex syllable structure) but they are syntactically different (in terms of BWO). This is a clear-cut negation of the so-called correlation between syllable structure and word order. The point is further discussed in the section below.

6.8 DISCUSSION

One of the main objectives of this study is to evaluate the validity of the correlation between syllable structure and word order. This correlation has been shown to be invalid in the case of Pashto and the comparison within this chapter has also validated the hypothesis that the syllable structure of an SOV language may not always be a simple one. This was explained by

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

comparing Pashto and English in the current chapter. The assumptions underlying the theoretic explanation are supported in this chapter. In particular, the analysis of syllable structures in both subject languages demonstrates that determining the syllable structure of a specific language is a matter of re-ranking the syllable constraints in the hierarchy of that language. These outcomes also support the suggestions given in the previous chapter about the irrelevance of word order as an influence on the syllable structure of a language.

Importantly, this distinction between word orders does not affect the syllable structure of a language. The crucial point is that syllable structures are organized in terms of the hierarchy of relevant constraints within a system which is acquired by native speakers in shaping their L1 phonology. I also suggest, on the basis of the treatment demonstrated above, that the patterns of a specific language are controlled by the hierarchy of the constraints available for that specific structure. These constraints are the regulations of basic patterns of a specific structure, while the hierarchy of them regulates the harmony of expression (agreement, in the general sense). Thus one can conclude that the basic patterns of any language are accordingly determined by the constraints available for them. These constraints may change from simple to powerful and vice versa, as per the nature of the hierarchy of the language, for specific morphological, phonological or syntactic structures.

In the case of Pashto and English, previous typological predictions were based on very limited data and scope. The purpose of analysing all constraints for a specific pattern in OT is that, in order to analyse a correlation, one must have all possible patterns for a structure. The subject correlation does not include all possibilities for SOV languages. If the SVO type includes the possibility of complex syllable structure as a straightforward sequence in its structure, so do other types such as SOV. All this means that the picture is more complex than a simple correlation between syllable structure and word order. In the next section, I summarize the main points of the comparison between Pashto and English as it relates to this supposed correlation.

6.8.1 Summary of the Comparison

The comparison in the current chapter indicates strong parallels between the syllable structures of closed and open syllables in the two languages under consideration. Table 6.7 summarizes the comparison of important points relating to the subject correlation.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6.5: Comparison between Pashto and English

	Pattern	Pashto	English
Syllable Structure	Syllable Category	Complex	Complex
		Both closed and open	Both closed and open
		(C)(C)(C)V(C)(C)	(C)(C)(C)(C)V(C)(C)(C)
	Two-consonant clusters		
	Word-initial	77	55
	Word-final	33	55
	Three-consonant clusters	8	49
	Four-consonant clusters	-	7
	Reverse sonority clusters	Found	-
	MAXST	CCCVCC	CCCCVCCC
	Word class mainly containing complex syllables	Content words	Content words
Word Order	Word Order	SOV	SVO
		Verb-final	Verb-initial
		Dependent-head	Head-dependent
		Both post-positional and pre-positional	Pre-positional

Table 6.5 illustrate very important points of similarity between Pashto and English with respect to consonant clusters and syllable structures. Whereas both languages have a large number of clear consonantal combinations, Pashto has syllabification with consonant clusters with reverse sonority clusters. A preliminary contrastive analysis based on the consonant combinations of Pashto and English indicates that the two languages have comparable clusters with respect to the correlation. On the other hand, serious diversity is found in terms of their word orders. This highlights the straightforward possibility of complex syllable structure for SOV-type

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

languages. Importantly, the usage of words with complex syllables is also very frequent in Pashto. The frequency of words with complex syllables was calculated for Pashto: from the sample of 5000 words collected from Pashto speakers living in rural areas, this frequency was 10% in content words. In both languages, such words are content words in frequent everyday use.

The claimed correlation also neglects the diverse patterns resulting from language change. When change occurs in small segments of structures, it can take a long time before it is clearly visible. For example, the change in the voicing of /t/ in some areas of American English is confined to word-medial sounds, as in *Minnesota*, and when it is followed by an unstressed vowel-contrast, as in *deter*, but is not evident if the unstressed syllable ends in /n/ – contrast *button* with *butter*. This phonological change has been in progress for at least a century (Lehmann, 1878). Changes in syntactic, morphological or semantic structures may require even longer periods.[32] Such changes are adopted by a speech community very slowly. But when these changes are considered, it is clear that languages are never static or completely regular or symmetrical. This lack of symmetry in one level or structure cannot be correlated with other structures. OT provides the logic for such a change – it is the rearrangement of the constraint hierarchy. On the basis of the comparison of Pashto and English in the current chapter, it is argued that this can also be understood through and supported by the analysis of the language of a new learner, in terms of the grammatical differences between his or her native and target languages. In the next section, I discuss the OT approach to typological studies.

6.8.2 OT Approach to Typological Studies

While explaining the goals of typological study, Lehmann (1978, p. 345) states:

> [typological study] is based on the view that linguistics is an empirical science. Theory is formulated in accordance with the data of languages, and tested by means of such data. In carrying out investigations, observation and description are certainly directed by one's theory, [which] provides "discovery procedures". But the strength of a theory lies in explanations it affords, or in understanding it contributes.

[32] Lehmann (1878) also gives the example of regularization of verbs such as *dive*/*dove* into *dive*/*dived*, a process which has been going on for the last millennium, to highlight the slow process of morphological changes.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Keeping in mind these objectives, OT as a framework of constraint hierarchy provides enough space to achieve such goals in order to carry out a cross-linguistic investigation. The term *explanation* has also changed its meaning in the field of linguistics since the nineteenth century. At one time it referred to the determination of an earlier form. Attempts were made to explain factors behind a specific form (e.g. an irregular verb like *sing*/*sang*/*sung*). Such explanations were attempted in the light of the diachronic study of the respective language. But there were a number of limitations to this understanding of *explanation*, as in the case of languages without a documented history it was not possible to provide such explanation. Thus, the explanatory aspect of historical linguistics was seriously in question. The term, more recently, was revolutionized by Chomsky's (1964) concept of transformational grammar. His goal was to explain language as a phenomenon, and specifically the treatment of linguistic structure by the human mind. In his own words:

> [such an explanation would] rest on the assumption that the concepts of grammatical structure and "significant generalization" made explicit in this theory constitute the set of tools used by the learner in constructing an internal representation of his language (i.e., a generative grammar), on the basis of presented linguistic data. (Chomsky, 1964, p. 928).

This approach based on the psychological explanation of language as a phenomenon proved very interesting in the beginning. One major contribution of transformational grammar was the discovery of the validity of the concept of abstract underlying linguistic structures. The notion of these structures ultimately led to the discovery of universal rules in linguistics. Later on, when some of the explanations offered by transformational grammar proved less satisfying, linguists started trying other approaches.

The typological approach also emerged initially as a theory for explaining and predicting the nature of human language. It involves the search for unmarked patterns of language, and these can easily be found through the framework of OT. OT is typological in its nature and it suggests that, for a specific pattern, one must know all the relevant constraints. Once these are determined, one can check from the data the unmarked structure of a pattern. My observations in Section 5.4.7 submit a number of points for careful typological investigation. Some principles are also formulated here for the typological investigation of languages.

It is evident from the data of Pashto and English presented in this chapter that the supposed correlation which is the subject of this study is merely an

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

attempt to visit the utopia of Gabelentz. It is beyond possibility to correlate the syntactic structure (word order) with the phonological structure (syllable pattern) of a language. Chapters 3 to 5 illustrate this with the help of data and analysis, and provide "proof of negation" for the subject correlation. The current chapter shows another aspect of typological study by comparing Pashto with English, and thus provides enough space for suggesting and formulating methodologies for cross-linguistic investigation.

Language theory must include all principles in terms of universal constraints for a specific structure. The hierarchy must also necessarily be explained. OT should be applied in accordance with other principles laid down by a typological approach. The principles mentioned above can explain an unmarked pattern for a given structure. This correlation has not yet been adequately investigated; the current study has pointed to a number of problems which have been overlooked by the previous studies on it.

The conclusion derived from such an OT typological study highlights certain issues with correlative implications and provides a guide for the further study of specific (e.g., syntactic or phonological) linguistic structures.

6.9 SUMMARY AND CONCLUSIONS

The comparison of Pashto and English in the current chapter illustrates how specific constraints govern language structure. Since the comparison deals primarily with the correlation between syllable structure and word order, the conclusions may suggest that any individual level of structure may not rigidly be confined to one single set of patterns. It is now clear from the ongoing discussion that complex syllable structure is not characteristic only of VO-type languages. Other language types, such as OV (including Pashto), can have such syllable structure. Similarly, it is also evident that simple syllable structure is not the only possible structure for OV languages: they can also have complex syllable structure.

The interrelationship between these two constructions is not seen logically as suggested by the so-called correlation. It is also observed that the selection of a framework for typological studies is especially important when the relationships between two different levels are examined. Additionally, while using frameworks such as OT, linguists must know all the related constraints which are active within a specific pattern of a language. The language should first be well documented. Such an

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

approach may lead towards improved an understanding and history of a language.

The main goal of linguistics is to explain the language phenomena, and this goal can be achieved through typological studies. To describe a theory of language (grammar) with "predictive capability" is the ultimate objective within the sphere of cross-linguistic studies. The power of OT, as reviewed above, is shown in its ability to predict and explain specific phenomena of language. The framework of OT can make possible the contribution of typological study to achieving the aims of linguistics. It is true that many of the structural characteristics of languages have been determined by linguists in the past, yet change leads to inconsistencies in structure. The possibility exists of change to the patterns of specific structures, whether in OV, VO or any other class of language, as shown in Chapter 5 in detail. I suggest that a deeper level of knowledge of constraints for a specific linguistic structure is essential for analysis, description and understanding of language structures. Thus, the OT framework best explores the fundamentals of the typological study of languages, whether it is synchronic or diachronic, theoretical or practical.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

CHAPTER 7

GLOBAL ANALYSIS AND CONCLUSIONS

7.1 OVERVIEW AND FINDINGS

At the beginning of the current study, in Section 1.2, the following main research questions were raised:

Q1: What are the patterns of syllabification in Pashto language? Q2: What is the basic word order of Pashto?

Q3: What are the similarities and differences between Pashto and English syllable structures?

Q4: What are the similarities and differences between Pashto and English word order? Q5: Is the nature of the syllabic pattern related to the word order of a language?

