Newcastle University

The Morphology of Loanwords in Urdu: the Persian, Arabic and English Strands

Submitted by:

Riaz Ahmed Islam

In fulfilment of the requirements for the degree of
Doctor of Philosophy

Supervised by:

Dr. S. J. Hannahs
&
Prof. Maggie Tallerman

School of English Literature, Language & Linguistics
Newcastle University

October, 2011

Abstract

The Morphology of Loanwords in Urdu: the Persian, Arabic and English Strands

Language contact and the influence of one language on another are very common phenomena. Persian, Arabic and English have influenced various languages globally. Urdu is one of the recipient languages from these three sources, and shows linguistic features borrowed from them. This study focuses on the Persian, Arabic and English loanword morphology in Urdu.

Loanwords from Persian/Arabic are far older than English loans, and function like native Urdu words. Therefore, native Urdu morphological structures and those from Persian and Arabic are treated as the patterns for English loans. The discussion describes the patterns and then the processes involving English loans in the light of these patterns. The hypothesis is that the affixation, whether inflectional or derivational, may be based on native Urdu patterns but that the compounding of English loans is more frequent with Persian and Arabic loanwords. This is a major factor, which needs to be established. It is equally vital to know whether Urdu also borrows any derivations of an English loan, as it did with Persian/Arabic loans with or without any morphological changes.

Almost nothing is written on the morphology of loanwords, from the three languages, into Urdu. Furthermore, there is no theory on loanwords specifically dealing with the morphological adaptation of loans. So, the present work is descriptive and deals with the characteristics of the morphological structures from native Urdu, Persian, Arabic and English. Due to space restrictions, the primary focus is on gender/number and case morphology, and derivation of by affixation and by compounding.

The study is divided into six chapters. The discussion begins in the first chapter with an introduction to the study and an overview of the sociolinguistic background of Urdu. It also

discusses the influence of English loanwords on South Asian languages in general and Urdu in particular. The chapter exemplifies pluralisation of English loans and compound verb forms with the words of recipient languages. The next three chapters focus on inflectional morphology, derivation by affixation and derivation by compounding. In chapter 2, only the gender, number and case morphology of the Urdu noun with relevance to the three sources comes under discussion. Chapter 3 focuses mainly on derivational affixes in Urdu from the three sources. Although a discussion of morphological issues is the main concern, some phonological and semantic issues with relevance to morphology are also included. However, phonological issues are only discussed in connection with Arabic loanword phonology, which shares in lexical creations e.g. *ɣʊnḍa* 'scoundrel'. Derivation of new words is also very frequent by means of compounding. Constituents from two different sources very often interact and are rather more frequent than normal, i.e. native + native, combinations in Urdu. Therefore, rather than looking at their source languages, Chapter 4 focuses on various types of compounds i.e. endocentric, exocentric and copulative etc.

Chapter 5 discusses the features of English loanwords adaptation in the light of Persian and Arabic loanwords adaptations. Morphological changes occur both on the inflectional and derivational level. On the inflectional level, the changes are more frequent and based on the native Urdu patterns. Derivational changes are seen in various loans, but the adaptability is limited in the derivation of other categories irrespective of native Urdu or Persian and Arabic patterns. It is far less frequent than the adaptation seen in Persian and Arabic loans. It is more frequent with native Urdu affixes, but the formation of compounds is more frequent with Persian and Arabic loan constituents. Thus, the hypothesis made in the beginning of the study is supported. English loan affixes have not found a place in formal Urdu, although they are used informally. The chapter draws some conclusions. Chapter 6 then presents a summary of the discussion made in the thesis, and presents the implications of the study.

Chapter 1

Morphology and Loanword Morphology

1.1. Introduction

Language contact and the influence of one language on the other are very common phenomena. Persian, Arabic and English are some of the great languages of the world. They have influenced various other languages globally. Although a great language in itself, Urdu is also one of the recipient languages. Lexical influence from various sources is a common occurrence, but it is significant to note that a recipient language shows the loan morphological and phonological features. The same can be seen in Urdu as well, but almost nothing has been studied on any of the linguistic aspects. A great many features are unexplored. This study attempts to bring some of these aspects to the surface. It focuses on the loanword morphology in Urdu and looks at loanwords from Persian, Arabic and English.

This study focuses equally on the morphology of loanwords from these languages. However, the loanwords from the former two are far older than English loans, and function like native Urdu words. They are far more morphologically integrated, which means that they have adopted various morphological changes. Therefore, in this perspective, the investigation into morphological adaptation of English loans is done by looking at the patterns of native Urdu morphological structures and the morphological structures of the older loans. Thus, the morphological structures of native Urdu, Persian and Arabic are treated as the model for the morphological adaptation of English loans, which is discussed in the final part of this study. As English loanwords are more recent, it is necessary to look briefly at some of the background of English loanwords in Urdu.

This is because the Urdu lexicon is basically an amalgamation of the native Urdu words and Persian and Arabic loans. Therefore, the morphological structures which are apparent as a whole are also an amalgamation of the morphological structures from these three sources. 'Native Urdu' is understood to mean a word belonging to Urdu, Hindi or even Sanskrit, provided that it is commonly used in Urdu. The reason for this is that linguists generally treat Urdu and Hindi as the same language, known by different names, i.e. Urdu, Hindi, Hindustani and Hindi-Urdu. According to Rai (2000: 11), "one man's Hindi is another man's Urdu". Linguists normally use the term Urdu-Hindi or Hindi-Urdu to study the common features despite the fact that some lexical, morphological and phonetic differences do exist. Sanskrit is considered to be the mother of both, but due to deep and distinct influences, Urdu is highly Persianised/Arabicised, while Hindi is highly Sanskritised.

Studying loanwords and loanword integration, the distinction and combination of two criteria of conformity, i.e. conformity of loanwords to the source language form and conformity to the target language system, help us better understand the diachronic processes of progressive integration of loanwords (Winter 2008: 156). By invoking conformity in both, one of two alternative possibilities must be addressed. (1) The properties of a loanword are unchanged in a borrowing language and they match those of the source language. (2) The properties of a loanword have changed, and they now match those of a target language system in that the loanword seems to be part of borrowing language vocabulary. Analysing conformity to the (SLF) Source Language Form and conformity to the (TLS) Target Language System, the main goal of this study is to establish first the relation between the Source Language (Persian and Arabic) and the Target Language (Native Urdu), with relevance to the morphology of loanwords. Then, treating the morphological structures of native Urdu, Persian and Arabic as the three faces of Urdu morphology as a whole, it is interesting to look at English loanword

morphology. The hypothesis is that the affixation, whether inflectional or derivational, may be on native Urdu patterns but that the compounding of English loans is more with the older loans (Persian and Arabic), which form a bigger part of the Urdu Lexicon. This is one of the factors which need to be established. Moreover, it is equally important to know whether Urdu also borrows any derivations, of an English loan, as it borrowed Persian and Arabic derivations with or without any morphological changes. The study generally observes the extent of the English loanword adaptation, in the light of Persian and Arabic loanword adaptation, and establishes if it is as complete as that of the two older loan influences. There are other related questions, e.g. what are the structural patterns after the loanword integration; how are the loanwords modified, and which of the elements are affected? It is also of interest to note whether Urdu borrows English affixes in the same way as it has borrowed Persian and Arabic affixes.

Two points in particular need to be considered. Regarding Persian and Arabic influence, there is extremely little written on Urdu linguistics in general and its morphology in particular. There is nothing written on English loanword influence on Urdu either. Therefore, the lack of any related literature is an obstacle. Moreover, there is no theory of loanwords dealing specifically with morphological adaptation of loans. In this context, the present work does not fit clearly into any specific theoretical framework. Therefore, there is need for good, careful descriptive work before serious theoretical analysis can be carried out. So, much of the current work is descriptive and deals with the characteristics of the morphological structures from the three prominent sources i.e. native Urdu, Persian and Arabic. Following this, the discussion of the morphological adaptation of English loans describes the processes involving English in the light of the three sources of morphological structures already shown.

Morphological adaptation includes all aspects of derivational and inflectional integration. However, due to space restrictions, the primary focus is on some major points, including gender/number and case morphology, the derivation of nouns, adjectives and verbs by affixation and by compounding. The proposed study is divided into six chapters. The examination starts in the first chapter with an introduction to English loanwords in Urdu, and more generally morphology and loanword morphology of the South Asian languages. A general picture is sketched regarding the status of English loanwords and how nouns and verbs in particular are adapted in some South Asian/ Indo Aryan languages. In the subsequent chapters, the focus is on the morphology of the Urdu noun itself (Chapter 2), derivation by means of affixation (Chapter 3) and by compounding (chapter 4). Chapter 5 takes a close look at the adaptation of English loans. The final chapter (Chapter 6) concludes the discussion summarising Urdu morphological features and those of English loans. The key point in this chapter is to have a look at the general implications of the study.

In order to discuss these points, the first factor required is to understand morphology or word structure, in particular inflection and derivation. Beyond that, it is equally important to understand the issues involved in loanword morphology.

1.2. Morphology

Morphology is the study of words and their internal structures mainly through the analysis of morphemes, affixation, reduplication (partial or full) and various sorts of compounding. Morphemes are various types of word parts, e.g. prefix, stem, infix and suffix. Thus, a single Urdu word, e.g. *la*ɪ*lmi* 'unawareness' can be broken into three morphemes, i.e. the Arabic prefix *la* 'un', Arabic stem *ɪlm* 'knowledge/awareness' and native Urdu suffix *-i*. The prefix *la* converts the noun *ɪlm* into an adjective *la*ɪ*lm* which is then converted into another noun *la*ɪ*lmi*

by the suffix *-i*. Sometimes, independently occurring words in a compound can also participate in morphological processes. For example, the Urdu word *kəmra-e-dʒəmaəʈ* 'classroom' is a compound word and both the constituents *kəmra* 'room' and *dʒəmaəʈ* 'class' are two independently occurring words. The function of the infix *-e-* is to link both the constituents and give possessive meaning. It replaces the semantic genitive *ka/ ki/ke*, which cannot occur in compounds.

Many morphologists make a distinction between inflection and derivation, which are often associated with affixation (the attachment of a morpheme). Inflection does not change the category but rather the form of the same word, e.g. *plæt* 'plate', *plætæ̃* 'plates', *plætõ* (plural oblique) *plæto* (plural vocative). In the given Urdu examples, *plæt* is the singular noun, *-æ̃* is a plural marker in the second, while *-õ* and *-o* show the oblique and vocative forms of the noun, respectively. Inflected forms are variants of one and the same word. Derivation, on the other hand, means new word formations from a base word. This may change the category of the base, but a derivation may also be in the same category as the base word. A noun, for instance *ʊsul* 'principle', can be converted into the adjective *ʊsuli* 'in principle', the noun and adjective *beʊsul* '(someone) unprincipled' and an adverb *ʊsulən* 'in principle'.

A loanword typically follows the morphological rules of a borrowing language. It may be helpful to know how languages influence each other and what morphological results occur.

1.2.1. Loanword Morphology

When speakers of a language want to identify with each other, they may find themselves adjusting their speech to eliminate the more obvious differences in pronunciation or vocabulary. But when multilingual speakers interact with each other, it often involves

language contact. This can be seen in Kupwar, India, where multilingualism has led to the convergence between the local dialects of two Indic languages, Marathi and Urdu, and two Dravidian languages, Kannada and Telugu (Thomason 2001: 45). Kupwar is a small town of 3000 inhabitants who speak the four languages. Due to centuries of contact, the varieties of the four languages are similar to each other syntactically and are distinct from the varieties spoken in other parts of India. Gumperz and Wilson (1971) note that historically, grammar is thought to be the most persistent, while lexicon is the most changeable component of a linguistic system, but in Kupwar the grammar is adaptable and the lexicon is the most persistent. The words of each language have been retained and coupled with a common syntax. They observe that the adaptations are far-reaching and multilateral, and each language has adopted some features from the others.

Bilingualism and education are two important factors. Presently, electronic media is also an important factor. Note the following two examples. The first is Greece, where most people are educated and monolingual (Tatsioka 2008: 129), and the other is the Indian Subcontinent where most people are bilingual and uneducated. In both these regions, only a small number of people know English, but most people frequently use English words and phrases. Expressions like *oh God!* are very common. This is mainly due to media influence.

The consistent use of such borrowings makes them part of the borrowing language. However, loanwords and simple borrowings are two different terms which distinguish them from each other. Haugen (1950: 212) uses the term "borrowing" for *the attempted reproduction* of some patterns of one language into the other, and the term "loanwords" for *one type of borrowing*. Loanwords are single words or compounds, but borrowings may be stems or full phrases. Loanwords are also different from code-switching, as Poplack et al (1989: 390) state that the

former follows the rules of the recipient language, while code-switching is a linguistic term denoting the concurrent use of more than one language, or language variety, in conversation.

The study of how a language reacts to the presence of foreign words – whether it rejects, translates or freely accepts them – may give us insight into its formal tendencies. The adaptation of a loanword depends both on its word class (verb, noun, adjective) and on its similarity to native words and to loanwords from other languages already existing in the borrowing language.

The native Urdu affixes are productive and form new words in combination with the words from a different source. But some of the loan affixes e.g. Persian *-ana → zalımana* 'tyrant' also derive a good number of words with roots from other sources e.g. Arabic in this case. The productivity of these affixes in general, although systematic, shows some irregularity as well. Seidenberg and Gonnerman (2000: 354) state that, "Morphological systems are quasi regular. They are productive and systematic but also show many seemingly irregular forms." They note that the inconsistencies in word structures reflect the uncertainty about the nature of morphological units and greatly complicate the task of interpreting empirical results; for example, *bakery* is related to *bake* and *cannery* to *can* but what is *groce* in *grocery*? While looking at similar inconsistencies, Ahmed (2008: 2) states that Urdu "lexical items, prefixes, and suffixes are difficult to study on the spoken level since their occurrence is quite unpredictable." Although Urdu morphology follows some patterns and rules, the rules may be complex and may cut across linguistic components e.g. morphology and phonology. It cannot be termed 'unpredictable', as Ahmed claims. Nevertheless, there are certain deviations from the rules of each pattern.

The patterns of words used as nouns differ from those used as verbs or adjectives, in relation to both morphological marking and words' incorporation into the borrowing language structure. The incorporation of loans also indicates both alterations in the phonology of the loanwords and subsequent adjustments in the phonology of the recipient language. For example, with respect to regions, phonetically, there are two alternations in the English noun *street*. Punjabi people say *sətrit*, but others pronounce it as *ɪstrit*. The loan also has two plural variants, *sətritẽ/ɪstritẽ*. The subsequent adjustments in the phonology of the recipient language can be observed in Urdu also due to Arabic loan velar fricatives /x/ and /ɣ/.

However, as the focus of this thesis is on the morphological integration of loanwords, rather than on loanword phonology, only some phonological aspects of relevance to morphology are discussed. Before discussing morphological adaptation, it is helpful to look at the sociolinguistic status of Urdu as compared with other languages of the Indian subcontinent.

1.3.1. Sociolinguistic Status of Urdu as Compared with other Languages

Urdu is an Indo-Aryan language of the Indo-Iranian branch from an Indo European family. Abbas (2002) traces its origin to the armies of Afghan emperor Mehmood Gaznavi in the 12th Century. However, as generally known, the term *Ordu* was first used for the language spoken by the soldiers of Mughal emperor Shahjehan (17th century) when he built the Red Fort in Delhi, and the surrounding town called *Ordu-e-Mu'alla*. Urdu abridged the several languages spoken by the soldiers in the Indian army and absorbed the loanwords from them. That's why it is often called *ləʃkəri* (army) *zəban* (language) or the language of the army. There have been four major dialects of Urdu i.e. 1) Dakhani or now called Daccani 2) Pinjari 3) Rekhta and 4) Modern Vernacular Urdu.

The modern vernacular language is based on the *Khaɽi Boli* dialect, of the Delhi region, spoken in 12th century. Urdu is most closely related to Assamese, Gujarati, Hindi, Marathi, Punjabi, Sindhi, Singhalese, and Romany. It is one of the largest languages of the Indian subcontinent with native speakers of over sixty million, and an official language in Pakistan and the Indian states of Uttar Pradesh, Kashmir and Delhi (Martindale (2009). It is one of the three most important languages in the modern Indian Subcontinent alongside Hindi and Bengali, while languages spoken throughout various regions are provincial languages. It is also widely spoken in Bangladesh, Afghanistan and Nepal and the Indian states of Bihar, Maharashtra, Andhra Pradesh and Karnataka. It is spoken as a medium of communication by a large number of communities in the Gulf, UK and many other countries of Europe, as well as Afghanistan, Thailand, Nepal, Turkey, USA, Australia, Fiji, South Africa, Zambia, Mauritius and Malawi (Junior, 2005). The eloquent Urdu poetry and Indian movies have given rapid popularity to this language. Then, the immigrant workers have spread it in all four corners of the world, and it is now one of the largest languages of the world. Rehman (2004) and Grimes (2000) consider Urdu-Hindi the second most spoken language of the world.

Urdu, Hindi, Hindustani and Hindi-Urdu are various terms used for the same languages. Linguists normally use the term Urdu-Hindi to study common features in both the languages. There are some morphological and phonological differences, but Urdu differs from Hindi mainly in its extensive Persian-Arabic borrowings along with Perso-Arabic script, while Hindi on the other hand is written in Devanagri script. Urdu has been a victim of cultural genocide at different periods of time (Azim (1975: 259) and was ruthlessly ignored by various governments. Perhaps, this is why Urdu borrows English loans more heavily than Hindi. A background of the lexical influence of English can help to understand loanword morphology.

1.3.2. Background of the Influence of English and the Status of the Loanwords

Lexical borrowing brings with it cultural implications, because it is not restricted to words alone, it includes cultural and social values (Hoffer 1980: 2). Borrowing normally involves conquest, cultural domination, disparity in social development, technological advancement and media etc. Lee (2004: 1) states, "Armed with technological, political, economic and cultural machines, the English language has made, and is making, a profound influence and impact on the enrichment or demise of world languages". All these factors have caused English to influence Urdu, which has borrowed a large number of its lexical items.

English is the language of science, technology, business and academic information. It is, therefore, globally thought to be a key to success. With the growth of English as an international language, there has been a rapid increase in the demand for teaching and learning English in Pakistan. It is the medium of instruction in all private schools and a compulsory subject from grade one in all public schools. It is also a popular subject of study in the universities. It is now an empowering language, and considered to have a high status and prestige as a foreign language. Rehman (998, 2005) considers, "English is a language of power and status symbol". The use of English loanwords in daily Urdu is not just a status symbol, it has become a social habit. Code-switching is the most frequent use of English loans in the media and by the elite. The reason for this is possibly the same as suggested by Takashi (1990) about Japanese that English borrowings are common because "they seem to convey a modernity and sophistication about the subject matter under discussion". The appearance of thousands of English loanwords is also due to new ideas and technologies. The borrowing of scientific and technical lexical items has increased the enrichment of Urdu also. The perspective in which these lexical items were borrowed can be understood by Baugh and Cable's (1978: 84) explanation of the power of English words that they were "so intimately

associated with an object or a concept that the acceptance of the thing [involved] was the acceptance also of the word". Scientific and technical lexical items are just an aspect of English loanwords being used in Urdu. Baumgardner (1990: 60, 1998) discusses word formation in Pakistani variety of English. He hightlights some forms of English usage eg. compounding *half pants* 'shorts', affixation *de-notify* and conversion *affectee* 'someone affected by something' etc. However, the use of such words is restricted to Englsih newspapers in Pakistan, as Baumgardner quotes.

In sum, like other world languages, Urdu is heavily influenced by English. It borrows a large number of English words, which is similar to the one noted by Carstensen (1986: 827) in some European languages. For these languages and for Urdu, English is a reservoir from which words can be taken at random. Loanwords related to almost every part of life are in common use, but crucially they are used under the morphological influence of Urdu. Morphological changes also occur both at the inflectional and derivational level. At the inflectional level, the changes are substantial and are based on the native Urdu patterns rather than those of Persian and Arabic loanwords. At the derivational level, there is an adaptation of a large number of loans, but the adaptability is limited in the derivation of words irrespective of the native Urdu or Persian and Arabic patterns. However, there is a difference of frequency in the derivation by affixation and by compounding. It seems more frequent by affixation when English loans correlate with native Urdu affixes. But the formation of compounds is more frequent with Persian and with Arabic loan constituents.

This is a general trend that is not just restricted to Urdu, but that is also noteworthy in many south Asian languages. A brief comparison of the morphological structures of some of these languages is given in the next section.

1.4. English Loanword Morphology in South Asian Languages

This section presents common and contrasting features of English loanword morphology in some of the South Asian languages. However, due to space restriction, the aim is to show only a sketch of the loanwords. Therefore, the focus is on just two aspects, the pluralisation of nouns and hybrid compound verbs in which the constituents belong to different languages.

It seems very probable that the psychological attitude of the speakers of a borrowing language towards linguistic material has much to do with its receptivity to foreign words. Some languages require structural changes in the loanwords. Sapir (1921: 2) argues that the nature and extent of borrowing depends entirely on the historical facts of cultural relation. On this basis, he compares the structural changes and contrast between German and English in their loanwords from Latin and French. A similar contrast with regard to the treatment of foreign English material can be seen in South Asia, where the Indo Aryan languages sometimes allow loans to be only phonetically restructured. At other times, inflectional or derivational changes are made by affixation (or by compounding). One element specific to their word structure is gender/number marking, e.g. in Urdu /a/ (m), /i/ (f) and /e/ (pl), in verbs and adjectives, not only nouns. It is vital in loanword morphology as well. The following subsections show some patterns of English loan nouns in five languages: Urdu, Marwaɽɪ, Sindhi, Punjabi and Pushto.

1.4.1. Pluralisation of English Loan Nouns

Of the five sub continent languages compared in this section, the morphology of four, i.e. Urdu, Marwaɽɪ, Sindhi and Punjabi is similar. Therefore, English loan nouns of one or two syllables may be phonetically different, but morphologically they are mostly the same as their original forms. For example, many loan nouns e.g. *plate*, *glass*, and *jug* are unchanged in Urdu, Punjabi, Sindhi, and Marwaɽi. However, the morphology of Pushto differs from that of

the rest. It exhibits no additional variations. This difference is clearly visible also in the English loan nouns, e.g. *pleɪt* is used with structural changes as *palæt*.

There are not many cases of three or more syllable loan nouns, as for example, *manager*. If and when such words are borrowed, they mostly have to be changed structurally. Table 1 below displays the comparison of the loan base forms in the five languages.

Table. 1: English Root Nouns in South Asian Languages

English	Pushto	Punjabi	Sindhi	Marwaṛɪ	Urdu
1-pleɪt	palæt	plæt	plæt	plæt	plæt
2-glas	gɪlas	glas	glas	glas	glas
3-bɪldɪŋ	bɪlding	bɪldɪŋ	bɪldɪŋ	bɪldɪŋ	bɪldɪŋ
4- bʊk	bʊk	bʊk	bʊk	bʊk	bʊk
5- məʃiːn	məʃiːn	məʃiːn	məʃiːn	məʃiːn	məʃiːn
6- kɪtʃin	kɪtʃiːn	kɪtʃin	kɪtʃin	kɪtʃin	kɪtʃin
7- kəmɪʃn	kəmæʃən	kəmɪʃən	kəmɪʃən	kəmɪʃən	kəmɪʃən
8- dʒəg	dʒəg	dʒəg	dʒəg	dʒəg	dʒəg
9- peɪpə	pæpə-ər	pæpər	pæpər	pæpər	pæpər
10- hospɪtəl	əspəṭal	əspəṭal	əspəṭal	əspəṭal	əspəṭal

Note that Pushto shows both phonological and morphological changes, and so it is a good example of structural contrast with the other regional languages. For example, in pluralisation, it shows no gender marking and takes the same suffix *-una* in both masculine and feminine plurals. As the illustrations of this are given in the table below, Pushto is the only singled out here. In the other four languages, the pluralisation of the loan nouns is based on the alternation of vowels at the suffix position, and the loan stems are affixed with the native

plural morphemes. Therefore, no morphological difference remains between the English loan

nouns and the native nouns.

Table 2a: The Masculine Plurals of English Loans in South Asian Languages

English	Pushto	Punjabi	Sindhi	Marwaɽɪ	Urdu
1. kɪtʃɪn	kɪtʃɪnoon*a*	kɪtʃɪn	kɪtʃɪn	kɪtʃɪn	kɪtʃɪn
2. kəmɪʃn	kəmæʃən<u>u</u>n*a*	kəmɪʃən	kəmɪʃən	kəmɪʃən	kəmɪʃən
3. dʒəg	dʒəgun*a*	dʒəg	dʒəg	dʒəg	dʒəg
4. peɪpə	pæpərun*a*	pæpər	pæpər	pæpər	pæpər
5. mænɪdʒər	menædʒərun*a*	mənɪdʒər	mənedʒər	menedʒər	menedʒər
6.*a*m 'mango'	*a*mun*a*	*a*m	*a*m	*a*m	*a*m

Table 2b: The Feminine Plurals of English Loans in South Asian Languages

English	Pushto	Punjabi	Sindhi	Marwaɽɪ	Urdu
1-pleɪt	p*a*lætun*a*	plæt*ã*	plæt*ũ*	plæt*ã*	plæt*ã*
2-gl*a*s	gɪl*a*sun*a*	gl*a*s*ã*	gl*a*s/gl*a*s*ũ*	gl*a*s*ã*	gl*a*s*æ̃*
3-bɪldɪŋ	bɪldɪŋgun*a*	bɪldɪŋg*ã*	bɪldɪŋg*ũ*	bɪldɪŋg*ã*	bɪldɪŋg*æ̃*
4- bʊk	bʊkun*a*	bʊk*ã*	bʊks/bʊk*ũ*	bʊk*ã*	bʊk*æ̃*
5- məʃiːn	məʃinun*a*	məʃin*ã*	məʃin*ũ*	məʃin*ã*	məʃin*æ̃*
6. mæz 'table	mæzun*a*	mæz*ã*	mæz*ũ*	mæz*ã*	mæz*æ̃*

The two tables exhibit some common and contrasting features in the plural morphology of some of the most commonly used English loan nouns in comparison with one native plural in the five languages. The nouns in (2a) are treated as masculine and those in (2b) are feminine. There may be morphological changes in masculine nouns as well, but the changes mainly occur in the feminine nouns. Pushto shows no gender distinction. In four of the languages, the pluralisation of the English loans is by suffixation, but the suffixation itself is no more than an alternation of vowels, which is the main element in their gender/number morphology. Phonological alternations come in many shapes and sizes, and the processes behind them are equally varied as are the kinds of factors which condition them (Davenport and Hannahs, 2005). Masculine loan nouns generally remain unchanged in their plurals, nevertheless a plural is sometimes formed by the alternation of vowels. For example, *deta* 'datum' which is treated as masculine singular is pluralised as *dete* in Urdu. But in Marwaṛī and Sindhi, it is taken as *deto* in its singular form and *deta* in its plural.

A feminine plural is based on the suffxal alternation of vowels with nasalisation. For example, a feminine nominative plural in Marwaṛī and Punjabi is *plætã*, but it is *plætẽ* in Urdu and *plætũ* in Sindhi. Note that there is just an insertion of a nasalised vowel in the loan base noun. A feminine plural for oblique and vocative forms is suffixal with the insertion of nasalised vowel i.e. õ and non nasalised vowel o respectively. Thus, the case morphology is also similar in these languages. Sindhi sometimes tends to use original plural patterns borrowed from English. Unlike the rest of the five languages, Pushto however discards gender distinction, and its pluralisation is also based on the consonant-vowel suffix *-una*, for both genders, rather than merely the alternation of vowels.

Generally, the pluralisation of a noun in the five languages is based on either alternation of vowels or suffixation. This is in sharp contrast to verb formation with English loans, which is done only by compounding.

1.4.2. Loan Verb Morphology

The verbs in these languages are very complex as compared to other syntactic categories. They inflect for tense, mood, gender and number. The verbal structure is in two forms. There are a large number of verbs already present in the basic lexical form. For example, a lexical verb *pi* 'drink' requires the suffix *-na* to form the infinitive *pina* 'to drink'. It takes other suffixes according to tense and aspect needs, e.g. *pija* 'drank' and *pirəha* 'drinking'. In the absence of lexical verbs, nouns and adjectives (particularly loans) combine with some dummy verbs to form compound verbs. A dummy verb is a kind of tool for the verb formation in the absence of base verbs in south Asian languages. In Urdu, it is generally a redesigned form of a Persian auxiliary. Schmidt (1999: 101) explains that Urdu verbs demonstrate a very regular conjugation with the exception of five verbs *ho* 'be', *kər* 'do', *ḍe* 'give', *le* 'take', and *dʒa* 'go'. These five verbs primarily function as main verbs and secondarily as dummy verbs (Versteegh 2001, 488) and light verbs (Butt, 1995) in complex predicates. The difference between a dummy verb and a light verb is not the issue here but, as concerns Urdu, it is explained in Chapter 4. For the present purpose, it is sufficient to note that a dummy verb is used as akind of tool to form a verbal compound in combination with a noun, e.g. N (*sãs* 'breath') + dV (*le* 'take') = V (*sãs le* 'breathe') or with an adjective, e.g. A (*xʊʃ* 'happy') + dV (*ho* 'be') = V (*xʊʃ ho* 'be happy'). It is needed when a lexical (base) verb does not exist. Versteegh (2001: 497) considers that the extensive use of dummy verbs in compound verbs is typical of Hindi-Urdu (Indo Aryan languages) generally.

The use of dummy verbs is particularly ubiquitous with loanwords for verb formation. Even if a loan verb enters Urdu, or other Indo Aryan languages in general, it ceases to function as a verb, and it has to combine with a dummy verb, which carries the inflectional information. The morphological functions of a dummy verb are the same. Not a loan lexical verb, but only a dummy verb shows any gender and number or tense and aspect changes. Thus, all the loans, whether they are nouns, adjectives or even verbs, must be integrated with dummy verbs to form compound verbs. For example, the English loan verb *provaid* 'provide' cannot function as a base verb in Urdu. Although borrowing verbs is frequent and convenient, the borrowed verbs cannot function as verbs unless they integrate with the dummy verbs. Therefore, to perform the function of a verb, the loan verb *provaid* is combined with a dummy verb *kərna* 'to do' (base form *kər*) to form a compound verb, i.e. *provaid kərna* 'to feel'. *provaid* is only used as a dysfunctional verb (used in the status of a noun). A loan verb does not function as a verb in Urdu. Its verbal capacity is taken by the dummy verbs probably due to the morphologically complex verbal systems of the Indo Aryan languages.

In light of this discussion, it can be claimed that verb construction with a loanword is in the form of a compound. More specifically, it is a hybrid compound. Various linguists have interpreted the term *hybrid* in different words. For example, Capuz (1997: 8) terms it *hybrid* or *loan blend*. The same terms are defined by Haugen (1950: 215) as those instances of lexical borrowing in which both 'importation' and 'substitution'" can be found. Kent (1999) interprets the term "hybrid" differently as a *pseudo loanword*. A hybrid combination of a native and a loan constituent is a general phenomenon in compounding. But as the Indo Aryan languages are heavily influenced by Persian-Arabic loanwords, it is also noteworthy that some dummy verbs themselves, e.g. *kərna,* are morphologically redesigned forms of Persian auxiliaries. It means that the verbal formations by them with English bases deviate from the

morphological rules of both the source language and the target language. The detailed discussion in Chapter 5 shows the situation in Urdu in this regard. Table 3 illustrates some compounds formed by dummy verbs in the five languages:

Table. 3: The Formation of Verbs with English Bases in South Asian Languages

English Root	Pushto	Punjabi	Sindhi	Marwaɽɪ	Urdu
1. provaid	kavəl	kərna	kərə̃ɽ	kərno	kərna
2. pruv	kavəl	kərna	kərə̃ɽ	kərno	kərna
3. əgri	kavəl	kərna	kərə̃ɽ	kərno	kərna
4. ərændʒ	kavəl	kərna	kərə̃ɽ	kərno	kərna
5. eksɛpt	kavəl	kərna	kərə̃ɽ	kərno	kərna
6. kritisaize	kavəl	kərna	kərə̃ɽ	kərno	kərna
7. rɪsponsəbɪləti:	haɣəʃʈəl	lẽɽã	wəʈʰə̃ɽ	lẽɽõ	lena
8. rɪspons	vərkavəl	ḍẽɽã	ḍejə̃ɽ	ḍẽɽõ	ḍena
9. hæpi	ʃo	hõɽã	ʈʰɪjə̃ɽ	hõɽõ	hona
10. æŋgri	ʃo	hõɽã	ʈʰɪjə̃ɽ	hõɽõ	hona

The patterns of verb formations are the same in all the five languages. The Indo Aryan languages require every English loan to be attached with a device for allowing verbal marking. This device is a dummy verb which forms a verbal structure. The dummy verbs, e.g. Urdu *kərna* 'do' (1-6), for transitive and *hona* 'to be' for intransitives and passive (9-10), construct infinitives e.g. *provaid **kərna*** 'to provide' (1) and *æŋgri **hona*** 'to be angry' (10). Similarly, for gender/number marking or tense/aspect morphology, it is only the dummy verb that undergoes morphological changes. The loan constituent, whatever the category, does not undergo any morphological changes. The only difference between the dummy verbs with regard to their sources is their (morpho) phonological structures. The dummy verbs used in

the four languages are similar to each other and take a nasalised syllable or an alveolar nasal + a vowel. On the other hand, Pushto differs again and mostly takes a lateral ending (1-8). There is no other significant difference, and therefore no further discussion is necessary at this point.

1.4.3. Loanword Morphology in South Asian Languages: A General View

The relevance of the foregoing discussion has been to show briefly two important aspects of loanword morphology in the five important Indo Aryan languages, i.e. the pluralisation of English loan nouns and the compound verbal structures formed by the English loanwords and the native dummy verbs. Generally (with the exception of Pushto), their gender/number morphology is based on the alternation of vowels. However, only the loan nouns treated as feminine are generally affected by the morphological changes. Their plural morphology is formed by the alternation of vowels. The vowels are nasalised. The nouns treated as masculine remain mostly unchanged. The gender marking of English loans is only seen in four of the languages; there is no gender distinction in loanwords in Pushto. Also, for pluralisation all loans undergo a change by suffixation and the loan nouns take the native plural marker -*una* e.g. gl*as**una** 'glasses'.

