

# Pashto Phonology



# Pashto Phonology:

## *An Evaluation of the Relationship between Syllable Structure and Word Order*

By

Muhammad Kamal Khan

Cambridge  
Scholars  
Publishing



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

*To my mother*

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



# 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	

Appendix B.....	221
Pashto Edge Phonotactics	
Appendix C.....	224
Pashto Consonant Clusters	
Appendix D .....	230
Pashto Words containing Consonant Clusters	



## 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

## 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

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

## 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!

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

## 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

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

## ABBREVIATIONS AND SYMBOLS

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



# 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

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).

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

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.

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.

## 1.4 METHODOLOGIES FOR PHONOLOGICAL STUDIES

Before 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

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

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



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.

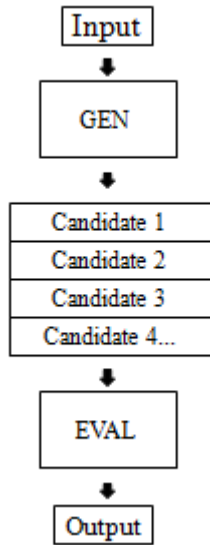


Figure 1.1. The Graphic Representation of OT

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:

## (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

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

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 field. 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:

(4)

- |    |         |         |
|----|---------|---------|
| a. | /wrəz/  | day     |
| b. | /wrot/  | brother |
| c. | /wla:t/ | 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 hierarchy. 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