To answer these questions, the present chapter provides a detailed evaluation of the theoretical findings of the preceding chapters and claims against the correlation between syllable structure and word order. The main insights of the preceding chapters were the following:

a. The analysis of the combinations of the various types of consonant clusters in Chapter 3 has enabled me to decide the maximum syllabic template (MAXST) for Pashto. It was defined as follows: the maximum Pashto syllable consists of three consonants in the onset and two consonants in the coda. The onset and the coda may both violate the Sonority Sequencing Principle (SSP).

b. The second main finding of this study, in Chapter 4, is the rigid verb-finality (SOV) of Pashto and its theoretical justification under the paradigm of OT.

c. On the basis of these findings, in Chapter 5, I proposed that the correlation between syllable structure and word order (as suggested by

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Lehmann, 1973, 1978; Donegan and Stampe, 1983, 2004; Donegan, 1993; Gil, 1986; Plank, 1998; Tokizaki, 2011; Tokizaki & Kuwana, 2012) does not exist, at least in the case of Pashto. Importantly, the usage of words with complex syllables is very frequent in both Pashto and English. The frequency of occurrence of words with complex syllable was calculated for the 5000 Pashto words collected by the researcher. This data showed frequent use of such content words in everyday language (10%; see appendex D for detail). Thus, it is argued that pattern of syllable structure is not correlated with word order.

d. This study also compares the syllable structure and basic word order of Pashto with those of English. This comparison was summarized in Table 6.7, which shows that despite similar complex syllable patterns, the two languages differ in terms of their word orders. The discussion was progressed by reviewing the claimed correlation in the light of the hypotheses outlined in the first chapter. The findings that Pashto is richer than English in terms of two-consonant clusters in the onset position (77 and 55 clusters, respectively), and the frequent combinations of three-consonant clusters (40) in the coda in English, indicate that the main feature in the categorization of syllable structure is the hierarchic schemata of the constraints given by OT. Importantly, word order pattern appears not to be correlated with – indeed, to be indifferent to – the syllabic structure. More interestingly, the comparison between Pashto and English does not confirm the predictions of the so-called correlation, according to which the syllable structure in Pashto should be the simpler one.

e. Finally, this comparison of the two languages reveals some important facts. The first is the similarity in the structure of syllable patterns. This similarity includes only the categorization given by Dryer (2005), which says that such patterns are called complex. However, other important aspects were also noted, such as the fact that English allows a maximum of four and Pashto a maximum of three consonants at the edge of the syllable. The second important finding is the difference in terms of the word orders of the languages.

In light of the above findings, I highlight the final conclusions and their implications in the next section.

7.2 FINAL CONCLUSIONS AND IMPLICATIONS

The current study makes a number of primary contributions to the field of typological linguistics. The first concerns the importance of examining the

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

active constraints for consonant clusters for the classification of language. The main findings of the study provide very interesting starting points for further detailed future work on related points. Importantly, the Pashto–English comparison (both syllabic and syntactic) and the optimality-theoretic treatment yield a number of relevant points for future studies. As the findings of the study run counter to the suggestions made by the earlier studies mentioned, it is important to carry out further work of the same nature on other language families. However, it may be necessary to investigate a large number of language families to better assess the effects of word order change in terms of syllable structure.

Further, in light of the strong similarity in terms of consonant clusters found by comparing Pashto with English in Chapter 6, it was clear that simple syllable structure is not correlated with SOV word order. The possibility of complex syllable structure within a language of any basic word order pattern is also predicted by the current study. As theoretical justification, it was suggested that complex syllable structure in a given language is the result of the specific hierarchy of syllabic constraints in CON. With the help of data from Pashto and English, it was concluded that complex syllable structure in a language may ultimately be attributed to the satisfaction of the top-level constraint FAITH. FAITH, as introduced in Section 3.8, requires that the output candidate does not alter the syllabification of the input form (observing both MAXIO and DEPIO). The comparison of Pashto and English and the ultimate similarity found in the case of complex syllable structure in Chapter 6 shows that this constraint is also relevant in the case of English. In short, the avoidance of a MAXST violation in Pashto requires that no syllable should have a cluster of more than three consonants in the onset, or more than two in the coda. Similarly, the assumption of the high-ranking FAITH constraint also provides the mechanism for explaining the unique type of Pashto syllabification of reverse sonority that has been very problematic for previous theories. Thus, both of the hypotheses are found to be confirmed. These were the following:

- The syllable structure of SOV languages is not always simple.
- The constraints on syllable structure are not directly related to the typology of word order in a particular language.

The investigation in the course of this study suggests that both of these statements are true in the case of Pashto. In other words, the nature of syllable structures in terms of their consonant clusters is not directly correlated with the basic word order of a language.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

The data for and comparison between Pashto and English in this study have consequences for the notion of human language itself. I am led to the view that language types are not formed on the basis of the classification of one structure (e.g. SOV type). One can perhaps regard a type as a family of variant subtypes which obey and share with other types at least one typological categorization. In this regard, one can conclude that the Greenbergian SVO/SOV/VSO trichotomy is no more logical in nature. In other words, it is fully possible, thanks to the typological nature of OT, that one can predict and justify an SOV language with complex syllable structure. OT provides for the possibility of variations across languages, and the "possible human language" can be explained in terms of a specific structure only when the range of varying patterns across languages is surveyed. For such an approach, one must search for the regularities underlying and explanations for the attested and non-attested variants of those structures. The case of Pashto is perhaps a non-attested case which can now, after the current study, be put into the attested category in order to provide clearer insight into the syllabic structure of SOV languages.

More importantly, the present study has also shown evidence that syllable structure is a different domain of linguistic investigation to, and cannot be correlated with, word order patterns. Having examined the correlation between syllable structure and word order in terms of its possibility in Pashto, and having compared it with English, one can predict that the simple syllable structure is not correlated with SOV. A number of analyses were conducted to test this correlation. Firstly, data on Pashto syllable structure and word order was analysed to determine the possibility of complex structure within an SOV language, and this possibility was found as expected. In the case of Pashto the presence of an SOV word order, significantly, was found to be unrelated to the likelihood of simple syllable structure. Secondly, the comparison of syllable structures in Chapter 6 shows that, despite their identical syllabic structure, Pashto and English differ significantly in terms of their syntactic structure (word order). Thus, both languages have complex structures for their syllables. Pashto two-consonant combinations with equally good scores syllabify in an English-like way. Finally, Pashto also shows a number of reverse sonority clusters in bisyllabic combinations, and this phenomenon was theoretically justified under the OT paradigm.

To sum up, the analysis of the results comparing both syntactic and phonological structures provides negative evidence against the correlation between word order and syllable structure. This highlights the global implication of the study, which challenges robustly the validity of a

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

number of studies mentioned already at different places throughout the work. The subject analysis has shown that phonological structure (syllabification patterns) is not necessarily correlated with syntactic structure (word order). This is explained with reference to data from two languages. This leads to the very plausible conclusion that the interaction of active constraints within a particular linguistic structure plays a significant role in the categorization of its pattern. Therefore, one may expect to find more evidence of this phenomenon in the study of a larger sample of other languages. The possible related directions for future study are given in the following section.

7.3 DIRECTIONS FOR FUTURE RESEARCH

This study has been devoted to the investigation of the correlation of syllable structure and word order. In the course of seeking to provide answers to the questions discussed at the start of this chapter, a number of new issues have been highlighted that deserve further detailed investigation in future. The current section concludes this work by outlining the most important directions for future investigation.

I have claimed that complex syllable structure is very much possible if a language keeps the FAITH (DEPIO and MAXIO) constraint at the top of its hierarchy. By doing so, a language may even allow codas and onsets to have reverse sonority combinations in their syllables. I hope that these claims will be vetted by scholars in the not-so-distant future. In the meantime, I suggest some of the interesting topics relating to the field.

- With respect to the similarity of the syllable structures and different word orders of Pashto and English, there are a number of questions. For instance, to what extent do these similar structures of segmental clusters prove helpful for Pashto learners of English in acquiring these patterns, and do they achieve the same level of productivity as native speakers? These questions could be investigated, for instance, by means of an study with a larger sample at a range of proficiency levels. Comparing the knowledge and productivity of the consonant clusters could allow a better understanding of the role and the process of acquisition of phonotactics as well. Such work could be conducted in the context of second language acquisition research.
- The OT perspective used in the study has made it possible to analyse the syllabification of Pashto, and especially the SSP-violating clusters in the language. I believe that this methodology

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

could further be applied in the investigation of sonority repair strategies from the morphological point of view, and it could have ramifications for synchronic analyses of the syllable structure of other languages spoken in the area.

- Finally, the study of other interesting typological aspects of Pashto phonotactic constraints could be based on the articulatory markedness or sonority sequencing preferences of Pashto, e.g. the constraint on stop + stop or fricative + stop onset sequences. Such constraints yield very interesting discussion on Pashto and further work needs to be carried out to determine if such factors have any role in phonology.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

BIBLIOGRAPHY

Alvar, M. (1967). *Linguistica Romanica* [originally by I. Iordan, reworked and heavily annotated by M. Alvar.]. Madrid: Ediciones Alcala.

Archangeli, D. B., & Langendoen, D. T. (1997). *Optimality theory: An overview* (Vol. 1). Oxford: Blackwell.

Babrakzai, F. (1999). *Topics in Pashto syntax. Doctoral dissertation*: Manoa: University of Hawai'i.

Bartsch, R., & Vennemann, T. (1973). *Semantic structures: A study in the relation between syntax and semantics*. Frankfurt: AthäenumVerlag.

Bečka, J. (1969). *A study in Pashto stress*. Prague: Oriental Institute in Academia.

Bell, A., & Hooper, J. B. (1978). *Syllables and segments*. Amsterdam: North-Holland Publ. Co.

Bell, A., & Saka, M. M. (1982). Reversed sonority in Pashto initial clusters. *The Journal of the Acoustical Society of America, 72*, S100.

Blevins, J. (ed.) (1995). The syllable in phonological theory. In J. Goldsmith (ed.), *Handbook of Phonological Theory*, 206-244. London: Basil Blackwell.

—. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.

Blevins, J., & Garrett, A. (2004). The evolution of metathesis. In Hayes, B., Kirchner, R, & Steriade, D. (eds.) *Phonetically driven phonology*, 117–156. Cambridge: CUP.

Bresnan, J. (2000). Pidgin genesis and optimality theory. In Siegel, J. (ed). *Processes of language contact: Case studies from Australia and the Pacific*, 145–173. Armidale NSW: Fides.

—. (2001). *Lexical-functional syntax* (Vol. 16). Blackwell textbooks in linguistics. Oxoford: Blackwell.

Brody, J. (2009). Some problems with the concept of basic word order. *Linguistics, 22*(5), 711–736.

Bukhari, K. (1960). On learning Pashto grammar for Urdu speakers. In Rehman (2009), *A comparative study of Pashto and English phonology for English language learning*. MS thesis, UMT Lahore, 59–60.

Burquest, D., & Payne, D. (1993). *Phonological analysis: a functional approach*. Dallas: Summer Institute of Linguistics.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Burton, S., & Grimshaw, J. (1992). Coordination and VP-internal subjects. *Linguistic Inquiry*, 23 (2), 305–313.