Verb formation with English loanwords in all five languages is achieved by compounding with dummy verbs. Even if a loan is borrowed in its basic verb form, it ceases to function as a verb. There are no morphological changes in the English loan constituent, and the dummy verb constituent takes the verbal capacity and gives the inflectional information. The combination thus forms a hybrid compound, and the formation is an interesting phenomenon in itself. The formation thus deviates from source language and target language rules.

Almost all the dummy verbs in the five languages are similar in morphological structures, the only significant difference being that the dummy verbs of four of the languages are based

19

either on nasal or nasalised vowels, while those of Pushto are mainly based on the occurrence of a lateral. There are various dummy verbs for transitive and intransitive verb formations and differ from each other according to their semantic expressions.

1.5. Summary of the Chapter

This chapter has briefly introduced the discussion which will be gone into in greater depth in the following chapters. The study is generally concerned with the adaptation of loanwords from Persian, Arabic and English. The loanwords from the former two languages are fully integrated and have formed certain morphological patterns in Urdu. Thus various morphological patterns borrowed functions parallel to native Urdu morphological patterns. In this perspective, the study focuses on the English loanwords whether they follow native Urdu morphology or Persian and Arabic loanword morphology.

For this, Urdu morphology itself requires due attention, in order to contextualise the adaptation of English loans. This cannot be done without paying proper attention to the three major sources of words in Urdu, i.e. native Urdu, Persian and Arabic. However, no literature is available in this respect. Therefore, much of the present work is descriptive and deals with the characteristics of Urdu morphological structures. Moreover, there is no theory available which specifically focuses on the adaptation of loanwords. So, there is also need for good, careful descriptive work before serious theoretical analysis can be carried out in future works. Thus, the discussion in general is descriptive and brings out features of the adaptation of loans from Persian, Arabic and in particular English, which are more recent and very important.

Due to its more recent status and its importance, the chapter has also given a brief background of the status of English in Pakistan and a sketch of English loanword adaptation in five South

Asian languages. The structural patterns of English loanwords may be different in one way or the other, the loanwords are generally similar. However, the major point of debate to which we will be returning is the adaptation of English loans with relevance to Urdu morphology.

In order to understand English loanword adaptation, it is significant to have an idea of Urdu morphology, which is an amalgamation of native Urdu, Persian and Arabic morphological structures. Therefore, their features require attention in inflectional morphology and derivational morphology. The next three chapters focus on inflectional morphology, derivation by affixation and derivation by compounding. However, regarding inflectional morphology, only the morphology of the Urdu nouns is discussed, due to space restriction, as the English loan noun is the major component in morphological adaptation.

****

Chapter 2

The Morphology of the Urdu Noun

2.1. Introduction

As introduced in the last chapter, the morphology of the Urdu noun is of relevance to the three sources of vocabulary i.e. native Urdu, Persian and Arabic. Native Urdu nouns behave in one way, the nouns borrowed from Persian behave another way and the nouns borrowed from Arabic behave in yet another way. From this perspective, the first point to be noted in this chapter is gender marking. This is because the major points of focus are based on the gender distinction in Urdu. There is then 1) the plural morphology and 2) the morphology associated with case. Although the loans generally follow their source language forms, there may be native Urdu gender and number marking, or vice versa.

Native Urdu distinguishes two genders (masculine and feminine), two numbers (singular and plural) and three cases i.e. nominative (also called direct case), oblique, and vocative. Most nouns are marked for gender, number and case. There are a tiny number of unmarked nouns, which cannot be changed structurally e.g. *gaõ* 'village/villages' (m) and *paõ* 'foot/feet' (m). Both are invariable and exhibit no morphological changes, although they are treated as masculine and can be used as singular and plural. With respect to morphological changes in the Urdu noun, Moizuddin (1989: 20) defines some criteria: 1) there are declinable and indeclinable nouns, which can and cannot be changed structurally; 2) many structural changes are correlated with gender and number agreement and 3) many other changes are based on a noun's relation to case markers. For example, the singular noun *ləɽka* 'boy' is pluralised as *ləɽke* 'boys', which is also used as the singular oblique form. The plural oblique form is *ləɽkõ*. The oblique forms whether singular or plural are always followed by a case marker e.g.

ergative *ne* as *ləɽke ne* (singular) or *ləɽkõ ne* (plural). These three types of inflections show various forms of morphological structures in Urdu nouns.

Hardie (2003: 35) states that, "Urdu inflection is based on suffixation; the suffixes are fusional, consisting overwhelmingly of a single syllable, or even a single vowel, that may mark multiple features e.g. gender, number and case are marked on nouns". The gender of a noun is generally indicated by the final vowel *-a* for masculine, which is the basic gender, and *-i* for feminine, which is derived by substituting *-i* for *-a*. They are the major gender markers, but there are some others as well:

Masculine Noun		Feminine Noun
1a. bet*a* 'son'		bet*i* 'daughter'
b. bhəŋg*i* 'toilet cleaner'		bhəŋg**ən**
c. ũt 'camel'	(all other (m) suffixes)	ũt**ni**

As the examples show, there seems to be a correlation between both the genders. If a masculine noun ends in *-a* e.g. *beta* 'son' (1a), the feminine form is derived by substituting *-i* for *-a* e.g. *beti* 'daughter'. Most native masculine nouns end with *-a*, but certain others end with *-i* as in *bhəŋgi* 'toilet cleaner' (1b), in which case the feminine is formed by substituting or affixing *-ən* i.e. *bhəŋgən* 'toilet cleaner'. Whether or not all other masculine nouns take any suffix, e.g. *bhuʈ* 'ghost', the feminine forms are derived by the substitution or affixation of *-ni* i.e. *bhuʈni*. The feminine gender thus shows more morphological alternants.

The two gender markers are applicable to both animate and inanimate nouns. Most inanimate nouns take the same markers. For example, *sona* 'gold' is masculine and *tʃanɖi* 'silver' is feminine. They are also often used to indicate the size i.e. masculine for bigger e.g. *tʃʰura*

'large knife' and feminine for smaller e.g. *tʃʰuri* 'small knife'. Rizvi (2007: 73) states, "There is no general rule for gender classification for inanimate nouns. Usually huge, heavy, powerful, dominant and big objects are masculine, while small, weak and light are feminine". In this context, a large number of plants and inanimate objects are generally assigned grammatical gender and must be either masculine or feminine.

There are also deviations from expectations. For example, the feminine of the noun *hathi* 'elephant' should be **hathən*, but this is ungrammatical. The correct form is *həth**ni***. There are various similar examples of not only native Urdu but also of loan nouns. Therefore, the focus on gender and number marking is specially in connection with the claim made by Ijaz and Moin (2003: 56) that there is a well-defined set of rules for gender and number changes, which take place by a series of processes, at the morphological and phonological level. Although there are certain rules and general phenomena, evidence is also available against their claim. There are identifiable rules or patterns, but also exceptions. Each morphological pattern shows deviation from the rules.

The deviations are particularly visible in the pluralisation of the nouns. The changes in the native nouns are presented first, as this helps in understanding the loan noun morphology.

2.2.1. Pluralisation of Native Nouns

The pluralisation of a noun is determined by the gender of that noun. If masculine and feminine nouns end in *-a* and *-i*, they take the plural markers *-e* and *-ɪjã* respectively. However, there are also other plural patterns correlated with gender, as given below:

2. **Nouns and noun pluralisation**

	Masculine	**Masc. Plural**	**Feminine**	**Fem. Plural**
a.	ləʈk*a* 'boy'	ləʈk**e**	ləʈki 'girl'	ləʈkɪj*ã*

b.	bhəŋgi 'toilet cleaner'	bhəŋgi	bhəŋgən	bəŋgənẽ
c.	mor 'peacock'	mor	morni	mornijã

(2a) describes the most general patterns of masculine and feminine plurals with Urdu native nouns. However, a large number of masculine nouns, irrespective of endings e.g. *bhəŋgi* 'toilet cleaner/toilet clean**ers**' (2b) and *mor* 'peacock' (2c), remain unchanged in their plurals. This is due to the fact that *bhəŋgi* doesn't end in *-a* and *mor* has no vowel ending. Therefore, regular pluralisation of masculine nouns affects only those nouns ending in *-a*.

In contrast, feminine nouns take plural markers with respect to their endings. Many feminine nouns ending in *-ni* e.g. *morni* 'peahen' (2c) or *həthni* 'elephant' (f) also take the plural marker *-jã* i.e. *mornɪjã/həthnɪjã*, which is the same as in (2a). All other feminine nouns e.g. *behn* 'sister' take the plural marker *-ẽ* e.g. *behnẽ* 'sisters'. In general, the plural rule for both genders depends on the ending of a noun. The sketch of native plurals can be described as:

Masculine Nouns   **Feminine Nouns**

3a) *-a* (sg) → *-e* (pl)   c) *-i* (sg) → *-ɪjã* (pl)

  b) all other endings → pl. unchanged.   d) *-ni* → *-ɪjã* (pl)

                                                        e) all other endings (sg) → *-ẽ* (pl)

In short, the plural form depends on the gender of the noun, as specific pluralisation strategies are associated with the gender of the noun. Kashif (2004: 91) notes some general features in the inanimate plurals, in connection with gender markers, all metals e.g. *sona* 'gold' and many food items e.g. *ḍehi* 'yogurt' are always treated as singular.

The gender and number phenomena described above show some regularity. There are also some deviations from the rules, which are discussed in the following subsection.

2.2.2. Irregularities and Deviations in Native Urdu Nouns

Although there are not many examples, native Urdu nouns do show some deviation from each of the regular patterns particularly feminine forms i.e. *-i, -ən* and *-ni* respectively. As noted in the irregular feminine form *həthni*, the same occurs with some other feminine forms. For example, the feminine of masculine singular *lohar* 'blacksmith' should be **loharni*, but it is ungrammatical and the correct form is *loharən*.

Moreover, there are deviations from plural norms. Generally, masculine plural ends with *-e* and feminine plural with *-ɪjã* and *-ẽ*. The examples in (4) illustrate deviation in gender numbeand number markingr:

Noun (m)	Plural	Feminine (f)	Plural
4a) bəkr*a* 'goat'	bəkr**e**	bəkr**i**	bəkrɪ**jã**
b) kʊtt*a* 'dog'	kʊtt**e**	kʊtt**i** /kʊttɪ**j*a***	kʊttɪ**jã**/ kʊttɪ**jaẽ**
c) tʃɪɽa 'cock-sparrow'	tʃɪɽe	tʃɪɽi/tʃɪɽja 'sparrow'	ʃɪɽjã / tʃɪɽjaẽ
d) dakɪja 'postman'	dakje	-	-

As shown in (4a), and noted above, the masculine *-a* and feminine *-i* are the major gender markers, while *-e* and *-ɪjã* form their plurals. But (4b) and (4c) show that there are sometimes deviations from the set patterns. Even if masculine ends in *-a*, its feminine marker can also be *-ɪja* rather than the regular marker *-i*. The feminine plural *-ɪjaẽ* also takes the form accordingly, besides *-ɪjã*. Despite the different backgrounds of the two nouns i.e. *kʊtta* comes from Praket and *tʃɪɽa* comes from Sanskrit, their gender and number markers show the same variations. The changes in the feminine singular/plural marking are the same in both the nouns. The deviations *-ɪja* (singular) and *-ɪjaẽ* (plural) do not however exist in large numbers. They are also used optionally.

The feminine singular marker *-ɪja* may be optional, but for some masculine nouns e.g. *dakɪja* 'postman' (4d), it is mandatory. This ending of the masculine singular is different from its primary form *-a*. It is possible in feminine nouns, but it is unusual in masculine nouns. The masculine noun does not have a feminine counterpart, which is even more unusual. This is not an exceptional case. Any other masculine nouns e.g. *tʃərɪja* 'mad' and *kalɪja* 'someone blackish in complexion', with this suffix, do not have a feminine counterpart. However, whether masculine singular ends in *-a* or *-ɪja*, the plural marking of both is the substitution with *-e*. This again shows that there are some complexities, and a few masculine nouns ending with *-ɪja* may be termed lexical exceptions.

The variations in feminine nouns are generally greater than those in masculine nouns. A masculine noun takes the plural marker *-e* e.g. *bete* 'sons' only if it ends with the vowel *-a*. Otherwise, it remains unchanged (4a-4c). Although the masculine singular in (4d) is different from those of (4a-4c), it takes the plural marker *-e.* Therefore, this is unlike the feminine plural markers (4b & 4c) that change according to the changes of feminine gender markers. There are more feminine gender markers i.e. *-i*, *-ɪja*, *-ən* and *-ni*, than the only masculine one i.e. *-a*. Feminine nouns also take two plural markers i.e. *-ɪjã* and *-ẽ* as compared to masculine nouns, which usually remain unchanged except for taking *-e*.

In short, the gender and number morphology of native Urdu nouns frequently shows irregularities. Although there are rules, the simple gender or number relation is often difficult to define due to frequent lexical exceptions in the morphological formations. While there are certain rules, there are also important exceptions.

This is also evident in Persian and Arabic loan nouns. For example, in Persian, nouns generally show no gender, but its loans in Urdu are obligatorily marked for gender. Therefore, there is likely to be a mismatch between the behaviour of nouns in the two languages.

2.3. Gender and Number Morphology of Persian Loan Nouns

This section discusses the gender and number morphology of the Persian loans and their deviations from both the native nouns and the Persian loans. Persian does not have gender morphology, and so the Persian nouns are assigned gender for Urdu. In addition, most loan nouns take the native Urdu plural markers. Therefore, it is imperative to focus first on these two points so that deviation from native Urdu morphology can be understood. Then, it is also significant to note that Urdu borrows more plural markers from Persian than from Arabic, though generally most of them are of lower frequency. Only two of them i.e. *-an* and *- aṯ* form a considerable number of plurals, which is discussed later in the section.

2.3.1. Persian Loans with Native Urdu Gender and Number Markers

Persian does not show grammatical gender. Megerdoomian (2000: 4) states that there is no gender distinction in Persian, but there are several plural morphemes some of which are loans from Arabic. This means that a Persian loan noun must be assigned gender, which is typically based on an equivalent native Urdu noun or in some cases on an Arabic equivalent. For example, a loan e.g. *əɖakar* 'actor' becomes *əɖakara(h)* 'actress' (f) in Urdu, although the noun is the same for both genders in Persian. The feminine marker is in fact influenced by the Arabic feminine marker e.g. *ʃaɪra(h)* 'poetess'. A brief sketch below illustrates the gender marking in Persian loans:

Masculine Noun   **Feminine Noun**

5a. bətʃ*a*(h) 'child' (m)   bətʃ**i** 'child' (f)

b. əḏ*aka*r (generally rhotic ending)   əḏ*aka*r*a*(**h**) 'actress'

c. ḏərz**i** 'tailor'   ḏərz**ən**

d. all other nouns e.g. ʃer   ʃer**ni** 'lion'

In many loan nouns e.g. *bətʃa*(*h*) 'child' (5a), the final [-h] is deleted. Therefore, they are treated the same as the native masculine nouns ending with -*a*. Keeping the correspondence between masculine and feminine, all masculine nouns with a final -*a* have the feminine marker -*i* e.g. *bətʃ**i*** (f) in Persian loans also. Thus, the gender morphology of the loans adopts the native patterns.

However, a problem arises when some loan nouns with the same suffix -*a*(h) e.g. *əḏakar**a*(**h**)* 'actress' (5b) are treated as feminine. This is in fact due to the Arabic influence on both languages, because these nouns are treated as feminine in Arabic from which such loans in Urdu e.g. *ʃaɪr**a*(**h**)* 'poetess' are also treated as feminine. They may nevertheless be borrowed either directly or via Persian. Rizvi (2007: 73) points out that -*a*(*h*) is used for mostly feminine nouns in Urdu. This means that some Persian loans are also assigned a feminine gender ending in **-***a*. Thus, semantics override what appears to be a grammatical marker. However, these feminine loan nouns are not numerous.

A similar contrast can be seen in the gender of masculine loan nouns, which end with -*i* e.g. ḏərz**i** 'tailor' (5c). These loans are feminised with -*ən* e.g. ḏərz**ən**. Most other masculine nouns with no specific ending, generally -*r* e.g. ʃer 'lion', take the feminine marker -*ni* e.g. ʃer**ni** 'lioness. Both the nouns in (5c & 5d) behave like native nouns, as in (1) seen above and show that the loans are fully integrated in Urdu morphology.

Like gender marking, plural marking mostly follows native morphology, although there are distinctive plural markers as well. Thus, a masculine loan noun ending in -*a* e.g. *bətʃa* 'child' (5a) is pluralised with -*e* i.e. *bətʃe*. All other masculine nouns e.g. *ʃer* 'lion' remain unchanged in the plurals. The feminine loan nouns with the same suffix -*a* are pluralised with -*ẽ* e.g. *əḍakaraẽ* 'actresses'. Some feminine loan nouns with no specific ending are also pluralised with -*ẽ* e.g. *ḍərzənẽ* 'seamstresses'. All other feminine loan nouns ending in -*i* or -*ni* e.g. *bətʃi* 'child' (f) are pluralised with -*ɪjã* e.g. *bətʃijã*.

Most of the loan nouns appearing with the native plural markers show this general tendency. Detailed discussion of native plural markers has been presented in the previous sections, so no further discussion is necessary at this stage. There are however many Persian loan plural markers, to which we now turn.

2.3.2. Persian Loan Plural Markers

Although most Persian loans are completely integrated with native Urdu plural markers, the morphology of many others is not restricted to this. Many loan nouns, treated as masculine, take Persian plural markers, but those treated as feminine are almost all devoid of the loan plural marking. There are five Persian loan plural markers, all are used for masculine nouns. Two of the loan plural markers -*an* and -*at̪* are frequent. Three others i.e.-*gan*, -*ha* and -*dʒat̪* cannot be fused with many bases. They may attach to the native Urdu nouns e.g. *bar**ha*** 'many times' and even to Arabic loan nouns e.g. *məsala**dʒat̪*** 'spices'. However, they are not very productive. Table 2 illustrates some examples:

Table 1 Plural Markers in Persian Loans

No.	Masculine Noun	Pl. Marker	Fem: Noun	Plural
1.	fərzənd 'son'	fərzənd*an*	dʊxt̪ər 'daughter'	dʊxt̪ər*an*
2.	məqal*a*(h) 'research paper'	məqal*at̪*		
3.	selab zəd*a*(h) 'sb flood affected'	selabzəd*gan*		
4.	həz*a*r 'a thousand'	həzar*ha*		
5.	məs*a*la 'spice'	məsala**dʒ***at̪*		

Of the the first two major plural markers, -*at̪* may be the result of Arabic influence on both Persian and Urdu. It is a plural marker for the feminine nouns in Arabic, but it doesn't show any gender distinction in Persian. In Urdu, -*at̪* is a feminine plural marker for Arabic loans e.g. *t̪alıbat̪* 'students' and a masculine plural marker for Persian loans e.g. *baɣat̪* 'gardens'.

In contrast to -*at̪*, -*an* is mainly used with masculine loans e.g. *fərzənd**an*** 'sons', but there are some examples of feminine loans as well e.g. *dʊxt̪ər**an*** 'daughters'. In semantics, it is parallel to -*ha* in Persian; while -*ha* is the most productive and can be attached to most nouns, -*an* is generally used with animate nouns in formal language, but can also pluralise some inanimate nouns (Kahnemuyipour, 2000). For example, some plurals e.g. *dʊxt̪ər**an*** 'daughters' and *dʊxt̪ər**ha*** are possible. In contrast, Urdu discards any difference, and -*an* is commonly used, as compared to -*ha*, which is restricted in Urdu plural marking. Therefore, the plurals like **dʊxt̪ər**ha*** are ungrammatical in Urdu, although they may be normal plurals in Persian.

However, -*ha* is distinctive in Urdu in the sense that it attaches to nouns and adjectives and its plurals are mainly used as adjectives and adverbs. For example, in the phrases, *bar**ha** məvaqe* 'many occasions' and *həzar**ha** vədʒuhat̪* 'thousands of reasons', -*ha* functions as the plural

marker of adjectives rather than those of nouns. Thus, *-ha* generally forms plural adjectives and adverbs by attaching itself to some native Urdu bases e.g. *bar* 'times'.

The examples of plural adjectives by *-ha* are a few, in spoken Urdu, and so are grouped with the less frequent Persian loan plural markers *-ha*, *-gan* and *-dʒaṯ*. Nonetheless, the latter two form plural nouns. The case of *-gan* is worthy to be noted, as its affixation causes phonological changes in the coda position of the bases. Any loan noun or adjective ending with the suffix *-ḍa(h)*, is often pluralised with *-gan* e.g. *bənḍa(h)* 'servant of God' → *bənḍəgan*. After the affixation of *-gan*, the base *bənḍa(h)* changes into *bənḍə*. Katamba (1993: 89) in discussing English divides affixes into two groups, neutral and non neutral. Neutral affixes have no phonological effect on the base to which they attach themselves e.g. *home/home**less***, in which *-less* is neutral, as it has no phonological effect on the base. In contrast, there are also some affixes that affect the base phonologically e.g. *morpheme* → *morphem**ic***. The suffix *-ic* is non-neutral, because it provokes stress shift in the base: in **mor**pheme the stress falls on the first syllable, in *mor**phe**mic* the stress falls on the second syllable. Therefore, *-ic* is termed non neutral affix. In this context, *-gan* is the only non neutral Persian loan suffix. The baseword is modified, and the suffix *-a(h)* is substituted by schwa *-ə-* along with the affixation of *-gan*. This is however different from nouns such as *fərzənḍ* 'son' which end in dental and require only affixation of *-an* for pluralisation with no change in the basewords. This is also different from the nouns which end in *-a(h)* e.g. *meva(h)* → *meve* 'dry fruits' and generally have the native plural markers *-e* for masculine and *-ẽ* for feminine. The difference of *-gan* in the two cases shows that it requires an ending of only *-ḍa(h)* in the bases. Despite the changes, *-gan* forms a regular pattern of pluralisation.

As compared to *-gan*, the use of *-dʒaṯ* is even less frequent. Some loans ending with *-a(h)* take this plural marker, e.g. *meva(h)* 'dry fruits'→ *mevadʒaṯ*. However, as noted above *-dʒaṯ* is substituted by *-aṯ* or native *-e*. Most loans generally have the native Urdu plural marker. Therefore, this is the rarest of the three less frequent loan plural markers even in written Urdu.

Both *meva(h)* and *məqala(h)* have the same endings and take the native plural marker. But the loan plural markers, which they take, are different i.e. *meva**dʒaṯ*** and *məqal**aṯ***. Similar usual plural pattern have been noted in the case of *-ha*, which attach itself only to adjectives to form plural adjectives rather than forming plural nouns.

This shows that there are frequent lexical exceptions in the loan plural patterns also. There are some deviations from Persian plural patterns and native plural patterns. This is discussed in detail in the following subsection:

2.3.3. Deviation from the Native Urdu and the Persian Loan Patterns

As noted, Persian loan nouns adopt the native genders, but they take both the native and Persian plural forms. There is often deviation from the set patterns in both the processes. Some loan nouns generally end with *-a(h)* e.g. *bətʃ**a(h)*** 'child' (m), and so with relevance to the native masculine marker *-a*, they are unarguably treated as masculine. However, some feminine loan nouns e.g. *əḍakar**a(h)*** also end with *-a(h)*. As noted above, Persian loans are also feminised following the Arabic patterns[1]. Thus, interestingly, the two Persian loans *bətʃa(h)* 'child' (m) and *əḍakara(h)* 'actress' (f) have the same origin but the gender markers are from two different sources i.e. native and Arabic.

---

[1]. The dictionary published by the **C**entre for **R**esearch in **U**rdu **L**anguage **P**rocessing, Islamabad. (CRULP)

The deviation in gender marking from both the native and the Persian rules also shows the deviation from native plural marking or vice versa. A loan feminine noun mostly takes the native plural marker -ẽ e.g. ɣɪza (h) 'meal/food'/ ɣɪzaẽ, but this is not always correct. Although rare, some feminine loans take the Persian plural marker -an e.g. ḏɤxṭəran. Two major points should be noted. First, this is a deviation from the native pluralisation of the Persian loan nouns treated as feminine. The native plural *ḏɤxṭərẽ is ungrammatical, although -ẽ attaches to most other feminine nouns. Secondly, as -an is mainly used as a masculine plural marker, its affixation with a feminine noun also shows deviation from usual plural marking. It is not an unusual phenomenon in many languages, but it is unusual in Urdu.

However the examples of -an attaching to both genders are only a few. In contrast, there are various examples of -aṭ which are very often used for both the genders. For example, nəvaḏraṭ 'valuables' is masculine, but ḏəraməḏaṭ 'imports' is feminine. It is in fact the result of the Arabic influence on both languages, due to which -aṭ generally differentiates the gender between Persian and Arabic loans. It is a masculine plural marker for Persian loans and feminine plural marker for Arabic loans.

However, many other nouns show deviation from only the usual masculine plurals, as the ones ending with -a(h), e.g. meva(h), generally take the native plural marker -e, i.e. meve, but the original Persian plural marker -dʒaṭ, e.g. mevadʒaṭ is also used, although only sometimes. Sometimes, it alternates with the Arabic loan plural marker -aṭ e.g. məqaladʒaṭ or məqalaṭ 'research papers'. This also shows variations.

Unlike nouns, the Persian loan adjectives show no gender distinction, which can also be seen in some Persian-Arabic hybrid adjectives e.g. selabzəḏa(h) 'sb flood affected' (A). Although it takes the masculine ending -a(h), the adjective is used for both the genders. The

pluralisation of an adjective is possible in Urdu, and with relevance to the ending, it takes the loan plural marker *-gan* i.e. *selabzədgan* 'flood affected (people)', although it is a hybrid adjective. The plural also represents both the genders. This is unlike native Urdu where an adjective whether singular or plural shows gender distinction. Thus, both the base adjective and the plural form show deviation from the rule. Generally, Persian loan nouns show various inflections often deviating from both the original Persian and the native Urdu. There is a case against Naim's (1999: 60) claim that Persian and Arabic plurals are never inflected in Urdu. This is to be discussed in the sections on Arabic loan morphology. All above examples show deviations from the rules of the source language and the target language.

In sum, most loan nouns follow the native Urdu morphology. Persian shows no gender distinction and the loans are treated as either masculine or feminine with the markers *-a* and *-i* respectively. However, some loan nouns, treated as feminine, also take the Arabic feminine marker *-a(h)*, which is in harmony with the masculine gender marker after the deletion of aspiration. It is thus often used as a common gender marker for both masculine e.g. *bətʃa(h)* 'child' and feminine e.g. *əḏakara(h)*, and shows no morphological difference. Many loans treated as masculine remain unchanged in plural forms, the same as native Urdu plurals. Some loan nouns nevertheless do take Persian plural markers. Two loan plural markers i.e. *-an* and *-aṭ* are important. Although *-an* is mainly related to the masculine gender, it attaches to certain feminine nouns as well. The other *-aṭ* is more frequent for both the genders. Thus, used for both the genders, these two plural markers show that a large number of loans deviate from native Urdu morphological rules. This is in fact due to the influence of the Arabic loan nouns in both languages, e.g. the Arabic feminine marker on Persian loans in Urdu.

There are a large number of Arabic loan nouns, but their gender and number marking is not necessarily by native Urdu or Arabic loan suffixes. Certain Arabic loan plurals are based on vowel changes, thus called broken plurals. They often take Urdu or Arabic plural markers.

2.4. Morphological Changes in Arabic Loan Nouns

This section discusses the gender and number morphology of Arabic loans in Urdu and their deviation from both native Urdu and Arabic patterns. Although deviation from source language form is less relevant, it may be helpful to understand the morphological changes. Citing Wegener (2004), Winter (2008: 168) terms it peripheral target language form, if there is a deviation from the rules of source language and those of target language. The adaptation is particularly interesting when Arabic loan plural markers attach to Arabic loan plurals based on vowel changes. This shows that these adaptations are more complex than the ones seen in the Persian loans. All Arabic loan nouns are used with three possibilities: 1) most loan nouns adopt the native Urdu gender and number markers. 2) both the loan singular and plural forms are borrowed and integrated, and this often occurs 3) some of the loan plurals, based on vowel changes, are also affixed with native Urdu or even loan plural markers.

The Arabic treatment of masculine gender as the basic gender already exists in Urdu, and the feminine nouns are derived forms. However, the gender treatment of a noun is not necessarily the same. For example, *kɪṯab* 'book' and *kɔrsi* 'chair' are masculine and *ʃəms* 'sun' and *buhera(h)* 'lake/sea' are feminine in Arabic, but the former loans are feminine and the latter ones are masculine in Urdu. The only Arabic loan gender marker is feminine *-a*(*h*) e.g. *ṯalɪb**a*** 'female student'. Rizvi (2007: 73) also points out that *-a*(*h*) is mostly used for feminine nouns. It is however not used in its original form. In Arabic, as Ryding (2005: 22) points out, it is presumed that a feminine noun ends with an underlying /ṯ/ e.g. *qabiilah(ṯ)* 'tribe', but in Urdu,

there is no such presumption at all. Feminine loan nouns simply end with -*a*(*h*), with even the deletion of aspiration. It generally attaches to the masculine loans ending with nasal /*m*/ or rhotic /*r*/ e.g. masculine *ʃaɪr* 'poet' becomes *ʃaɪra*(***h***) 'poetess'. The nouns are similar to the native Urdu nouns ending with rhotic /*m*/, /*r*/, but they are feminised with *-ni* e.g. *lohar* 'blacksmith' becomes *lohar**ni***. Although **-*a*(*h*)** attaches to Persian loans also, it is nevertheless limited to a few Persian and Arabic loan nouns, which are however frequent in spoken and written Urdu. Most singular loan nouns appear with the native gender markers.

However, considerable number of loan plurals are formed with the Arabic plural markers *-in* and *-aṯ*. Linguists e.g. Hafez (1996: 393) and Ryding (2005: 84) describe three forms of Arabic plurals i.e. dual plurals, sound plurals and broken plurals. Hafez (1996) states, "there are two plural inflections. One of them is suffixal and regular, traditionally termed 'sound plural' with the two realizations *-in* for masculine and *-aṯ* for feminine…The alternative is the 'broken plural' inflection, usually leading to insertion, deletion, and vowel change, and often an indication of maximal integration of a word into the language." There is no place of the Arabic dual plural in Urdu, but both the loan sound plural and the broken plurals are very common. Arabic terms of sound plural and broken plural are also irrelevant, and the use of these terms is only to differentiate the two loan plural forms. The loan sound plurals have entered Urdu with both the suffixes *-in* e.g. *mahɪrin* 'experts' (m) and *-aṯ* e.g. *ṯalɪbaṯ* 'students' (f). However, for any specific nouns, only rarely are both the gender forms borrowed. It is either masculine e.g. *mahɪrin* or feminine e.g. *ṯalɪbaṯ* is in general usage. The latter are however more common than the former. Frequently, the two loan plural forms in parallel use with the native plural forms e.g. *ṯalɪbaṯ* 'students' (f) and *ṯalɪbaẽ* respectively.

Arabic plurals are also formed by vowel change e.g. *qəlb* 'heart' is pluralised as *qʊlub*, which is also used in the same way in Urdu. However, not necessarily all forms of a loan noun are used. Besides, a loan broken plural may also undergo a nativisation process and take the native Urdu or even Arabic plural markers. The following table illustrates the patterns:

Table 2. Arabic Loan Plurals Used in Urdu

No.	Masculine Nouns	Plural	Feminine Nouns	Plural
1	*mahɪr* 'expert'	*mahɪr**in***	*ṭalɪba*(h) 'student'	*ṭalɪba**ṭ***
2	*same* 'listener'	*same**in***	*baqi* 'remaining'	*baqia**ṭ***
3	*mərz* 'disease'	*əm**r**az*	*qɪsṭ* 'instalment'	*əqsaṭ*
4	*nəfɪl* 'additional prayer'	*nə**vaf**ɪl*	*xaṭun* 'woman'	*xəvaṭin*
5	*vəkil* 'lawyer'	*vʊkl**a***	*ɔrəṭ* 'woman'	*ɔrəṭẽ*
6	*məqsəḏ* 'aim/purpose'	*məq**a**sɪḏ*	*mədʒlɪs* 'assembly'	*mədʒalɪs*
7	*qəlb* 'heart'	*qʊl**ub***	*səṭər* (*səṭər) 'line'	*səṭur*

Among the seven loan forms, the first two represent the simple plurals formed by the affixation and are termed sound plurals, and the rest are various loan patterns of the so called broken plurals. Both the two Arabic loan plural markers *-in* (m) and *-aṭ* (f) attach to the loanwords, which are basically adjectives but also used as nouns. The feminine plural marker is nevertheless more nativised and shows more morphological changes.

2.4.1. Nativisation of Arabic Sound Plurals

The feminine plurals are formed by simply affixing *-aṭ* e.g. *məsṭur**aṭ*** 'women', but they are not always so simple. Sometimes, a base noun ending with *-jəṭ* undergoes an alternation of vowels, and the schwa *-ə-* is substituted by the long *-a-* for the pluralisation:

6a. *xasɪjəṭ* 'quality' → *xasɪjaṭ,* (N)

38

b. *kefɪjəṭ* 'state/condition' → *keɪfjaṭ*   (N)

c. *ajəṭ* 'verse'   → *ajaṭ*   (N)

d) *baqi* 'remaining'   → *baqɪjaṭ* 'remains' (A)

The three feminine plurals (6a-6c) are as usual, although the singular forms e.g. *xasɪjəṭ* 'quality' have endings other than the normal loan Arabic feminine marker *-a*(*h*). Saying **xasɪjah* is ungrammatical. Ryding (2005: 22) points out the concept of underlying /ṭ/ in Arabic. There is nothing like this in Urdu. However, as the examples show, it is pronounced in some loanwords, and so it is in sharp contrast to the original forms in the source language. This marker *-aṭ* derives feminine plurals from certain loan adjectives as well e.g. *baqi* 'remaining' → *baqɪjaṭ* 'remains' (6d). It is not unusual but generally it is the function of the masculine plural marker *-in*. There are various examples of this, as shown below:

7a. *same* 'listener' (A) → *same**in*** 'listeners' (N)

 b. *hazɪr* 'present' (A) → *hazɪr**in*** 'spectators/those who are present' (N)

 c. *nazɪr* 'watcher/seer' (A/N) → *nazɪr**in*** 'audience' (N)

As examples (7a-7c) show, masculine plurals are generally formed by adjectives. An adjective being used as a noun is not native Urdu. It is only an Arabic loan form. However, it is only plural adjective being used as a noun. A singular adjective remains an adjective in Urdu, unlike that in Arabic. Moreover, plurals are only used as nouns rather than as adjectives. The process of pluralisation is simply by the suffixation of *-in* or *-aṭ*. Some bases e.g. *nazɪr* 'watcher/seer' (7c) are mainly adjectives but may be used as nouns. The gender and number morphology of adjectives is a feature of Urdu, but gender distinction in a plural

adjective is not traditional. This is also in contrast to Persian loan adjectives which show no gender distinction.