Caroe, O. (1985). *The Pathans: 550 B.C.–A.D. 1957* (4th ed.). Karachi: Oxford.

Choi, H. W. (1996). *Optimizing structure in context: Scrambling and information structure*. Stanford, CA: Stanford University.

—. (1999). *Optimizing structure in context: Scrambling and information structure*. Stanford: CSLI.

Chomsky, N. & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.

Clements, G. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. Beckmann (eds), *Papers in laboratory phonology I*, 283–333: Cambridge: Cambridge University Press.

Clements, G. N. (1985). The geometry of phonological features. *Phonology, 2*(1), 225–252. Clements, G. N., & Keyser, S. J. (1983). *CV Phonology. A generative theory of the syllable*. Cambridge , MA: MIT Press.

Columas, F. (1999). *Sociolinguistics*. Philadelphia: Open University Press.

Costa, J. (1997). Word order typology in Optimality Theory. In Legendre, G., Grimshaw, J. & Vikner, S (eds). *Optimality-theoretic syntax*. Cambridge, MA: MIT Press.

—. (1998). *Word order variation: A constraint-based approach* (Vol. 14). Holland: Academic Graphics.

—. (2001a). The emergence of unmarked word order. In G. Legendre, J. B. Grimshaw and S. Vikner (eds), *Optimality-theoretic syntax*. Cambridge, MA: The MIT Press.

—. (2001b). Marked versus unmarked inversion and optimality theory. In *Subject inversion and the theory of universal grammar*, 91–106. Oxford: Oxford University Press.

Croft, W. (1990). *Typology and universals; Cambridge textbooks in linguistics* (2nd ed.). Cambridge: Cambridge University Press.

Crosswhite, K. (1999a). Intra-paradigmatic homophony avoidance in two dialects of Slavic. *UCLA working papers in linguistics, 1*, 48–67.

—. *Vowel reduction in optimality theory.* PhD dissertation, UCLA. Crystal, D. (1987). *The Cambridge encyclopedia of language*. Cambridge and New York: Cambridge University Press.

Davenport, M., & Hannahs, S. J. (2005). *Introducing phonetics and phonology* (2nd ed.). London: Hodder Arnold.

Derbyshire, D. C. (1979). A diachronic explanation for the origin of OVS in some Carib languages. *Work Papers of the Summer Institute of Linguistics, University of North Dakota, 23*, 35–46.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Derbyshire, D. C., & Pullum, G. K. (1979). A select bibliography of Guiana Carib languages. *International Journal of American Linguistics, 45*(3), 271–276.

—. (1981). Object-initial languages. *International Journal of American Linguistics, 47*(3), 192–214.

Devine, A. M., & Stephens, L. D. (1994). *The prosody of Greek speech*. New York: Oxford University Press.

Dik, S. C. (1978). *Functional grammar*. Dordrecht: Foris.

—. (1997). *The theory of functional grammar* (Vol. 1). In Hengeveld, K.(ed). Berlin/New York: Mouton deGruyter.

Donegan, P. (1993). Rhythm and vocalic drift in Munda and Mon-Khmer. *Linguistics of the Tibeto-Burman Area, 16*(1), 1–43.

Donegan, P. J., & Stampe, D. (1983). Rhythm and the holistic organization of language structure. *CLS Parasession on the Interplay of Phonology, Morphology and Syntax, 19*, 337–353.

Donegan, P., & Stampe, D. (2004). Rhythm and the synthetic drift of Munda. *The Yearbook of South Asian Languages and Linguistics, 2004*, 3–36.

Dost, A. (2004). *Sonority distance constraints in syllable evaluation*. Qualifying paper, University of California, Santa Cruz Embic.

Dryer, M. S. (1991). SVO languages and the OV:VO typology. *Journal of Linguistics, 27*(2), 443–482.

—. (1992). The Greenbergian word order correlations. *Language*, 68(1), 81–138.

—. (2005). Order of object and verb. In: Haspelmath et al. (eds), *World atlas of language structure*. Oxford: Oxford University Press.

Duanmu, S. (2005). The tone-syntax interface in Chinese; some recent controversies. *Cross-Linguistic Studies of Tonal Phenomena, Historical Development, Tone-Syntax Interface & Descriptive Studies,* edited by S. Kaji. Institute for the Study of Languages & Cultures of Asia & Africa, Tokyo University of Foreign Studies.

—. (2009). *Syllable structure. The limits of variation*. Oxford: Oxford University Press.

Elfenbein, J. (1997). *Pashto phonology.* In *Phonologies of Asia and Africa (including the Caucasus)* (Vol. 2). Winona Lake, IN: Eisenbrauns.

Fenk-Oczlon, G., & Fenk, A. (2004). Systemic typology and crosslinguistic regularities. In V. Solovyev & V. Polyakov. (eds), *Text processing and cognitive technologies*, 229–234. Moscow: MISA.

—. (2005). Crosslinguistic correlations between size of syllables, number of cases, and adposition order. *Sprache und Natürlichkeit: Gedenkbandfür WilliMayerthaler, 483*, 75.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

—. (2008). Complexity trade-offs between the subsystems of language. In M. Miestamo, K. Sinnemäki, & F. Karlsson (Eds.), *Language complexity: Typology, contact, change*, 43–65. Amsterdam: John Benjamins.

Fries, C. C. (1940). On the development of the structural use of word-order in Modern English. *Language*, 16(3), 199–208.

Gabelentz, G. v. d. (1901). *Die sprachwissenschaft: Ihreaufgaben, methoden und bisherigenergebnisse*. Leipzig: Chr. H. Tauchnitz.

Gell-Mann, M., & Ruhlen, M. (2011). The origin and evolution of word order. *Proceedings of the National Academy of Sciences, 108*(42), 17,290–17,295.

Giegerich, H. J. (1992). *English phonology: An introduction*. Cambridge: Cambridge University Press.

Gil, D. (1986). A prosodic typology of language. *Folia linguistica, 20*, 165–232.

Givon, T. (1977). The drift from VSO to SVO in Biblical Hebrew: The pragmatics of tense- aspect. *Mechanisms of syntactic change*, 181–254. Austin: University of Texas Press

Greenberg, J. (ed.) (1966). *Some universals of grammar with particular reference to the order of meaningful elements* (2nd ed.). Cambridge, MA: MIT Press.

—. (1974). *Language typology: A historical and analytic overview* (Vol. 184). Walter de Gruyter.

Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. *Universals of language, 2*, 58–90.

—. (1972). Numeral classifiers and substantival number: Problems in the genesis of a linguistic type. *Working Papers on Language Universals, 9*.

Grierson, G. A. (1904). *Linguistic Survey of India, III*. Calcutta: Superintendent of Government Printing, India.

Grierson, S. G. A. (1921). *Linguistic survey of India: Specimens of languages of the Eranian family*. Calcutta: Superintendent of Government Printing,India.

Grimshaw, J. (1997a). The best clitic: Constraint conflict in morphosyntax. In Liliane Haegeman (ed.). *Elements ofgrammar*, 169–196. Kluwer Academic Publishers.

—. (1997b). Projection, heads, and optimality. *Linguistic Inquiry*, 28(1), 373–422.

—. (2001). Optimal clitic positions and the lexicon in Romance clitic systems.*Optimality-theoretic syntax*, 205–240.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Grimshaw, J., & Samek-Lodovici, V. (1998). Optimal subjects and subject universals. In *Is the best good enough*, 193–219.

Hai, M. A. (1966). *Bengali language handbook*. Washington, DC: Center for Applied Lingistics.

Hallberg, G. (1992). *Sociolinguistic survey of Northern Pakistan, volume 4: Pashto, Waneci, Ormuri*. Quaid-i-Azam University Islamabad: National Institute of Pakistan Studies.

Hankamer, J., & Aissen, J. (1974). The sonority hierarchy. *Papers from the parasession on natural phonology*. Chicago.

Harris, A. C., & Campbell, L. (1995). *Historical syntax in cross-linguistic perspective* (Vol. 74). Cambridge University Press.

Harris, J. W. (1983). *Syllable structure and stress in Spanish: a nonlinear analysis*. Linguistic Inquiry Monograph, 8. Cambridge, MA: MIT Press.

Hashimoto, M. (1978). *Gengoruikeichiri-ron* (Typological and geographical linguistics). Tokyo: Kobundo. Also in Hashimoto Mantaro Chosaku-shu vol. 1, 29–190. Tokyo: Uchiyama- shoten,.

Haspelmath, M., & Bibiko, H. J. (2005). *The world atlas of language structures* (Vol. 1). New York: Oxford University Press.

Haspelmath, M., Dryer, M. S., Gil, D., & Comrie, B. (2005). *The world atlas of language structures*. Oxford: Oxford University Press.

Haspelmath, M., König, E., Oesterreicher, W., Raible, W., & eds (2001). *Language typology and language universals. An international handbook*. Berlin and New York: Mouton De Gruyter.

Hawkins, J. A. (1983). *Word order universals*. New York: Academic Press.

—. (1990). A parsing theory of word order universals. *Linguistic Inquiry, 21*(2), 223–261.

Henderson, M. M. T. (1983). Four varieties of Pashto. *Journal of the American Oriental Society*, 595–597.

Hernanz, M. L., & Brucart, J. M. (1987). *La sintaxis. La oración simple*. Madrid: Editorial Critica.

Hiersche, R., De Saussure, F., Von der Gabelentz, H. G., & Durkheim, E. (1972). *Ferdinand de Saussures langue-parole-Konzeption und sein Verhaltniszu Durkheim und von der Gabelentz*. Institut fur Vergleichende Sprachwissenschaft der Universitat.

Holt, D. E. (2006). Optimality Theory and language change in Spanish. *Optimality-Theoretic Studies in Spanish Phonology*, 378–396.

Hooper, J. B. (1976). *An introduction to natural generative phonology*. New York: Academic Press.