Many of the loan sound plurals are parallel to their nativised counterparts. There may or may not be changes after nativisation, but changes are mostly in feminine plurals, as shown below:

Masculine Loans	Nativised Singular Loan	Plural Form	Nativised Plural
8a) ṯalıb 'seeker/student'	none	ṯalıb**an**	ṯəlb*a*
b) nafir 'Publisher'	none	nafir**in**	nafir (unchanged)
c) mahır 'experts'	none	mahır**in**	mahır (unchanged)

Feminine Loans			
d) ṯalıba(**hṯ**) 'student'	ṯalıba(**h**)	ṯalıbaṯ	ṯalıb*a*ẽ
e) nafıra(**hṯ**) 'Publisher'	nafıra(**h**)	*nafıraṯ	nafıra*ẽ*

The nativised forms of the masculine loans generally remain unchanged, but there is a possibility of morphological changes. The masculine plural *ṯalıban* (8a) is unusual because the loan plural marker is generally *-in*. The loan may have entered Urdu via Persian, as *-an* is a Persian plural marker used for the loans treated as masculine. The nativised plural *ṯəlba*(*h*) (8a) is also in sharp contrast to the Urdu pattern, which is generally unchanged (8b & 8c). This shows its deviation from both the source language and target language forms.

The nativised forms of feminine loan nouns e.g. *nafıra*(***h***) (8e) often undergo morphological changes, but there may be certain deviations from the normal changes. Its nativised plural e.g. *nafıraẽ* is possible, although a female publisher is untraditional. Its loan plural **nafıraṯ* is

ungrammatical, but another feminine plural *ṭalɪbaṭ* 'female student' in (8d) is grammatical. This also shows that there are certain lexical exceptions in the nativisation.

In sum, most Arabic loan nouns must adopt native plural forms. The nativisation may however vary depending on the borrowed patterns. The two plural markers i.e. *-in* and *-aṭ* may attach to a large number of adjectives, but it is mainly the masculine plural marker that does the job. The nativised masculine plurals generally remain unchanged and most changes take place in the feminine nativised plurals, which are preferred to the loan ones.

The two loan plural markers are just one aspect of Arabic loan pluralisation, and they are in limited number. However, they are in daily usage. A large number of loan plurals are based on vowel changes. They are formed the same way as in Arabic depending on the modification of base. However, they are simply 'borrowed' forms. Many masculine loan nouns adopt this plural pattern, whereas feminine nouns take native plural markers.

2.4.2. Nativisation of Arabic Broken Plurals

Arabic plurals based on vowel changes are termed broken plurals. They are the most challenging plurals, as they are complicated and follow various patterns. They primarily involve internal modification of the base by vowel change e.g. *nəfs* 'soul'→ *nəfus*. Booij (2005) terms the modification of base *revowelling*. He states that the inflection in broken plural involves vowel change. Radcliffe (1990: 113) notes 27 patterns of broken plurals. Urdu has borrowed hardly some patterns. Although used the same as the original forms in the source language, broken plurals are merely borrowed forms in Urdu. Therefore, the complexity of their structure is irrelevant. Without any discussion, Moizuddin (1989: 40) terms them "examples of transplanted Arabic morphology". This pluralisation is in sharp

contrast to the native Urdu pluralisation, also used parallel to it. For example, the masculine *nəfs* 'self conscious' remains unchanged in the native form and has a parallel use.

These plurals are examples of synchronic borrowing, and are used widely, but many of them have diachronically adapted themselves. An interesting element of the adaptation is that some of the non Arabic loans e.g. Turkish *xəvaʈin* 'women' also adopt the broken plural forms. In addition, although rare, some loan nouns are normally affixed with plural markers in Arabic, but take the broken plural forms in Urdu. This is a unique feature in the whole pluralisation process, and therefore it is the focus of this section. Abdallah (2006) points out various patterns of broken plurals in Arabic from which Urdu has borrowed many forms. The following table illustrates some major patterns and their nativised Urdu counterparts:

Table 3: Major Patterns of the Loan Broken Plurals and their Native Counterparts

MasculineNoun	Loan/Nat. Pl	Feminine Noun	Loan/Native: Pl
1.dʒohər 'quality'	dʒəvahır/dʒohər	xaʈun 'woman'	xəvaʈin / *xaʈunẽ
2. hakım 'rulers'	hʊkam/ hakım		
3. mərz 'disease'	əmraz / mərz	ɣərz 'interest'	əɣraz / ɣərzẽ
4. mʊlk 'country'	məmalık / mʊlk	mənzıl'destination'	mənazıl/ mənzılẽ
5. ʃək 'doubt'	ʃʊkuk. / ʃək	rəsəm 'tradition'	i)rəsum (ii)rəsəmẽ (iii) rəsumaʈ
6. vəkil 'lawyer'	vʊkla / vəkil		

The table shows the six major patterns. For each noun given, there are at least two plurals. The one on the left is the loan broken plural, while the one on the right is the nativised form. Firstly, an interesting element to see is the Turkish feminine noun *xaʈun* 'woman' (1) that only adopts the broken plural form *xəvaʈin* rather than the native form. The other loans have nativised counterparts as well, but they have not integrated with the native Urdu plural marker

-ẽ. Therefore, the nativised plural *xaṯunẽ* is ungrammatical. This shows the possibility that some non Arabic loanwords may also be pluralised the same as Arabic loan broken plurals, whether or not they take nativised plural forms. Moreover, though the Arabic singular counterpart *ɔraṯ* 'woman' is more common, it only takes the native plural marker i.e. *ɔraṯẽ* rather than forming the broken plural **ɔraṯ*, which is ungrammatical. It is an interesting element that Urdu adopts a non Arabic noun in broken plural form rather than plural noun from the original source. Their semantics is the only difference between the two singular nouns, and the Turkish *xaṯun* expresses more formality and respect. Arabic pluralisation with a Turkish loan noun may be because Arabic has also influenced Turkish.

Evidence is also available from the Persian masculine loan noun *dʒohər* 'quality/precious stone' → *dʒəvahır* (1). Its native plural form remains unchanged. Both loan plurals are influenced by Arabic, so their formation may have occurred in their respective languages. However, the two are identical in the plural formation and have the common infix *-əva-* in combination with quite different bases coming from different sources. The same infix *-əva-* in both non Arabic loans suggests that their formation occurred within Urdu. Nevertheless, they are not a native pattern, but signal the nativisation of broken plurals as a pattern emerging.

This is further witnessed in *hakım* 'ruler' → *hɔkam* (2) in the table. It shows a broken plural, which it is not used in Arabic. Besides Turkish and Persian examples, this is concrete evidence that there is a formation of broken plurals within Urdu, although it is Arabic feature.

In sum, the patterns of the so called Arabic broken plurals are complex and the most challenging types of loans. However, some bases of sound plurals and non Arabic nouns show that broken plural as a plural form may be surfacing in Urdu. Two different bases taking the same infix *-əva-* also show the complexity. This is further explicitly seen in some loan broken

plurals taking Arabic plural markers i.e. *-in* or *-aṯ*. Thus, sometimes there may be three plurals i.e. the loan broken plural, the native plural, and the plural deviating from both forms.

2.4.3. Affixation of Arabic Plural Markers with Loan Broken Plurals

Although broken plurals are frequent, the speakers generally use an adapted plural. The loanword integration with frequent deviation from the rules is interesting. Most importantly the loan broken plurals take the loan plural markers *-in* or *-aṯ*, and so they are generally used as sound plurals e.g. *vədʒah* 'reason'→ *vədʒuh* → *vədʒuhaṯ*. This illustrates two points: 1) while functioning like native affixes, *-in* or *-aṯ* adopt the native status. 2) the loan broken plurals adapt into the peripheral nativised plurals, because *-in* and *-aṯ* are not the original native plural markers. The two tables give examples of the phenomenon:

Table 4a. Feminine Nouns

Noun	Broken Pl.	Peripheral Nativised Pl.	Native Plural
1. rəqəm 'amount'	rəqum	rəqumaṯ	rəqəmẽ
2. rəsəm 'custom'	rəsum	rəsumaṯ	rəsəmẽ
3. vədʒa(h) 'reason'	vədʒuh	vədʒuhaṯ	vədʒa(h)ẽ
4.ṯəqrıb 'ceremony'	ṯəqarib	ṯəqrıbaṯ	ṯəqrıbẽ

Table 4b. Masculine Nouns

Noun	Broken Pl.	Peripheral Nativised Pl.	Native Plural
1. rʊkən 'member'	ərkan	ərakin	rʊkən
2. məʃrıq 'east'	məʃarıq	məʃrıqen	məʃrıq
3. dʒohər 'ruby' (Persian	dʒəvahır	dʒəvahıraṯ	dʒohər
4. ləqəb	əlqab	əlqabaṯ	ləqəb
5. lazım (adj) 'necessary'	ləvazım	ləvazımaṯ	-

There are loan broken plurals, the native Urdu plurals and the peripheral nativised plurals. The Persian masculine loan *dʒohər* 'quality/ruby' (3) adopts a broken plural form *dʒəvahɪr*, which further takes the plural marker -*aṯ* i.e. *dʒəvahɪraṯ* (3). The plural marker -*aṯ* attaches to feminine loans from Arabic, but attaches to masculine loans from Persian. The affixation of -*aṯ* with the broken plural form certainly shows the nativisation of the two loan plural markers. This sort of pluralisation seems frequent both in masculine (3,4,5) and feminine nouns (1,2,3). Many adjectives, e.g. (5) in (4b), also undergo a similar adaptation. Thus, the rule deviation is not just seen in the recipient language but also in the donor language.

Some affixations e.g. *ṯəqrɪbaṯ* 'ceremonies' (f) (4) and *məʃrɪqen* 'east and west' (m) (2) occur with singular nouns rather than broken plurals. They are thus seen in two directions, one with the broken plurals and the other with their bases. Both are unusual in the source language and in the target language. This means that there are certainly complexities in the adaptation process and deviation from the set patterns which need to be discussed in detail.

2.5. Deviation from Morphological Patterns in General

In addition to the broken plurals of non Arabic loan nouns, a few distinctive factors have been noted. Some loans normally take plural markers in Arabic but they may adopt broken plural forms in Urdu, although this is rare. Moreover, Arabic loan plural markers -*in* and -*aṯ* attach to the loan broken plurals. In some cases, they attach to their singular bases instead. Further, some loan adjectives taking broken plural forms also undergo the affixation. Thus, though Arabic loans generally follow the rules of the two languages, they also often deviate from both. Consider the three loans and their plurals:

9a. hakɪm 'ruler' (m) → i) hakɪm**in** / ii) hʊkam / iii) hakɪm 'rulers/authorities' (N)

  b. xaḏɪm 'servant' (m) → i) xʊḏam / ii) xaḏm**in** / iii) xaḏɪm 'servants' (N)

c. lazım '(something) essential' → i) ləvazım / ii) ləvazm**aṯ** 'essentials'   (A)

The first two (9a & 9b) are masculine nouns and the third (9c) is an adjective, but their plurals are also treated as masculine plurals. Each of the three examples has at least two plurals of which the first is the Arabic loan and the second (and the third) is the adapted form. Although all singular forms are identical, morphologically, their plurals are different and deviate from both the language patterns. The plurals (i) and (iii) follow Arabic and native Urdu patterns respectively, but (ii) deviates from both. In (9a-i), there is normal Arabic loan sound plural, while (9a-ii) shows a broken plural form, not used in Arabic. In (9b), the pattern of pluralisation is a reversed from the one in (9a). The normal Arabic plural is broken plural (9b-i) and the peripheral nativised plural is sound plural (9b-ii). The adjective in (9c-ii) shows that its broken plural itself takes a plural marker. The newly formed plural adjective *ləvazmaṯ* 'essentials' thus shows a deviation from Arabic rules while at the same time it does not follow Urdu rules for pluralisation. There are two contrasting elements. Unlike Persian adjectives, it takes gender and number marking, which is a feature of Urdu. Like Persian masculine plurals, it takes -*aṯ* for pluralisation. All the integrations of loans show a variety of changes and deviations from the rules.

The patterns of change are evidence that there are certain rules and patterns but there is also deviation seen in almost every pattern. The examples also illustrate that a loan noun (or adjective) may have more than one plural. The adaptation may be partial and full. For example, *hʊkam* (9a-ii) and *xaḏmin* (9b-ii) are partially adapted, because the integration is not exactly according to the native pattern. The examples above show the three forms of morphological integration, as defined by Winter (2008: 158). The first plural in each is according to the non native pattern. The second is peripherally integrated plural, and the third

follows a target language pattern. Therefore, it can be claimed that Urdu morphology is not restricted to native Urdu patterns. Rather, there are elements of Persian and Arabic morphology as well. There is often amalgamation of the three morphological structures.

Sometimes, the amalgamation is to the extent that it is difficult to decide between the native and the loans. This is because many Arabic loan nouns have entered Urdu via Persian, and undergo significant morphological changes. These changes are not restricted to either of the two origins. The root word *xəbər* 'news' for example, is an Arabic loan and entered Urdu via Persian. The morphological patterns are different from those of the origins. This is compared in the modified examples from Riaz, (2007: 5):

10a. xəbər 'news'   (Urdu, Persian and Arabic root form)

   b. əxbar 'newspaper'   (Urdu and Persian, but broken plural of 'news' in Arabic)

   c. əxbar**aṯ** 'newspaper**s**'   (Urdu and Persian)

   d. xəbr**ẽ** 'news' (plural)   (native Urdu)

The tri-consonantal Arabic loan base entered Urdu via Persian and undergoes various morphological changes in Persian and then in Urdu. The Arabic broken plural *əxbar* is used as a singular noun with different semantics, 'newspaper' in Urdu and Persian. After the affixation of *-aṯ*, its plural *əxbaraṯ* is also common in both languages, but it is alien in the Arabic lexicon. The native Urdu plural *xəbrẽ* is strange in both Arabic and Persian. All the changes make it difficult to differentiate between the origins of the forms.

The morphological changes in the Arabic loans, in short, show a variety of features. Urdu has borrowed three major points i.e. the feminine gender marker *-a(h)*, the sound plural markers *-in* (m) and *-aṯ* (f) and various patterns of broken plurals. The deletion of the final [*-h*] leaves

the feminine marker *-a*, which is homophonous with the native masculine marker *-a*. However, it is limited to very few Arabic and Persian loans. The loan nouns are pluralised in the Arabic manner, but often take the native plural markers, which is the nativisation of the loans. In the nativisation process, there are two important points. Some non Arabic loans and the bases of sound plurals adopt the broken plural forms. Moreover, some Arabic loan broken plurals undergo further morphological changes and are also affixed with the Arabic sound plural markers *-in* and *-aṱ*.

All of these changes show that the loans often deviate from the native rules. However, whatever the source of the nouns, the morphological changes according to case mainly follow the native Urdu rules, although there is a possibility that a loan noun e.g. *əxbaraṱ* 'newspapers' may remain mostly unchanged in various case forms.

2.6. Morphological Changes of Nouns Correlated with Case

A grammatical word form is a form of word that may express a semantic relationship with the other constituents of a sentence. Thus, Urdu word forms *a* 'come', *ana* 'to come' and *aja* 'came' are all grammatical forms of the base verb *a* 'come', as they express semantic relationships with the other constituents (e.g. subject) in a sentence. Aronoff (1996) explains that a grammatical word is a lexeme in a particular syntactic context, where it is provided with morpho-syntactic features (like case and number). One of the word forms is the form of a noun based on case. This section discusses the singular and plural morphology of the Urdu noun in relation to case. In this regard, the morphological structures with relevance to the three cases i.e. nominative, oblique and vocative are the major focus. McGregor (1972: 1) treats vocative as a special form of the oblique case, while in his division of the Urdu nouns, Moizuddin (1989: 31) terms them the three variants.

The Urdu nominative (also called direct case) is phonologically null, as it does not bear any case marker. The oblique and vocative forms are always followed by various case markers. The two forms of a singular masculine noun are identical to nominative plural, but are different if a noun is plural. The oblique plural ends with a nasalized vowel, while vocative plural is not nasalized. There is no other difference between the two. In fact the vocative form is rare and used only in poetry.

A noun is incomprehensible unless the gender and number morphemes *-a/i/e* are affixed to it. Consider the stem form *ləɽk* 'boy/girl' that cannot be understood without gender and number markers. The difference between the three forms of the masculine noun *ləɽka* 'boy' is shown by Mohanan (1994: 80) in her table:

Table 5: Case forms of native masculine noun (ending in *-a*)

Function	Singular	Plural
Nominative	ləɽk**a** 'boy'	ləɽk**e** 'boy**s**'
Oblique	ləɽk**e** 'boy'	ləɽk**õ** 'boy**s**'
Vocative	ləɽk**e** 'boy'	ləɽk**o** 'boy**s**'

(Source. T. Mohanan)

Table 5 summarises the structural changes of a noun in relation with a case. The singular oblique and the vocative forms i.e. *ləɽke* are identical to its plural nominative form *ləɽke*. The stem forms of masculine nouns ending in *-a* are always inflected when used as non nominative arguments. This is also shown in the oblique and vocative forms of the plural. The difference between both is that of nasalization. If a masculine noun ends in something other than *-a*, no changes occur in the singular nouns with relevance to case:

Table 6: Case forms of native masculine noun (all other endings)

Function	Singular	Plural
Nominative	*nokər* 'servant'	*nokər* 'servant**s**'
Oblique	*nokər* 'servant'	*nokər***õ** 'servant**s**'
Vocative	*nokər* 'servant'	*nokər***o** 'servant**s**'

The changes in Table 6 show some differences from those seen in the previous table. The singular nominative, oblique and vocative forms are identical, if a masculine noun ends in other than -*a*, e.g. *nokər* 'servant'. The nominative plural also remains unchanged, and so it is identical to the three singular forms. However, the oblique and the vocative plurals end in -*õ* and -*o* respectively. They are identical to the ones seen in Table 5.

The three singular forms of a feminine noun are also the same. The structural changes of a feminine noun are usually more interesting and different from those of masculine nouns. Consider the feminine noun *ləɽki* 'girl':

Table 7: Case forms of native feminine noun (ending in -*i*)

Function	Singular	Plural
Nominative	ləɽk**i** 'girl'	ləɽkɪj*ã* 'girl**s**'
Oblique	ləɽk**i** 'girl'	ləɽkɪjõ 'girl**s**'
Vocative	ləɽk**i** 'girl'	ləɽkɪjo 'girl**s**'

Compare Table 7 with Table 5, as they show the changes in the masculine noun *ləɽka* and feminine *ləɽki* respectively. The feminine singular, in sharp contrast to the masculine singular, remains unchanged in its nominative, oblique and vocative forms, which are the same as those of the other noun *nokər* in Table 6. Moreover, the three forms of the plural nouns are also notable. The nominative plural is different from those of the masculine nouns in the two previous tables but interestingly the oblique and the vocative plurals are the same i.e. -*õ* and

-o. The two case forms remain the same with all nouns, which can also be witnessed in Table 8, which shows the case forms of a feminine noun with zero marking:

Table 8: Case forms of native feminine noun (all other endings)

Function	Singular	Plural
Nominative	behn 'sister'	behnẽ 'sisters'
Oblique	behn 'sister'	behnõ 'sisters'
Vocative	behn 'sister'	behno 'sisters'

The changes in Table 8 represent all other native Urdu feminine nouns with an ending different from *-i* (or *-ni*). The feminine singular noun remains unchanged in the three case forms. However, the plural nominative, oblique and vocative forms show different endings. The nominative plurals end with *-ẽ*, while the oblique and vocative plurals are the same as those of all other nouns, noted in the three tables, i.e. *-õ* and *-o* respectively. Unlike masculine nominative plurals, which are identical to the masculine oblique and vocative singulars in Tables 5 & 6, the feminine nominative plurals in Tables 7 & 8 are different from the feminine oblique and vocative singulars. This shows that the feminine plural reflects case differences but there is no structural change in the feminine singular noun based on case.

The case morphology generally shows gender distinction. Masculine nominative plural is the same as the masculine oblique and vocative singular whatever the ending is. However, the oblique and vocative forms of masculine plurals are different from each other and from those of masculine singular nouns. The case morphology of feminine singular nouns remains unchanged irrespective of any ending. However, the nominative form of feminine plurals is different from the oblique and vocative singular nouns. The oblique and vocative forms of the feminine plurals are also different from those of singular nouns.

The case morphology of both genders remains unchanged in singular forms, except that a masculine noun ends in -*a*. Moreover, if masculine and feminine nouns end in -*a* and -*i*, there is a morphological difference between the nominative plurals of and feminine nouns i.e. *ləɽke* 'boy**s**' and *ləɽkɪjã* 'girls' respectively. But there is no difference in their oblique and vocative plurals, which end in -*õ* and -*o* respectively. The two latter forms are the same in all nouns.

The case morphology of a loan noun generally follows the native morphology, but it may also adopt a case without any major changes regarding gender distinction, as noted below:

Table 9: Case forms with a Persian masculine noun

Function	Singular	Plural
Nominative	mehman 'guest'	mehman**an**
Oblique	mehman 'guest'	mehman**an**
Vocative	mehman 'guest'	mehman**an**

Table 10: Case forms with an Arabic feminine noun

Function	Singular	Plural
Nominative	səṯər 'line'	səṯur 'lines'
Oblique	səṯər 'line'	səṯur 'lines'
Vocative	səṯər 'line'	səṯur 'lines'

It is noteworthy that there is no gender distinction, and the case morphology remains in a similar pattern. The distinction of the Persian masculine noun *mehman* 'guest' and the Arabic feminine noun *səṯər* 'line' does not make any difference, and singular and plural loan nouns undergo no structural changes in the three case forms, if they do not adopt native morphology.

In short, the case morphology of a singular noun is generally the same for the three cases except that a masculine noun ends in -*a*. The oblique and vocative plurals are also the same, in

both the genders, with the difference of nasalization. However, the nominative plurals have differences depending on the endings and on the gender. Gender distinction again disappears in a loan noun. All three forms of case are the same for singular and plural nouns respectively. The loan forms may also be used, but the case morphology is mainly based on native patterns.

2.7. Summary of the Chapter

This chapter has discussed the gender and number morphology of the Urdu nouns with respect to the three sources i.e. native Urdu, Persian and Arabic. The gender markers of the native Urdu nouns are generally -*a* and -*i*. Masculine is the basic gender, and feminine gender is derived with relevance to the masculine noun. If a masculine noun ends in -*a*, the feminine form ends in -*i*. The masculine nouns other than -*a* are not many but they do exist. If a masculine noun ends in -*i*, the feminine form ends in -*ən*. Any other masculine ending takes the feminine form -*ni*. A native masculine noun ending in -*a* takes the native plural marker -*e*. All other masculine nouns are unchanged in plurals irrespective of their endings. All feminine nouns ending in -*i* are affixed with the plural marker -*jã*, and -*i* is substituted with the short vowel -*ɪ*-. All other feminine nouns take the plural marker -*ẽ*.

The Persian loan nouns have no gender marker of their origin. They are treated as masculine in most cases. However, the gender morphology of some feminine nouns is influenced by the Arabic feminine marker -*a(h)*. Persian loans take various plural markers i.e. -*an*, -*aṯ*, -*gan*, -*ha*, -*dʒaṯ*, but only the first two are frequently used with masculine nouns. The loan nouns take the plural patterns of their origin, but they are mainly adapted in the native Urdu morphology. For example, the masculine nouns ending in other than -*a* e.g. *əḍakar* 'actor/actors' do not bring plural inflection at all. This could be attributed to the conformity to

the plural system of the target language. Moreover, some Persian loans e.g. *əḓakara(h)* 'actress' (f) show the Arabic gender, but they form their plural in native manner *əḓakaraẽ*.

The entrance of Arabic feminine marker -*a*(*h*) in Urdu is due to some Arabic feminine loan nouns. However, gender treatment of most Arabic loans follows native rules. It is also noteworthy that many nouns e.g. *kɔrsi* 'chair' are masculine in Arabic, but they are feminine in Urdu. The loans have brought only the feminine gender marker -*a*(*h*), but two Arabic loan plurals i.e. the sound plurals ending in -*in* and -*aṭ* and the broken plurals based on the alternation of vowels. Although they are no more than borrowed forms in Urdu, some base nouns of the sound plurals and some non Arabic nouns show the emergence of broken plurals within Urdu. Moreover, many loan plurals deviate from both the source language and the target language patterns, -*in* and -*aṭ* attach to even loan broken plural forms.

This is also due to the fact that many Arabic loans entered Urdu via Persian. Sometimes, two similar nouns e.g. *lohar* 'black smith' and *ʃaɪr* 'poet' adopt different plural patterns e.g. *lohar* and *ʃɔra*. This shows that Urdu nouns have some gender and number rules, but there is often divergence from the rules. Although some loan gender/number patterns exist, the loans often adopt native forms. Such integration causes more than one plural forms, which sometimes follow neither native rules nor loan plural rules. So, they are peripheral target language forms.

There are variations in the gender and number formations. This is because loanwords may take native affixes or loan affixes may also sometimes attach to native bases. This is also frequent in derivational morphology, because it is also based on the three sources of words.

****

Chapter 3

Derivational Affixes in Urdu

3.1. Introduction

One of the major word formation processes is affixation. The three sources words i.e. native Urdu and Arabic also function in Urdu derivational affixes. Ahmed (2008: 2) states that Urdu has also borrowed a large number of lexical items and many derivational affixes from Persian and Arabic. Therefore, the language of origin plays a vital role in derivational morphology as well. There are various patterns from each source and many variations within each pattern. This chapter focuses on Urdu derivational affixation with reference to these three sources.

The interactions between bases and derivational affixes allow us to study derivation not only from a structural point of view but also from semantic and phonological ones. Although a discussion on structural issues is the main concern, some phonological and semantic issues with relevance to morphological structures are also included. However, phonological issues are only relevant in connection with Arabic loanword phonology, which has also shared in word formations, e.g. *ɣɒnɖa* 'scoundrel', a product of two distinctive sounds i.e. the voiced velar fricative / ɣ / from Arabic and the voiced retroflex stop /ɖ/ from native Urdu.

The derivation of most words is rule-governed, but there are also words whose affixations deviate from the rules in differing degrees. A derivation of word within a derivational pattern may be grammatical, but it may be ungrammatical for another word in the same pattern. For example, applying a rule, a noun *sətʃ* 'truth' can change into an adjective and then a noun again. But its antonym *dʒut* 'lie' does not follow the same rule, as shown below:

1a. sətʃ   'truth'   (N) →   sətʃ*a*   'truthful' (A)   → sətʃ*ai*   'truth'   (N)

 b. dʒut   'lie'   (N) →   dʒhut*a*   'liar'   (A)   → *dʒut*ai*   (N)

Note that, the noun *sətʃai* (1a) is grammatical in its second stage of derivation. But unlike it, the other noun **dʒutai* (1b) is ungrammatical. There are various similar examples, as Urdu derives a large number of words from Persian and Arabic loans. Derivation forms all major categories e.g. nouns, verbs, adjectives, adverbs and even pronouns through both prefixes and suffixes. Although there are fewer suffixes than prefixes, most words are derived by means of suffixation. There are a number of derivational patterns and many variations within each pattern. Many affixes are productive, and can attach to a variety of words to derive new formations. Some affixes are ancestral i.e. derived from latter stages of Sanskrit and Prakrit. For example, *-a* and *-va* create transitive and causative verb roots from intransitive verb roots e.g. *h̃əs* (i): *h̃əs* 'laugh' + *a* → *h̃əsa* 'make (somebody) laugh' and (ii) *h̃əs* 'laugh' + *va* → *h̃əsva* 'get (somebody) laughed'. A causative shows the possibility of a semantic expression.

Aronoff (1976) expresses his doubt about the assumption that morphemes are minimal meaning-bearing units. He states that a morpheme e.g. *-mit* in per**mit**, sub**mit**, and re**mit,** has no sense in isolation. It expresses meaning only after its affixation. Following him Kashif (2004: 91) claims that most Urdu affixes have no meaning of their own and cannot be understood in isolation. However, the same is not true of the affixes borrowed from Persian and Arabic e.g. *na* 'no' + *ehl* 'capable'→ **na***ehl* 'incapable'. There are various similar affixes. Kashif is nevertheless right about the native Urdu affixes which do not generally express semantics in isolation. The semantics is discussed only with relevance to structural issues.

Of the various affixes from each source, it is only necessary to discuss some of the most important ones in order to show the various derivational patterns. The native Urdu affixes derive most new words (even) with loanwords, and therefore the discussion starts with them.

3.2.1. Native Urdu Affixes

There are few native Urdu affixes, e.g. *-a*, *-i* and the negative markers *ə*, *ən*. As they form the base of Urdu morphology, they demand particular attention. The formers are vowel based and homonyms of native Urdu gender markers, but also derive new words with the bases of various source languages. They do not show gender morphology with loan adjectives. Derivational patterns by them seem to be systematic and rule based.

There are also some less frequent affixes. Although, in general, native Urdu affixes do not express any meaning, it is not seen in these less frequent ones, as illustrated below:

Table 1a: Native Urdu Affixes

Pref/Suff	Word	Derivation	Word	Derivation
1. *-a/-i*	*dʒʰut* 'lie' (N)	*dʒʰuta/i* 'liar' (A/N)	*mehnət* (N) 'diligence'	*mehnti* 'diligent' (A/N
2. *-i*	*xʊʃ* 'happy'(A	*xʊʃi* 'happiness' (N	*bʊra* 'bad' (A	*bʊrai* 'evil' (N)
3. *-ɪja*	*dak* 'mail' (N)	*dakɪja* 'postman' (N)	*kəbaɽ* (N) 'wreckage'	*kəbaɽɪja* 'ragpicker' (N)
4. *ə*	*mər* 'die' (V)	*əmər* (A) 'immortal'	*tut* (V) 'break'	*ətut* (A) 'unbreakable'
5. *ən* 'not'	*pəɽh* 'read' (V)	*ənpəɽh* illiterate' (N)	*dʒan* 'know' (V)	*əndʒan* 'stranger' (N)
6. *tʃo* 'four'	*paja* 'foot' (N)	*tʃopaja* 'four footed 'animal' (N	*rah* 'path' (N)	*tʃoraha* 'crossroad' (N

The suffixes *-a* and *-i* are quite distinctive even among the most productive affixes. Plag (1999) defines morphological productivity as "the property of a given word formation process to be used to derive a new word in a systematic fashion." This systemaic fashion is observed in these affixes, and so they are quite productive to derive new words. They attach to nouns and mainly derive adjectives. The derivation expresses gender, if the base noun is native. If the base noun is a loan, only *-i* attaches, and the derived adjective expresses no gender.

2. N + *-a*/-*i* → A/(N)   (Gender distinction with native noun)

   a. sətʃ 'truth'   (native) → sətʃ**a** (m) / sətʃ**i** (f)   'true/ truthful person' (A/N)

   b. dʒʱut 'lie'   (native) → dʒʱut**a** (m) / dʒʱut**i** (f)   'lying/liar' (A/N)

   c. mehnət̪ 'hard work'   (loan) → mehnt̪**i** (N/A) 'diligent'

   d. dʒənnət̪ 'heaven'   (loan) → dʒənnət̪**i** (N/A) 'the one rewarded with Heaven'

The nouns (2a & 2b) take either *-a* or *-i* and show gender marking. The newly derived words are typically adjectives e.g. *sətʃa kam* 'true work' (2a) and *dʒʱuti xat̪un* 'a lying woman' (2b). But they may also be used as nouns e.g. *ek dʒʱuta* 'a liar'. The ones in (2c & 2d) are loans and so take only *-i* without specifying any gender. Both the derived adjectives can be interpreted as either masculine or feminine, although the borrowed nouns i.e. *mehnət̪* 'hard work' and *dʒənnət̪* 'heaven' are treated as feminine. Most of the base nouns that *-i* attaches to, are common nouns and abstract nouns and are treated as feminine. The native and the loan nouns differ in that the derivations from the native nouns are mainly adjectives but may also be used as nouns. The derivations from the loans are only adjectives and can rarely be used as nouns.

Morphology sometimes interacts with phonology. In the process of derivation, there may also be some minor structural changes. For example, *-ə-* is deleted and the original coda becomes

an onset after the change of category from *mehnət* to *mehnti* (2c). This is because the final-*i* allows the -*ṭ*- to function as an onset, which it cannot do when word-final. This is also because the nasal /*n*/ is pre-aspirated [ʰn] without requiring any vowel. The structural changes illustrate what was said earlier that the affixations sometimes deviate from expected patterns of morphology, which is seen in almost every pattern of word formation.

The suffix -*i* is more comprehensive in its role. It also transforms a large number of adjectives into nouns, which is quite a contrast to the previous function. It often attaches to a native or loan adjective without any distinction. Consider the examples below:

3. A + - *i* → N

   a. xʊʃ 'happy'   (A) → xʊʃ**i** 'happiness' (N)   (Persian)

   b. mehrum 'deprived' (A) → mehrum**i** 'deprivation' (N)   (Arabic)

   c. bʊr*a* 'bad'   (A) → bʊr*a***i** 'evil' (N)   (Native)

   d. ətʃ*a* 'good'   (A) → ətʃ*a***i** 'goodness' (N)   (Native)

   e. bhuk 'hunger' → bhuk*a* 'hungry'→ *bhuk*a***i**   (Native)

Note that the derivation from adjective to noun is possible with both native and loan bases. If the base is a loanword, the suffixation of -*i* is enough without differentiating the source language, as in (3a & 3b). But a native adjective shows gender marking, and so -*i* is only possible with masculine bases, as in (3c & 3d). The affixation of -*i* with a feminine base means doubling the final -*i* e.g. *bʊri**i** and so it is ungrammatical, although a word-final long vowel may be accepted in other instances. This shows that there are two suffixes -*i*, with different effects and that they both have the same pronunciation is in a way accidental. However, the affixation of -*i* to derive noun from adjective is possible only if the base is

originally an adjective. It is generallay ungrammatical, if the base is already derived from a noun as in (3e), although there may be exceptional cases.