Hyman, L. M. (2008). Directional asymmetries in the morphology and

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

phonology of words, with special reference to Bantu. *Linguistics, 46*(2), 309.
Ijaz, M. (2003). Phonemic inventory of Pashto (pp. 83–88). CRULP-Lahore.
Itô, J. (1986). *Syllable theory in prosodic phonology.* PhD dissertation, University of Massachusetts at Amherst.
—. (1989). A prosodic theory of epenthesis. *Natural Language & Linguistic Theory, 7*(2), 217–259.
Itô, J., & Mester, A. (1998). The phonological lexicon. *Optimality Theory in Phonology*, 552–568.
Jakobson, R. (1962). *The phonemic concept of distinctive features*. In R. Jakobson, *Selected writings I: Phonological studies*, 221–230.
—. (1971). *Studies on child language and aphasia*. The Hague: Mouton De Gruyter.
—. (1972). *Child language aphasia and phonological universals*. The Hague: Mouton De Gruyter.
Jesperson, O. (1904). *Lehrbuch der Phonetik*. Leipzig and Berlin.
Jones, C. (1976). Some constraints on medial consonant clusters. *Language*, 121–130. Jones, D. (1964). *An outline of English phonetics* (9th ed.). Cambridge: Heffer.
Kager, R. (1999). *Optimality theory*. Cambridge: Cambridge University Press.
Kahn, D. (1980). Syllable-structure specifications in phonological rules. *Juncture*, 91–105. Kayne, R. S. (1994). *The antisymmetry of syntax* (Vol. 25). Cambridge, MA: MIT Press.
Kenstowicz, M. (2005). Paradigmatic uniformity and contrast. *Paradigms in phonological theory*, 145–169.
Kenstowicz, M. J. (1994). *Phonology in generative grammar* (Vol. 7). Wiley-Blackwell.
Kenstowicz, M. J., & Kisseberth, C. W. (1979). *Generative phonology.* San Diego, CA: Academic Press
Khairullah, Q. (1905). *Khair-Ul-Lughat*. Lahore: Shor Gas Printing Press.
Khan, M. K. (2009). *A comparative study of Pashto and English linguistic taboos*. Unpublished MPhil dissertation, University of Management and Technology, Lahore.
Khattak, K. P., & Ishaq, M. M. (1974). *Pashto Zaba* (Pashto language). Peshawar Pashto Academy, Peshawar University.
Kitagawa, Y. (1986). *Subjects in English and Japanese*. PhD dissertation, University of Massachusetts, Amherst.
Klimov, G. A. (1974). On the character of languages of active typology. *Linguistics, 131*, 11–25.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Koontz-Garboden, A. (2001). *A stochastic OT approach to word order variation in Korlai Portuguese*.

Koopman, H., & Sportiche, D. (1991). The position of subjects. *Lingua, 85*(2–3), 211–258. Ladefoged, P. (1967). *Linguistic phonetics* (Vol. 6). Phonetics Laboratory, University of California.

Ladefoged, P. (1990). Some reflections on the IPA. *UCLA Working Papers in Phonetics, 74*, 61–76.

—. *A Course in Phonetics* (4th ed.). New York: Harcourt College Publisher.

—. (2004). *Phonetics and phonology in the last 50 years*. New York: Harcourt College Publisher.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford: Basil- Blackwell.

—. (2001). Vowels of the world's languages. *Phonology: critical concepts, 2*, 5.

Lahiri, A., & Plank, F. (2009). What linguistic universals can be true of. *Universals of language today*, 31–58.

Legendre, G. (2001). An introduction to Optimality Theory in syntax. In G. Legendre, J. B. Grimshaw, & S. Vikner (eds), *Optimality-theoretic syntax*, 1–27. Cambridge, MA: MIT Press.

Lehmann, W. P. (1971). On the rise of SOV patterns in New High German. In K. G. Schweisthal (ed.), *Grammatik, Kybernetik, Kommunikation (Festschrift Alfred Hoppe),19*, 24.

—. (1973). A structural principle of language and its implications. *Language*, 47–66.

—. (1978). *Syntactic typology: Studies in the phenomenology of language* (Vol. 10). Austin: University of Texas Press.

Lehmann, C. (1992). Word order change by grammaticalization. *Internal and external factors in syntactic change*. Berlin & New York: Mouton de Gruyter (*Trends in Linguistics, 61)*, 395–416.

Levi, S. V. (2004). *The representation of underlying glides: A cross-linguistic study.* Unpublished PhD dissertation, University of Washington, Seattle.

—. (2005). Reconsidering the variable status of glottals in nasal harmony. *Chicago Linguistic Society, 41*(1), 299–312.

Li, C. N. (1977). *Mechanisms of syntactic change* (Vol. 1976). Austin: University of Texas Press.

Li, C. N., & Thompson, S. A. (1974). An explanation of word order change SVO → SOV. *Foundations of Language, 12*(2), 201–214.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Lorimer, D. L. R. (1915). *Pashtu part 1: Syntax of colloquial Pashtu, with chapters on the Persian and Indian elements in the modern language*. Clarendon Press.

MacKenzie, D. N. (1958). Pashto verse. *Bulletin of the School of Oriental and African Studies, 21*(2), 319–333.

—. (1959). A standard Pashto. *Bulletin of the School of Oriental and African Studies, University of London, 22*(1/3), 231–235.

—. (1987). Pashto. In B. Comrie (ed.), *The world's major languages*. New York: Oxford University Press.

Maddieson, I. (2005). Syllable structure. In Haspelmath et al. (eds), 54–55.

McCarthy, J. (2008). *Doing Optimality Theory*. Cambridge: Cambridge University Press. McCarthy, J., & Prince, A. (1994). The emergence of the unmarked: Optimality in prosodic morphology. *NELS, 24*, 333–379.

McCarthy, J. J. (2002). *A thematic guide to Optimality Theory*. Cambridge: Cambridge University Press.

McCarthy, J. J., & Prince, A. (1995). Faithfulness and reduplicative identity. *University of Massachusetts Occasional Papers in Linguistics, 18: Papers in Optimality Theory*.

McClean, V. A. (1990). *Ethnolinguistic groups in Afghanistan by Central Intelligence Agency (Map number 724842 (R00434) 4–90)*.

McNally, L. (1992). VP coordination and the VP-internal subject hypothesis. *Linguistic Inquiry* Add volume (and issue?) number. 336–341.

Meillet, A. (1912). L'évolution des formesgrammaticales. *Scientia, 12*(26), 6.

Mengal. (ed.) (2009). *The dictionary of nine languages*. Lahore: Urdu Science Board. Mohammad, J. (1993). Second-position clitics in Pashto. MS dissertation, University of Arizona, Tucson.

Mohanan, T. (1989). Syllable structure in Malayalam. *Linguistic Inquiry*, 589–625.

Momind, Q., & Sehrai, F. (eds) (1994). *Pashto dictionary*. Daryab Peshawar: NWFP Text Book Board.

Morgenstierne, G. (1926). *Report on a linguistic mission to Afghanistan*. Instituttet for Sammenlignen de Kulturforskinning, Serie CI-2.

Morgenstierne, G. (1927). *An etymological vocabulary of Pashto*. Oslo: J. Dybwad.

—. (1940). Archaisms and innovations in Pashto morphology. *Norsk Tidsskrift for Sprogvidenskap, XII*, 88–114.

Morimoto, Y. (2001). *Deriving the directionality parameter in OT-LFG. Editors' Note, 238*. Stanford, CA: SCLI Publications.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

—. (2002). Prominence mismatches and differential object marking in Bantu. Paper presented at the Proceedings of the LFG02 Conference. National Technical University of Athens, Athens: CSLI Publications.
Muller, G. (1999). Optimality, markedness, and word order in German. *Linguistics, 37*(5; ISSU 363), 777–818.
Murray, R. W., & Vennemann, T. (1983). Sound change and syllable structure in Germanic phonology. *Language*, 59(3), 514–528.
Ordoñez, F., & Treviño, E. (1995). Los sujetosy objetospreverbales en español. Paper presented at the 5th Colloquium on Generative Grammar, Coruña, Spain.
Payne, J. (1987). Iranian languages. In B. Comrie (ed.), *The world's major languages*, 514–522: Oxford: Oxford University Press.
Penzl, H. (1955). *A grammar of Pashto: A descriptive study of the dialect of Khandahar, Afghnistan*. Washington DC: American Council of Learned Society.
Plank, F. (2009a). The co-variation of phonology with morphology and syntax: A hopeful history. *Linguistic Typology, 2*(2), 195–230.
—. (2009b). Prädikativ und Koprädikativ. *Zeitschriftfürgermanistische Linguistik, 13*(2), 154–185.
—. (2009c). WALS values evaluated. *Linguistic Typology, 13*(1), 41–75.
Price, David P. (1976). Southern Nambiquara phonology. *International Journal of Anthropological Linguistics, 42*, 338–348.
Prince, A., & Smolensky, P. (1993). Optimality Theory: Constraint Interaction in Generative Grammar. Ms., Rutgers University and University of Colorado, Boulder. Rutgers Optimality Archive ROA-537.
—. (2004). *Optimality Theory: Constraint interaction in generative grammar*. Wiley Online Library.
Rahman, T. (2003). *Language policy, multilingualism and language vitality in Pakistan*. Islamabad: Quid-i-Azam University.
Ramasamy, M. D. (2010). *Topics in the morphophonology of standard spoken Tamil (SST): An optimality theoretic study*. Unpublished PhD dissertation, Newcastle University, Newcastle-upon-Tyne.
Rastorgueva, V. S. (1964). *A short sketch of the grammar of Persian* (translated by S. P. Hill; edited by H. H. Paper). Bloomington: Indiana University.
Raverty, C. (ed.) (1859). *Pukhto, Pushto or the language of the Afghan* (2001 ed.). Lahore: Sang-e-Meel Publications.
Ray, P. S., Hai, M. A., & Ray, L. (1966). *Bengali language handbook*. Washington, DC: Center for Applied Linguistics.
Rehman, G. (2009). *A comparative study of Pashto and English phonology*

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

for English language learning. Unpublished MPhil thesis, University of Management and Technology, Lahore.

—. (2012). *Verbal clitics in Pashto.* PhD dissertation in progress, University of Azad Jammu and Kashmir, Muzaffarabad.

Roach, P. (2002). *English Phonetics and Phonology. A practical course*. Cambridge: Cambridge University Press.

Roberts, T. (1997). The optimal second position in Pashto. In G. Booij and J. van der Weijer (eds), *Phonology in Progress-Progress in Phonology: HIL Phonology Papers III*, 367–401. The Hague (Holland).

—. (2000). *Clitics and agreement.* Unpublished PhD dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Robson, B., & Tegey, H. (2009). Pashto. In G. Windfuhr (ed.), *The Iranian languages*. Oxford: Routledge.

Romaine, S. (1988). *Pidgin and creole languages*. London: Longman.

—. (1994). Germanic creoles. In E. König and J. van der Auwera (eds), *The Germanic Languages*. London: Routledge.

Samek-Lodovici, V. (1996). Constraints on subjects: An optimality-theoretic analysis. *Glot international, 2*(8), 11–14.

Samek-Lodovici, V. (2007). An invitation to OT syntax and semantics. *Lingua, 117*(9), 1513–1521.

Schlegel, F. von (1808). *Uber die Sprache und Weisheit der Indier*. Heidelberg: Moht & Zimmer. (New edition by E. F. K. Koerner, 1977, Amsterdam: John Benjamin).

Seiler, H. (1990). Language typology in the UNITYP model. In W. Bahner et al. (eds), F*ourteenth International Conference of Linguists*. Berlin: Akademie Verlag.

Selkirk, E. O. (1980). The role of prosodic categories in English word stress. *Linguistic Inquiry, 11*(3), 563–605.