This discussion has shown that *-i* is a productive suffix, and derives both nouns and adjectives. It derives adjectives, from native nouns, showing feminine gender. It is thus parallel to *-a*, which derives adjectives showing masculine gender. It sometimes requires the base noun to undergo structural changes in the coda. Bauer (2003: 70) observes, "a morphological process can be said to be more or less productive according to the number of new words which it is used to form." As the two affixes can conveniently aattach to native and loan bases, it can be said that both *-a* and *-i* are productive affixes.

A parallel of *-a* is *-ɪja*, which derives masculine agentive nouns and has no feminine counterpart. This thing distinguishes it from *-a*. Moreover, it is very limited in word formation as compared to *-a* and *-i*, which seem to derive most words. Unlike *-a*, it is distinctive in that it may derive new words with loan bases as well. The examples below illustrate this:

4. N + *-ɪja* → N

 a. dak 'mail' → dak**ɪj*a*** 'postman'

 b. kəbaɽ 'wreckage' → kəbaɽ**ɪj*a*** 'wreckage buyer'

 c. fr*a*d 'fraud' → fr*a*d**ɪj*a*** 'cheater'

The suffix *-ɪja* derives only agentive masculine nouns from nouns e.g. *dakɪja* 'postman' in (4a). The derivations are very small in number and have no feminine counterpart, but they are the same in form as some feminine nouns e.g. *kɔʈʈɪja* 'bitch'. As discussed in the previous chapter, the feminine nouns are base nouns and not derivations. Therefore, the two are different. The affixation of *-ɪja* with loanwords e.g. in (4c) is also possible. It is

ungrammatical to attach a feminine marker e.g. *dakɪjai* or treat the masculine derivation also as feminine. Thus, it also shows irregularity and it becomes evident that Urdu morphological patterns are complex and often seem to deviate from the rules, as all the three suffixes also show deviations from their patterns in one way or the other.

There are no other important suffixes whereas among the prefixes the negative markers ə and ən are noteworthy. Although not very productive, they do form some new words. Both ə and ən seem to function in exactly the same way. They mainly derive adjectives from verbs and the derivation is an antonym variant i.e. it is from a different category. The derivation from ən may also be used as an adverb. Consider the following examples:

5. ə + V (base form)   →   A(dj)

a. **ə** + mər 'die'   →   əmər 'immortal' (A)

b. **ə** + tut 'break'   →   ətut 'unbreakable' (A)

6. ən + V (base form)   →   A (may be used as Adv or Noun)

a. ən + pəɽh 'read/study'   →   **ən**pəɽh 'illiterate/uneducated' (A/N)

b. ən + ʈʰək 'exhaust'   →   **ən**ʈʰək 'untiring/untiringly' (A/Adv)

Both ə and ən attach to verbs but are different in derivation. They form deverbal adjectives, but the formations with the latter may be used as nouns or adverbs as in (6). Both are restricted to some specific root verbs and cannot replace each other e.g. *əʈʰək* or ***ən**mər* are ungrammatical. Infinitives like *mərna* 'to die' are not their bases.

In sum, the main native Urdu affixes -*a*, -*i*, -*ɪja* and the prefixal negative markers ə, ən are only few, but they derive various new category words. Although the prefixes playing the role of negative markers are important, the suffixes have the key role in the derivation. They attach

to nouns and adjectives and derive all major category words. These affixes are fewer than loan counterparts, but they derive nouns and adjectives from Persian and Arabic loan bases also. Gender morphology with loan adjectives is irrelevant. This means that there may be various exceptions particularly in case of loanwords. The irregularity is nevertheless not in the sense of being 'unpredictable'. The affixes certainly show some rules, patterns and a system.

It is also sensible that the major native affixes do not show semantics in isolation, although not about the less frequent ones e.g. *tʃo* 'four'. Moreover, Persian and Arabic loan affixes also deny this. Persian affixes are the most of all and generally express semantics in isolation.

3.3. Persian Loan Affixes

Naim (1999) states that most of the commonly used words in Urdu are of Persian and Arabic origin, otherwise these languages have exerted little or no influence on grammatical structure. He terms these words Perso-Arabic, as many Arabic loans entered via Persian, which makes extensive word building, by derivational agglutination. There is a large number of Persian loan affixes. They are productive in that they may attach to words other than those of Persian itself. However, they derive fewer words than do the native affixes. Suffixes predominate in Persian itself, but prefixes are the dominant loans in Urdu. The majority are negative markers and express from absolute to partial negation. Some negatives e.g. *na* 'no/not' derive mostly adjectives from nouns and nouns from nouns. Although suffixes are not very frequent, some e.g. *-əgi*, *-ana* and *-amez* interestingly derive more words with non Persian bases. Thus, they show complete integration, and seem more native affixes than Persian loans.

This section discusses only a few loan affixes, giving a glimpse into their integration with words from various sources. Some parallel examples with non Persian bases are given below:

Table 2: Persian Loan Affixes in Urdu

Prefix/Suffix	Word	Derivation	Word	Derivation
1. na 'no/not'	bina 'clear-sighted' (N/A	**na**bina 'blind' (N	mərḍ 'man'(N	**na**mərḍ 'impotent man' (N)
2.be 'without	qərar (N) 'rest'	**be**qərar (A) 'in panic'	hısab (N) calculation'	**be**hısab (A) 'countless'
3. kədʒ 'little'	fehem 'sense' (N)	**kədʒ**fehem 'credulous' (A)	rəftar 'speed' (N)	**kədʒ**rəftar 'slow' (A)
4. kəm 'little'	ılm (N) knowledge	**kəm**ılm 'less educated' (A)	əqəl 'wit' (N)	'**kəm**əqəl' 'stupid' (A)
5. pəst 'little/mean'	hımmət 'courage' (N)	**pəst** hımmət 'coward' (A)	fıtrət 'nature' (N)	**pəst**fıtrət 'mean nature' (A)
6. nɔ 'new'	dʒəvan (N) 'of mature age'	**nɔ**dʒəvan (N) 'young man'	ḍolət (n) 'wealth'	**nɔ**ḍolətia (N) 'newly rich'
7. pʊr 'full of'	dʒoʃ 'excitment' (N)	**pʊr**dʒoʃ (A) 'excited'	məɣəz (N) 'mind/brain'	**pʊr**məɣəz (A 'thoughtful'
8. bala 'above/up'	xana (N) 'house'	**bala** xana (N) 'upstairs'	ḍəst (N) 'hand'	**bala**ḍəst (Adj) 'strong'
9. ba 'having'	ənḍaza (N) guess	**ba**ənḍaza (Adv) with guess	vədʒuḍ (N) existence	**ba**vədʒuḍ (Adv) 'despite'
10. zuḍ (sth that does sth)	əsər 'effect' (N)	**zuḍ**əsər 'effective' (A)	həzəm (N) 'digestion'	**zuḍ**həzəm 'digestive' (A)

11. -(ə)gi	zɪnḑa 'alive' (A)	zɪnḑəgi 'life' (N)	hævan 'animal' (N)	hævangi 'animal' (N)
12. -ana	mərḑ 'man' (N)	mərḑana 'manly'(A/Adv)	bətʃa 'child' (N)	bətʃkana 'childish' (A)
13.-amez 'full of'	ʃɪkajəṭ'complaint' (N)	ʃɪkajəṭamez 'plaintive' (A)	rəŋ 'colour' (N)	rəŋ amez 'dyed' (A)
14. -ʃʊḑa 'done'	gʊm 'lose' (V)	gʊmʃʊḑa 'the lost one' (A)	xʊʃk 'dry'(A)	xʊʃkʃʊḑa 'dried' (A)
15. -ḑar 'showing'	ɪzzəṭ (N) 'respect'	ɪzzəṭdar (A)'respectable'	vəfa (N) 'loyalty'	vəfaḑar (A) 'loyal'
16. -ɪsṭan 'place/land'	qəbər 'grave' (N)	qəbrʊsṭan 'graveyard' (N)	pak (n) 'pure'	Pakɪsṭan 'Pakistan' (N)

Of the sixteen affixes, ten are prefixes (1-10) of which five (1-5) are negative markers. The first two *na* 'no/not', and *be* 'without' express absolute negation and the others i.e. *kəd*ʒ 'little', *kəm* 'little' and *pəsṭ* 'little/mean' express partial negation. In general, *na* and *kəm* are the most frequent both in spoken and written Urdu.

Used as a noun in rare cases, *na* has a variety of functions as a negative marker. Functionally, it is the same as the native negatives *ə* and *ən* and Arabic *la*, but there are major structural differences. Unlike native *ə* and *ən* which attach to verbs only and yield mainly adjectives, Persian *na* attaches to adjectives, nouns and even verbs and derives adjectives and nouns (*na* + A/N/(V) → A/N). Moreover, its affixation with nouns may not necessarily change the

category of the derivation. For example, it derives a noun **na**aʃna 'stranger' with a noun aʃna 'familiar person', but its affixation with another noun e.g. ʊmiḏ 'hope' forms an adjective i.e. **na**ʊmiḏ 'hopeless'. Such patterns of changes with the same category of words in two different directions justify irregularity. There are nevertheless some observations. When *na* affixes animate common nouns e.g. aʃna 'familiar person', it doesn't change the category. When it interlinks with abstract nouns e.g. ʊmiḏ 'hope', it transforms the base nouns into adjectives.

Although a few examples are found, the affixation of *na* is possible with Arabic loan nouns e.g. **na**mehrəm '(for woman) a non blood relative with whom marriage is allowed' and adjectives e.g. **na**maqul 'inept'. It does not change the category but shows some complex formations. As seen in the examples, some formations show no changes in the bases, but in some others the affixation of *na* requires the deletion of final vowel in the stems. Some base adjectives, e.g. raz**i** 'consented', are prefixed with *na* to derive an antonym **na**raz 'angry', which must lose the final vowel. The derivation is ungrammatical ***na**raz**i** 'angry', if the deletion does not occur. The derived words keep two syllables *na.raz*, and the third *-i* is deleted. Nonetheless, this is not the general feature in the affixation of *na* with Arabic loans. It only shows complexities in the affixations.

Unlike the Persian and Arabic bases, the native Urdu stems are base verb froms e.g. *səməd͡ʒh* 'understand', although there are very few of them. The formations with them are nouns or adjectives (*na* + V→ A/N) e.g. **na**səməd͡ʒh 'credulous'. However, the use of *na* in this formation type is restricted to few derivations, because most verbs are affixed with native Urdu native negative markers *ə* and *ən*, which convert the verbs into nouns and adjectives. Nevertheless, the affixation of *na* strengthens the perception that some patterns, in each case, are complex in rules and often show deviations.

The alternations shown by *na* are in the majority of all Persian loan prefixes. They seem even more complex by some suffixes, which cause morphological changes in the bases. Although most suffixes derive adjectives, some derivations may be nouns. *-gi* needs particular attention, as it attaches to both adjectives and nouns but only derives nouns. Non Persian base nouns sometimes show no morphological changes. Consider the following:

8a. zɪnd̪*a*(h) 'alive'   + -gi → zɪnd̪**əgi** 'life'   (A + -gi → N) (Persian)

 b. d̪ərɪnd̪*a*(h) 'beast'   + -gi → d̪ərɪnd̪*ə*gi 'beastliness'   (N + -gi → N) (Persian)

 c. ʃərmɪnd̪*a*(h) 'shameful' + -gi → ʃərmɪnd̪*ə*gi 'shamefulness'   (A + -gi → N)   (Arabic)

 d. hev*a*n 'animal'   + -gi → hev*a*ngi 'wildness/beastliness' (N + -gi → N) (Arabic)

The four examples show the affixation of *-gi* with both nouns and adjectives, but the derivations are only nouns. The basewords in (8a-8b) are Persian loans, but those in (8c-8d) are Arabic loans. The affixation generally causes some phonological alternations in the basewords, although (8d) shows no change. The coda vowel *-a(h)* is substituted by the schwa *-ə-* along with the affixation of *-gi* for the derivation. The Persian loan suffix *-gi* is a non neutral suffix, according to Katamba's (1993: 89) definition, as it causes structural changes in the adjectives and nouns ending with *-a*(*h*). Merely the affixation of *-gi* to Persian or Arabic bases is not enough to transform them into nouns. Therefore, *zɪnd̪*agi* 'life' and *ʃərmɪnd̪*agi* 'shamefulness' are ungrammatical. The final vowel must be replaced with *-ə-*, and then the morpheme **-**gi is attached to the basewords, e.g. zɪnd̪*ə*gi 'life' (8a) or ʃərmɪnd̪*ə*gi (8c). This shows no distinction of the source language, and the derivation refers to the nativisation.

The affixation rule applies the same way. However, it may cause morphological differences in the derived nouns, as seen in Arabic *hevangi* (8d) and the Persian *d̪ərɪnd̪əgi* (8d). Like the adjective, the Arabic noun is not itself a derived form. Moreover, it does not end with *-a*(*h*),

and so its closed coda only takes the suffix -*gi*. Further, -*gi* also shows that Arabic loans are the convenient bases for the Persian affixes which can attach to the bases of a variety of source languages. Ending with -ḍa(*h*) is the major cause of the changes by the Persian -*gi*. The affixation with two base endings also refers to the previous argument that although Urdu derivations follow some patterns they may be complex and may cut across linguistic components. The affixes are thus completely integrated and nativised.

Some other suffixes e.g. -*ana*, -*amez* and -ḍar also show some interesting nativisation, because they mostly attach to non Persian bases. Among them, -*ana* is particularly flexible in the affixation and can freely attach to a base of any of the major source languages discussed. The following example illustrates the process properly:

9a. mərḍ 'man'   + -*ana* → mərḍ**ana**   'manly'   (N + -*ana* → A) (Persian)

 b. radʒpuṭ (a surname) + -*ana* → radʒpuṭ**ana** 'like a radʒpuṭ' (N + -*ana* → A) (Native)

 c. zalım 'tyrant'   + -*ana* → zalım**ana** 'tyrant'   (A/N + -*ana* → A) (Arabic)

 d. fəlsəfı 'philosopher' + -*ana* → fəlsəfıj**ana** 'philosophical'   (N + -*ana* → A) (English)

The four derivations by Persian -*ana* show its affixation with the base nouns of four source languages. The affixation with a Persian base (9a) is not an unusual matter. Although on rare occasions it can also attach to native Urdu bases as in (9b). This is also expected to be a result of the diachronic process. However, its derivations with Arabic loans (9c) and particularly English loans (9d) are surprising and show the nativisation of Persian affixes.

Despite being large in number and having the capacity to attach to a variety of bases, Persian loan affixes derive fewer words than those from native ones. There are more loan prefixes than suffixes, and most of them are negative markers. Some of the suffixes e.g. -*ana* show

various derivations with even non Persian bases. This again refers to the complexities in Urdu derivational morphology. Although the derivation follows some patterns and rules, the rules are complex and may cut across linguistic components. Therefore, Naim (1999) rightly claims that there is frequent derivational agglutination and the formations deviate from Urdu and Persian patterns. The loan affixes are so completely integrated that they seem more native affixes than Persian loans. Various such derivational patterns with many variations are seen.

This is also because some Persian loan affixes easily attach to Arabic loan bases. A large number of Arabic loanwords can be seen in Urdu though Arabic affixes are the rarest of the three source languages. Moreover, the interaction of morphology and phonology seen in (8) may be an exceptional case in Persian loans, whilst it is often seen in Arabic loans.

3.4. Derivation in Arabic Loans by Affixation and by Phonological Changes

Sound change in loanwords is a normal process. It becomes interesting, when it affects the formation of new words. The derivational morphology of Arabic loans in Urdu is not restricted to affixation and compounding. Rather, it shows that the modification of base by internal vowel change, e.g. a noun *qətəl* 'murder', can change into another noun *qaʈɪl* 'killer'. Thus, there is often an interaction of morphology and phonology. Affixation also shows this interaction many times. For example, when used word medially, the Arabic loan definite article *ʊl/əl* follows the source language rule, and the lateral /l/ is deleted, which results in the doubling of the following consonant e.g. *kɪʈabʊ*(*l*)***nn***ur* 'the book of light'. However, the Arabic rule is discarded word initially e.g. *əlnur* 'the light'. Kaisse (2005) concludes that "the relation between word formation and phonology is complex." Arabic loanword derivation also shows various complex features in the interaction between morphology and phonology. Although the interaction is not the issue here, and the main concern is morphology, Arabic

loanwords often show this interaction. So, it becomes necessary to see in the end of this chapter some major features of Arabic loanword phonology. Some key interesting features of this include rise and fall of Arabic phonemes.

There are four major points discussed in this chapter: 1) derivation by Arabic loan affixes 2) derivation by the modification of Arabic loan bases (by phonological changes in them) 3) the nativisation of loan morphemes in a manner that they affix even with many words derived by phonological changes and 4) loanword phonology. This section includes some discussion on the neutralisation of many Arabic loan phonemes (losing their phonetic value) and their rise and reversal from native Urdu words.

An interaction between morphology and phonology may cause inflectional and derivational changes. Derivational changes can be in two ways i.e. by affixation and by modification of base. Some affixes, e.g. the Arabic loan definite article *əl/ɒl*, also show this interaction.

3.4.1.1. Derivation by Affixation

Although there are a large number of Arabic loanwords, there are not as many derivational affixes, as those of Persian. But they are used very frequently. Most of them are negative markers. Structurally and functionally, the negatives are different from native and Persian negative markers. They are also different from each other. For example, *la* attaches to nouns and forms both nouns and adjectives. It refers to absolute negation. *bɪla* also attaches to nouns, but it forms adverbs. It refers to a lack of something. Among other loan affixes are the adverbial suffix -*ən* and the definite article *əl/ɒl*, which is structurally similar to its appearance in Arabic but functionally different in Urdu. This section describes only the structures of a few affixes. It shows in particular that *əl/ɒl* connects two independent nouns into a single

whole. It also partly supports Naim's (1999) claim that *əl/ʊl* is mostly found in nominal and adjectival compounds, e.g. *ḏarʊ̱lʊlum* 'educational institute' in which the second constituent is made definite. Consider the table below illustrating some important affixes:

Table 3: Arabic Loan Affixes in Urdu

Pref/Suff	Word	Derivation	Word	Derivation
1. la 'not'	sani 'second' (A)	**la**sani (A) 'incomparable'	ʃəur 'consciousness'	**la**ʃəur (N) 'unconsciousness'
2.bɪla 'without'	zərurəṯ (N) 'necessity'	**bɪla**zərurəṯ Adv) 'unnecessarily'	dʒəvaz (N) 'reason'	**bɪla**dʒəvaz (Adv) 'unreasonably'
3. ɣær 'without'	zəruri (A) 'necessary'	**ɣær**zəruri (A) '**un**necessary'	hazɪr (A) 'present'	**ɣær**hazɪr (A) 'absent'
4. -ən	ɪṯṯfaq (N) 'coincidence'	ɪṯṯfa**qən** (Adv) 'coincidentally'	əmanəṯ 'trust'	əmanəṯ**ən** (Adv) 'by trust'
5.zi 'having	ʃəur 'wit' (N)	**zi**ʃəur 'wise' (N)	ruh (N) 'soul'	**zi**ruh 'animate' (A
6. əl/ʊl	Nurʊ(**l**) ḏḏin			

There are three loan negative markers i.e. *la* 'no/not', *bɪla* 'without' and *ɣær* 'stranger/without', the definite article *əl/ʊl* and a few other suffixes e.g. adverbial -*ən*. The negative *ɣær* is among the most frequent affixes in spoken and written Urdu. It derives adjectives from adjectives and nouns from nouns, as shown below:

10a. zəruri 'necessary' (A) → **ɣær**zəruri '**un**necessary' (A)

  b. mərḏ 'man' (N) → **ɣær**mərḏ '(for woman in a religious sense) non relative man' (N)

In deriving adjectives from adjectives and nouns from nouns, *ɣær* shows some interesting structural and semantic features in comparison with the Persian *na*. Unlike *na*, it forms only antonyms. Like *na*, it can also construct a noun from a noun, but there are not many examples. Both can attach to Persian or Arabic nouns but their affixation with native nouns is rarely seen. In rare cases, both have interchangeability e.g. **ɣær**məsavi/ **na**məsavi 'unbalanced'. Sometimes, the interchanging may also cause semantic change e.g. in two synonyms i.e. *aḏmi* 'man' (from Hebrew via Persian), and *mərḏ* 'man' (Persian).

11a. *a*ḏmi 'human/man' (N)   → **ɣær**aḏmi 'stranger' (N)

  b. mərḏ 'man'   (N)   → **ɣær**mərḏ '(for woman in a religious sense) non relative' (N)

  c. *a*ḏmi 'human/man' (N)   → *****na**-*a*ḏmi 'stranger' (N)

  d. mərḏ 'man'   (N)   → **na**mərḏ 'impotent' (A/N)

As (11a & 11b) show this, *ɣær* attaches to both *aḏmi* and *mərḏ* and forms antonyms with two different meanings. *na* is ungrammatical with *aḏmi* as in (11c) and affixing with *mərḏ* 'man', it shows a quite a different meaning as in (11d). Comparing (11b) and (11d) shows the interesting combination of two negative markers with the same noun. The first derives a noun, as in (11b), and the other an adjective that is also used as a noun, as in (11d).

The affixation of both with the three nouns i.e. *məsavi* 'balanced', *mərḏ* 'man' and *aḏmi* 'human/man' shows three features. First, they may be interchangeable with no semantic change, e.g. **ɣær**məsavi/**na**məsavi. Second, their interchanging may also cause semantic changes, e.g. (11b & 11d), and third their interchange ability is limited and not possible in every case, e.g. (11a & 11c). Although both function as native negative markers and share few factors in common, the differences are dominant. The three different results of the same process show that there are some rules, assumptions and possibilities while at the same time

there are frequent deviations from the rules. These deviations exist because both negative markers are from two different sources. As noted in the previous section, *na* is wide and comprehensive. It attaches to nouns, adjectives and some verbs. Moreover, it forms both nouns and adjectives. By contrast, *ɣær* is limited. It mainly derives adjectives from adjectives, but sometimes it derives adjectives also from nouns, whether bases may be Persian.

Like *ɣær*, other Arabic negative markers *la* 'no/not' and *bɪla* 'without' can also attach to Persian and native bases. However, they are not interchangeable with the native or Persian loan affixes. They are comparatively more restricted in use, but they show structural variations. *la* is more frequent, and attaches to nouns and adjectives and derives antonyms:

12a. ʃəur 'consciousness' (N)   (Arabic) → **la**ʃəur 'unconsciousness'   (N)

  b. sani 'second'   (A)   (Arabic) → **la**sani 'incomparable'   (A)

  c. pərva 'care'   (N)   (Persian) → **la**pərva 'careless'   (A)

  d. zərurət̪ 'need'   (N)   (Arabic) → **bɪla**zərurət̪ 'unnecessarily'   (Adv)

  e. tʃuk 'mistake'   (N)   (Native) → **bɪla**tʃuk 'unmistakingly'   (Adv)

  f. tʃ̃ũ-v-tʃəurr*a* 'objection' (N)   (Persian) → **bɪla**tʃ̃ũ-v-tʃəurr*a* 'without objection' (Adv)

la attaches to both nouns and adjectives and forms the same categories as in (12a) and (12b). But (12c) shows that the derived category may be different from that of bases, although in some cases. For example, the Persian base *pərva* is noun but the derivation *la**pərva** is adjective. As seen, *la* can affix with Persian or native Urdu bases as well. In general, it shows fewer word formations than *ɣær*. It can also be observed in *bɪla* 'without' (12d), which is the least frequent in word formations in all negative markers. It attaches to a noun but forms an adverb (12d). Like some other Arabic negatives, it can also conveniently attach to native and Persian bases e.g. ***bɪla**tʃuk* (12e) and ***bɪla**tʃ̃ũ-v-tʃəurra* (12f) respectively.

Thus, generally seen, Arabic negative markers are prefixes and function like native affixes. They are only a few and restricted in function. However, they show a variety of affixation with native and Persian bases. Moreover, like other negatives, they show some interchange ability with Persian affixes.

Other than negative markers, Arabic loan affixes also include the definite article *əl/ʊl*. Used mostly as a prefix or an infix, it has mostly lost its original status. It occurs mainly at a prefix position, and may attach to noun and adjective. As an infix, it can connect two nouns to form a proper noun e.g. *ʃəmʃʊ(l)ḏḏin* 'the light of the religion'. The two constituents *ʃəmʃ* 'sun' and *ḏin* 'light' may seem to be two separate nouns, but the proper noun functions as a single unit by virtue of some morpho-phonological features shown by *ʊl*. The lateral */l/* is deleted and the following consonant is doubled. As a result, the so called definite article becomes a binding force that connects the two separate nouns into a single whole, which is only taken as single name, not a compound. There is no first or second name in it, which means that both the definite article and the second constituent '*ḏin*', have been completely incorporated. With such names as this, a pserson can be addressed as *ʃəmʃ* but not as **ʊ(l)ḏḏin*, which is ungrammatical. This shows that *əl/ʊl* has lost its status as a definite article *ʊl*. However, this is partially true, and *əl/ʊl* is found in nominal and adjectival compounds e.g. *ḏarʊlʊlum* 'educational institute', in which the second constituent is made definite. The semantic role of the definite article is not the issue of concern here. As an infix, it may form proper nouns as single units or compounds. It has no suffixing role.

There are nevertheless some important loan suffixes, e.g. the adverbial *-ən*. It is is distinctive among loan affixes in that it does not convey any meaning at all. It derives adverbs from nouns e.g. *iṭṭfaqən* 'coincidentally'. But its affixation with nouns is restricted to mainly

dental, nasal and uvular stops at coda position i.e. /t̪/, /d̪/, /n/, /m/ and /q/ respectively. It

cannot attach to any noun with an open coda, as shown below:

13a. ɪrad̪a(h) 'intention' + **ən** → *ɪrad̪aən

   b. ɪrad̪a(h) 'intention' + **ən** → *ɪrad̪ən

   c. ɪrad̪a(h) 'intention' + **ən** → ɪrad̪at̪ən 'intentionally'

Example (13) exhibits morphophonological harmony of -*ən* with the base nouns. Many base

nouns e.g. *ɪrad̪a(h)* 'intention' end with an open coda, which is the result of the deletion of

final aspiration in the noun *ɪrad̪a(h)*. As noted in chapter 2, the deletion of the final aspiration

is a common feature in Arabic loans. However, in this case, only affixation of -*ən* causes

ungrammaticality, as shown in (13a). Its affixation requires dental, nasal and uvular stops at

coda position, as said above. There are two options here then, either the deletion of final

vowel -*a* or its affixation along with one of the consonant required. In the former case, the

derivation is ungrammatical again as shown in (13b). It in fact requires an insertion of

voiceless dental /t̪/, which becomes necessary along with its affixation. Why a voiceless

dental /t̪/ should necessarily be inserted is only because many Arabic nouns treated as

feminine originally end with an underlying voiceless dental /t̪/, but in Urdu the dental sound is

not pronounced at all. This is quite contrasting situation. The dental /t̪/ and the aspiration /h/

are deleted and not pronounced in many nouns, but in the transformation of noun to adverb

the dental is reinserted after the affixation of the adverbial -*ən*, which causes the reinsertion

and the modification of base noun.

This formation of adverbs by -*ən* with a noun is parallel to compounding by

ablative/instrumental marker *se* with a noun e.g. *ɪt̪t̪ɪfaq* **se** 'coincidentally'. Although both can

sometimes interchange with each other, there is a major difference in that *-ən* is limited in deriving adverbs, and it attaches to mostly nouns with certain coda consonants:

14a. ɪt̪t̪ɪfa**q** 'chance'   + **ən** → ɪt̪t̪ɪfa**qən** 'coincidentally' /ɪt̪t̪ɪfaq **se** 'coincidentally'

  b. dʒəld̪i 'quickness' + **ən** → ***** dʒəld̪i**ən**   / dʒəld̪i 'quickness' + se → dʒəld̪i **se**

  c. əmanət̪ 'trust'   + **ən** → əmanət̪**ən** 'with/by trust'/*əmanət̪ **se**

Both the instrumental *se* and adverbial suffix *-ən* can sometimes interchange, and both can substitute for each other, as in (14a). But it is not always possible, as shown the ungrammaticality of *-ən* with an open coda in (14b). There are also certain places where the affixation of *-ən* is possible but that of the ablative *se* is ungrammatical, as in (14c), due to a semantic constraint. *əmanət̪ se* is morphologically correct, but there is no semantic expression. *əmanət̪* is culturally taken as something kept (for a certain period) by somebody to return it to its owner at a proper time. As *se* describes only manner of action, it cannot give the correct semantic expression, while *-ən* can. Thus, *əmanət̪ən* means something kept as a trust to be returned to its owner. In general it can be seen that *-ən* is sometimes interchangeable with *se*, but at other times one of them shows ungrammaticality.

In every pattern, Urdu affixation shows some rules, though there are also deviations from the rules. As seen, *-ən* cannot affix with an open coda base but this rule is irrelevant if a noun ends with *-a*(*h*). Although the final aspiration is deleted, it allows dental /t̪/ along with the affixation. This shows that there sometimes its affixation also requires modification of base. A large number of Arabic loans do not require affixes for derivation. Rather, they undergo some modification in their base form.

3.4.1.2. Derivation by Modification of Base (Phonological Changes)

The Arabic loan affixes discussed derive only a limited number of formations. Most loanword derivations are based on the modification of base. Bauer (2003) states that where affixes are not used for creating new words, the most widely attested method is to make some kind of phonological change to base. This involves discontinuous affixes and discontinuous bases.

Arabic derives various categories of words by making phonological changes in tri consonantal bases e.g. *zlm*. The derivations are however no more than loans, as the process of deriving paradigms from trilateral roots does not occur in Urdu. In addition, the loan derivations are restricted to some patterns and each loan pattern contains only a few derived words. They are used in the same way as the original forms. For example, a noun e.g. *zɒlm* 'tyranny' can become an adjective i.e. *zalɪm* 'tyrant' or *məzlʊm* 'tyrannised'. The three forms are known and commonly used, but many other derivations are uncommon. Moreover, each loan derivation from a tri consonantal root may undergo a nativisation process by some affixation. Thus, by this affixation, there are some native Urdu derivations, not part of Arabic derivation by the modification of base or phonological changes. The three loan derivations are thus nativised, as *zɒlmi* 'cruel person', *zalɪmana* 'tyrannical' or *məzlumana* 'tyrannised'. There are thus both the processes, the modification of the base first and then the affixation. The latter is both by native and Persian affixes. The Arabic patterns of derivation by modification of base and their nativisation by the affixation go together. Co-occurrence of affixation and the modification of base is a common phenomenon and is termed *transfixes*, by Bauer (2003).

All derivations, whether loanwords or nativised forms, derive two major categories of words i.e. nouns and adjectives. Table 6 below shows only some loanword patterns of derivations. It gives a brief explanation of all loan patterns and then nativised derivational forms in Urdu.

Table 4: Some Arabic derivational Patterns in Urdu

Derivation	Root 1 = żlm	Root 2= ʕʃq	Root 3= ð�envlt̯	Root 4= ʕlt̯	Root 5= ṣd̯q
a.	zʊlm (N) 'tyranny'	ɪʃq (N) 'adoration'	zɪllət̯ (N) 'disgrace'	ɪllət̯ (N) 'illness'	sɪd̯q (N) 'truth'
b.	**za**lɪm (N/A) 'tyrant'	**a**ʃɪq (N/A) 'lover'	zəlil (A) 'disgraced'	əlil (A) 'sick'	səd̯aqət̯ (N) 'truth'
c.	məzl**um** (N/A) 'tyrannised'	maʃuq (N/A) 'beloved'	zəl**a**lət̯ (N) 'baseness'	əl**a**lət̯ (N) 'sickness'	t̯əsd̯iq (N) 'verification'
d.			t̯əzlil (N) 'disgrace'	t̯**a**lil (N) 'cause'	mʊsəd̯d̯ɪq**a** 'verified' (A)

Note: All emphatic sounds in brackets are neutralised and spoken as: ż/z, ð̱/z, ṣ/s & ʕ/vowel.

The concept of triconsonantal root, as known in Arabic, is itself absolutely unknown in Urdu. The derivational patterns in Arabic are pronounced with Arabic emphatic sound, but an emphatic sound in the loanwords in Urdu must be modified to be pronounceable. Therefore, the derivations from triconsonantal roots are shown with native sounds. In addition to the modification of base, there is sometimes an insertion of a consonant-vowel prefix in almost every root pattern e.g. *mə* in **mə**zlum 'tyranised'. Generally a prefix is formed from /m/ or /t̯/ in combination with a vowel. Almost every root pattern contains at least one derivation with such prefixes. This prefixation is used in the same way as in the source language.

The patterns of derivations are generally used the same as in Arabic, and so they are no more than borrowed forms. Arabic may have various derivations from a root pattern, but they are not necessarily used in Urdu. For example, a derivation *t̯əssəd̯d̯ʊq* 'alms giving' (N) from the root *ṣd̯q* is not used with the original meaning, but it is used as a proper noun. There are various similar loan patterns. Moreover, sometimes Arabic derivations by modification of base take native Urdu, Persian or even Arabic affixes to form nativised derivations. For

example, the noun *zʊlm* 'tyranny' can be converted into an agentive noun *zʊlmi* 'cruel man' by a native suffix *-i*. However, it may be expected. But it is interesting to see when a loan suffix from one source attaches to a loanword from another source.