—. (1984). On the major class features and syllable theory. In M. Aronoff and R. T. Oehrle (eds), *Language sound structure*, 107–36: Cambridge, MA: MIT Press.

Sells, P. (1999). Constituent ordering as alignment. *Harvard Studies in Korean Linguistics, 8*, 546–560.

—. (2001). *Structure, alignment and optimality in Swedish*. Stanford, CA: CSLI Publications.

Shafeev, D. A. (1964). *A short grammatical outline of Pashto* (ed. and trans. H. H. Paper). The Hague: Mouton and Co.

Shibatani, M., & Bynon, T. (1995). Approaches to language typology: A conspectus. In Shibatani and Bynon (eds).*Approaches to language typology, 1-26*. Oxford: Oxford University Press.

Shopen, T. (2007). *Language typology and linguistic description*.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Cambridge: Cambridge University Press.
Sievers, E. (1881). *Grundzüge der Phonetik*. Leipzig: Breitkopf and Härtel.
Skalička, V. (1935). *Zurungarischen Grammatik*. Prague: Universitas Caroliensis. Skjærvø, P. O. (1989). Pashto. In Schmitt, R. (ed). *Compendium Linguarum Iranicarum*, 384–410. Wiesbaden: Reichert.
Smolensky, P. (1995). On the internal structure of the constraint component Con of UG. Handout of talk, UCLA, Aptil 7.
—. (1996). The initial state and "richness of the base" in Optimality Theory. Technical report JHU-CogSci-96-4, Department of Cognitive Science, The Johns Hopkins University, Baltimore, Md.
Speas, M. (1997). Optimality theory and syntax: null pronouns and control. In Archangeli, D. and Langendoen, D. T. (eds.). *Optimality Theory: An Overview,* 171–199. . Blackwell: Oxford.Spencer, A. (2002). *Phonology*. Oxford: Blackwell.
Tair, M. N. (1989). *Pashto language* (9th ed.). Peshawar: Pashto Academy.
Tegey, H. (1977). *The grammar of clitics: evidence from Pashto and other languages*. PhD dissertation. Urbana: University of Illinois.
—. (1979). Ergativity in Pushto (Afghani). *Linguistic method: essays in honor of Herbert Penzl, 79*, 369.
Tegey, H., & Robson, B. (1996). *A reference grammar of Pashto*. Washington DC: Center for Applied Linguistics.
Tesar, B., Grimshaw, J., & Prince, A. (1999). Linguistic and cognitive explanation in Optimality Theory. In Lepore, E. and Pylyshyn, Z. (eds.). What is Cognitive Science?. Blackwell: 295-326.
Thomason, S. G. (2008). Pidgins/Creoles and historical linguistics. In Kouwenberg, S. and Singler, J.V. (eds.).*The Handbook of Pidgin and Creole Studies*, 242–262.
Tokizaki, H. (2008). Symmetry and asymmetry in the syntax–phonology interface. *On-in Kenkyu (Phonological Studies), 11*, 123–130.
—. (2011). Correlation between word order and phonology: Variation in dialects and languages. *Variation and Typology: New trends in Syntactic Research*, 68. Conference paper, Helsinki.
Tokizaki, H., & Kuwana, Y. (2007). *Do OV languages have simple syllable structure?* Conference paper, Association for Linguistic Typology 4, Paris.
—. (2012). Limited consonant clusters in OV languages. In P. Hoole et al. (eds), *Consonant clusters and structural complexity, 71-91*. Berlin: Mouton de Gruyter.
Trubetzkoy, N. S. (1958). *Grundzüge der phonologie*. Gottingen:

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Vandenhoeck & Ruprecht.
Trudgill, P. (1987). *Sociolinguistics – An introduction to language and society*. Gateshead: Northemberland.
Vennemann, T. (1972). Analogy in generative grammar: The origin of word order. Paper presented at the Proceedings of the Eleventh International Congress of Linguists.
—. (1973). *Topics, sentence accent, ellipsis: A proposal for their formal treatment*. LAUT.
—. (1974). Topics, subjects and word order: from SXV to SVX via TVX. *Historical linguistics, 1*, 339–376.
—. (1975). An explanation of drift. In, Li, CN (ed) *Word order and word order change, 1974*, 269–305. Austin: University of Texas Press.
—. (1988). *Preference Laws for Syllable Structure*. Berlin: Mouton de Guyter. Wardhaugh, R. (2002). *Pidgin and Creole languages. An introduction to sociolinguistics*. Oxford: Blackwell.
Yartseva, V. (1979). The Semantic and Structural Parameters of a Language Type. *Acta Linguistica Academiae Scientiarum Hungaricae, Tomus* 29 (3-4), 275-187.
Zagona, K. (1982). *Government and proper government of verbal projections*. PhDdissertation, University of Washington, Seattle.
Zaman, J. (2002). Syllabification rules for Pashto. *CRULP, 161*, 166–174.
Zubizarreta, M. L. (1995). *Word order, prosody, and focus*. Los Angeles: University of Southern California.

ELECTRONIC SOURCES

http://wals.info/ (accessed on 20 August 2011)
http://typo.uni-konstanz.de/archive/intro/index.php (accessed on 20 July 2010)
http://sswl.railsplayground.net/browse/languages/Pashto (accessed on 26 September 2011)

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

APPENDIX A

GLOSSARY OF CONCEPTS AND CONSTRAINTS

The following concepts and constraints are used in the current study. Though they are explained where required at the relevant point in the text, a complete glossary is given here.

ONSET *[σ V Syllables must have an onset

NOCODA *C] σ Syllables are open

*COMPLEXONS *[σCC Onsets are simple

*COMPLEXCOD *CC]σ Codas are simple

DEP-IO Output segments must have input correspondents (no epenthesis)

MAX-IO Input segments must have output correspondents (no deletion)

*CLUSTER No complex syllable codas and onsets

FAITH Faithfulness constraints (DEP-IO and MAX-IO)

SON-SEQ Complex onsets rise in sonority and complex codas fall in sonority Syllabic Consonant Constraint * V [+ consonantal] No syllabic consonant allowed MAXST Maximum syllabic template in Pashto (i.e. CCCVCC). MAXST consists of three consonants in the onset and two consonants in the coda. Both onset and coda may violate the SSP.

HD-RT The head is rightmost in its projections

HD-LT The head is leftmost in its projections

SUBJ-LT A subject aligns left in the clause

SO Subject precedes object

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

STAY Trace is not allowed (Economy of Movement)

SSP Sonority Sequencing Principle: between any member of a syllable and the syllable peak, only sounds of higher sonority rank are permitted

SOV The specific order of three basic syntactic elements (subject + object + verb)

OT Optimality Theory: the framework of constraints interaction

GEN (or Generator) The function of OT which generates a set of unique output candidates for an input. Out of these candidates, one is identical to the input while the rest of the candidates are somewhat modified in their structure.

EVAL (or Evaluator) Another function of OT. It evaluates the candidates and chooses the optimal one that best satisfies the set of specially ranked constraints depending on the nature of the violation.

Markedness Constraints Markedness constraints ensure that the output candidate is well-formed, prohibiting forms which are difficult to produce or comprehend. Such constraints ensure the utterance of certain segments and impose restrictions on others. NoCoda and Onset are examples of markedness constraints.

Faithfulness Constraints Faithfulness constraints ensure the similarity of the output to the input. For example, Dep-IO and Max-IO are faithfulness constraints.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

APPENDIX B

PASHTO EDGE PHONOTACTICS

The following data is used in the study to determine the edge phonotactics of the Pashto language. The sounds given here are just a sample representation of the phonological features of Pashto. For exhaustive exploration, dictionaries and digital resources may be used.

WORD-INITIAL

Bipartite

Plosives + Liquids
#pr #pɽ #pl #br #bɽ #bl #tr
#tɽ #tl #dr #dɽ #kr #kɽ #kl
#gr #gɽ

Plosives + Glides
#pj #bj #dw #dj #gw #tw #kw
#qj

Fricatives + Liquids
#sr #zɽ #ʃr #ʃʈ #ʃl #xr #xɽ

#ɣr #ɣɽ #ɣl

Fricatives + Plosives
#sk #sp #st #ʃk #ʃp #xk #xp
#zg #zb #zd

Fricatives + Glides
#sj #sw #zj #zw #xj #xw #ɣw

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Fricatives + Fricatives	#zɣ	#ʃx	#sx
Fricatives + Nasals	#zm	#ʃm	#ʃn
Nasal + Plosive	#ng		
Nasal + Fricative	#nɣ		
Nasals + Liquids	#ml	#mr	#mɽ
Nasals + Glides	#mj	#nw	#nj
Nasal + Nasal	#nm		
Liquid + Fricative	#rɣ		
Liquid + Nasal	#lm		
Liquid + Glide	#lw		
Glides + Liquids	#wr	#wɽ	#wl
Affricate + Liquid	#dʒr		
Affricates + Glides	#dʒw	#tsw	
Affricate + Fricative	#dʒɣ		
Affricate + Plosive	#tsk		
Tripartite			
Fricative + Glide + Liquid		#xwl	#xwɽ
Fricative + Glide + Plosive		#xwd	

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Fricative + Fricative + Glide #sxw #ʃxw

Fricative + Plosive + Glide #skw #xkw

Nasal + Plosive + Liquid #ndr

WORD-FINAL

Bipartite

Fricatives + Plosives
xt# ʃt# ʃq# ʃk# sk# st# Nasals + Plosives nd# nɖ# mp# nk#

Liquids + Plosives
rg# rq# rd# rt# rp# ɽd#

Plosives + Fricatives bs# ks# bz# qʃ# Glide + Fricative wz#

Nasals + Fricatives mz# nz# ns# nʃ#

Liquids + Fricatives	rs#	rz#	rk#	ɽx#
Fricative + Fricative	xs#			
Liquids + Nasals	rm#	rn#		
Nasal + Affricate	ndʒ#			

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

APPENDIX C

PASHTO CONSONANT CLUSTERS

The following list is used in the study to highlight the consonantal clusters of the language. All of the representative examples containing clusters have been taken from the following sources unless otherwise noted.

- *Daryab Pashto dictionary*, Momind & Sehrai (1994)
- *The dictionary of Nine languages*, Mengal (2009)
- *A dictionary of the Pukhto, Pushto or Language of the Afghan*, compiled by Captain Raverty (first published 1860, republished 2001)
- A recorded collection of 5000 words spoken by native speakers of the language (see Appendix D for details)
- Being a native speaker of the language, the observation and intuition of the researcher.

However, this appendix cannot be considered exhaustive. Note that in analysing the data, the pronunciation of the words has been adapted to the phonemic inventory given in Section 3.2. In addition to the native speaker's intuition about the clusters (or reverse sonority sequencing principle where applied), software such as PRAAT has also been used.