3.4.1.3. Affixation with the Derivations by Modification of Bases

Arabic loanwords are often adaptable enough to take at least a few specific affixes from not just native Urdu, e.g. *-i* but also from Persian, e.g. *-ana* and *-amez* and even from Arabic, e.g. *-ɪjəṯ*. In the process of this nativisation, *-ɪjəṯ* is different from its original function in Arabic.

The affixation of the morphemes from the three sources with particularly the derivations by modification of bases is a major interesting phenomenon and requires attention. The table below illustrates some nativised derivations:

Table 5: Arabic loan bases and derivations, by modification of bases, with Urdu Affixes

Affixes	Derivation A	Derivation B	Derivation C	Derivation D
1.-*i* (Nat)	*zʊlmi* 'cruel man'(A/N	*aʃɪqi* 'love' (N)		*ṯəsḍiqi* 'verified' (A)
2.-*ana* (Per)	*zalɪmana* 'oppressive'	*aʃɪqana* 'loving'	*zəlilana*	
	məzlumana 'like an oppressed' (A)	*maʃuqana* 'like beloved'		
3.-amez (Per			*zɪlləṯamez* (A)	
			zəlaləṯamez-A	
4.-ʃʊḍa (Per)				*ṯəsḍiqʃʊḍa* 'verified'(A

5. -ɪjət̪ (Ar)	məzlumɪjət̪ 'the state of oppressed' (N)'			
6. -ɪja(h) (Ar)		ɪʃqɪja(h) 'of love		
7. -a(h) (Ar)				mʊsəd̪d̪ɪqqa(h)'verified'

The table shows the possibility of various affixes, from the three sources, to attach to the Arabic loanword derivations by phonological changes in bases. There may be others, but the affixes mostly attached are native -*i*, Persian -*ana*, *ʃʊd̪a* and -*amez*, and Arabic -*ɪjət̪*, -*a(h)*/ -*ɪja(h)*. However, these affixes attach only to a few loanwords which are adaptable enough. They cannot freely attach to any of the loanwords. Most Arabic loans are adaptable with the native Urdu affix -*i* and the Persian -*ana*, which also seem to be the most productive with them. The native -*i* has a variety of formations e.g. *zɔlmi* (N/A), *aʃɪqi* (N) and *t̪əsd̪iqi* (A), as in (1), although it attaches only to nouns. The formations verify its productivity and the adaptability of the Arabic loans.

The adaptability is also often seen with Persian affixes which generally form adjectives. There may be even more formations with the Persian -*ana*. However, there are not a variety of formations. Arabic qualitative nouns, often treated as adjectives, e.g. *aʃɪq* 'lover' as in (2), can be converted into adjectives i.e. *aʃɪqana*. The base nouns describe the characteristics of a person, but their derivations modify the action. Only some derivations by the phonological changes in the bases show their adaptability with some affixes, but many others do not. For example, *zəlilana* 'base' is possible, but *zɪllət̪ana* is ungrammatical, although the formation with -*amez* instead may be grammatical *zɪllət̪amez* 'disgraceful'.

The use of Arabic affixes here is also not the same as the way they are used in Arabic loanwords. For example, *-a(h)* is generally used as a loan feminine marker in Urdu, but presently it functions as a derivational affix and derives nativised adjectives from adjectives as in (7). This sort of affixation sometimes also causes semantic differences. These Arabic derivations taking the Arabic affixes *-a(h)*, *-ɪja(h)* and *-ɪjət* are though even less frequent than those with Persian and native affixes, they derive formations that are unknown in Arabic morphologically or semantically. For example, the loan *mʊsəɖɖɪq* is originally used as a derived Arabic noun/adjective in Arabic, but it is used as a male proper noun in Urdu. The affixation of *-a(h)* converts it into an adjective i.e. *mʊsəɖɖɪqa(h)* 'verified', which is also not used in Arabic. The former is not common in Urdu, unlike the latter that is frequent. Of the other Arabic affixes, *-ɪjət* forms the nativised nouns e.g. *masumɪjət* 'innocence' in combination with the loan derived adjectives i.e. *masum* 'innocent'. The other *-ɪja(h)* derives adjectives from nouns *ɪʃqɪja(h)* 'of love'.

These nativised Urdu formations show some common features in the nativisation process. First, they are only possible when some Arabic loan derivations are adaptable enough to adjust with native and the loan affixes in Urdu. However, not all derivations are adaptable and the affixes cannot attach to them freely. The changes by the affixes show some rules, but the patterns are sensitive to both structure and semantics. There are various other factors involved. It cannot be predicted which affixes attach to which loanwords. For example, *-i* attaches to an abstract noun, e.g. *zʊlm* 'tyranny' and forms mainly noun and adjective *zʊlmi* 'cruel man', but with other abstract nouns, e.g. *t̪əsd̪iq* 'verification' it may form only an adjective, e.g. *t̪əsd̪iqi* 'verified'. Sometimes, the semantics of the (Persian) affixes also affect the formation. For example, *-ana* describes the way something is done, but *-ʃʊda* describes the

completion of a job. Therefore, both attach only to the loans relevant to their semantics. From this perspective, formations like *zalımʃʊɖa and *t̪əsd̪iqana are ungrammatical.

Although the affixation with the Arabic loan derivations by the phonological changes in bases is not on a large scale, the formations neither conform to the source language rule nor to the target language rule. Therefore, they are on peripheral target language patterns. The affixation with these derivations shows some irregularities. Thus, they deviate from the general native Urdu and all the loan patterns. There are certainly some rules in native Urdu, but the deviations cannot be ignored particularly when the loanword morphology interlinks with the native morphology. This is because rules from the three sources of words function together.

The three sources have formed some patterns of derivations which sometimes accommodate each other's affixes but which at other times do not. Arabic loanwords, particularly the derivations by phonological changes, show adaptability, and so they take affixes from the various sources. As Dressler (2003: 39) points out, Wurzel (1984) considers that adaptability shows fitting properties of loanwords with the native affixes, and so their integration into native morphological systems occurs. Arabic loanwords are well-suited in the Urdu derivational system, as they have fitting properties. The derivational affixes from the three sources are productive enough to accommodate them without requiring any major changes, i.e. deletion or substitution of a vowel. This shows that Arabic bases can be extended by the addition of new items (the Urdu derivational affixes in general), and therefore seem to have been fully integrated into the Urdu derivational process.

Arabic loanword derivations in sum are in two forms, by affixation and by modification of base. Both loan types are affected by phonological changes e.g. neutralisation or rise and fall of Arabic phonemes in Urdu. It is necessary to discuss some phonological changes also, as

they often interact with morphological changes. Arabic loan affixes are however less affected by them. They are not as many as Persian and the native Urdu affixes, and the derivations from them are also fewer, they show a variety of derivations. Most commonly used affixes are the negative markers *ɣær*, *la bɪla*, the definite article *əl*/*ɒl* and the adverbial -*ən*.

There are only a few patterns of loan derivation by the phonological changes in bases, but they are the same as in the original language. However, they are more affected by nativisation than the derivations by affixation. They take various suffixes without any distinction of the three sources in focus i.e. native Urdu, Persian and Arabic. The affixes from these sources can attach to the derivations by phonological changes. This shows that Arabic loan derivations by the phonological changes in the bases are adaptable and that some affixes from the three sources are productive. This is why interaction of both occurs and further words are derived.

Although the patterns of derivation by phonological changes are the same in Urdu, with some changes in the Arabic loanwords phonology, some Arabic phonemes function in Urdu the same as the native Urdu phonemes. Therefore, it is imperative that there be a brief glance at Arabic loanword phonology.

3.4.2. Arabic Loanword Phonology in Urdu

Urdu has various phonemes that are also available in many other languages spoken in the world. Many of the segments found in Urdu are loans and to a degree Urdu thus shares a segment inventory with the languages it borrows from, i.e. in this instance Arabic. As Watson (2002: 1) points out, Semitic languages (including Arabic) are marked by a limited vocalic system and rich consonantal system e.g. uvular, velar fricatives, pharyngeal, and pharyngealized (emphatic) sounds. Urdu has borrowed a large number of words in whom some consonant sounds e.g. uvular and velar fricatives have been retained while many others,

e.g. pharyngealized (emphatic) sounds, have lost their phonetic value and are substituted by the native sounds. Despite the sound changes, all loanwords can be recognised both by the mutual segment inventory and by common words in the two lexicons. Interestingly, children learning the Quran get special training for the correct pronunciation of words, which they ignore when speaking Urdu. This is because the same is the case with religious scholars using original Arabic words in speaking Arabic, but using the same as loans in Urdu. Such as they correctly pronunce emphatic dentals in Arabic, but they ignore this while speaking Urdu.

Naim (1999) notes, most Arabic loans are easily recognisable due to dentals and emphatic dentals. They contrast in Arabic, but there is no difference perceived in Urdu relative to these sounds. For example, words containing ط/t̀/ are treated the same as those with ث/t̪/. Similarly, other words containing ض /d̀/, ذ /ð̀/ and ظ / ż/ are treated like those with ز /z/; ص /s̀/, ث /θ/ with س /s/, ح/ħ/ with ہ /h/ and ع/ʕ/ with the following vowel. This shows the phonological adaptation of the words, which are different from how they originally were in Arabic.

3.4.2.1. Consonantal Integration

Some Arabic loan phonemes are accommodated in the Urdu sound system. There are many sound segments, which do not show changes and are adopted without any modification. However, many times the loanwords undergo important alternations of sounds and the foreign sound segments are replaced with more 'native essences'. Hardie (2003: 34) observes that Urdu's consonant inventory is somewhat different from that of Arabic; distinctions that existed in Arabic are neutralised (the sounds lose phonetic value) disregarding the position whether word initially, medially or finally. The neutralisation of the loan sounds is mainly seen in the dentals and emphatic dentals, as illustrated in the following table:

Table 6: Arabic Loanword Phonology in Urdu: The Change of Emphatic Sounds

Phonetic Change	Arabic Origin	Urdu Adaptation	Gloss
1. /ḍ/ → /z/	ħa**d**ır	ha**z**ır	present
2. /ż/ → /z/	**ż**alım	**z**alım	tyrant
3. /ð̀/ → /z/	ð̀ıkər	**z**ıkər	mentioning
4. /θ/ → /s/	va**r**ı**θ**	varır**s**	heir
5. /ṣ/ → /s/	xa**ṣ**	X**a**s	special
6. /ṭ/ → /t̪/	ṭə**l**aq	t̪ə**l**aq	divorce
7. /ħ/ → /h/	**ħ**a**d**ır	**h**a**z**ır	present
8. /ʕ/ → following vowel	**ʕ**ılm	**ı**lm	knowledge
	fa**ʕ**al	fa'**a**l	Active

The most important point is to see the substitution of Arabic emphatic dentals by native sounds. The native Urdu alveolar fricatives /z/ and /s/ substitute for two or more loan emphatic sounds. The Arabic dental /ḍ/ is neutralised and substituted by the alveolar voiced fricative /z/, as shown in (1). Therefore, words like *haḍır* 'present' are pronounced as *hazır*. The alveolar fricative /z/ also substitutes for voiced strident fricative /ż/ e.g. *żalım* 'tyrant' (2) and dental fricative /ð̀/ e.g. *ð̀ıkər* 'mention' (3). Thus, the two words are pronounced as *zalım* and *zıkər* respectively. The voiceless alveolar fricative /s/ also substitutes for two loan phonemes i.e. voiceless dental fricative /θ/ and the emphatic sounds /ṣ/. For example, words like *varıθ* 'heir' (4) and *xaṣ* 'special' (5) are pronounced as *varıs* and *xas* respectively. The last emphatic dental sound /ṭ/ as in *ṭəlaq* 'divorce' (6) is also changed with the voiceless dental stop /t̪/. The word is now pronounced as *t̪əlaq*.

The Arabic voiceless pharyngeal fricative /ħ/ (7) does not exist in Urdu. It is pronounced as aspiration /h/. Thus, there is no phonological difference between /h/ and ħ in loans like *haḍi* 'leader/guide' and *ħadʒi* 'pilgrim'. Both words begin with /h/.

Finally, Arabic voiced pharyngeal fricative /ʕ/ (8) is often considered a controversial subject. It is mostly deleted and the loanwords are pronounced with the following vowel. Hardie (2003: 34) observes that /ʕ/ is not found or pronounced as a glottal stop, as zero, as *ā* or *a*, depending on various factors including its environment. The Urdu pronunciation of most loanwords shows that it is deleted and loans, e.g. *ʕʊmer* 'age' are pronounced with the following vowels /a/, /ɒ/, /ɪ/, /ə/ and /ʊ/, instead of, i.e *ʊmer* in this case. Paradis and LaCharité (2001: 258) state, "The non-availability of the Pharyngeal node also explains systematic deletion of the pharyngeal and laryngeal gutturals in Arabic loanwords in French." The same non-availability of the pharyngeal node also results in systematic deletion of the pharyngeal gutturals e.g. /ʕ/ in Arabic loanwords in Urdu. Even if /ʕ/ is accepted as the glottal stop, the presence of glottal stop in Urdu is itself an issue of argument. Some writers, e.g. Hussain (2004: 3 & 5), (Nawaz 2002: 227) and Shahid (2002) agree that it is controversial whether glottal stop exists in Urdu or not. Shahid (2002: 6) comes to the conclusion that no rule completely describes the existence or removal of the glottal stop, while Hussain admits that its existence is a controversial matter and a point of open debate[2]. It is generally neutralised, does lose its phonetic value and replaced by following vowel.

Neutralisation however does not mean that all the phonemes have lost their phonetic values. Some Arabic sounds, e.g. velar fricatives /x/ and /ɣ/ are a permanent part of the Urdu phonetic inventory. They substitute for native phonemes in native words. Moreover, they also create

---

[2] Email correspondence (November, 2007)

new lexical items in combination with native phonemes. However, as recently seen, there seems a reversal of the sounds and some of them are confused with native phonemes.

3.4.2.2. Rise and Fall of Arabic Loan Phonemes

Although some Arabic phonemes are neutralised, and they lose their phonetic value, many others act like native phonemes and are commonly used in native Urdu words e.g. *ʈərxa* 'reject humbly' instead of *ʈərka*. The recent data nevertheless show that some of them are substituted by native sounds. Thus, there appears to be a rise and fall of Arabic loan phonemes, which is the focus of this section. The rise of Arabic sounds here means that they act as if they were the original native sounds. Some Arabic loan phonemes substitute for native Urdu phonemes, e.g. the Sanskrit origin word *ʃakʰ* 'branch' becomes *ʃax* in Urdu. The aspirated velar stop is thus substituted by the velar fricative i.e. /kʰ/ → /x/. Shapiro (2003: 261) claims that it is *hypercorrect pronunciation*. But his observation is restricted to Hindi. Although Hindi and Urdu are both treated as the same language by linguists, some lexical, morphological and phonemic differences however exist. Looking at this view, *ʃakʰ* becoming *ʃax* does not seem to be merely hypercorrect pronunciation. Both voiceless and voiced velar fricatives /x/ and /ɣ/ in fact take the status of normal Urdu phonemes and function the same as original native phonemes. They do not merely cause the substitution but also the creation of new lexical items e.g. *pəʈaxa* 'firecracker'. Thus, their role is distinctive.

In recent years after the rapid growth of electronic media in the Indian subcontinent, it is observed that some loan phonemes sometimes undergo a reverse process. Although on rare occasions words like *xali* 'empty', in normal Urdu, are pronounced as *kʰali*. This is a reverse process /x/ → /kʰ/, however it may be small. It is also possible that *kʰali* existed all along, but

was restricted to certain regional/social varieties, and is now being used more widely. In such a case, the variation between /x/ and /kʰ/ from the start cannot be denied.

However, certain other examples show that a process of reversal of sounds is going on. Phonetic reversal is the process of reversing the phonemes of a word. Wohlin (2010: 38) notes this in Indo-European and Germanic languages, from plosives to fricatives to plosives again.

3.4.2.2.1. Substitution and Creation Processes

The substitution and creation as well as the reversal are interesting and are a contrast to each other, and therefore some discussion is necessary. Consider the proess in the ctable below:

Table 7: The Substitution of Native Phonemes by Arabic Phonemes Urdu

Change	Example	Word origin	Word used in Urdu
1. k → x	ʃəlak*a*	Sanskrit	səla**x** 'rod / iron bar' (N)
2. k → x	ʈɪnkrək*a*	Sanskrit	ʈə**x**n*a* 'ankle' (N)
3. kʰ → x	ʃa**kʰ**	Skrt/(via)Pers.	ʃa**x** 'branch' (N)
4. k → x	bə̀tə**k**	(via Persian)	bəʈə**x** 'duck/swan' (N)
5. k → x	ʈərk*a*	Hindi	ʈər**x***a* 'reject humbly' (V)
6. k → x	ʈɪʈ**k***a*ri	Hindi	ʈɪʈ**x***a*ri 'clacking of tongue' (N)
7. /x/	pəʈ*a***x***a*	Urdu	*pəʈa**x**a* 'firecracker'
8. k → x	tʃəʈə**k**ni	Hindi	tʃəʈə**x**ni 'bolt to lock the door' (N)
9. k → x	tʃəʈə**k**	Hindi	tʃəʈə**x** 'crackle' (V)
10.k → x	pəʈə**k**	Hindi	pəʈə**x** 'throw down violently' (V)
11.ʃ → x	**ʃ**ʊʃ**k**	Sanskrit/Ostai?	**x**ʊʃ**k** 'dry' (A)
12.k → q	**k**əl*a*	Sanskrit	**q**əla*baz* 'Juggler'(N)

13. none	ḍ*a***k**	Prakrit	ḍ*a***k** 'mail' (N)
14. none	kəsə**k**	Prakrit	kəsə**k** 'pain or anxiety' (N)
15. none	l*a***kʰ**	Prakrit	l*a***kʰ** '100,000' (N)

12: The word k*ə*la, common in Hindi, is used in Urdu as qəla*baz* with only the Persian suffix *baz*.

The sound changes are historical and mostly occur in nouns and then in base verbs and adjectives. Although there are other changes also, one is stable i.e. velar stop /k/ or sometimes, aspirated velar stop /kʰ/ is substituted by velar fricative /x/. It remains consistent on the same pattern. In some rare examples, post-alveolar fricative /ʃ/ is also substituted by velar fricative /x/ i.e. ʃ → x (11) or by alveolar fricative ʃ → s (1), but it is unimportant. Various other changes do not seem to show a general rule because the changes are not with relevance to, or affected by, the neighbouring segments. Sometimes, there are also alternations of vowels along with the substitution of phonemes but none are consistent, which may be called the cause of the native-loan sound swap. The only morphemes consistently appearing in some cases are *tʃəṭ* / *pəṭ* at the prefix position in Hindi origin words (8-10).

The velar fricative /x/ also appears word initially, e.g. *xəṭra* 'danger' and word medially e.g. *saxṭ* 'structure', but it is no more than loan phonology from Persian and Arabic. However, its word final appearance in the native words, e.g. *ʃax* shows the substitution. The change /k/ (or /kʰ/) → /x/ generally takes place in most cases (1-10), but in many others (12-14) it does not. Then, what is the cause of this variation and deviation from a pattern emerging? A closer look reveals that the substitution occurs, if the origin of the word is Hindi or Sanskrit. It does not occur, if a word is from Prakrit, which is however a vernacular dialect of Sanskrit[3]. This also shows that the change is not consistent.

---

[3]. A dictionary produced by CRULP (**C**entre for **R**esearch in **U**rdu Language **P**rocessing)

Thus, the origin of a word does not affect the substitution process much. A few words seem to have entered Urdu via Persian, i.e. Sanskrit → Persian → Urdu, but the changes are the same, e.g. *ʃax* (3) and *səlax* (1). This is interesting, as Sanskrit words entered Urdu either directly or via Hindi. It may be claimed that the one in (3) is a simple loan from Persian, or from Arabic via Persian, but the other in (1) shows substitution of /k/ by /x/ in the native words.

There is no transformation of sounds except for a few, e.g. *bət̀ək* → *bət̪əx* 'duck / swan' in (4). The example exhibits two contrasting processes. The Arabic emphatic dental /t̀/ is substituted by the Urdu dental /t̪/ word medially. Although it is Urdu sound, the velar stop /k/ is also substituted by the loan velar fricative /x/. Similar changes are also noteworthy in (11), in which /x/ substitutes for the native post-alveolar fricative /ʃ/ i.e. *ʃʊʃk* → *xʊʃk* 'dry', this time, but the velar stop remains unchanged. Moreover, /ʃ/ is substituted by /x/ word initially but is retained word medially. However, such contrasting processes are rarely seen.

Some words, e.g. *qəlabaz* 'Juggler' (12), show phonological shuffling, while also showing a mixture of morpho-phonemic processes. Their examples are however rare. The baseword *qəla* 'trick' is a phonological reshaping of Sanskrit *kəla*. Although both *qəla* and *kəla* are obsolete, the derivation *qəlabaz* is commonly used. It shows the substitution of native velar stop /k/ by the Arabic uvular stop /q/, but also shows lexical creation by a suffix *-baz* (from Persian *baxt̪ən*), which is nevertheless not the point of discussion at this stage. Like various other examples, no specific rule can be derived here.

Shapiro (2003: 261), terms the substitution process *hypercorrect pronunciations*. However, he does not discuss in detail and only notes that in some instances they "are heard in which Hindi words are pronounced with loan phonemes, as though the items were foreign borrowings e.g. *fɪr* 'then' instead of *phir*." Urdu perspective is quite different from this. Arabic /q/, /x/ and /ɣ/

may originally be loans. Nevertheless, they are part of the Urdu phonetic inventory, whether or not they are present in Hindi. This is evident in some words, e.g. *pəʈaxa* 'firecracker' (7). The example shows that two contrasting sounds i.e. the native retroflex stop /ʈ/ and the loan velar fricative /x/ create a new lexical item. However, the creation in general is not on a large scale. It is even less by /x/ and /ʈ/, but more by the voiced velar fricative /ɣ/ in association with the native retroflex stops /ʈ/, /ɖ/ and /ɽ/:

Word	**Origin**		**Use in Urdu**
15a. ɣəʈ ɣəʈ	(Urdu)	-	ɣəʈ ɣəʈ 'successive gulps' (N)
b. ɣʊrrahəʈ	(Urdu)	-	ɣʊrr*a*həʈ 'sound made by a beast in anger' (N)
c. ɣəɽəp	(Urdu)	-	ɣəɽəp 'a sound made by drowning/sinking' (N)
d. ɣʊnɖ*a*	(Sanskrit ???)[4]	-	ɣʊnɖ*a* 'scoundrel' (N)

All creations by the combination are nouns. Although the number of creations by each retroflex individually may be very small, the loan fricative actively participates in the creation. It always appears word initially. Most formations are with /ʈ/, but there are also examples with flap /ɽ/ e.g. *ɣəɽəp* 'a sound made by drowning/sinking' (15c). The formation by /ɣ/ and /ɖ/ in *ɣʊnɖa* 'scoundrel' (15d) is doubtful, whether it is a creation or a substitution. The expression of a distinctive meaning in this creation is rare, but most creations contain similar semantic expressions and refer to the names of a specific sound e.g. *ɣəʈ ɣəʈ* 'successive gulps' or the factors producing these sounds e.g. *pəʈaxa* 'cracker'.

The examples illustrate that some Arabic loan phonemes do exist and function like normal native phonemes in the Urdu phonetic inventory. However, the substitution of native

---

[4]. This is according to CRULP's spelling. I assume, /ɣ/ must be either /g/ or /k/ here, as /ɣ/ is not a Sanskrit sound. If this is the case, it must be a native Urdu word, I have pointed out this to CRULP. If the sound is really /ɣ/, the word must be a lexical creation within Urdu.

phonemes is far more frequent than the creation of new words. In either case, the loan phonemes show their presence.

Although Urdu and Hindi are considered to be the same language by linguists, they exhibit lexical, morphological and phonemic differences. Hindi is highly Sanskritised, and Arabic phonemes in a large number of loans are replaced by native Hindi phonemes. So, words like *ɣʊlam* 'slave' and *xaḍɪm* 'servant' are used as *gʊlam* and *kʰaḍɪm*. Therefore, from Hindi perspective, Shapiro (2003: 261) may be right in his argument, as the loan phonemes are not used even in Arabic loans. But the phonemes are part of the Urdu phonetic inventory. They are commonly used. They not only substitute for native phonemes in native words but also create new lexical items. It is not merely hypercorrect pronunciation, but a higher function instead. It should be termed a diachronic renovation of Urdu lexicon.

However, in recent years, some surprising observations reveal that there is a reversal of the loan phonemes, although on a very small scale. Arabic loan sounds i.e. /q/, /x/ and /ɣ/ are substituted by the native sounds /k/, /kʰ/ and /g/ even in Arabic loanwords. The loans, e.g. *xali* 'empty' are pronounced as *kʰali*. This is interesting and does require attention, because this is in sharp contrast to what we have noted, namely, that Arabic sounds entered the Urdu phonetic inventory as a result of diachronic process.

3.4.2.2.2. Reversal of Arabic Phonemes: q/k, x/kʰ and γ/g

Generally, Urdu tends to accommodate loanwords, following some adaptation rules like other recipient languages. The heavy influx of some loan types has caused the Arabic phonemes, e.g. /q/, /x/ and /ɣ/, to enter the Urdu phonetic inventory. However, as noted in recent years, some people are confused when using /k/, /kʰ/ and /g/ for /q/, /x/ and /ɣ/ respectively and cannot differentiate between them. There may be some other factors, but they are particularly

influenced by a flood of Indian movies and TV serials. People's confusion between uvular and velar stops /q/ and /k/ seems very easy. Although different from each other, both the sounds are very close in their place of articulation. There is also a reversal of the velar fricatives /x/ and /ɣ/. They seem to not only be substituted by the native velar stops /k/, /kʰ/ and /g/ in the native words they entered initially but also in the Arabic loans they originally belong to. Consider the table illustrating some examples:

Table 8: The Reversal of Arabic Loan Phonemes

Sound Change	Word Origin			Urd. Adaptation	Glossary
1. q → k	qaʈɪl	(N)	(Arabic)	kaʈɪl	killer
2. q → k	məqsəd̪	(N)	(Arabic)	məksəd̪	aim
3. q → k	rɪzq	(N)	(Arabic)	rɪzk	subsistence
4. x → kʰ	tʃətəx	(V)	(Native)	tʃətəkʰ	crackle
5. x → kʰ	xɑli	(A)	(Arabic)	kʰɑli	empty
6. ɣ → g	ɣəʈ ɣəʈ	(N)	(Native)	gəʈ gəʈ	straight gulps
7. ɣ → g	bɑɣ	(N)	(Arabic)	bɑg	garden

Note the three nouns (1-3) showing /q/ → /k/. There is no difference in the appearance of /q/ whether word initially, medially or finally. Although their place of articulation is different, the two sounds are close to each other, and so some people are confused by them. The words, e.g. *qaʈɪl* 'killer', are pronounced with /k/as *kaʈɪl*.

The case of /ɣ/ and /x/ is important, as they previously substituted for native phonemes and on a certain scale are also used along with native Urdu phoneme in newly created lexical items. Unlike the case of their entrance, there is no difference in their appearance in any position, for the reversal. Voiceless velar fricative /x/ is particularly notable. It diachronically substituted

for the native velar stop /k/ in some native words, as in *tʃəʈəx* 'crackle' (9) in Table 4. In some words, it also replaced /kʰ/ and /ʃ/. However, in the reverse process, it is now exchanged with the native aspirated voiceless stop /kʰ/ only, as in (4) Table 5. Moreover, /x/ is replaced by /kʰ/ even in the Arabic loans e.g. (5), which is however a natural process. The voiced velar fricative /ɣ/ is a little different. It diachronically created some words e.g. **ɣəʈ ɣəʈ** 'successive gulps', as in (10). However, the same words are now pronounced with voiced velar stop /g/ as in ***gəʈ gəʈ***. Moreover, Arabic loans are also used with /g/ as *bag* 'garden' (7) in Table 8.

It is often observed in the conversation of some people on TV street shows. It is also noted in the communication between my wife and daughter of 14. There is nevertheless no evidence available in literature, as the reversal is not on a large scale. Note that the Arabic phonemes are confused with the native ones, and their reversal from Urdu is noteworthy.

Wohlin (2010: 38) notes a similar reversal of sounds in the Indo-European and Germanic languages, e.g. from plosives to fricatives to plosives again. He states that a difference of phonemes for the same words between neighbouring ethnic groups could be interpreted as a manifestation of a polarizing force, whose counter force is the reciprocal loans. He terms counter force "the self-preservation" of a language and sees a historical correlation between the development of phonemes and the social divergence-convergence. The geographical spread of humans happens together with the spread of phonemes from plosives to fricatives, voiceless or front phonemes, which then turn to the original phonemes, when some groups turn back towards the areas of their language. Citing Grimm's[5] (1827) and Verner's[6] (1875)

---

[5]. A sound law, first formulated by German philologist Jakob Grimm in 1827, relating German consonants in other Indo-European languages. It states a chain reaction: aspirated voice stops become regular voiced stops. voiced stops in turn become voiceless stops, and voiceless stops become fricatives.

[6]. Karl Verner (1875), Danish philologist, explained a category of seeming exceptions to Grimm's Law and observed that voiceless fricatives in Germanic became voiced if the preceding syllable was unstressed.

laws, he states that linguists talk about the circular pathway in the displacement from plosives to fricatives to plosives again in the development of Indo-European → Germanic languages; perhaps this is just one example of how more original forms of words may reappear.

Of the two possible causes, Wohlin (2010: 38) points out, i.e. the neighbouring communities and historical development, the former (via electronic media) seems to influence the reversal of the Arabic loan phonemes in Urdu. However, the extent of the sound shift that he sees in the Indo-European and Germanic language families cannot be observed here. It is limited to a few specific cases and the shifting back is also not the same, as shown below:

Loan	Example	Reversal	Example	Glossary
16a. k → x (Substitution)	tʃətək / tʃətəx	x → kʰ	tʃətəkʰ	crackle
b. → ɣ (Creation)	ɣəʈ ɣəʈ	ɣ → g	gəʈ gəʈ	successive gulps
c. k → q (Substitution)	qəlabaz	q → k	kəlabaz	juggler

There is some difference in shifting back shown in (16a), although (16a-16b) seem the same. The loan voiceless velar fricative /x/ substituted for the native voiceless velar stop /k/ in the past generally but the shift back shows its substitution by the aspirated voiceless velar stop /kʰ/. Therefore, there is an obvious divergence in the reversal process. Secondly, the voiced velar fricative /ɣ/ created some lexical items but is replaced by the native voiced velar stop /g/ in the same words, as in (16b) and in Table 5. But, this may be expected as a natural process. Similar to this, (16c) shows the loan uvular stop /q/ substituted by the native velar stop /k/.

Although there are some differences, the shifting back of the phonemes in Arabic loanwords is noteworthy. Therefore, to an extent, the data seem to support Wohlin's (2010: 31) claim regarding Uralic languages that most sound shifts occur between phoneme types within the

group with the same articulation site: Velar plosives ↔ velar fricatives, alveolar plosives ↔ alveolar fricatives; the sound shifts within a group of tongue sounds appear as natural process.

From the Urdu perspective, some differences are noted. There is a shifting back to native velar stops from loan velar fricatives, but it shows aspiration in the voiceless stop i.e. x → k^h. Moreover, there is a difference in groups of sounds shifting /q/ to /k/, the former is uvular and the latter is velar, in contrast to Wohlin's observation 'phoneme types within the group'.

A third difference is the cause of the rise and fall of Arabic loan sounds. Wohlin's (2010: 38) hypothesis suggests that the geographical spread of humans is the major cause of the rise of loan phonemes in Germanic and Indo European languages; when some groups turn back towards the areas of their language, the shift back also occurs and the loan sounds are substituted by the original sounds. Although he agrees that various examples may contradict this hypothesis, he emphasises that it could be valid even if it is only one out of 10 ratios. In contrast, it was the geographical spread of Arabs, instead, which caused the introduction of Arabic loan sounds in native Urdu phonetic inventory. In this context, the grounds of the shift back also differ from Wohlin's (2010: 38) hypothesis. It is not the return of native speakers of Urdu, but the strong influence of media instead.

The hypothesis suggests that the different social and geographic conditions involved in the reversals of the loan sounds in some languages could also be involved in the reversal within a word. Reversal within a word conveys that there is sometimes an alternation of two neighbouring consonants which swap with each other. He states that reversal of phonemes adjacent to or away from each other, called inter version or metathesis, is a sound change interpreted in terms of switching directions. A slip of the tongue in children's language e.g. Swedish *bräda* 'board' → *bärda* and correspondingly English *broad* → *board* could be good

examples. Although not very frequent, some Arabic loans show similar metathesis i.e. the consonants swapping may be adjacent or a vowel may be in-between them, as shown below.

17a.   /f/   ⇔   /l/

    qʊ**fl**i   qʊ**lf**i   'a jelly made of milk'

  b.   /r/   ⇔   /z/

    gʊ**re**z   'avoidance'   V   n*a*gʊ**zi**r 'unavoidable'

Example (17a) shows a simple metathesis. The fricative /f/ and the lateral /l/ have diachronically swapped positions within the Arabic loan. The word was actually spoken as *qʊfli* initially but after the swapping of the two sounds the word is now pronounced as *qulfi*. Thus the word index has changed from $C_1V_1C_2C_3V_2$ to $C_1V_1C_3C_2V_2$.

Example (17b) is different and shows a morpho-phonemic process. The metathesis is the result of the affixation of the Persian negative marker *na*, and the consonants /z/ and /r/ are not swapped diachronically. Moreover, there is a vowel in-between, and it also changes from /e/ to /i/, although the position remains intact. This confirms Wohlin's (2010: 30) observation that the consonants swapping may be adjacent or distant from each other. The word index thus also shows extreme changes from $C_1V_1C_2V_2C_3$ to $C_1V_1C_2V_2C_4V_3C_3$. The reversal of the sounds and the change of the word index is however associated with morpho-phonology and the examples like those following a negative are rare.

The reversal is nevertheless not the dominant feature in standard Urdu. Arabic phonemes are an important part of the Urdu phonetic inventory and distinctive from Hindi phonemes used in the same words. Like the substitution of the native phonemes and the creation of new words they undergo various other phonological processes e.g. assimilation, elision, gemination.