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

WORD-INITIAL

Bipartite (CC)

Type	Cluster	Examples	English Translation
Plosives + Liquids			
	pr	prəde	Stranger
	pɽ	pɽək	Moment/glitter
	pl	pla:r	Father
	br	brəg	Spotted/piebald
	bɽ	bɽa:s	heat/exhalation/steam
	bl	blarba	Pregnant
	tr	tror	Aunt
	tɽ	tɽək	Breaking noise
	tl	tlәl	Going
	dr	drənə	Heavy
	ɖr	ɖrənka	Harsh, braying noise
	kr	krəka	Disgusting
	kɽ	kɽəp	Bang/slam
	kl	klək	Tight
	gr	gra:n	Precious
	gɽ	gɽəng	Bang/snotty-nosed
	pj	pja:z	Onion
Plosives + Glides			
	bj	bja:	Then/again
	dw	dwə	Two
	dj	dja:rləs	Thirteen
	gw	gwəɽa	Brown sugar
	tw	twa:n	Power/ability
	kw	kwənɖə	Widow
	qj	qja:mət	Day of Judgment
	sr	srə	Gold
Fricatives + Liquids			
	zɽ	zɽə	Heart
	ʃr	ʃrəng	Clink/jingling
	ʃɽ	ʃɽək	Thunder/swift
	ʃl	ʃla:nɖə	Frog/heavy
	xr	xrə	Donkey
	xɽ	xɽi:s	Large and bulky

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

ɣr	ɣrət	A sudden blow
ɣɽ	ɣɽap	Quick gulp
ɣl	ɣla:	Theft
sk	skor	Charcoal

Fricatives + Plosives

sp	spa:i	Soldier
st	stən	Wooden pillar
ʃk	ʃku:ɳ	Porcupine
ʃp	ʃpəg	Six
xk	xka:r	Hunting
xp	xpəl	Relative
zg	zgerəl	Moan/sew
zb	zba:dəl	To prove
zd	zdakol	To learn

Fricatives + Glides

sj	sja:l	Equal
sw	swəzedəl	To burn
zj	zja:t	Much
zw	zwa:n	Young
xj	xja:l	Care
xw	xwa:	Side
ɣw	ɣwa:	Cow

Fricatives + Fricatives

zɣ	zɣaməl	To tolerate
ʃx	ʃxəɽa:	Dispute
sx	sxər	Father-in-law

Fricatives + Nasals

zm	zməka	Land/earth
ʃm	ʃmer	Counting
ʃn	ʃna	Greenery

Nasal + Plosive

ng	ngor	Daughter-in-law

Nasal + Fricative

nɣ	nɣəre	Hearth/stove

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Nasals + Liquids			
	ml	mla	Waist
	mr	mrəsta:	Help
	mɽ	mɽawe	Faded
Nasals + Glides			
	mj	mja:ʃt	Month
	nw	nwəri:	Clothes
	nj	nja:	Grandmother
Nasal + Nasal			
	nm	nməsi:	Grandchildren
Liquid + Fricative			
	rɣ	rɣəɽi:	Rolls
Liquid + Nasal			
	lm	lmər	Sun
Liquid + Glide			
	lw	lweʃt	Span
Glides + Liquids	**wr**	wra:rə	Nephew
	wɽ	wɽəl	Carry
	wl	wluna	Cleaning
Affricate + Liquid	**dʒr**	dʒrənda	Watermill
Affricates + Glides	**dʒw**	dʒwənd	Life
	tsw	tswa:rləs	Fourteen
Affricate + Fricative	**dʒɣ**	dʒɣorəl	To preserve/keep
Affricate + Plosive	**tsk**	tska:k	Drinking
Tripartite (CCC)			
Fricative + Glide + Liquid			
	xwl	xwlə	Mouth
	xwɽ	xwɽi:n	Softened
Fricative + Glide + Plosive			
	xwd	xwda:i	God

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Fricative + Fricative + Glide

sxw	sxwəndər	Bull
ʃxw	ʃxwənd	Chewing

Fricative + Plosive + Glide

skw	skwərəʈə	Ember
xkw	xkwəle	Beautiful

Nasal + Plosive + Liquid

ndr	ndror	Sister-in-law

WORD-FINAL

Bipartite (CC)

Fricatives + Plosives

xt	ba:ləxt	Pillow
ʃt	səlweʃt	Forty
ʃq	Iʃq	Love
ʃk	əʃk	Tear
sk	həsk	Cloth
st	drəst	Right

Nasals + Plosives

nd	dʒwənd	Life
nɖ	ɖənɖ	Pool
mp	pəmp	Pump
nk	benk	Bank

Liquids + Plosives

rg	tʃərg	Chicken
rq	ərq	Essence
rd	dərd	Pain
rt	pərt	Property record letter
rp	tərp	Radish
ɽd	ʃa:gəɽd	Student

Plosives + Fricatives

bs	həbs	Humidity
ks	əks	Contrary/reverse
bz	sərsəbz	Lush green
qʃ	nəqʃ	Copy/drawing

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

Glide + Fricative			
	wz	Iəwz	Promise
Nasals + Fricatives			
	mz	rəmz	Secrecy
	nz	Munz	Prayer
	ns	həns	bird
	nʃ	benʃ	beam
Liquids + Fricatives			
	rs	tərs	Pity
	rz	tərz	Form
	rk	tərk	Cleft
	ɽx	əɽx	Aspect
Fricative + Fricative			
	xs	ʃəxs	Person
Liquids + Nasals			
	rm	fa:rm	Form
	rn	qərn	Century
Nasal + Affricate			
	ndʒ	torəndʒ	Citron tree

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

APPENDIX D

PASHTO WORDS CONTAINING CONSONANT CLUSTERS

The researcher collected 5000 different words from the native speakers of Pashto (from his native town) where the Yousafzai dialect is spoken. Out of the collection, the following list shows the words that have consonant clusters at any position. The main purpose of the collection of these words was to show the Pashto words which are frequently used by the native speakers in their daily life. Their syllable boundaries and English translation are also given.

Pashto Words	**Syllabification**	**English Translation**
əzmext	VC.CVCC	Test
əmpʈʌkəke	VC.CCV.CV.CV	Tip-cat
Ndror	CCCVC	Sister-in-law
xpa	CCV	Foot
ərt	VCC	Wide open
ərtaxpa	VCC.CCV	First visit to in-laws after engagement
əʈx	VCC	Side
ərz	VCC	The earth/request
ʌŋənɳ	VC.VCC	Building
ʌzmext	VC.CVCC	Test
ʌrhənɖ	VC.CVCC VCC	Castor oil tree
u:gd	VCC	Long
ʌʃk	VCC	Tear
əsp	VCC	Horse
aqəlmənd	V.CVC.CVCC	Wise
bazʊbʌnd	CV.CV.CVCC	Arm ornament
badrʌnŋ	CV.CVCC	Cucumber
balʌxt	CV.CVCC	Pillow
bambənɖ	CVC.CVCC	Brahman
bʊtparəst	CVC.CV.CVCC	Idol worshipper

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

blarb	CCVCC	Pregnant
barbʌst	CVC.CVCC	Goods
bərbənɖ	CVC.CVCC	Naked
bʊrdʒ	CVCC	Tower
bərdaʃt	CVC.CVCC	Patience
brʌɣ	CCVC	Fold
bɽəŋgedʌl	CCVC.CV.CVC	Humming
blɑ:rb	CCVCC	Pregnant
bləwsəwʌl	CCVC.CV.CVC	Rub/scuff/scrape
bɑdʒri	CV.CCV	Mortar/grout
bjʌrtə	CCVC.CV	Far away
bjɑ:stə	CCVC.CV	Rope
bjələwʌl	CCV.CV.CVC	To separate
brəg	CCVC	Spotted
brəndu:	CCVC.CV	A river in Buner
brənd	CCVCC	To challenge
bərɣənɖ	CVC.CVCC	Cut wood
brəɣgedəl	CCVC.CV.CVC	Being folded
brəmtah	CCVC.CVC	Seizing of a village
brexna	CCVC.CV	Lustre
brexawal	CCV.CV.CVC	To illuminate
bret	CCVC	Mustache
bɽa:s	CCVC	Vapour
bɽʌstən	CCVC.CVC	Quilt
bɪlʌnd	CV.CVCC	High
blarb	CCVCC	Pregnant
bləwsedəl	CCVC.CV.CVC	To distress/afflict
bʌndʒ	CVCC	Chesnut tree
bənd	CVCC	Fastening/ban
bʌndubʌst	CVC.CV.CVCC	Settlement
bənŋ	CVCC	Hemp
bɪhɪʃt	CV.CVCC	Paradise
bagənɳ	CV.CVCC	Fly
bənɳ	CVCC	Forest
betərs	CV.CVCC	Fearless
bedəst	CV.CVCC	Helpless
bedreɣ	CV.CCVC	Pitiless
benʌnɳ	CV.CVCC	Shameless

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

berʌnɳ	CV.CVCC	Colorless
bja	CCV	Again
bja:ban	CCV.CVC	Desert/heath
bjartja	CCVC.CV	Back again
bexrəte	CV.CCV.CV	Unreasonabale
birənɳ	CV.CVCC	A type of grass
bjəl	CCVC	Separate
benʃ	CVCC	Timber beam
bza	CCV	Goat
pɽɑŋ	CCVC	Leopard
pedʒəndgəlo	CV.CVCC.CV.CV	Familarity
prʌtʃ	CCVC	Saucer
prədæ	CCV.CV	Stranger/alien/foreigner
prekəwʌl	CCV.CV.CVC	Cut/bob/crop/pollard/prune
prexəwdʌl	CCV.CVC.CVC	Give up/abandon/leave
prot	CCVC	Lying, lying down
pɽəkedʌl	CCV.CV.CVC	To shine
plʌn	CCVC	Wide
pjɑ:ɽmə	CCVC.CV	Webbed part at one end of a bedstead
pərtʃanŋ	CVC.CVCC	Ladder
pətang	CV.CVCC	Moth
paləng	CV.CVCC	Bed
pətənɳ	CV.CVCC	Lower part of a spinning wheel
puʃt	CVCC	Back
puʃtgir	CVCC.CVC	Supporter
paʃm	CVCC	Wool
paganɳ	CV.CVCC	11th month of the Indian calendar
pəndʒ	CVCC	Five
pənɖ	CVCC	Load
pənd	CVCC	Advice
post	CVCC	Poppy head
pɒha:nd	CV.CVCC	Scholar/professor
pahm	CVCC	Sense/acquaintance
pedajext	CV.CV.CVCC	Birth
pedʒəndgali	CV.CVCC.CV.CV	Acquaintanceship
pewəst	CV.CVCC	Connected
pewənd	CV.CVCC	Joint
pəndtja	CVCC.CCV	Artfulness
pʊnɖ	CVCC	Fat
pj:az	CCVC	Onion
pla	CCV	Stubborn/sticky