3.4.2.3. Assimilation, Elision and Gemination

A sound is assimilated by adopting some feature(s) of an adjacent sound. Elision is simply the deletion of sound due to a certain process, while gemination may occur as a result of elision. Although the three processes are frequent in Arabic loanwords, they are restricted to a few specific loans. Therefore, the discussion in this section is restricted to some major points.

The place assimilation rule for nasals (found in many languages like English, Spanish and Urdu with a little variation) states that a nasal has the same place of articulation as the following consonant (Shah, 2002). In Arabic loans in Urdu, assimilation is generally seen in the nasal /n/ and the lateral /l/ sounds. It is influenced by some following consonants. For example, when the nasal /n/ precedes bilabial voiced stop /b/, it gets labialised and becomes /m/ i.e. /n/ → [m]/ __b e.g. ʌnbʌr 'a rich perfume' → ʌmbʌr. The alveolar nasal /n/ loses its phonetic value, as the place of its articulation is bilabial due to the following bilabial voiced stop. Therefore, the loan is pronounced differently.

In contrast, elision is the complete omission of a sound due to the effect of a following consonant. This is often seen in the Arabic loan definite article əl/ɔl, when the lateral /l/ disappears, while the following consonant is geminated e.g. nurɔ(l)d̪d̪in 'the light of the religion'. The article generally occupies word initial and medial positions and functions in connection with some consonants divided into two groups based on whether or not /l/ is geminated with them. Although the Arabic distinction in terms of the 'solar letters' /t̪/, /θ/, /d̪/, /ð/, /r/, /z/, /s/, /ʃ/, /ṡ/, /ḍ/, /t̀/, /ð̀/, /l/, /n/ and the 'lunar letters' /b/, /ǧ/, /ḥ/, /ḫ/, /ġ/, /f/, /q/, /k/, /m/, /w/, /y/, /h/ is unknown and irrelevant in Urdu, the elision of /l/ follows the same rule to a certain extent, and so a brief explanation of these terms is necessary. Solar letters represent coronal consonants in the classical Arabic (the Quranic language), but the moon letters are

not. Since the article *əl/ʊl* ends in a coronal consonant, it loses phonetic value with these sounds. Following the Arabic rule, the lateral /l/ is elided in Urdu also, when followed by solar letters, e.g. *nurʊ(l)ḍḍin* 'the light of the religion' and is not otherwise e.g. *nurʊlkɪtab* 'the light of the book'. However, in many other loanwords, the lateral /l/ in the Arabic definite article shows phonological differences with respect to the position of its appearance in the words and compounds. Unlike in Arabic, there is a restriction of position in Urdu and the elision and gemination occur only word medially. The process does not occur word initially, although the following consonant may be a solar letter. For example, the Arabic rule does not apply when any noun begins with the Arabic loan definite article. Consider the following:

18a. kɪtabʊ[l]**n**nur

   b. əlnur

   c. *ə[l]nnur

The application of the Arabic rule is possible only when the definite article appears word medially, as in (18a). The lateral /l/ is elided as it precedes a solar sound /*n*/ which is geminated, and thus the word is grammatical. But the elision does not occur if the loan Arabic definite article appears word initially as in (18b), although it may be ungrammatical in Arabic. The elision the lateral /l/ and the germination of the following nasal are in fact ungrammatical in this position, as shown in (18c). This is in sharp contrast to how it is used in the source language. Elision and gemination are not features of Urdu. Therefore, a rule may be applied in one case (18a), but it is not applicable in other acse (18b- 18c).

In short, in addition to the neutralisation and rise and fall of the loan sounds, Arabic loanwords show assimilation, elision and gemination of the sounds. Both assimilation and elision are caused by the effect of the following consonants but the two rules function in

different environments and have different results. For example, the nasal /n/ is assimilated and becomes /m/ when followed by bilabial stop /b/. Elision is the deletion of a sound, e.g. lateral /l/ of the loan definite article *əl/ɒl*, if it is followed by a solar letter. The elision rule in Urdu is functional only if *əl/ɒl* appears word medially. As a result of the elision the following consonant is geminated i.e. doubled e.g. *səlahɒ[l]ḏḏin* (proper noun) and *kɪʈabɒ[l]nnur* 'book of light'. The process of gemination is not found in native Urdu words.

However, the elision and consonantal doubling are specific with only the Arabic loan definite article, not in any native Urdu or Persian loan affixes. They are among the major aspects of the consonantal sound changes in the loanwords. Although the consonantal sounds show more important features the changes are less than those in the vocalic sounds. Sometimes, the deletion of a consonant results in vowel changes, e.g. vowel lengthening or creation of diphthongs in the loanwords. Vowel changes may occur in various places.

3.4.2.4. Vowel Integration in Arabic Loans

Change of vowels in the loanwords is a general phenomenon and can be seen in the Arabic loanwords as well. However, considering the morphological focus of this work, this section discusses briefly some major aspects e.g. vowel lengthening, shortening caused by the deletion of pharyngeal fricative, by the pluralisation and by derivational phenomena.

3.4.2.4.1. Vowel Changes Caused by the Neutralisation of Pharyngeal Fricative /ʕ/

As noted in the beginning of Section 3.4, many of the Arabic loan sounds are neutralised and lose their phonetic value. The neutralisation of pharyngeal fricative /ʕ/ causes a word to be pronounced with the following vowel, which may take one of three different forms. It may lengthen, become a diphthong or retain its original form. For example, when a loanword, e.g. *fa(ʕ)al* 'active' is neutralised, the word is often spoken with a long vowel as *fal*, but when /ʕ/

in between two different vowels e.g. *fa(ʕ)ıl* 'subject/agent' is deleted, there is the creation of a diphthong, as [aɪ] in *faıl*. Languages differ in the length of diphthongs. In languages with phonemically short and long vowels, diphthongs typically behave like long vowels. Waqar and Waqar (2002: 20) agree that the deletion of some consonants i.e. glottal stop /ʔ/, palatal approximant /j/, labiodental /v/ and pharyngeal fricative /ʕ/ results in the formation of diphthongs whose most underlying vowels are long. This means that the deletion correlates with the lengthening of the vowels, which are sometimes diphthongs. However, there are certain cases in which the first vowel may be long while the second is generally short if taken individually. For example, the word **fail* is ungrammatical if both the vowels are long.

Moreover, neutralisation of /ʕ/ does not always result in the formation of a diphthong or the lengthening of a vowel. Sometimes, a vowel may remain unchanged.

19. (ʕ)ılm   → ılm   'knowledge'

There is neutralisation of /ʕ/ in the example, but there is no lengthening of the vowel or the formation of a diphthong as a result. The short vowel /ı/ remains intact, which means that the neutralisation may result in three different changes. There may be: lengthening of the vowel, the formation of a diphthong which itself may function as long vowel, or a vowel may remain unchanged. This exhibits complexities also in the phonological features in the Arabic loanwords. A close look reveals that pharyngeal fricative appears either word initially or medially. It is rare word finally. Therefore, the changes in the two positions depend on which vowel /ʕ/ appears with. Consider the following:

20a. (ʕ)alım 'scholar'   → *aalım*

  b. (ʕ)ıʃq 'love'   → *ıʃq*

  c. mə(ʕ)ıjət̪ 'companionship' → *məıjət̪* 'companionship/funeral'

The lengthening of the vowel only occurs when /ʕ/ appears with /a/ irrespective of the position, whether word initially as in (20a) or word medially as in *faal*, noted above. Otherwise, no vowel changes occur word initially, as in (20b), and there is a formation of a diphthong word medially, as in (20c). The last change shows an interesting element. After the neutralisation of /ʕ/, there is no difference left in the pronunciation of the word with that of another word having the meaning 'funeral'.

Therefore, this sort of change sometimes causes confusion. Although the examples show certain regularity, the three results of the same consonantal disappearance confirm the complexities of phonological changes in the various patterns of the loanwords. These intricacies can be observed also in the shortening vowel, explained in the next section.

3.4.2.4.2. Shortening of Vowels

Vowel shortening in the Arabic loans is mainly the result of the pluralisation of feminine loan nouns with native Urdu plural markers and the formation of adjectives from nouns. If a loan noun ends in *-i*, it is affixed with the plural marker *-jã*. But at the same time *-i* has to be substituted by the short front vowel *-ɪ-*. Consider the following examples:

Nouns	**Plurals**

21a. *kɒrsi* 'chair'   (f) → *kɒrsɪjã*

  b. *hazri* 'presence' (f) → *hazrɪjã*

The noun in (21a) takes the feminine plural marker *-jã* while at the same time the long vowel *-i* is also substituted by the short vowel *-ɪ-*. The morphology here is in conformity with phonology because it is not possible for long *-i* to precede *-jã*, a suffix with a glide on onset position. The noun *hazri* in (21b) is the natively derived feminine noun from the loan

noun/adjective *hazɪr* '(someone) present'. The feminine marker *-i* fulfils the requirements of the pluralisation, which ultimately shows the substitution by *-ɪ-*. Naim (1999) states that Urdu long vowels are inherently long, similar to the vowels in *tall* and *lord* respectively and all long vowels are pure, as they are not pronounced with a following glide. The feminine plural marker *-jã* starts with a glide and for this reason the process of pluralisation requires the short vowel *-ɪ-* which substitutes for the feminine marker *-i*.

The same rule is applied in the formation of nouns from other nouns or adjectives, both of which end in a long *-i*. The adjectives are themselves nativised forms and originally derived from the Arabic loan nouns, e.g. *ɪlm* 'knowledge/information' → *ɪlm**i***. After the suffixation of *-jəʈ*, the adjectives become nouns, but there is also a vowel change from long to short, as in *ɪlmɪjəʈ* in this case. The derivational suffix *-jəʈ* begins with a glide and therefore the rule remains the same. The formation of a noun requires the long *-i* to change into a short *-ɪ-*.

The suffix *-i* also derives nouns from base adjectives e.g. *mehrum* 'deprived' → *mehrum**i*** 'deprivation', which in turn can be transformed into another noun *mehrumɪjəʈ*. The process of vowel shortening is the same. This shows that there is some regulation in the shortening of vowel, although it may not be seen in other sound changes.

The discussion on loanword phonology can be summarised as follows; many Arabic loan phonemes are neutralised but some others retain their phonetic value. Some of these latter ones, e.g. /x/, /ɣ/ and /q/ not only substituted for some native phonemes diachronically but also created new lexical items in correlation with them. It cannot simply be termed hypercorrect pronunciation, as Shapiro claims about Hindi. The two processes, substitution and creation, in Urdu show that the three phonemes are part of the Urdu phonetic inventory. However, as observed recently, the loan phonemes shift back not only from the native words

they previously entered but are also substituted by native counterparts in the loanwords, which is not unnatural. The reversal of sounds shows some features different from those noted by Wohlin (2010: 31) in Uralic languages, although data in Urdu supports his observation. There is a shifting back to native velar stops from loan velar fricatives e.g. /x/ → /kʰ/, but it also shows aspiration in a voiceless stop. Moreover, the reversal of phonemes occurs in different groups of sounds, i.e. in /k/ and /q/, the former is a velar stop and the latter is a uvular stop. There is also a difference of the rise and fall of the sounds from what Wohlin notes. However, as for the reversal within a word, the data from Arabic loan phonology confirm his claim that two sounds within a word can swap even if they are distant.

3.5. Summary of the Chapter

The main focus of the chapter has been to describe derivation by affixation. There are more prefixes than suffixes, but they derive fewer words. Affixes from the three sources in general give various patterns, but frequent deviation from each pattern is also noteworthy. The vast majority of affixes in Urdu come from Persian, though the native affixes are productive and derive most words. Thus, there is a distinction between the number of forms and frequency of use. Native affixes, e.g.-*a* and -*i,* do not express their meaning.

Loan affixes however express various semantics and thus give evidence against his claim. Although Persian loan affixes are the most, they derive only a few words. There are nevertheless some productive affixes, e.g. *na* 'no/not' and -*ana*. The number of derivations from them is lower than that of native Urdu affixes and higher than that of Arabic affixes.

The derivational changes in Arabic loanwords are achieved both by affixation and by the modification of bases. The affixes, e.g. *la* 'no/not' and *ɣær* 'not/without' are few, but they are productive and derive some words. Like Persian affixes, they express their meanings as well.

But the derivations by affixation are not as many as derivation by vowel changes. The words derived by the modification of bases are large in number and in various patterns, but they are no more than loans in Urdu. Moreover, they are nativised by taking affixes from native Urdu, Persian and even from Arabic. The affixation with these derivations has some irregularities, which show that there is often deviation from the rules particularly regarding loanwords.

The intricacies in the loanword morphology from Arabic are often affected by phonological changes in the Arabic loans. Therefore, it was necessary to have an idea of Arabic loanword phonology, as Arabic phonemes function like native Urdu phonemes and also play a role in the creation of lexical items. The most important feature is the neutralisation and the rise and fall of some Arabic phonemes. Some loan phonemes not only substituted for native phonemes but also created new lexical items diachronically. Therefore, Shapiro's (2003: 261) term *hypercorrect pronunciation* is not justifiable in respect of Urdu. Arabic loan phonemes are important members of the Urdu phonetic inventory and function like native phonemes. However, a reverse process is also seen and borrowed velar fricatives /x/ and /γ/ are being replaced by the velar stops / k^h/ and /g/. The data presented confirms Wohlin's (2010: 38) observation of reversal of sounds, but the features shown in the case of Urdu are different.

The phonological process often interacts with morphology and affects inflectional and derivational morphology. Derivational morphology, by affixation, has been discussed in this chapter. Derivation is not restricted to affixation and many words are derived by compounding also. Urdu compounds have some key features. Most compounds are formed by some distinctive compounding words, while a large number of compounds take infixes. Compound formation is discussed in the next chapter.

****

Chapter 4

Derivation by Compounding

4.1. Introduction

Two aspects of the derivation discussed in the last chapter, with respect to the three major sources of words in Urdu, were the affixation of various morphemes to bases and the modification of bases themselves. The derivation of new words is also very frequent by compounding two words having independent lexical status. However, compounding is not necessarily made by two constituents from the same source language. There is a possibility that both constituents are from two different sources. Such compounds are called hybrid compounds. They are heavily dominant in Urdu and more frequent than normal, i.e. native + native, combinations. They form the same types of compounds and show the same features, which normal native based compounds show. Therefore, this chapter will focus on the various types of compounds rather than look at the compounds from a specific source language. These types include, endocentric, exocentric, copulative, postpositional and verbal compounds.

Although these combinations are specific with their features, e.g. an endocentric compound is a hyponym of one of its constituents and a copulative has two independent constituents, they are sometimes ambiguous and confusing. Sogaard (2004: 322) exemplifies that although the Mandarin Chinese compound *fumu* (father-mother) 'parents' is a copulative compound, it can be given the endocentric interpretation 'a mother who acts like a father'. This ambiguity, often noted in Urdu compounds, may be due to the two different source languages of the constituents. For example, when two constituents are synonymous or near synonymous e.g. *hʊsən* 'beauty'+ *dʒəmal* 'beauty' → *hʊsən-o-dʒəmal* '(kinds of) beauties', it is generally an endocentric compound. Nevertheless, constituents are synonyms in many other compounds e.g. *bal bətʃe* 'family', but the formations may be other types e.g. copulative in this case. The

constituents in *hʊsən-o-dʒəmal* are both Arabic, but the constituents in *bal bətʃe* are from native and Persian sources. The former compound is endocentric and simpler to understand due to the same source constituents. However, the latter compound is copulative and ambiguous, as the constituents are from two sources, although they are synonymous. Thus, the change of the source language of the two constituents may cause ambiguity in the semantic of a compound.

Ambiguity also arises due to the structural nature of the Urdu compounds. Like those in Germanic languages, Urdu compounds may be left-branching (modifiers come before the head) e.g. *kʊʈʊb* 'books' + *mela* 'fair' → *kʊʈʊb mela* 'bookfair' whereas like those in Romance languages, they may also be right branching (the modifiers come after the head) e.g. *nur* 'light' + *xʊɖa* 'God' → *nur-**e**-xʊɖa* 'the light of God'. Of the two compounds *kʊʈʊb mela* and *nur-e-xʊɖa*, an infix plays key role in the formation of the latter, and former is formed by a distinctive compounding word. These are two major elements in Urdu compounds. There are three infixes, *-e-*, *-o-*, and *-b-*, which play various grammatical and morphological roles in the compounds, due to which Durani (2007) terms them determinative compounds. The role of *-b-* in some determinative compounds needs separate discussion, due to the semantic ambiguity that the compounds may be exocentric or copulative. Most compounds are formed by distinctive compounding words, each of which may derive various compounds. For example *xʊʃ* 'happy' derives a large number of formations e.g. *xʊʃ* + *bu* 'smell'→ *xʊʃbu* 'fragrance', *xʊʃ* + *nəsib* 'fate' → *xʊʃnəsib* 'lucky'. Sometimes, some constituents also conjoin randomly to form a compound, e.g. s*otʃa* 'thought' + *səmdʒha* 'understood' → s*otʃasəmdʒha* 'well planned'. Neither of the two verbs derives any other compound. Such casual compounds are not numerous, and so not the focus in this work.

By classification type and number of word formations, most compounds are endocentric and verbal compounds, but they also include exocentric and copulative compounds as well as postpositional nexuses. The formations include AN (Adjective + Noun), NA (Noun + Adjective), NN (Noun + Noun), AV (Adjective + Verb), NV (Noun + Verb) and the resulting compounds are mostly adjectives, nouns and verbs, but there are also adverbs, pronouns and postpositions. This discussion begins with some issues in the endocentric class followed by exocentric and copulative compounds.

4.2.1. Endocentric Compounds

An endocentric compound is a type of compound in which one member functions as the head, the other as its modifier, and the compound is the hyponym of the head constituent. For example, the Urdu compound *kʊʈʊb mela* 'bookfair' is an endocentric compound in which *kʊʈʊb* 'books' is the modifier and *mela* 'fair' is the head, and the compound *kʊʈʊb mela* is the hyponym of the head constituent. Endocentric compounds seem to be the most frequent kind of compounds found in Urdu. They occur primarily in NN and AN, NA and AA combinations. Consider the various configurations.

1a. Noun + Noun

      i) nur 'light' + xʊɖa 'God' → nur-**e**-xʊɖa (N) 'the light of God'

      ii) hʊsən 'beauty' + dʒəmal 'beauty' → hʊsən-**o**-dʒəmal '(kinds of) beauties'

      iii) kʊʈʊb 'books' + mela 'fair' → kʊʈʊb mel*a* 'bookfair'

b. Adjective + Noun

      i) ədʒib 'strange' + ɣərib 'poor' → ədʒib-**o**-ɣərib (A) 'very strange'

      ii) pak 'pure' + ɖɪl 'heart' → pak ɖɪl (N) 'sincere/true-hearted'

c. Noun + Adjective

i) ədalət̪ 'court' + ʊzma 'supreme/great' → <u>ədalət̪</u>-**e**-ʊzma 'supreme court'

d. Adjective + Adjective

i) xʊʃ 'happy' + xʊrrəm 'very happy' → <u>xʊʃ</u>-**o**-xʊrrəm (A) 'very happy'

ii) həlka 'light' phʊlka 'light' → <u>həlka</u> phʊlka (A) 'very light'

e. Pronoun + Pronoun

i) t̪ʊm 'you' + xʊd̪ → <u>t̪ʊm</u><u>xʊd̪</u> (P_{RO}) 'yourself'

As the examples above show, there are five structural configurations, and each has some variations. For example, NN compounds contain three variations i.e. with and without the infixes -*e*- and -*o*-. The AN and AA compounds also have two variations i.e. with and without the infix -*o*-. Similarly, there are combinations of NA and (Pro) (Pro). Each of these fulfils various functions conveying various semantics. The focus here is mainly on NN and AA combinations with relevance to the functions of the two infixes -e- and -o-, which show interesting semantic features.

4.2.1.1. Infixes or Interfixes in the Compounds

The formation of new words in Urdu by compounding is seen in almost all categories of words, but the compounds taking an infix are mostly found in the endocentric class. Other classes of compounds are rarely seen with any infix. Bauer (2003) terms an infix in a compound *interfix* and describes it as follows, "A special kind of infix that appears between the two elements of a compound is an interfix. This is found in many of the Germanic languages e.g. *tag* + *reise* → *tag*-**e**-*reise* 'day's journey'." The use of such an interfix in Urdu is a Persian borrowing. Naim (1999) states that it is the most commonly used Persian grammatical feature in Urdu. There are two main interfixes i.e. -*e*- and -*o*- whose function is to structurally link the two constituents of a compound. Moreover, they describe the

grammatical relations between them. Bogel et al (2008) and David et al (2009) term *-e-ezafah* 'increase/addition' and state that theoretically it can only join Perso-Arabic loanwords, but in spoken usage it is occasionally used with words of Indic origin as well. It represents the genitive marker *ka* e.g. *nur* 'light' + *xʊḍa* 'God' → *nur-e-xʊḍa* 'the light of God' and the agentive particle *vala*, which also expresses possession. Neither the genitive *ka* nor the agentive *vala* themselves are used in NN compounds. However, Mohanan (1994: 115) mentions some uses of *vala* in some N + V+ *vala* combinations. A third function of *-e-* is also to link NA constituents phonologically, in which it plays no grammatical role.

Unlike -e-, the interfix *-o-* is in fact a phonetic substitution of *-v-* that is often used to connect two nouns or two adjectives, e.g. *Ahmed-v-Hamidi*, and conveys the conjunctive sense of *ɔr* 'and'/'an addition to something'. Historically, *-v-* is a reduced and shorter form of *ɔr*, like many other Sanskrit origin words, and the development is described in Paniniyan grammar[7]. Therefore, in compounds also, it plays the same conjunctive role and replaces the conjunction *ɔr* 'and', which is never used as an interfix. Spelling draws on the knowledge of sound-letter correspondences, syntactic rules, orthographic rules and on the knowledge of morphology (Carlisle 1987). Orthography has also played a part in the shifting from -v- to -o-, and thus phonology is influenced by spelling as well. The interfix *-v-* is written with a single letter, 'و' called *vao*, which is also the representation of the vowel sound *-o-* in various words. In ordinary conversation, speakers generally pronounce *vao* in the compounds with an *-o-* sound rather than a *-v-* sound. Orthography in particular has influenced the way children read such compounds. As the letter *vao* is the same for both *-v-* and *-o-*, they pronounce the words with an *-o-* sound instead of a *-v-* sound. This is possibly because the *-o-* sound in the compounds

---

[7]. *Panini's Grammar and Paniniyas' Tradition*- A lecture on May 04, 2011 at Uniwersytet Im. Adam. Mickiewicza, Poznan, Poland: By Professor Boris Zakharin from Moscow State University. The author of this work was present in the lecture.

is more rhythemic and more convenient to use than *-v-* sound. Therefore, *-v-* is restricted to spelling and *-o-* is the typical pronunciation. But, orthography-phonology correspondence is not the point of discussion here.

The presence of the two interfixes *-e-* and *-o-* in the compounds expresses some grammatical function. Durani (2007) terms such compounds *determinative compounds*: "there are *descriptive compounds* in which the modifier is used in attributive, appositional or adverbial manner, and there are also *determinative compounds* in which one is not an attribute to the other; it is rather related to the other in a way corresponding to one of its grammatical cases". Consider his division of compounds:

2a. Descriptive compounds

    həlka (A) 'light' + phʊlka (A) 'light' → <u>həlka</u> phʊlka (A) 'very light'

 b. Determinative compounds

   nur (N)   'light' + xʊḓa (N) 'God' → <u>nur</u>-**e**-xʊḓa (N) 'the light of God'

Example (2a) simply describes the configuration in which the modifier is used in an attributive, appositional or adverbial manner. Such compounds are termed descriptive compounds by him. They also include reflexive pronouns. Example (2b) *<u>nur</u>-**e**-xʊ̃ḓa* 'the light of God' shows an interfix *-e-*. They are determinative compounds, as the infix *-e-* plays a grammatical role between the two constituents. It represents the genitive marker *ka*/*ki* or *ke* and expresses its semantics. It thus shows a possessive relation between the nominal elements. It is only semantically related and not a phonological reduction of the genitive marker *ka/ki/ke*. As mentioned above, the genitive has no use in compounds. Therefore, its semantics must be expressed by *-e-*. With the genitive marker, the grammatical relation is expressed as *xʊ̃ḓa **ka** nur* 'the light of God', which is a phrase and not a compound.

Durani's (2007) explanation of the role of an interfix is limited to -e- and it is with just one semantic feature i.e. the representation of the genitive *ka*. The role of -e- is wider than what he claims, and there are other structures with a different function. Moreover, as noted above, there is another interfix, -o-, that represents the conjunctive particle *ɔr* 'and'. Giving the sense of plurality and emphasis, it also forms the structures which do not behave in the way Durani claims. There are five functions the two interfixes perform.

3. The interfixes (from 3b. to 3e.) representing other than genitive

a. r*a*h (N)'path' + həq (N) 'truth' → r<u>ah</u>-**e**-həq (N) 'the right path'

   (-e- represents the genitive marker *ka/ki/ke*)

b. s*a*hıb (N) 'someone/Mr.' + ılm (N) 'knowledge'→ <u>sahıb</u>-**e**-ılm (N) 'scholar'

   (-e- represents the agentive particle *vala* showing 'someone having something')

c. ədĩalət̪ (N) 'court' + *a*lıj*a* (A) 'high' → <u>ədalət̪</u>-**e**-alıj*a* (N) 'high court'

   (-e- describes no semantics, but links the two constituents phonologically)

d. əsər (N) 'influence' + rəsux (N) 'influence' → <u>əsər</u>-**o**-rəsux (N) 'resources'

   (-o- represents the conjunctive particle *ɔr* 'and')

e. xʊʃ (A) 'happy' + xʊ̃rrəm (A) 'happy' → <u>xʊʃ</u>-**o**-xʊ̃rrəm (A) 'very happy'

   (The interfix -o-conveys emphasis)

Example (3) shows the structural varieties of the Urdu compounds with interfixes showing various semantic interpretations. Example (3a) is the same as example (2b). The two constituents are nouns, where N_1 is a head and N_2 is a modifier, and the interfix represents the genitive *ka.* In (3b), the two constituents are again two nouns but the interfix -e- does not represent the grammatical case. Rather, it is a morphological element representing the agentive particle *vala* showing possession 'someone having something'. In (3c), the two

111

constituents are noun and adjective, and *-e-* expresses no semantics but shows only phonological relations between the two constituents. Examples (3d) and (3e) show the interfix *-o-* in the structures of NN and AA. It represents the conjunctive particle ɔr 'and'. When it appears with two nouns, as in (3d), it shows plurality but when it appears with two adjectives, as in (3e), it denotes emphasis. The examples from (3a) to (3e) are thus evidence that the two affixes *-e-* and *-o-* in fact form four structures and perform five functions. This is in sharp contrast to Durani's (2007) semantic interpretation, which is limited to the genitive case shown by *-e-*. This contrast thus leads to a discussion of what functions the affixes play and what semantics they convey by showing grammatical, morphological and phonological relations between the constituents.

4.2.1.2. Functions Performed by the Two Interfixes: *-e-* and *-o-*

As noted, (3a) describes possession, and *-e-* represents the grammatical case i.e. the genitive *ka*. Constituent N$_1$ *rah* 'path' functions as a possession and N$_2$ *hɔq* 'truth' as a possessor. Although (3b) also describes possession, the function of *-e-*contrasts with the one stated in (3a) and the expression of semantics is also different. Accordingly, the roles of N$_1$ and N$_2$ also change, as shown in (4).

4a. r*a*h-**e**-hɔq (N) 'the right path' =   hɔq **ki** r*a*h 'the right path'

     (*-e-* represents the grammatical case i.e. the genitive *ka*)

 b. s*a*hɪb-**e**-ɪlm (N) 'scholar'   =   *ɪlm **k*a*** s*a*hib / ɪlm **v*ala*** 'scholar'

     (*-e-* doesn't represent genitive *ka* but the morphological particle *vala*)

Although the interfix *-e-* and the two nominal constituents in (4b) seem to be structurally the same as in (4a), they are functionally and semantically different. *-e-* does not represent the grammatical case, here, but refers to the morphological particle *vala* 'someone having

something.' Like the genitive *ka*, the agentive *vala* also shows possessive relation between the nominal elements, but the role of the two constituents is quite the reverse here. N₁ *sahıb* 'someone' now functions as the possessor and N₂ *ılm* 'knowledge' as the possession. Thus the compound <u>*sahıb*-**e**-*ılm*</u> means *ılm* **vala** 'scholar/the man of knowledge'. Treating *-e-* as if it were representing the genitive *ka/ki* with no change of role by N₁ and N₂ would cause the ungrammaticality of the compound.

Among the individual particles, *vala/vali* can be understood in isolation and functions like a suffix to a noun, an adjective and an infinitive, as the examples below show.

5a. d̤ud̤ v*ala* (N) 'milkman'

 b. bʊri v*a*li (A) 'the bad one'

 c. ron**e** v*ala* (Infinitive) 'about to weep'.

Like the genitive marker *ka/ki/ke*, the agentive particle *vala* also shows gender/number marking with possessive meaning. Mohanan (1994: 116) describes that *vala* may be used in four forms i.e. as agentive marker, a relative clause marker, an aspect marker indicating immediate future and in some lexical compounds N + V + *vala*. However, our concern is not *vala* but the inetrfix *-e-* that is used to show its semantic in the compounds. So, it is enough to undertand that *vala* is lexical and phrasal, and it may mean a possessor, seller, agent or distributor (5a), or it may have the sense of 'one with'. Hardie (2004: 8) terms it occupational particle. It is used to convey the value or price (5b). It expresses a marker of the immediate future also (5c). Like the genitive *ka*, it is independent and has its own status. Similarly, *vala* does not have any role in compounds. The only relation between the interfix *-e-* and the particle *vala* is that *-e-* represents *vala* in the compounds exactly the same as it represents the

genitive *ka*. This representation is however seen in a few nominal compounds, while the representation of the genitive is broad enough.

By this representation, the interfix *-e-* thus performs a dual role which is morphological as well as grammatical. The roles of the two nouns are also changed. Now the question arises as to how to recognise the two roles of the possessor and the possession performed by N_1N_2 with the same interfix *-e-* representing the genitive *ka* in some compounds and the particle *vala* in some others. It is only possible through an understanding of the wider semantics of the two nominal constituents i.e. supposing what relations they can have in the compounds or which one of them can be the possessor and which one the possession. *-e-* does not express any other semantic role in addition to those of the genitive marker *or* the agentive particle.

However, it does appear in some compounds, and its absence causes ungrammaticality. Its appearance sometimes involves phonological relations between the two constituents. The main cause of phonological relations lies behind the fact that there are some NA compounds, in which the second constituent is an adjective. Seidenberg and Gonnerman (2000: 356) consider that morphemes are discrete units, and morphology is correlated with other types of lexical information, including spelling, sound, and meaning. Effects attributed to morphology may be due to these correlated factors. Consider the examples below:

6a. *əḓaləṯ alɪja/ *vəzir *az*əm

  b. əḓaləṯ (N) 'court' + *al*ɪj*a* (A) 'high' → əḓaləṯ-*e-al*ɪj*a* (N) 'high court'

  c. vəzir (N) 'minister' + *az*əm (A) 'prime/great'→ v̱əzir-e-*az*əm (N) 'prime minister'

As (6a) shows, the removal of *-e-* from such structures causes the ungrammaticality of a compound. Examples (6b) and (6c) show that the function of *-e-* is wider than merely a

representation of grammatical and morphological elements. In the NA combination, -e- does not perform any possessive role because the representation of the genitive *ka* and the agentive *vala* is impossible. This means that the interfix -e- must play some other role. It is in fact a phonological correlation between the two constituents. The aim of these few lines is not to focus on all phonological aspects, but to point out a further role that -e- may play. Therefore, the discussion on this aspect is left to future researchers, as the focus here is mainly on the morphological functions of the interfixes with relevance to their semantic expressions.

The three functions of -e- present a case against Durani's (2007) claim that -e- creates a relation between the two constituents in a way corresponding to the grammatical case i.e. the genitive. In addition to the three functions of -e-, there are two roles, i.e. plurality and emphasis, played by the interfix -o-, this time representing the conjunction *ɔr* 'and'. Although there are certain other structures, for the sake of space economy, what is discussed here are those with NN and AA constituents, both often being synonyms. Consider example (7):

7a. əsər (N)'influence' + rəsux (N) 'influence'→ əsər-**o**-rəsux (N) 'much influence'

  b. hʊsən (N) 'beauty' dʒəmal (N) 'beauty' → hʊsən-o-dʒəmal (N) 'all beauty'

Both the constituents are synonyms in the NN compounds, which convey plurality and emphasis. Such a formation is a collective noun, and it conveys emphasis on some collective actions. It forms agreement with a singular verb. It can be observed in the following sentence.

8a. ʊs **ka/*ke**   əsər-**o**-rəsux   zɪjaḓa   **he**
    his/her.gen.**s**   influence.interfix.influence.**s**.m   much.   **is**
    '(S)he has much influence'

 b. ʊs ***ka/ke**   vəsaɪl   zɪjaḓa   **hẽ**
    his/her.gen   resources.**pl**.m   many   **are**
    '(S)he has many resources'

c. ʊs **ka/ke**   zʊlm-**o**-sɪt̪əm   məʃhur   **he/hẽ**
   his/her.gen   tyranny.**s/pl**.m   famous   **is/are**
   1) 'His/Her tyranny is famous.'   2) 'His/Her tyrannies are famous.'

Sentence (8a) shows that the compound _əsər-**o**-rəsux_ is a collective noun and expresses plurality and emphasis. Therefore, it agrees with a singular verb. The genitive marker _ka_, agreeing with the compound, also gives evidence that it is a singular noun. One possible synonym of the compound is _vəsaɪl_ 'resources' (8b) that is a plural noun agreeing plural verb.