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

plar	CCVC	Father
plarmәʈe	CCVC.CV.CV	Orphan
plalə	CCV.CV	Rice straw
plarganəi	CCVC.CV.CV	Paternal relative
planedəl	CCV.CV.CVC	To grow wide
plənawal	CCV.CV.CVC	To render broad
plorəl	CCV.CVC	To sell
pranastəl	CCV.CVC.CVC	To open
prebande	CCV.CVC.CV	On
prət	CCVC	Absurd
prəta	CCV.CV	Idol
prəde	CCV.CV	Strange
pregdena	CCVC.CV.CV	Leaving
prəndʒəkol	CCVCC.CV.CVC	To sneeze
prənɳ	CCVCC	Sour milk
prɒt	CCVC	Lying
prexwəl	CCV.CCVC	To discontinue
prexodəl	CCV.CV.CVC	To leave
pregdəl	CCVC.CVC	Abandon
prewan	CCV.CVC	The afterbirth placenta
prewana	CCV.CV.CV	In abundance
prewatəl	CCV.CV.CVC	To fall
prewata	CCV.CV.CV	Falling
prewəstal	CCV.CVC.CVC	To throw
prewenzal	CCV.CVC.CVC	To wash with
prejasta	CCV.CVC.CV	Push
pʈəpiʃ	CCV.CVC	Wild cat
pʈang	CCVCC	Tiger
psan	CCVC	A sharpening stone tool
psarle	CCV.CV	Spring
trɪw	CCVC	Sour
trɪx	CCVC	Bitter
trʌx	CCVC	Armpit
trəŋgəwʌl	CCVC.CV.CVC	Play upon musical instruments
tlʌl	CCVC	To go
tjɑrʌ	CCV.CV	Darkness/dark
taxt	CVCC	Throne
traʈ	CCVC	Whip
trak	CCVC	Crack
trək	CCVC	Stroke
trap	CCVC	Jump

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

trax	CCVC	Armpit
tarz	CVCC	Manners
tars	CVCC	Fear
turʃ	CVCC	Sour
tərkanɳ	CVC.CVCC	Carpenter
tərk	CVCC	Leaving/crack
tərgme	CVCC.CV	Dim light in the morning before dawn
təʊrəndʒ	CV.CVCC	Citron
trək	CCVC	Stroke/wooden scale
trəndʒəbiːn	CCVCC.CV.CVC	Manna
trəndʒuke	CCVCC.CV.CV	Twisted silk
trəndʒ	CCVCC	Delicate
trənkʊtse	CCVCC.CV.CV	Sling
trəŋ	CCVC	String of a musical instrument
trəŋgawəl	CCVC.CV.CVC	To play a musical instrument
trəŋgəƫ	CCVC.CVC	A net used for carrying grass
trɒr	CCVC	Aunt
truke	CCV.CV	A kind of sorrel
trə	CCV	Paternal uncle
tre	CCV	From
triːx	CCVC	Bitter
triːxe	CCV.CV	Gall-bladder
triːw	CCVC	Sour
tɽək	CCVC	Knock
tɽəp	CCVC	Pat
tɽəpawəl	CCV.CV.CVC	To jump/to agitate
təʃt	CVCC	Flat vessel
txəl	CCVC	To fly from
tləl	CCVC	To depart
tlə	CCV	Departure
tuːɳ	CVCC	A vessel with a long
təng	CVCC	Narrow/neck/tight/confined
təngsja	CVC.CCV	Helplessness
tənɳ	CVCC	Roar of lion
twaːn	CCVC	Power
twaːm	CCVC	Food
tasənd	CV.CVCC CV.CVCC	Dry
tʊləndˈ	CV.CVCC CCV.CV	Partridge gait
təwang	CCVC	Basket
tjarə		Darkness
tjaːk		Inflated

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

ʈang	CVCC	Fixed period
ʈɽaq	CCVC	Bang
ʈanɖ	CVCC	Fat
ʈangʈu:ng	CVCC.CVCC CVCC	Musical sound
ʈaunɳ		Imitation
dʒrəndʌ	CCVC.CV	Lock
dʒrʌndə	CCVC.CV	Watermill
dʒwɑ:r	CCVC	Maize
dʒwənd	CCVCC	Life
dʒəndʒ	CVCC	Marriage guest
dʒu:xt	CVCC	Fixed/tight
dʒrəndgəʈe	CCVCC.CV.CV	Miller
dʒrənda	CCVC.CV	Watermill
dʒreb	CCVC	Measure of land
dʒʊft	CVCC	Even
dʒəlbəng	CVC.CVCC	A stinging
dʒəld	CVCC	Soon
dʒins	CVCC	Thing
dʒəng	CVCC	Battle
tʃərx	CVCC	Wheel
tʃware	CCV.CV	Dry fruit
tsa:ʃt	CVCC	Time between sunrise and noon
tsa:ng	CVCC	Wing
tsxedəl	CCV.CVC	To creep
tsrəp	CCVC	Flap
tsərgənd	CVC.CVCC CCVC	Apparent
tsxa:k	CCVC	Drink
tskəl	CCV.CVC	To drink
tskəwəl		To cause to drink
tskʊnɖəl	CCVC.CVC	To pinch/squeeze
tskedəl	CCV.CVC	Crawl/creep
tsəlwext	CVC.CVCC	Forty
tsənɖ	CVCC	Shake
tsəng	CVCC CCVC.CVC	Side
tswarləs	CCVC.CV	Fourteen
tswərle		An auger
xpemanə	CCV.CV.CV	Repentant
xwlʌ	CCCV	Mouth
xwde	CCCV	God
xkwʌlæ	CCCV.CV	Pretty/beautiful
xɑ:rʌxt	CV.CVCC CCVCC	Mange

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

xwʌnd	CCVC	Taste/flavour
xpʌl	CCV.CV.CVC	Relative/own
xpələwʌl		To own/possess/win/to be relative
xrɑːʈə	CCV.CV CCVC.CV	Skylark
xrɑːʃkæ	CCV	Sputum/phlegm/gob
xrʌ	CCV.CV.CV	She-ass
xrezəkæ		Stubble
xrəjʌl	CCV.CVC	Shave
xrɑŋge	CCVC.CV	Ham/back part of the thigh
xʊʃt	CVCC	Drenched/wet/well beaten
xlɑːs	CCVC	Finished/completed/ended/free
xwɑːr	CCVC	Poor/skint/penurious/indigent
xwɑˈri	CCV.CV	Poverty/construction/indigence
xwɑjəxedʌl	CCV.CV.CV.CVC	Gloat over/enjoy revenge/malicious acts
xwɑːst	CCVCC	Request/appeal/begging/solicitation
xwɑːxi	CCV.CV	Mother-in-law
xwərzʌ	CCVC.CV	Niece/sister's daughter
xwərəwʌl	CCV.CV.CVC	Spread/scatter/disperse/extend/strew
xwʌʈ	CCVC	Creek/brook/rill/stream/rivelet
xwəʈʌl	CCV.CVC	To eat
xwʈɪn	CCCVC	Cooked
xwəzəwʌl	CCV.CV.CVC	Shove/move/stir/push
xwəjəwʌl	CCV.CV.CVC	To slide/to fester
xɑjest	CV.CVCC	Beauty
xpʌ	CCV	Foot
xkɑːr	CCVC	Hunt
xkɑrʌ	CCV.CV	Obvious/open/clear/vivid
xkʌtə	CCV.CV	Lower/down
xkʌr	CCVC	Horn
xkwələwʌl	CCCV.CV.CVC	To kiss
xna	CCV	Piece/bit of something
xwəredʌl	CCV.CV.CVC	To spread
xweidʌl	CCV.CVC	To slip
drənɑwæ	CCV.CV.CV	Honour/respect/reverence
drənʌ	CCV.CV	Heavy/boasting
drəbəwʌl	CCV.CV.CVC	Beating/raping/tapping/throbbing/wal
drəzedʌl	CCV.CV.CVC	lop Beating (heart
drənəwʌl	CCV.CV.CVC	etc.)/throbbing/shuddering To make
drʌŋ	CCVC	heavy/take seriously Delay/short while/mine

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

drʊnd	CCVCC	Heavy
dʌst	CVCC	Diarrhea
dwɑ:nə	CCV.CV	Both
djɑ:rləs	CCVC.CVC	Thirteen
djɑrləsʌm	CCVCV.CV.CVC	Thirteenth
derʃt	CVCCC	Thirty
ɖʌnɖ	CVCC	Pond/puddle
rɑzɣələwʌl	CV.CCV.CV.CVC	Scoot/to make run away
rɑtlʌl	VCC.CVC	Come
reʃtja, rextja	CV.CCV.CVC.CV	Truth
rɣəʈəwʌl	CCV.CV.CVC	To roll
rʌnz	CVCC	Disease/pain
rɑnzʈʌ	CV.CCV	Diseased (female)/patient/ill
rost	CVCC	Rotten/stale/decomposed
roʃnɑˈji	CV.CCV.CV	Brightness/ink/light
rwerʌ	CCV.CV	Niece, brother's daughter
ʈwɑ:nde	CCVC.CV	Before/ahead
ʈʊnd	CVCC	Blind
zbərg	CCVCC	Saint
zbɑ:d	CCVC	Remind others of the good one has done them
zbexʌl	CCV.CVC	Rub/tread
zdəkəwʌl	CCV.CV.CVC	Learn
zʈʌ	CCV	Heart
zʈəswɑ:ndæ	CCV.CCVC.CV	Sympathatic
zxɑ:xtʌl	CCVC.CVC	To run
zɣʌrə	CCV.CV	Panoply
zɣələwʌl	CCV.CV.CVC	To make run
zɣʌm	CCVC	Patience/endurance/tolerance
zɣəmʌl	CCV.CVC	Bear/endure/tolerate
zməræ	CCV.CV	Lion
zmʌkə	CCV.CV	Earth
zwʌʈənd	CCV.CVCC	Suspended/hanging
zwʌg	CCVC	Noise
zwæ	CCV	Son
zjɑ:t	CCV	Plenty/great in number
zjɑtəwʌl	CCV.CV.CVC	Increase
zjɑte	CCV.CV	Injustice/cruelty/excess
zjɑ:n	CCVC	Loss
zjʌʈ	CCVC	Yellow/pale
zjəʈæ	CCV.CV	Hepatitis