Many other NN compounds, as in (8c), have synonymous constituents and express plurality or a collective sense of plurality. The compound in (8c) is in fact a hybrid formation of the Arabic noun _zʊlm_ 'tyranny' and its Persian counterpart _sɪt̪əm_ 'tyranny/tyrannies'. The plural of the Arabic _zʊlm_ is _məzalɪm_ 'tyrannies', but the plural of the Persian _sɪt̪əm_ is the same. Note that apparently there is no head-modifier relation between the two constituents. However, we may see a superordinate or a subordinate in the two constituents, as in (8a), but it is not necessarily found in all NN compounds. The two constituents in (8c) do not show this, as they are from two different sources and synonyms of each other. N$_2$ _sɪt̪əm_ is equally compatible with N$_1$ _zʊlm_ in daily usage. The compound is used in both senses i.e. it shows plurality; and it also shows a collective sense of plurality. Therefore, it may agree with a singular or plural verb. Even if it is in a singular sense, it tries to convey a series of events with an emphasis.

Emphasis is particularly conveyed by the synonymous constituents of AA compounds. The role of _-o-_ structurally remains the same i.e. connecting the two AA constituents, but it shows changes semantically. It now no longer represents the conjunctive particle _ɔr_ 'and', but rather conveys an emphasis in the expression, as shown by the following examples:

9a. xʊʃ (A) 'happy' + xʊrrəm (A) 'happy' → xʊʃ-**o**-xʊrrəm (A) 'very happy'

 b. həsin (A) 'beautiful' +d̪ʒəmil (A) 'beautiful'→həsin-o- d̪ʒəmil (A) 'very beautiful'

Like those in (8a & 8c), the two constituents in (9a) and (9b) are synonyms. There is no head modifier relation between the constitutents, and so the resulting compound may be termed a hyponym of both the constituents. However, there is a superordinate, traditionally the first constituent i.e. *xɵʃ* 'happy' (9a) and *həsin* 'beautiful' (9b). The subordinates i.e. *xɵrrəm* 'happy' and *d̺ʒəmil* 'beautiful' also mean the same as the superordinates, but their meaning is only an additive to the semantics in the compound.

Thus, it is clear that among the two constituents, whether NN or AA in all compounds with the interfix *-o-*, it is generally the second constituent which plays as a semantic additive. However, there is a positional binding. The positions cannot be reordered whether or not they have superordinate/subordinate relations. This is mainly because of the wide use of $N_1/A_{1,}$ and so it is superordinate. On the other hand, N_2/A_2 is secondary and in little use.

In sum, there are four major formations of endocentric compounds, i.e. NN, AA, AN and NA, showing various structures mostly taking the interfixes *-e-* and *-o-*. Both the interfixes express some grammatical relations and help to convey possession, plurality and emphasis. The discussion was restricted to the functions of the interfixes. Durani (2007) limits it to just the representation of the genitive *ka* by *-e-*, and ignores all other functions by both the interfixes *-e-* and *-o-*. The former performs the additional function of representing the agentive *vala.* Although both representations of the genitive *ka* and the agentive particle *vala* by *-e-* give the semantic expression of possession, they are different and cannot be substituted by each other.

Playing a structural role as a synthesizer of the two constituents means that *-e-* is purely morphological in its function. It is neither a phonetic reduction nor a phonological substitution of any morpheme, unlike *-o-*, which is a phonological substitution of *-v-* that is itself a

phonetic reduction of the conjunctive particle ɔr 'and'. The discussion on -o- focuses on only two formations i.e. NN and AA (while both constituents being synonyms). The constituents do not play head-modifier roles in the traditional sense, but usually N1/A1 is superordinate according to the frequency of use. The compounds express plurality and emphasis.

Discussion of endocentric compounds was restricted to mainly the combinations of NN and AA in which the two constituents are often synonyms. The compounds are simple in the sense that they are unambiguous and thus easy to understand. Some exocentric compounds also show synonymy, but they are not as simple as endocentric compounds.

4.3. Exocentric Compounds

As Lieber and Scalise (2006) agree that, exocentric compounding has increased scholars' interest especially in relation to formations like En[glish] *blue eyed*; such constructions, though exocentric, represent a challenge to morphological Lexicalism. *Blue-eyed* soul (also known as white soul or pop soul) is a media term used to describe rhythm and blues and soul music, influenced by pop music. This specific term was first used in 1960s for the performance made by white artists. Thus, although the compound seems to point to someone having blue eyes, it may also refer to a specific piece of music. The Urdu exocentric compounds also show similar semantic ambiguity. Like those in endocentric compounds, the constituents in them are often (near) synonyms, but the compounds are sometimes ambiguous. For example, the two constituents *bal* 'child' and *bətʃtʃe* 'children' in *bal bətʃtʃe* 'family' are near synonyms, but the compound is not a hyponym. Exocentricity arises when it is difficult to take a decision on semantic headedness. However, as Katamba (1993: 321) points out, such compounds are syntactically headed in that inflections are added to the right-hand edge of a word. A left hand constituent functioning as modifier cannot be pluralised in an NN

compound, as in a) *birds-brain 'idiot' b) birds-brain**s**. Therefore, compound (a) is ungrammatical and compound (b) is grammatical. Bauer (2005) notes that this is in contrast to the situation in Spanish as reported in Olsen (2000: 912), where the plural of *actor-bailarín* 'actor-dancer' is *actores-bailarín**es***, thus apparently confirming the dual-headedness of such constructions in Spanish.

In Urdu, however, these compounds are not dual headed, although they may be right headed or left headed. This section discusses basic characterization and the forms of exocentric compounds. Some semantic issues with relevance to structures are also the focus.

4.3.1. The Structures

All Urdu exocentric compounds are descriptive in nature and, unlike endocentric compounds, contain no interfix that expresses any grammatical function. They are mostly restricted to AN, but there are also examples of NN, AA, NV and VN combinations. The constituents may be synonyms or near synonyms. Durani (2007: 12) considers that (Urdu) exocentric compounds "occur more often in adjectives rather than nouns." The adjectival formations are mostly composed of AN/AA. There are also some nominal formations, but others may be both nouns and adjectives. Like in endocentric compounds, constituents do no udergo structural changes.

10a. AN

 i) zəbər'top/great' (A) + ḍəsṱ 'hand' (N) → zəbər ḍəsṱ 'strong/vigorous' (A)

 ii) ṱəŋ 'narrow' (A) + nəzər 'sight' (N) → ṱəŋ nəzər 'narrow minded' (A)

 iii) tʃərb 'sharp' (A) + zəban 'tongue'(N) → tʃərb zəban 'talkative' (A)

 iv) mɪtʰi 'sweet' (A) + tʃʰʊri 'knife' (N) → mɪtʰi tʃʰʊri 'charming words to flatter' (N)

 v) xʊʃ 'happy' (A) + lɪbas 'clothing' (N) → xʊʃ lɪbas 'one who wears good clothes' (N)

  b. NN

i) bal 'child' (N) + bətʃtʃe 'children' → bal bətʃ-tʃe 'family' (N)

ii) ʃah (N) 'king' + xərtʃ (N) 'expense/cost' → ʃah xərtʃ 'spendthrift' (N)

 c. AA

i) mota 'fat' (A) + ʈəgṛa 'strong/fat' (A) → mota ʈəgṛa 'healthy' (A)

ii) ɖʊbla 'thin' (A) + pəʈla 'thin' (A) → ɖʊbla pəʈla 'weak' (A)

 d. NV

i) tʃɪɽi 'sparrow' (N) + mɑr 'kill/beat' (V) → tʃɪɽi mɑr 'fowler/ boaster' (N)

 e. VN

ii) hə̃s 'laugh' (V) + mʊkh 'face' (N) → hə̃s mʊkh 'cheerful/jolly' (A)

All the examples show various exocentric compounds, which are nevertheless not in a large number. Most formations are AN, but some examples of NN, AA, NV and VN are also seen. If the constituents are AN, the derivation may be a noun or an adjective. If both the constituents are nouns or adjectives, the derivations may remain unchanged in its category. Although there are not many examples, the combinations of NV/VN form either noun or adjective, or the compound may also be used as both an adjective and a noun. Bauer (2005) describes the Romance type of exocentric compound in English, made up of a verb and a noun which functions as the direct object of that verb; however, it is difficult to say in some cases that the first element is typically a verb or a noun. In contrast, the verb in Urdu exocentric compounds is in base form whatever the constituent position it takes. Moreover, disregarding the category of words, the right hand constituent is generally the morphological head. There is no example of left headedness in AN (the highest number among all combinations). NN, NV or VN formations are also right headed. Only AA combinations, e.g. *mota ʈəgṛa* 'healthy', (10c-i) are left headed and the modifier e.g. *ʈəgṛa* is nothing but a

semantic additive. The same relationship can be observed in other compounds e.g. ɖ̥o̽bla pə̱t̪la 'weak' (10c-ii). These compounds are nevertheless rare.

4.3.2. Semantic Expressions of the Structures

The headedness is however morphological, not semantic. As Bauer (2005: 7) points out, some linguists (e.g. Plag 2003: 146) consider that, while these compounds are semantically exocentric in that a *redcap* is not a kind of *cap* nor a *birdbrain* a type of *brain*, they are nevertheless morphologically headed: in *redcap*, *red* modifies *cap*, and in *birdbrain*, *bird* modifies *brain*. This also applies to Urdu exocentric compounds. The constituents show head modifier relation morphologically, but the compounds are headless semantically, as shown:

11a. mɪt̪ʰi 'sweet' (A) + tʃʰʊri 'knife' (N) → mɪt̪ʰi tʃʰʊri 'charming words to flatter' (N)

  b. t̪əŋ 'narrow' (A) + nəzər 'sight' (N) → t̪əŋ nəzər 'narrow minded' (A)

The adjective constituent is the modifier and the noun is the head in both the compounds. But semantically both are headless. The adjective *mɪt̪ʰi* 'sweet' in (11a) refers to something soft and alluring and the noun *tʃʰʊri* 'knife' refers to something sharp. The compound *mɪt̪ʰi tʃʰʊri* thus refers to charming/alluring words used to mislead somebody. Similarly, although the adjective *t̪əŋ* in (11b) means something physically narrow, it semantically refers to the approach of a person. The noun *nəzər* refers to a person who shows such an approach. The compound refers to a narrow minded person or an approach. The exocentric class contains many other compounds specifically referring to the nature of a person or his manner of speaking. The constituents are simple lexemes individually. However, in the compounds, as Bauer (2005) states, they do not denote the object to which they apparently refer. He calls the phenomenon synecdoche in which a lexeme, e.g. those in the compounds (11), is interpreted according to a well-known figure of speech. Figures of speech are productive, but some

become established as idioms or the words involved become polysemous. Thus, as he states, a syntactic (morphological) tree is so-called by a metaphor, and the tree becomes polysemous as a result of the usage. We can thus say that *ʈəŋ nəzər* and *mɪʈʰi tʃʰʊri* are established idioms.

Benczes (2004) terms such compounds "non-transparent" or "creative" compounds for metaphorical (and/or metonymical) noun–noun combinations. He states that with the help of cognitive linguistic tools, e.g. metaphor/metonymy, blending, these compounds are just as easily analysable and transparent as endocentric ones. Thus, it is a more imaginative word formation process. Urdu exocentric compounds are from various sources unlike Romance type English exocentric compounds which are only really exocentric compounds, as Bauer (2005) states. They are not restricted to a single source or some specific structure. Although most are from native Urdu, the non native origin of a constituent is also frequent. The origin of a constituent from diverse sources also gives the compounds various metaphorical meanings. For example, take *ʈəŋ nəzər* 'narrow minded' and *ʈəŋ dəsʈ* 'poor'. Both compounds contain the native modifier *ʈəŋ* with the meaning '(something physically) narrow'. It expresses various semantics depending on the head noun. In the former, with the Arabic head *nəzər* 'sight,' it refers to the approach of a person. But in the latter, with the Persian head *dəsʈ* 'hand', it refers to poverty. *ʈəŋ* is a distinctive compound word and, as the examples show, forms various compounds. Aronoff (1976: 45) points out that the existence of one of these (lexical items) with a particular meaning does not appear to block its existence with another unrelated meaning; this is important because lack of blocking is usually considered to be a sign of a productive process. This means that the distinctive compound words in Urdu are productive in nature, because each of them forms various compounds expressing distinctive semantics, as also in exocentric compounds. Qureshi and Akram (2008: 8) verify that word

compounding is a productive process to form new words in Urdu; one meaning of a constituent does not block its appearance with another, unrelated meaning.

The two constituents may be synonyms or near synonyms, e.g. *bal bətʃtʃe* 'family' (11b-i), disregarding any word. Keane (2001) states that Hindi (Urdu), like other Indian languages, is rich in exocentric compounds comprising two synonyms or near synonyms:

12a. ḍʊbla (A) 'thin' + pət̪la (A) 'thin' → ḍʊbla pət̪la 'weak' (A)

  b. bal 'child' (N) + bətʃ-tʃe (N) 'children' → bal bətʃ-tʃe (N) 'family'

The constituents in (12a) have the same meaning literally, but semantically they refer to different objects. ḍʊ*bla* is related to an animate object, and *pət̪la* to an inanimate object. However, it semantically functions as an additive to ḍʊ*bla*. The compound conveys an emphasis in the meaning 'weak' for a man. N_2 in (12b) is not exactly the same in meaning but refers to the plurality of N_1. This also shows the complex nature of Urdu exocentric compounds. There are variations in structures, but there are also variations in the constituents' semantic expressions. A modifying constituent may be an additive to the meaning of a head constituent (12a), it may show a singular form of the plural head (12b), and it may also change semantics depending on the source of a head.

In sum, there are not a large number of Urdu exocentric compounds, but they form various combinations, e.g. AN, NN, AA, NV and VN. They are metaphorical and creative in semantics, a feature which Bauer (2005) and Benczes (2004) note in English exocentric compounds. However, unlike the English ones which are only from Romance languages, they are not restricted to a single source of their constituents or only to one type of combination. Each constituent keeps its lexical status, and unlike some other compounds the modifier is not

incorporated and fused with the head constituent. Incorporation of a constituent literally means the merger of one constituent into the other, and losing its lexical identity. Unlike the previous two types of compounds discussed, the incorporation of a constituent is only seen in Urdu copulative compounds, which are also in small number but in various formations.

4.4. Copulative Compounds

A copulative compound shows no head modifier relation but a coordination between the two constituents, e.g. *actor-manager* 'actor and manager'. The two constituents are independent, and thus the compound has two semantic heads. Urdu copulative compounds are relatively few in number as compared to endocentric and exocentric compounds. Nonetheless, they show a greater variety and more structural changes. Unlike those of endocentric and exocentric compounds, their constituents are fused, and one is completely incorporated into the other. Incorporation literally means the merger of one constituent into the other, but here it may also convey the merger of both constituents, and a compound seems to be a single word.

This section discusses copulative compounds focusing on the claims made by Koul (2008: 73) that Hindi (Urdu) copulatives are composed of semantically-related nouns; each noun behaves as an independent constituent in the sense that each may be separately inflected for gender and number, though not for a postposition. There are two points to argue: 1) the composition of copulatives is not made by nouns only 2) there is a possibility of only second constituent taking gender and number marking. It is also agued against Durani (2007) that Urdu copulatives "refer to two or more morphemes connected in a sense by a conjunction."

4.4.1. The Structures

Copulative formations are both with and without interfixes. However, the interfix is only *-o-*. There are thus two forms of structures i.e. with the interfix *-o-* and without any infix. The role

of *-o-* in copulatives is structurally the same as noted in Section 2, as it functions as a conjunctive particle. But, there are semantic differences in its use in copulative compounds. Copulatives show structural differences from endocentric compounds. Most combinations are NN, but some of AA and NA formations are also seen, as the examples illustrated below:

13a. NN

  i). sər (N) 'head' + ṭadʒ (N) 'crown'→ sər ṭadʒ 'husband'   (N)

  ii) zəmin 'earth' + asman 'sky'→ zəmin-**o**-asman   'universe'   (N)

  b. AA

  i) xas 'special' + am 'ordinary' → xas-**o**-am (Pro) 'everyone'   (Pro)

  ii) sejah 'black' + səfeḍ 'white' → sejah-**o**- səfeḍ 'everything'   (Pro)

  c. NA

  i) ḍɪl 'heart' ʃəspã 'stick' → ḍɪltʃəsp 'interesting'   (A)

  d. NV

  i) sər 'head' + phɪr 'turn' → sər phɪr*a* (m) sər phɪr*i* (f) 'mad/crazy'   (N/A)

There are various formations of NN, AA, NA and NV. The resulting compounds are nouns, pronouns, adjectives, adverbs and pronouns. The formations may be with or without interfixes, which mostly appear in NN and in AA. There may be structural changes in the compounds without the interfixes, but the changes are often caused by *-o-* and by gender and number. Many changes are not merely by affixation, but there is a complete incorporation of a constituent, e.g. in (13c), as explained in the following subsections.

4.4.1.1. Functions of -o- in Copulatives

Although the functions of **-o-** seem to be simply an expression of collectivity, it shows some comprehensiveness not restricted to the object mentioned. Consider the following:

14 a. zəmin (N) 'earth' + *a*sm*a*n (N) 'sky'→ zəmin-**o**-*a*sm*a*n (N) 'universe'

   b. sej*a*h (N/A) 'black' + səfe*ḍ* (N/A) 'white'→sej*a*h-**o**-səfe*ḍ* (Pro)'(ownership) everything'

   c. x*a*s (A) 'special' + *a*m (A) 'ordinary' → x*a*s-**o**-*a*m (Pro) 'everyone'

The appearance of *-o-* is limited in copulatives, as compared to that in endocentric compounds. The interfix *-o-* in (14a) conveys collectivity and implies all and everything in the universe. The compound is formed from the Arabic *zəmin* 'earth' and Persian *asman* 'sky'. This shows that Urdu copulatives may also be hybrid formations. Similar collectivity is also seen in (14b). The compound *sejah-**o**-sfəfeḍ* also shows emphasis on somebody's ownership of a certain thing e.g. a business with control of everything related to it, and gives the sense of absolute authority, which nobody can share. Although it expresses simple collectivity, the same as in endocentric compounds, the functions of *-o-* in the copulatives show different semantic expressions from those of the former. The plurality or collectivity expressed in the endocentric compounds is in a specific sense and is limited to certain factors. But the plurality or collectivity expressed in copulatives gives a comprehensive sense. Recall that in endocentric compounds, if NN constituents are synonymous, the compound is a collective noun, e.g. *əsər-**o**-rəsux* 'much influence'; it conveys emphasis on some collective actions, and it agrees with a singular verb. However, a copulatives formation may be a noun, but the two constituents are not usually synonyms, and therefore the compound may agree with a singular or plural verb. In both the formations of NN and AA, the constituents may be nouns and adjectives and the resulting compounds are generally nouns and pronouns i.e. N-o-N or A-o-A → [P] / [N]. When the infix *-o-* appears with two nouns, i.e. N-o-N compounds, it shows plurality/collectivity and sometimes emphasis, but when it appears with two adjectives i.e. A-o-A compounds, it shows mainly emphasis, as in (14c). Thus, the role of the infix *-o-* here is

distinctive. However, many other compounds do not take any infix and even then their constituenst are incorporated.

4.4.1.2. Copulatives without Interfixes

Booij (1992: 14) considers regarding Dutch that, "There are copulative adjectival compounds such as *rood-wit-blauw* 'red-white- blue' and *Duits-Frans* 'German-French', but it is hard to decide whether they should be considered normal compounds with a right head, or as cases of word-internal asyndetic[8] coordination". Urdu copulative compounds may sometimes be termed the cases of word internal asyndetic coordination, e.g. *sər ʈadʒ* 'husband' (16a), as they need no interfix. Moreover, the two constituents often show complete incorporation without requiring any interfix. However, there may be structural differences:

16a. sər (N) 'head' + ʈadʒ (N) 'crown'→ sər ʈadʒ 'husband'   (N)

  b. ɖɪl (N) 'heart' + ʃəspã (V) 'stick' → ɖɪltʃəsp 'interesting'   (A)

  c. sər (N) 'head' + fəroʃ (V) 'sell' → sər fəroʃ 'very brave'   (A)

  d. sər (N) 'head' + phɪr (V) 'turn' → sər phɪr*a*[9] (m) sər phɪr**i** (f) 'mad/cr*a*zy'   (N/V)

The various examples present two cases of incorporation. Unlike those in (16b-16d), the two constituents in (16a) are both nouns, and are completely incorporated into each other, and the compound is never treated as a compound but as a single lexical item. There is also a possibility that the genitive *ka* may be inserted, but the combination then no longer remains a compound. Rather, it is a phrase *sər **ka** ʈadʒ* 'husband' without any effect on the meaning. Thus, breaking apart the two constituents restores their independent lexical status, while in the compound both lose their individuality. Examples (16b-16d) show a complete incorporation

---

[8] Having no conjunction.
[9] phɪr → phɪr*a*/ phɪr**i** (past participle form)

in which a word loses its distinctive identity and shows a merger into the other. The second constituent *ʃəsp* 'stick' in (16b) is the reduced form of the verbal noun *ʃəspã* that is itself a restructured form of the Persian infinitive *ʃəspɪɖən*. *ʃəspã* is no longer used in Urdu, and due to its merger into the first constituent *ɖɪl* its distinctive identity cannot be seen. Although the compound *ɖɪltʃəsp* literally means 'something that sticks to heart', it refers to something interesting. The compound is not treated as a compound. It is used as a single word instead. The second constituent in (16d) shows gender marking.

Katamba (1993: 286) notes that there may be a choice between using independent words and incorporating words such as object NPs, verbs or prepositions into verbs to form complex words. He considers that incorporation is affected by rules that are essentially the same as syntactic movement rules which shift around the constituent of sentences. Morphological incorporation puts one word inside another word rather than shift it to another place. Thus, there are important morphological changes. This is the case with Urdu copulative *ɖɪltʃəsp*. One constituent *ʃəspã* ends up inside the word *ɖɪltʃəsp* that is in fact a compound. However, in contrast to what Katamba notes, a verbal constituent is incorporated into a nominal one.

This is also the case with the verbs *fəroʃ* in (16c) and *phɪr* in (16d), though there is a little difference. The second constituent is a verb, and there is a noun-verb relation. Katamba (1993: 285) states that in such a relation the noun functions as an object of the verb, and once incorporated "the object NP becomes an integral part of the verb and loses its separate identity." Such incorporation, he states, creates complex words. The two compounds (16c and 16d) are also different from Katamba's approach in the sense that the incorporated constituent is the verb. Like *tʃəsp* (16b), the second constituents *fəroʃ* (16c) and *phɪr* (16d) are incorporated into the nominal constituents. This shows a difference from Katamba's

observation. In his case, the formation is in fact a verb, and the object NP is incorporated to be part of the verb. But in Urdu, the formation is a noun or an adjective, although the literal meaning shows that the nominal constituent functions as an object of the verb. These compounds may also show that the second constituent is incorporated.

The incorporation of the second constituent also shows differences from the observations made by Koul (2008: 73). Some important aspects noted in the Urdu copulatives are in sharp contrast to his approach to Hindi (Urdu) copulatives. His description[10] shows two weaknesses i.e. the limitation of copulatives to NN and gender/number marking to both constituents.

4.4.2. Arguments against Koul and Durani

Koul (2008: 73) claims, "Hindi (Urdu) copulative compounds, also known as co-compounds, are composed of semantically-related nouns. Each noun behaves as an independent constituent in the sense that each may be separately inflected for gender and number, though not for a postposition." This examination is divisible into two parts: 1) Hindi (Urdu) copulative compounds are composed of semantically-related nouns 2) each constituent may be separately inflected for gender and number. The first part of the claim is partially true. Urdu copulatives are not necessarily composed of NN only, but also of AA, NA and NV:

17a. xas (A) 'special' + am (A) 'ordinary' → xas-o-am (Pro) 'everyone'

   b. ḍɪl (N) 'heart' ʃəsp (A) 'sticky' → ḍɪltʃəsp (A) 'interesting'

The various combinations include NN, AA and NA. Although his claim of semantic relation of nouns seems true in many of the examples, e.g. in (13a.ii), (13bi & 13b.ii) where the constituents seem antonyms, it is not justified in other compounds e.g. *ḍɪltʃəsp* (17b) in which

---

[10]. Koul (2008) has just described Hindi (Urdu) copulative in a few lines. Unfortunately, he gives no data.

no direct semantic relation between the constituents is seen. So, it can be argued that a direct semantic relation is not necessary. However, a relation is seen only after the incorporation.

Secondly, Koul (2008: 73) observes that each (constituent) may be "separately inflected for gender and number." Various examples show that only the second constituent is inflected.

18a. sər phɪr*a* (m) / sər phɪr**i** (f) 'mad/crazy' → sər phɪr**e** / sər phɪr**ɪjã**

  b. ɖɪltʃəsp 'interesting' → ɖɪltʃəsp**i** 'interest' (f) → ɖɪltʃəsp**ɪjã** 'interests' (f)

Example (18a) is noun, and (18b) is an adjective, which is converted into a feminine noun after affixation of -*i*. The compounds show the gender and number marking in the second constituent only. There is no change in the first constituent. sər *sər phɪra* (18a) is originally a masculine noun. It is changed into a femine noun and the plural forms by the alternation of coda vowel i.e. -*i*, -*e*, and -*ɪjã* respectively in the second constituent. The original compound in (18b) is derived from the Persian loans '*ɖɪl*' and *ʃəspɪɖən* 'to stick'. Both entered the Urdu lexicon at different periods of time, and so they do not show gender distinction. However, after derivation of the abstract noun, *ɖɪltʃəspi*, it shows the feminine gender in the second constituent. Then by the alternation of final vowel -*i* with -*ɪjã*, it shows feminine plural marking in the second constituent. All the examples show gender and number marking in the second constituent only. If Koul's (2008: 73) claim were to be accepted, the first constituent in the two compounds should be inflected for gender/number: (a) **səraphɪra* (b) **ɖɪiltʃəspi*. But the compounds are ungrammatical. This proves his claim to be invalid.

Durani's (2007: 14) observation is even more restricted. Interfixes in Urdu copulatives may also function like a conjunction, but certain points do not validate his claim that Urdu copulatives "refer to two or more morphemes that can be connected in a sense by a

conjunction." Although he means to convey that the two constituents are interpreted as if a conjunction were present, this interpretation is not always necessary. As noted above, the interpretation of a conjunction is possible with the interfix -o-. Consider the following:

19a. sejah 'black' + səfeḏ 'white' → sejah-**o-** səfeḏ (N) 'everything'

   c. sər 'head' + ṭadʒ 'crown'→ sər ṭadʒ 'husband'

A copulative may appear in various forms. It may show an interfix -o- functioning like a conjunction ɔr 'and' with its full semantics, as in (19a). But example (19b) denies the first possibility and shows that the presence of an interfix is not necessary. In fact, he limits structural and semantic variety of Urdu copulatives. The two constituents are incorporated and connected without any conjunction or any interfix representing a conjunction.

Olsen (2000) states that it is unclear from traditional discussion, whether copulative compounds form an independent class of compounds or determinative compounds or if they are simply a formal and/or sub types of more basic determinative patterns. Like Urdu endocentric compounds, some copulatives show grammatical relations, and so they are determinative compounds. However, they are distinctive in various other aspects, e.g. some of them contain no interfix showing any grammatical relation. Therefore, they should be treated as an independent class. This work has tried to clarify two major structural aspects of the copulatives with and without the interfix -o- with various formations of NN, NA, AA and NV. Whether or not they take interfixes, they are distinctive in that their constituents are often completely incorporated. The interfix -o- functions like the conjunction ɔr 'and', which has no role in the compounds. However, Urdu copulatives do not necessarily take any conjunctive interfix, which is in contrast to Durani's (2007) claim. Contra Koul (2008: 73), Urdu copulatives are not restricted to semantically related nominal constituents. They may show

gender and number distinctions but only by marking the second constituent, and this is as predicted by the right-hand head rule. The semantics of *-o-* is wider than that in endocentric.

More unique is the interfix *-b-*, whose structural role, in some determinative compounds, is similar to those of *-e-* and *-o-*, but its semantic link is seen with the second constituent. *-b-* functions in the manner of postpositions in compounds. These compounds cannot be semantically substituted by any postpositional phrase. This is unlike the compounds formed by *-e-* and *-o-*, which have the semantic substitutes of the phrases formed by the genitive *ka*, the conjunctive *ɔr* 'and' and particle *vala.*.

4.4.3. Structural and Grammatical Role of *-b-* in Some Determinative Compounds

This section discusses some general determinative compounds focussing on the features of the third infix *-b-*. In all such compounds, which are few as compared to others, the common element is the Persian loan noun *sər* 'head' in the N_1 position. The second constituent in N_2 position may be a native noun or a loan noun, e.g. Arabic *sɪdʒd̪a* '(head down) in prayers', as shown below, and the formation is also a noun.

15a. sər 'head'   + sɪdʒd̪a '(head down) in prayers' → sər**b**sɪdʒd̪a 'showing respect'

  b. sər 'head'   + rah 'path/way'   → sər**b**rah 'leader'

  c. sər 'head'   + d̪əst̪ 'hand'   → sər**b**d̪əst̪ 'bravely (lit: one's head on his hand)'

The infix *-b-* appears after one noun, and so it seems allomorphic. However, it does not have any other variant. It[11] seems affixed more with the second constituent morphologically and functions in the manner of postpositions to indicate *the position above/supported by/in contact with* (something). Example (15a) literally means 'one's head down in prayers',

---

[11]. Explanation given by the dictionary from CRULP (*Centre for Research in Urdu Language Processing*, Islamabad)

though, it symbolically refers to showing respect to God, particularly, or even to a human. Similarly, (15b) literally means someone 'bringing heads (people) on the path' i.e. the leader, while (15c) literally means one 'having his head in his hands', i.e. it refers to someone doing something while putting his life in danger.

Thus, by functioning in a postpositional manner, -b- actually performs a grammatical role. Therefore, removing it from the compounds renders them ungrammatical e.g. *sər-rah and *sər-sɪdʒd̪a. However, it does not mean that it is a replacement of postpositions, e.g. mẽ 'in' in (15a). Like the genitive ka/ki and conjunctive particle ɔr 'and', postpositions do not function as affixes in the compounds; they do not have any morphological roles and only function syntactically. Therefore, their semantic meaning in the compounds is only conveyed by -b-. Moreover, unlike those of -e- and -o-, the grammatical use of -b- in compounds cannot be substituted for by postpositions in phrasal forms. Some phrases can be formed by the genitive ka/ki and the conjunctive ɔr, but there can be no phrases formed by postpositions e.g. pər 'on' in *rah pər sər, which gives the semantic sense of sərbrah 'leader'. This is unlike the functions of -e- and -o-, whose substitutes, i.e. the genitive ka/ki and the conjunctive particle ɔr are commonly used in phrasal forms. Thus, the grammatical role performed by -b- is unique in this sense.

In addition to the grammatical role, -b- performs one more peculiar function not performed by -e- and -o-. The two constituents separated by -e- and -o- keep their distinctive identity, and so their combination is presumed to be a compound. In contrast, -b- helps to fuse the two constituents completely   together into a word, so much so, that the compound seems to be a single entity. The compound is not taken as a compound; rather, it is treated as a single lexical item functioning as a noun, adjective or an adverb, as in (15a-15c).

There are several reasons to treat it as a single word. The first of them is morphological. The constituents are completely incorporated, and the second constituent particularly loses its lexical status due to -*b*-. It is this function that makes -*b*- distinctive and unique. As said earlier, removing -*b*- is not only ungrammatical but also results in a nonsense phrase e.g. *sər-ḏəsṭ. This is also unlike the compounds with -*e*- and -*o*-, e.g. *sejah-o-sfəfeḏ*, which retain some meaning after the removal of the infixes.

Secondly, unlike those with -*e*- and -*o*-, the compounds with -*b*- are orthographically written as single word. -*b*- seems affixed more with the second constituent. In this status, if separated apart, e.g. ***b**rah* form *sər*, it makes no sense. Only the first constituent seems to constitute a single word written.

The third reason is to do with the semantics. -*b*- is semantically linked in the postpositional manner with the second constituent. A formation, e.g. *sər**b**rah* (15b), literally means 'one who brings heads *on the path*', but it semantically refers to a leader. Generally speakers of Urdu are not aware of literal meaning of the compound and know only the surface meaning, considering the compounds as single words.

-*b*- is thus distinctive from the two other infixes in its grammatical role, incorporation and semantics. The Ffirst constituent *sər* is a Persian loan and the second constituent may be native or Arabic, but the source of any constituent does not affect the function of -b-. Its use is limited to some compounds, but it plays a vital role in the incorporation. Unlike those of -*e*- and -*o*-, the compounds by -*b*- do not have substitutes in the form of postpositional phrases.

Urdu postpositions themselves form certain compounds and show some distinctive features, but such compounds are not numerous.