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

zjʌm	CCVC	Wetness/moisture/gloomy
swali:	CCV.CV	Borrowed
swəl	CCVC	To fire
srə	CCV	Gold
spʊnʈæ	CCVC.CV	Herpes/shingles/a skin disease
skɑrʌ	CCV.CV	Wood coal/charcoal
sʌxt	CVCC	Hard
sɑ:xt	CVCC	Sun-up/sunrise/first blush of the day
skʌl	CCVC	To drink
skundʌl	CCVC.CVC	Pinching
səlwext	CVC.CVCC	Forty
slext	CCVCC	Glue/sticking substance used to bond paper
swɑrleɪ	CCVC.CVV	Vehicle/conveyance
spɑrʌl	CCV.CVC	Bury/assign/hand over
spərlæ	CCVC.CV	Spring
spəʈdʌl	CCVC.CVC	Open/bloom/dehisce
spəgdʒʌn	CCVC.CVC	Full of lice
spʌk	CCVC	Light/mean
spor	CCVC CCVC.CVV	Dry/not oily
spogmeɪ	CCV	Moon
spæ		Dog
spegma spi:n	CCVC.CV CCVC	Nostril White Peel/pare
spɪnəwʌl	CCV.CV.CVC	Your/yours
stɑ:so	CCV.CV	
stɑjʌnə	CCV.CV.CV	Praise
stərɣəli	CCVC.CV.CV	Eyeball
stʌrgə	CCVC.CV	Eye
stərge	CCVC.CV	Eyes
stʌʈæ	CCV.CV	Tired
stʌɣ	CCVC	Difficult/tough
stʌm	CCVC	Sighing/moaning
stʌn	CCVC	Needle
storæ	CCV.CV	Star
stomɑ:nə	CCV.CV.CV CCVC	Regretful/repentant
stu:n	CCV	Returned/retaliated Dirty
sxɑ:		
sxɑ:t	CCVC CCV.CV.CV	Alms after death
sxɑ:wʌdʒə	CCCVC.CVC	A plant like ginger with a bad smell
sxwəndʌr		Bull/ox

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

sxæ	CCV	Calf
skərwʌʈə	CCVC.CV.CV	Ember
Slex	CCVC	Glue/a sticking substace used to paste things
Soxt	CVCC	Angry esp. for taking revenge
swəzəwʌl	CCV.CV.CVC	Burn
Sporal	CCV.CVC	To rebuke
ʃɑ:gʌʈd	CV.CVCC	Student
ʃpɑ:ʈəs	CCV.CVC	Sixteen
ʃpʌg	CCVC	Six
ʃpol	CCVC	Fence/pale/hedge
ʃpənkæ	CCVC.CV	Shepherd
ʃpʌ	CCV	Night
ʃpetʌ	CCV.CV	Sixty
ʃpetæ	CCV.CV	Wedge/dowel/chunk of wood to join or strengthen
ʃpeʃtʌræ	CCVC.CV.CV	A type of vegetable/greens
ʃpelæ	CCV.CV	Whistle
ʃpɒl	CCVC	Heath
ʃxʌʈə	CCV.CV	Clash/combat/fray/conflict/nonsense
ʃxwʌnd	CCCVCC	Chewing
ʃʈənŋgəwʌl	CCVC.CV.CVC	Jingle
ʃkor	CCVC	Paten, a tray for keeping breads
ʃkʊnɖ	CCVCC	Porcupine/hedgehog
ʃlʌʈ	CCVC	Sluggish/lazy/idle
ʃledʌl	CCV.CVC	Tear/wear out
ʃmɑrʌl	CCV.CVC	Count
ʃnʌ	CCV	Green
ʃwʌ	CCV	Became (feminine)
ʃpa:	CCV	Night
ʃxwənd	CCCVCC	Chewing
ɣərbʌnɖ	CVC.CVCC	Heavy load
ɣbərg	CCVCC	Twofold
ɣwənɖ	CCVCC	Round
ɣwʌnɖ	CCVCC	Whole/round
ɣbʌrg	CCVCC	Together
ɣwəsəwʌl	CCV.CV.CVC	Gash/slash/scratch
ɣwaxa	CCV.CV	Meat
ɣrʌb	CCVC	Sound of falling things

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

ɣrambəɳa	CCVC.CV.CV	Flame
ɣræb	CCVC	Sobbing/stopping of breath while weeping
ɣwəʈɑskæ	CCV.CVC.CV	A plant (dodonoea burman niana)
ɣʈʌp	CCVC	Swallow/the sound of swallowing
ɣʈəmbɑ:r	CCVC.CVC	Roaring of animals/sound of springing water
ɣwəʈəndæ	CCV.CVC.CV	Tether
ɣwəʈəwʌl	CCV.CV.CVC	Lubricate/anoint/to oil
ɣlɑ:	CCV	Theft
ɣwəlɑ:nzə	CCV.CVCC.CV	Udder/teat
ɣwələwʌl	CCV.CV.CVC	Wheedle
ɣwəndoskæ	CCVC.CVC.CV	Ball/any round thing
ɣwʌnɖə	CCVC.CV	Assembly/gathering
ɣwɑ:	CCV	Cow
ɣwədʒʌl	CCV.CVC	Cowshed/stable/cattle pen
ɣwʌr	CCVC	Goitre
ɣwʌrə	CCV.CV	Favourite/best
ɣwʌʈ	CCVC	Oily
ɣwəʈəwʌl	CCV.CV.CVC	To oil
ɣwəʈi	CCV.CV	Ghee
ɣwəʃɑ:n	CCV.CVC	Cowpat/dung excreted by cow or baffalo
ɣwʌg	CCVC	Ear/plectrum
ɣoxt	CVCC	Millet/plants grown for their seeds
ɣwʌxə	CCV.CV	Meat
ɣwemʌnɖ	CCV.CVCC	Uproar/turmoil/commotion/noise
ɣwara	CCVCV	Better
ɣwəz	CCVC	Walnut
qwəlbʌ	CCVC.CV	Yoked/harnessed animals to plough
qwələŋgi	CCV.CVC.CV	Fighting cock
qendʒ	CVCC	Fraud/pride
qʈʌte	CCV.CV	Drivel/blather
krʌtʃæ	CCV.CV	Cartilage
krəpəwʌl	CCV.CV.CVC	Crunch
krʌkə	CCV.CV	Abhorence/dislike/abomination
klʌk	CCVC	Hard/tough
kwʌnɖ	CCVCC	Widower
kənɖkəpʌr	CVCC.CV.CVC	Broken/not smooth
kwʌr	CCVC	Grapes
kwəræ	CCV.CV	Berberis lycium

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

kwedʒdʌn	CCVC.CVC	Engagement/wedding
gwədəwʌl	CCV.CV.CVC	To cripple
grɑ:n	CCVC	Costly/expensive
gwrʌt	CCCVC	Joint between fingers
grʌm	CCVC	Responsible/guilty/liable
grewɑ:n	CCV.CVC	Neckline
gəʈwʌndʒ	CVC.CVCC	A place to keep a pitcher, usually made of wood
gumʌnz	CV.CVCC	Comb
gwɑʈʌ	CCV.CV	A heap of woodchips protected from rain
gwɑ:x	CCVC	Separation/end war
gwʌtə	CCV.CV	Finger/deception
gʈaŋ	CCVC	Rocky area
gʈʌŋ	CCVC	A very loud sound
lʌnɖ	CVCC	Short/short in length
lwɑ:ʈ	CCVC	Long/tall
lwəʈəwʌl	CCV.CV.CVC	Lengthen
lwʌʈ	CCVC	Long/great
lwʌgə	CCV.CV	Hunger
lwəstʌl	CCVC.CVC	Read
lwəʃʌl	CCV.CVC	To milk
Lu:nd	CVCC	Wet/drenched
mərg	CVCC	Death
məʈəqwʌr	CV.CV.CCVC	Small-sized but aged/something small but not easily cooked
mrʌstə	CCVC.CV	Help/take someone's side
mroʈəl	CCV.CVC	Twist/twitch
mʈɑ:we	CCV.CV	Withered/dry
mʈa	CCV	Dead (female)/all dead people or things
mlɑ:	CCV	Back
mlɑ:st	CCVCC	Lying
məlxwʌze	CVC.CCV.CV	A grass
mʌnz	CVCC	Middle
mjənz	CCVCC	Middle
məngwəræ	CV.CCV.CV	A small extremely poisonous snake
məŋgwʌn	CVC.CCVC	Bug
məŋgwʌl	CVC.CCVC	Hand excluding wrist/the hand's width
mu:nz	CVCC	Prayer

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

mʊnɖ	CVCC	Root/base/bottom/origin
mjɑ:ʃt	CCVCC	Month
melməstjɑ	CVC.CVC.CCV	Hospitality
mjənzəgʌʈe	CCV.CV.CV.CV	Arbitrator/mediator
nɑ:st	CVCC	Sitting
nɑdʒoɽtjɑ	CV.CVC.CCV	Illness
ndror	CCCVC	Sister-in-law
nekbʌxt	CVC.CVCC	Lucky
nərɣwʌʈ	CVC.CCVC	Tom/male cat
ngor	CCVC	Daughter-in-law
nxətʌl	CCV.CVC	Attach/cling
nxʌtə	CCV.CV	Tussle/fight/hindrance/hurdle
nxələwʌl	CCV.CV.CVC	To attach/cling
nɣɑ:ʈʌl	CCV.CVC	Wrap
ngwʌlæ	CCCV.CV	Curry
nmɑzʌl	CCV.CVC	Celebrate
nnəwətʌl	CCV.CV.CVC	Enter
nwʌr	CCVC	Sun
nwəʈʌz	CCV.CVC	Coturnix/quail/bob-white/partridge
nwəsæ	CCV.CV	Grandson
nɪmgəʈtjɑ	CVC.CVC.CCV	Incompleteness
Nanŋ	CVCC	Unity
waɣɪst	CV.CVCC	Purchased
wrɪt	CCVC	Fried/roasted
wrʌk	CCVC	Lost
wʈumbæ	CCVC.CV	First
wrʊn	CCVC	Thigh
wʌxt	CVCC	Time
wrɑrʌ	CCV.CV	Nephew/brother's son
wrɑ:n	CCVC	Out of order/naughty
wrɑ:jə	CCV.CV	From a distance
wərpʈɑnestʌl	CVC.CCV.CVC.CVC	Untieng
wərtlʌl	CVC.CCVC	Going
wredʒe	CCV.CV	Rice
wrʌz	CCVC	Day
wrusto	CCVC.CV	After
wrəstʌ	CCVC.CV	Rotten
wʌrkʈə	CVC.CCV	Giving
wʌrkʈərɑ:kʈə	CVC.CCV.CV.CCV	Give and take
wrʌgə	CCV.CV	Plea
wrəndɑ:r	CCVC.CVC	Sister-in-law/brother's wife

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use

wror	CCVC	Brother
wrʊzə	CCV.CV	Eyebrow
wreɪ	CCVV	Gum
wrerʌ	CCV.CV	Niece/brother's daughter
wɽande	CCVC.CV	Before/in the past
wɽʌl	CCVC	Carry
wɽʊkæ	CCV.CV	Small/younger
hɪndwɑnʌ	CVCC.CV.CV	Watermelon
jʌwɪʃt	CV.CVCC	Twenty-one

EBSCOhost - printed on 2/10/2023 4:03 AM via . All use subject to https://www.ebsco.com/terms-of-use