4.5. Postpositional Compounds

There are only a few postpositions in Urdu, but they play a key role in some compound constructions. They include the instrumental/ablative *se*, the fused forms of the genitives i.e. *ke* and *ki*, *ṯərəf* 'towards' and *xɪlaf* 'against'. As discussed by Bauer (2005), a phrasal type of exocentric compound in English, e.g. *showoff,* may be a noun that is made up of a verb + a preposition/adverb/particle; there is no noun involved in *showoff*, yet the construction as a whole is a noun. He acknowledges that people are reluctant to accept this analysis, as phrasal verbs are often considered to be syntactic constructions, while nouns are single lexemes. Whether or not the verbs are lexemes, they are like listemes, in Bauer's words, the same as nouns. All morphemes, irregular forms, and idioms are listemes that must be memorized because their sounds or meanings do not conform to some general rule. Urdu postpositional compounds are different from the English phrasal type compounds in that they are neither noun formations nor do they involve verbs. Rather, they are adjectives and adverbs. Moreover, a postposition has a fixed position in such compounds. For example, the instrumental/ablative *se* always appears as the second constituent and the genitive *ke* or *ki* appears as the first constituent. Their compounds convey various semantics, as shown below:

Table 1. Postpositional Compounds

Constituent 1	I	Constituent 2	Compounds
1. *ara*m 'rest' (N)		se (instru/ ablative)	*ara*m se 'peacefully' (Adv)
2. genitive ka/ki/ke		lɪje 'concerning'	ke lɪje 'for'
3. ḍo 'two' (N)		ṯərəf 'side/towards'(PP)	ḍo ṯərfa 'bilateral' (A/Adv)
4.mənḍərdʒ 'inserted' (A)		zel 'below' (PP)	mənḍərdʒa zel 'given below'(A)

Note: pp —postposition (3,4), I - Interfix

The most important postpositions are *se* and *ke/ki*, and therefore they are the focus here. They form a large number of compounds with nouns, adjectives and postpositions, and express various semantics. *ke* and *ki* mainly form postpositions and *se* mainly forms adverbs. It is noteworthy that *ke* and *ki* are homophonous to the genitives *ka/ke/ki* and seem the same, but in fact they are different from genitives. Genitive markers *ka*, *ke* and *k*i show gender and number distinction, while *ke* and *ki* do not. As Thakur (1997: 14) points out, Vajpeyi (1958: 30) is the first Hindi grammarian to differentiate the genitive case-markers, i.e. *ka/ke/ki* from *ke* which occurs with compound postpositions. He notes that /-a/, /-i / and /-e/, in the genitive *ka/ke/ki*, are "agreement markers and *k* in them is the genitive case marker"; postpositions denote different cases and, on this basis, grammarians have categorized *ka*, *ke* and *ki* as postpositions which denote the genitive case. Thakur (1997: 39) considers that *ke* is the fused form and functions like the genitive, but it expresses no gender and number distinction. His observation is nevertheless limited to *ke*. Another fused form, *ki*, also functions in the same manner. A comparison with the genitive markers reveals an interesting picture:

20. **Group A.**　　　　　　　　　　　　　　　　**Group B**.
　a. ləkir　　k**a**　　　sɪr**a**　　　　　　　d. ləkir　　ke　　upper
　　　line.　　gen (m)　edge (m)　　　　　　line.　　ke.　above
　　　'the edge of the line'　　　　　　　　　　'above the line'

　b. ləkir　　k**i**　　ləmba-**i**　　　　　　e. Joe　ki　　t̪ərah　　se
　　　line.　　gen (f).　length (f)　　　　　　Joe　ki　　like.　　instr
　　　'the length of the line'　　　　　　　　　'like Joe'

　c. ləkir　　k**e**　　nɪʃan**at̪**　　　　　　f. Anne　　ke　　　bare　　mẽ
　　　line.　　gen (pl)　marks (pl)　　Anne　　ke　　　about　loc.(in)
　　　'the marks of the line'　　　　　　　　　'about Anne'

As *ki* functions the same as *ke*, it should also be treated as the fused form of the genitive. However, it forms fewer compounds, e.g. *ki t̪ərəf* 'towards' and *ki dʒanɪb* 'towards'. The genitives *ka*/*ke*/*ki* in Group A and their fused forms *ke* and *ki* in Group B are different from

each other. The genitives *ka*/*ki*/*ke* in Group A show possession and gender/number distinction due to their agreement with the following nouns i.e. *sɪra* 'edge' (m), *ləmba-i* 'length' (f) and *nɪʃanaʈ* 'marks' (pl). The combinations are simple phrases and not taken as compounds. The fused forms *ke* and *ki* in Group B do not show possession and gender/number distinction. They are followed by an adjective, postposition and noun (in rare cases) with which they form compounds. None of the three compounds are grammatical without *ke/ki*, e.g. **ləkir upper*. Not showing gender, number and agreement morphology means that they must be playing some other role. This is to form some kind of relation between the preceding noun and the following element, usually postposition or adjective. Without them, the position of a noun cannot be specified e.g. **Joe ʈərah se* (20e).

There are two major differences between the genitives *ka*/*ke*/*ki* and their fused forms *ke*/*ki*. Genitives show gender and number morphology in agreement with the following noun. They are not followed by any other category word. The fused forms *ke* and *ki* do not show possession, gender and number morphology. They are followed by adjectives, postpositions, and nouns (in rare cases) with which they form compounds as shown below:

21a.   ke + bare 'about' + mẽ (locative)   → ke bare mẽ 'about'

   b.   ki + ʈərah 'like' + se (instrumental) → ki ʈərah se 'like'

Both *ke* and *ki* combine with adjectives in (21a) and postposition (21b). The combinations with postpositions also involve the ablative/instrumental *se* and the locative *mẽ*. The major point to explain is that genitives involve the relation of NP with an NP, but the fused forms generally make some relation between the preceding noun and the following element, usually a postposition or adjective. The homonyms involve postpositions and adjectives which are nevertheless in a limited number.

Although they express some semantics, e.g. purpose, function, reference, *ke* and *ki* are deficient in productivity as compared to the instrumental *se* with which they combine and form compounds. The instrumental/ablative *se* in contrast is a productive element in Urdu, as it derives a lot of new words from native and loanwords, and conveys meaning. As the name shows, it functions as instrumental, e.g. *zəban se* 'with tongue' and source of action, e.g. *Karachi se* 'from Karachi'. Ahmed (2008) states that instrument markers in South Asian languages show syncretism, i.e. the merger of two or more originally different inflectional forms with different markers. There is in fact no syncretism pattern; some South Asian languages have the examples of comitative instrument (marking companions). Urdu/Hindi uses the postposition *saath* 'with' as the comitative marker, e.g. *mere saath* 'with me'. But, *se* shows secondary comitative marking.

The ablative/instrumental *se* forms adverbs of manner and comparative/superlative degrees of adjectives. The formation of compound adverbs is simple and is only made by attaching to nouns. *se* sometimes links two lexemes, particularly two adjectives to construct an adverb.

22a.   zor 'force' (N   → zor **se** 'forcefully'   (Adv)

  b.   məkkari 'cunningness' (N)   → məkkari **se** 'cunningly' (Adv)

  c.   ətʃʰi 'good' (A) + t̪ərah 'like' (A) → ətʃʰi t̪ərah **se** 'well'   (Adv)

Both (22a & 22b) show simple adverbial constructions, involving nouns and instrumental, but (22c) shows a combination of three elements i.e. A + A + Inst → Adv. Such formations are however rare. The adverb formations are thus in two categories, although the manner of forming adverbs is the same.

However, there are two ways to form comparative/superlative degrees. In the first, like adjectives in English, Urdu adjectives also take comparative and superlative degree markers, -t̪ər, -t̪ərin, which are borrowed from Persian and attach to adjectives in the same manner:

23a. sərd̪ 'cold'   → sərd̪t̪ər 'colder'   → sərd̪ t̪ərin 'coldest'   (Persian)

   b. bʰond̪a 'odd'   → bʰond̪at̪ər 'odder'   → bʰond̪at̪ərin 'oddest'   (Native)

   c. bəd̪sʊrərt̪ 'ugly' → bəd̪sʊrərt̪t̪ər 'uglier' → bəd̪sʊrərt̪ t̪ərin 'ugliest'   (Arabic)

Although most formations may be Persian borrowings, as in (23a), the Persian degree markers may also form comparitives and superlatives with native (23b) and Arabic loanwords (23c). Example (23c) is particularly interesting. The base form is already a hybrid compound of the Persian *bəd̪* 'bad' and the Arabic *sʊrərt̪* 'face'. Moreover, it takes Persian degree markers for comparitives and superlatives. Persian formations may be of the greatest frequency individually, but the formations with native and Arabic words as a whole are certainly more numerous than them. These formations are primarily inflectional except for a few hybrid compounds, e.g. *bəd̪sʊrərt̪* (23c), which take the degree markers.

Degrees of adjectives are natively formed derivationally by using the instrumental/ablative *se* in combination with the Arabic loan adjective *zɪjad̪a* 'much' for comparatives and an additional native adverb *səb* 'all' for superlatives. Such compounds are also not entirely native formations, and there is no distinction of the source of base form, as shown below:

24a. bəɽa 'big/old'   →   Wania **se** (*ziad̪a*) bəɽa 'older than Wania'

                   →   səb **se** (*ziad̪a*) bəɽa 'oldest'   (Native)

   b. paed̪ar 'durable' → lohe **se** (*zɪjad̪a*) paed̪ar 'more durable than iron'

                   →   səb **se** (*ziad̪a*) paed̪ar   (Persian)

c. həsin 'pretty'   →   Asifa **se** (ziaḏa) həsin 'prettier than Asifa'

              →   səb **se** ziaḏa həsin 'prettiest'   (Arabic)

The three base forms are taken from the three sources. *se* functions differently but conveys the sense of comparative/superlative degrees. The Arabic adjective *zijaḏa* 'much' is optional and often ignored, when two objects of comparison are mentioned. Thus in such a case, *se* for comparative and the phrase s*ə*b *se* (s*ə*b 'all') 'among all' for superlative are sufficient for the degree formations.

Both the ablative *se* and the fused forms of the genitives *ka*/*ki* play distinctive roles, but their formations are not as varied as the other compounds, e.g. endocentric compounds. *ka* and *ki* are different from the gentitive markers *ka*, *ki* and *ke* in that they do not show possession and gender/number marking. Unlike genitive markers, which only form a phrase, they play a key role in compound formation, but the compounds are relatively few in number. The ablative/instrumental marker in contrast is not so limited in compound formation. It forms qualitative adverbs and degrees of adjectives, for which it combines with the various base verbs of the three sources. Thus, it seems to be productive enough to form certain compounds.

There are a variety of Urdu postpositional compounds. They are unlike the phrasal type of exocentric compounds in English e.g. *showoff*, which are actually made up of a verb + a preposition/adverb/particle, and the derivations are nouns. But, there are also verbal compounds, in which one constituent is a noun, adjective or a verbal noun but the other must be a dummy verb.

4.6. Verbal Compounds

Urdu verbs come in two types. A large number of verbs are already present in the base form. For example, a lexical verb *pi* 'drink' requires the suffix -*na* to form the infinitive *pina* 'to

drink'. It takes other suffixes according to tense and aspect needs, e.g. *pija* 'drank' and *pirəha* 'drinking'. However, many times verbs do not exist in their verbal capacity to describe an action. Therefore in contrast to those inflectional in nature, a large number of verbs are derivationally constructed in compound forms from nouns and adjectives (particularly loans) in combination with some dummy verbs.

A dummy verb in Indo Aryan languages is in fact a main verb but also functions as a tool, in combination with a noun or an adjective, to form a compound verb in the absence of a base verb. In Urdu, it is often a nativised form of a Persian auxiliary. For example, *kərna* (from *kərḍən*) 'to do' and *hona* (from *buḍən/ʃoḍən*) 'to be' show the nativisation of Persian loans. The Persian suffix *-ən* shows the infinitive marking, but its nativisation is in two different ways. Urdu borrows a large number of Persian verbs but doesn't accept the infinitive marker *-ən*, which is replaced mostly either by a dummy verb, e.g. *foruxṯən* becomes *fəroxṯ **kərna*** 'to sell', or with the native infinitive marker *-na,* e.g. a verb *xəriḍən* becomes *xəriḍ**na*** 'to buy'. Discussing the Arabic influence on Urdu, Versteegh (2001: 497) states that it appears that all these formations follow the pattern of Arabic loans in Persian. This is possibly because the Arabic loans are commonly used with the Persian verbs *kərḍən* 'to do' and *buḍən/ʃoḍən* 'to be' etc. However, as noted, some dummy verbs are themselves modified forms of Persian loans, which shows that a compound verb may be a hybrid compound in Urdu.

Moreover, a loan verb alone cannot function as a verb. It is dysfunctional; it merely functions as a verbal noun, and it has to combine with a dummy verb, which carries the inflectional information. The morphological functions of a dummy verb are the same, and only a dummy verb shows gender and number or tense and aspect changes. Thus, all loans, whether nouns, adjectives or even verbs, must be integrated with dummy verbs to form compound verbs. For

example, an Arabic loan *ɪzzəṭ* 'respect' is used in the verb form with a dummy verb as *ɪzzəṭ kərna* 'to respect' (infinitve) and *ɪzzəṭ ki* 'respected' (past). Thus, a loanword functioning as a verb must be used in a compound form.

Sometimes, a verbal base exists and is fully functional, but it is also used as a verbal noun and is combined with a dummy verb for the formation of a compound verb. However, there is a change of semantics. For example, *ḍɔɽ* 'run' is a base verb, but it is also used as a noun. The infinitive form of the base verb is *ḍɔɽna* 'to run'. But the noun *ḍɔɽ* 'a run' combines with a dummy verb *ləgana* 'to touch' to form a simple verbal structure in the infinitive form *ḍɔɽ ləgana* 'to run'. It is neverthelss semantically different from *ḍɔɽna*, as it involves a volitional act of running. The dummy verb *ləga* is a transitive verb and shows subject's intention. So, in *ḍɔɽ ləgana*, it expresses subject's willingness for the act of running, which is not seen in *ḍɔɽna*. The use of a verbal noun with a dummy verb in this manner also shows that a loan verb depends on a dummy verb to form compound verbs.

The number of dummy verbs is judged to be between 12 to 15, but the most common verbs used as dummy verbs are *kərna* (from *kərḍən*) 'to do', *hona* (from *buḍən/ʃõḍən*) 'to be', *lena* 'to take', *ḍena* 'to give', *dʒana* 'to go' and *ana* 'to come' etc. Schmidt (1999: 101) explains that in general, Urdu verbs demonstrate a very regular conjugation with the exception of five verbs *ho* 'be', *kər* 'do', *ḍe* 'give', *le* 'take', and *dʒa* 'go'. These five verbs primarily function as main verbs, and secondarily as dummy verbs (Versteegh 2001: 488) for the formation of compound verbs and light verbs[12] (Butt, 1995) in complex predicates.

---

[12]. For the purpose of morphological study, dummy verbs have been discussed here. The term light verb is used in syntax, which is not the interest of this work. However, the difference between the two terms has been explained.

There is a controversy over whether dummy verbs should be termed dummy verbs or light verbs. The two terms are different functionally. A light verb forms a complex predicate, which is a verbal structure of two or more verbs in a sentence in which the main verb is mostly in lexical form, and the light verb forms the tense morphology. It can be said that V (*kha* 'eat') + lV (*lɪja* 'taken') = complex predicate (*kha lɪja* 'has/had eaten'). Note that V is the main verb, and it is followed by a light verb (lV) in syntactic order. Tense formation depends not on the main verb but on the light verb (Butt, 1995), as seen in the present example. Thus, the function of a light verb is to support a main verb to form a complex verbal construction V_1V_2 or complex predicates in other words. By contrast, the function of a dummy verb is to combine with a noun, adjective or a verbal noun to derive a compound verb NV/AV combination. Bukhari (2009: 80) claims that, "creating new verbs by a productive process of combining a nominal (or adjective) with *a light verb* is a common feature of Indo Aryan languages." He states that a verbal structure of N + V or A + V is a light verb, as it combines with a noun or adjective to form complex predicates. There is a discrepancy between a dummy verb and a light verb, which Bukhari overlooks and it is necessary for this to be clarified. If N +V or A + V is a productive process of creating new verbs, a verb in such a construction must be a dummy verb. Neither the construction N+ V or A+ V is a complex predicate nor is V a light verb in this structure. Dummy verbs are used with their lexical meanings. Light verbs do not express their lexical meanings. Rather, they modify the semantics of a sentence. In simple terms, both types of verbs are in fact main verbs and can be used as dummy verbs and light verbs in two different constructions. In VV formation, they function as light verbs but do not express their lexical meaning. In NV or AV formation, they function as dummy verbs, expressing lexical meaning, to derive compound verbs. The difference is illustrated below:

25a. Wania ne   Azan ko   məʃvəra   **d̪ɪja**
    Wania.f.s.erg   Azan.m.s.acc   advice.m.s   give.**pst**.m.s
    'Wania advised Azan.'

  b. Wania ne   Azan ko   məʃvəra   **d̪e**   **d̪ɪja**   **(he/th*a*)**
    Wania.f.s.erg Azan.m.s.acc   advice.m.s   give.m.s.**bse**   give.m.s.**perf**   (be.3.s.pre/pst)
    'Wania has/had advised Azan'

In examples (25a) and (25b), the verb *d̪e* 'give' has been used in two different roles, as a

dummy verb and as a light verb. In (25a), functioning as a dummy verb with its lexical

meaning in the past form, it constructs a simple verbal structure in combination with the noun

məʃvəra 'advice'. The construction thus expresses the semantics of a past verb. With a basic

lexical form *d̪e* 'give', the dummy verb retains this role in (25b) and correlates with the noun

məʃvəra 'advice' to form again a simple verbal construction that functions as the main verb of

the sentence. However, used in perfective form *d̪ɪja* in a different syntactic position in the

same sentence, it takes the role of a light verb and expresses the tense/ aspect morphology. In

this role, it now modifies the semantics of the main verb *məʃvəra d̪e* 'advise'. The dummy

verb *d̪ɪja* (simple past)' in (25a) and the light verb *d̪ɪja* (perfective) in (25b) are identical. The

perfective aspect in (25b) is only possible when the main verb *məʃvəra d̪e* (used in base form)

combines with the light verb *d̪ɪja* (expressing the tense morphology). Thus, VV as a whole is

a complex verbal construction or complex predicate in Butt's (1995) terms.

In short, the function of the light verb is syntactic in nature, as it forms a complex predicate,

i.e. V_1V_2 in which the main verb (V_1) is mostly in its base form and the light verb (V_2) shows

the tense morphology. The function of the dummy verb is to exhibit morphological

information. It derives a compound verb in combination with a noun or adjective, i.e. N

V/AV, which is needed when a lexical verb does not exist. Versteegh (2001: 497) considers

that the extensive use of dummy verbs in compounds is typical of Hindi-Urdu.

Bauer (2003) defines two types of compounds: 1) root compounds or primary compounds discussed in different classes of compounds, e.g. endocentric, exocentric (see above) and 2) synthetic compounds or verbal nexus in which the verb is the head. As noted, Urdu compound verbs are the combination of dummy verbs with nouns, adjectives or verbal nouns. They are different from Bauer's (2003) synthetic compounds, e.g. *dish washer*, in which a verb is the head element and fully incorporated. The head element in an Urdu compound verb is a noun, adjective or a verbal noun. A dummy verb remains a modifier. It does not incorporate and maintains its distinctive status. It shares the meaning in the formation of a compound verb.

4.7. Summary of the Chapter

Urdu compounds show various features. Most compounds are formed by some distinctive compounding words each of which derives a large number of compounds without distinguishing the source language of the constituents. Therefore the formation of hybrid compounds is dominant, and so the major types of compounds were the focus in this chapter rather than distinguishing them with reference to their source language. These various types include endocentric, exocentric, copulative, postpositional and verbal compounds. The majority of compounds are endocentric verbal compounds, but there are also formations of copulatives, genitives like *ke* and *ki*, and the ablative/instrumental *se*. The resulting compounds are mostly adjectives, nouns and verbs, but there are also adverbs.

Urdu compounds are often ambiguous due to their structural nature, as they may be left branching (modifiers come before head) e.g. *xoʃnəsib* 'lucky' or right branching (modifiers come after head) e.g. *əḍalət-e-ozma* 'supreme court'. They are also ambiguous because the two constituents are often synonyms, and the compounds may be an endocentric formation,

e.g. *rah-**e**-hǝq* 'the right path' or an exocentric one, e.g. *balbǝtʃe* 'family'. The four examples show that Urdu compounds are formed with and without interfixes. Those with interfixes, found in endocentric and copulative classes, show interesting structural and semantic properties. The interfixes in endocentric and copulative compounds, *-e-* and *-o-*, function as genitives *ka* and the conjunctive *ɔr* 'and'. They convey possession, plurality and emphasis. But the interfix *-b-*, in a few determinative compounds, express the semantics of various postpositions, e.g. often *pǝr* 'on'. The compounds without interfixes are mainly found in exocentric and verbal compounds. Exocentric compounds are metaphorical and creative in semantics, but they are not in a large number. Verbal compounds are only hybrid formations in the sense that their key constituents are dummy verbs, often modified forms of Persian auxiliaries. However, there is no distinction in the source languages of the nominal and adjectival constituents.

Taking the morphological structures from the three sources of Urdu words, i.e. native, Persian and Arabic as the base of Urdu morphology was necessary throughout the last four chapters, because there are major changes in Urdu morphology with respect to the sources. Therefore, now, it may be helpful to understand the morphological structures of English loans, and the discussion on the more recent loans can be started. It is necessary to see to what extent English loans follow the same patterns as native or Persian and Arabic loanwords. This will be the focus of the next chapter.

****

Chapter 5

Morphological Adaptation of English Loans in Urdu

5.1. Introduction

In the previous chapters, Urdu morphological changes have been discussed with respect to three sources of words i.e. native Urdu, Persian and Arabic. The discussion in this chapter is in the light of Winter's (2008: 156) discussion of the integration of loans. It looks at some forms of the heavy integration of English loans, and how the adaptations take place. The discussion focuses on two points, i) that the English loans are in principle subject to the same processes as Persian and Arabic loans, but ii) that they are less well integrated than these older Persian and Arabic loans. It is necessary to see the productivity of native Urdu, Persian and Arabic affixes attaching to English loans, and the adaptability of English loans with which of the three source language affixes, the loans are more frequent.

Some aspects e.g. gender, number, case morphology, derivation with various categories of the same word borrowed, formation of hybrid compounds with or without any morphological changes in the English constituents, English affixes in Urdu and creation of new lexical items are taken into consideration. Although the creation of new lexical items is not expected, it is nevertheless possible. As the loanwords are simply more recent, they are felt by general speakers to be 'more foreign'. There is a strong correlation between the length of time a loanword is seen in a language and the degree of integration that takes place. Crawford (2008: 64) notes the same correlation in Japanese, "The degree of nativisation of a loanword also correlates to the length of time it has been attested since 19th century showing more nativisation than those words that have only been attested recently." This is seen in the discussion ahead. Some sub conclusions are drawn based on these two points and the morphological structures from the three sources.

5.1.1. Extent of English Loans and Morphological Changes in them

New words are not created from nowhere, but are either borrowed or formed by combining words or parts of words. Borrowing may vary in degree from casual to heavy lexical borrowing, and from slight to the significant incorporation of structural features. Winford (2003: 17) states that borrowing begins with vocabulary, and the incorporation of structures comes only after the substantial importation of loanwords. Based on his scale of borrowing, the contact situation between English and Urdu can be illustrated as:

Table 1: Language Contact Situation between Urdu and English

	English → Urdu	**Urdu → English**
Lexicon	very strong	minimal
Phonology	weak	strong
Morphosyntax	weak	strong

English here refers to the Urdu speakers of L2 English. In this context, the influence of English words on Urdu is strong. Conversely, the influence of Urdu words on English is minimal, but the influence of Urdu phonology and morphosyntax seems to be strong. This is natural because the speakers of a language use loanwords which are influenced by their first language, and therefore the phonological changes take place. Winford (2003: 22) considers some major outcomes after the importation of loanwords over a certain period of time. There are three stages of contact situations, i.e. (A) language maintenance (B) language shift and (C) language creation. Despite heavy borrowing from English, Urdu can still be seen to be in the language maintenance stage and has not yet reached the language shift, as has occurred under the Persian and Arabic influence. However, the use of the term language shift here means a community of speakers no longer using their traditional language but instead shifting to the

use of another language. Urdu tends to turn to Arabic and Persian, while Hindi tends to turn to Sanskrit for new words. Due to this fact, the two languages have consistently given a shape to Urdu with the passage of time, and their influence has gone into the creation of words, phrases, and terminologies. With regard to English, the influence is nonetheless not to the same extent. However, the lexical influence and the structural fusion of English loans are considerable, and they need to be examined.

King (2000: 173) observes that the influence of English on PEI (Prince Edward Island) French has been essentially lexical and these lexical innovations "have triggered particular language-internal changes, resulting in the emergence of a number of structural differences in PEI French". Sankoff (2001: 654) states that King bases this analysis, as English *back* corresponds in meaning to one meaning of the French affix *re-* 'return to a former state or place', as in *venir back* 'to come back' (*revenir*); the borrowing of English *wh*-words also ensues changes in French relative clauses. Similar to this, the lexical influence of English loans on Urdu is deep, and there are some structural changes as well. Nevertheless, the language internal changes are not so many.

Winter's (2008: 157) discussion on the issue in light of Haugen's (1950) *Importation and Substitution* of loanwords is useful here. A word is said to be imported, e.g. the English loanword *faithful* in Urdu, if it is not changed structurally and conforms to the source language form. However, a loanword is said to be substituted, e.g. (English) *frad* 'fraud' → *fradıa / fradi* 'one who gets benefits by treacherous means' (Urdu), when it is structurally changed in the target language and replaces the original one. The words *fradıa* and *fradi* are in fact new lexical creations in Urdu, and there is no parallel derivation for their meaning in the English lexicon. Although such new creations are not many, they mirror the start of the

similar influence Persian and Arabic have had on Urdu. The morphological changes are not widespread, but the degrees of change are notable. To understand the degrees of change, it is necessary first to see the general types of English loans and the patterns of adaptations.

5.1.2. Types of Loans and the Patterns of Adaptation

Although there are a large number of adjectives and verbs as well, loan nouns are dominant and can be classified into three groups. Almost all scientific names and terminologies, e.g. *ozone*, *oxygen*, *carbon*, *computer*, *TV* and *engine*, have no Urdu substitutes and so are a part of the Urdu lexicon. Others have Urdu substitutes but they are not very commonly used. English counterparts are more frequent e.g. *lɪbas* 'suit', *pʌʈlun* 'pants' and *bæ̃tʰʌk* 'drawing room' etc. Still others have common Urdu substitutes, and people use them as well. However, using English loans instead is gradually becoming a fashion, e.g. *ɪmʈɛhan* 'exam', *dɒsʈ* (m)/*sʌhɪlɪ* (f) 'friend' and *kaɣʌz* 'paper'. Native equivalents to English loans are frequently supplemented or replaced. In some instances, the long spelling and difficult pronunciation of Urdu words, e.g. *bavərtʃɪxana* 'kitchen' and *bæt̪ɒlxəla* 'toilet' have facilitated the choice of English loans as a compromise. Heavy borrowing has had so much influence that many Urdu words are getting close to becoming obsolete, as the latter of the two examples. There are a huge number of loans used in their original form with no morphological changes. Some lexical items and compounds are shown in the table below:

Table 2: Patterns of English Loans with no Changes

Lexical Items	Compounds
hotel, nursery, agenda, vote, certificate, consultant, dictator, parking, propaganda, ideal, challenge, committee, president,	all rounder, governing body, income tax, local government, civil society, chief engineer, supreme court, chain smoker,

navy, security, administrator, target, formula, record, control, appeal, judge, bureaucrat, market, publicity, bus, power, criticise, start, programme, contract, media, budget, national.	deputy commissioner, whole sale, swimming pool, work permit, selection committee, network, result card, service centre, homework, book shop, registration form, adhoc judge, super star.

Source: Urdu *Daily Express* (07/02/2010)

The loans given in the table are very frequent; in some cases they are even more common than their native counterparts and are the only options e.g. *security* and *parking*. It is an irony of fate with Urdu that the names of the organisations and institutions, e.g. the *Daily Express* (Urdu newspaper) and *Urdu Dictionary Board*, related to promoting the Urdu language are themselves in English.

The integration of loans causes various forms of morphological changes, as illustrated below:

Table 3. Patterns of English Loans with Morphological Changes

Pattern of Loans	Examples
1. Gender Distinction	daktər (m) 'doctor' → d*a*ktər**ni** (f) 'lady doctor'
2. Pluralisation	bʊk 'book'→bʊk**ẽ** 'books'
3. Derivation	fr*a*d 'fraud' → fr*a*d**ɪa**/fr*a*d**i** 'one who gets benefits by treachery'
4. Hybrid compounds	phəl 'fruit' + frut 'fruit' → phəl frut '(a variety of) fruits'
5.Acronyms (English)	HEC - Higher Education Commission
6. Acronyms (Urdu)	PTI - Pakistan Terik-e-Ins*af* 'Pakistan Justice Movement'
7. Reduplication	t*a*ɪm t*a*ɪm (emphasis) 'with changing time'
8. Echo Reduplication	**t**a*ɪm **v**a*ɪm 'time and the like'
9. Affixes	-*ean*: Lahor**ean** 'citizen of Lahore', -*ite*: Karachi**ite**

These loans seem well integrated and should be measured by the extent to which they are affixed with Urdu morphological markers. Various forms of morphological changes include initials, signatures, and acronyms e.g. HEC (Higher Education Commission), abbreviated forms of long English loans (Prof.), shortening of two word compounds, e.g. Indo-Pak (India-Pakistan) and reduplication e.g. *time time* 'with the passage of time'. Urdu acronyms are not traditional, but the English loan forms of acronyms are popular, e.g. *HEC* (*Higher Education Commission*) *Pakistan.* Even more interesting are the Urdu compound nouns, which follow the pattern of English acronyms e.g. *PTI* (*Pakistan Tehreek-e-Insaaf*) 'Pakistan Justice Movement' (a political party led by a cricketer turned politician, Imran Khan).

These forms of lexical use of English loans are not exhaustive but suggestive. Morphologically, they have had a minimal effect upon Urdu. It is the native affixes that establish a pattern, according to which a loanword is adapted. English loans in Urdu are restricted in regard to the addition of new items. For example, the loan nouns very often take the native Urdu gender and number markers, but they are unaffected by the Persian and Arabic gender and number marking. The derivational changes are also very limited to some loans, although the forms of changes may be various. The frequency of each change varies depending on the situation involved. The integrated loans and the adaptation depend both on the word class (verb, noun, adjective) and on similarity to loanwords from other languages already existing in the borrowing language. Certainly, loanwords used as nouns differ from those used as verbs, or adjectives, in relation to both morphological marking and their incorporation into borrowing language structures. Reed (1948) discusses the adaptation of English loans in Pennsylvania German Morphology. There are many similarities between the adaptation of English loans in the two languages. One common feature is the gender marking

of the borrowing languages. This is necessary, because both Urdu and Pennsylvania German have overt gender marking, whereas English does not.

5.2. Adaptation of Loans: Gender and Number Changes

5.2.1. Gender Treatment and Gender Marking of Loans

This subsection discusses two points, firstly some general principles that determine the gender of English loan nouns in Urdu, and secondly the formation of feminine noun from masculine noun following a rule may be grammatical in some cases but ungrammatical in others. Urdu gender morphology comes from three sources i.e. native Urdu and Persian and Arabic. Any loan noun can be expected to adopt a pattern from either of these sources. The abundance of English loan nouns shows that they adopt the native Urdu gender morphology, which is based on not only sound but also meaning. General gender markers are *-a* and *-i*, as noted in Chapter 2. Therefore, all loan nouns that end with *-**a*** (m), e.g. *sof**a*** 'sofa' and *-**i*** (f) e.g. *dɪgr**i*** 'degree' are treated as masculine and feminine respectively. However, many other loan nouns do not end with *-a/-i*. In this case, the gender treatment is based on how their Urdu counter parts are taken, whether masculine or feminine. This is particularly necessary, if the loan nouns are inanimate nouns. For example, English loan *bed* is masculine and *tebɪl* is feminine only because their Urdu counterparts, *bɪsṭər* and *mez*, are masculine and feminine. However, it is difficult to determinie, if Urdu counter parts are not available and English loans are the only words used for certain objects. This is the option particularly for scientific objects and terms. In this case, Rzivi's (2007: 73) resolution, as noted in Chapter 2, is also helpful that huge, heavy, powerful, dominant and big objects are masculine, while small, weak and light ones are feminine. Therefore, *refrigerator*, *computer*, *TV* and *engine* are masculine and *pin*, *ozone*, *oxygen* and *carbon* are feminine in Urdu. These inanimate nouns do not have the general

native Urdu suffixes, -*a* and -*i*, for determining gender, and therefore their image emerging in mind is the only option left for it.

In addition to -*a* and -*i*, many English loan nouns may also have two other forms of gender, although limited in their use. Thus, there are three forms of nativised gender marking.

Masculine Nouns		**Feminine Nouns**
1a. det*a* 'data'	→	dɪkʃnəri 'dictionary'
b. d*a*ktər (m) 'doctor'	→	daktər**ni** (f) 'lady doctor'
c. fr*a*d**i** /fr*a*dɪ*ja* (m) (from *fraud*) 'cheater' →		fr*a*d**ən** (f)

The loan nouns in (1a) are inanimate nouns and treated as masculine and feminine, based on their endings which are the primary gender markers in native Urdu. Examples (1b-1c) show that a larger number of English loan nouns, treated as masculine, have a zero suffix and a few may also have -*i*, both following the secondary native patterns of gender. The feminine patterns in (1b and 1c) are also derived forms in accordance with the native Urdu masculine patterns, and so end with -*ni* and -*ən* respectively. However, the masculine gender suffix -*i* and the two feminine forms -*ni* and -*ən* attach to only a few English loans, as they are unproductive in native Urdu also. It is noteworthy that the formation of some feminine nouns following a rule may be grammatical in some cases but ungrammatical in others.

2a) daktər (m) 'doctor'   → daktər**ni** (f)

 b) mastər (m) 'teacher'   → m*a*stər**ni** (f)

3a) titʃər (m) 'teacher'   → *titʃər**ni** (f)

 b) mænɪdʒər (m) 'manager'   → *mænɪdʒər**ni** (f)

The feminine nouns in (2) take the gender marker *-ni* in accordance with the masculine nouns, which have a zero suffix for gender marking, as examplified above. However, this rule is not applicable in some other loan nouns, as in (3). This shows that *-ni* is not as productive as the major gender markers *-a* and *-i*, although there are a larger number of English loan nouns treated as masculine with a zero suffix. Secondly, Crawford's (2008: 64) view is also helpful here that the degree of nativisation of a loan correlates with the length of time it has been in a language and its frequency of use. The loan nouns in (2) are quite older than those in (3). They are not treated as loans, but as native words. This is why, possibly, they are used with the gender distinction. But the loans mouns in (3) are more recent and still treated as loans. Therefore, they are used without any gender distinction.

There is another aspect of the case. Note that the nouns in (2) have retroflex sound in the final syllable, while the nouns in (3) are devoid of it. Therefore, the former seem more adaptable and are more in harmony with the native Urdu nouns, for which a retroflex sound is a feature.

The feminine nouns in (3) do not seem to be adaptable, although they are morphologically similar to those in (2). However, it is not the case that if a loan noun is adaptable, the other loan with a similar morphology is also adaptable. The examples in general show the adaptability of some English loan nouns for native gender marking. Thus, the gender morphology of English loans is in conformity with the TLS. In the examples, there are two different processes at work here: one is the gender assignment to a loan word; the other is whether the loanwords can act as the base for further formations, through derivation, i.e. the change of gender from masculine to feminine. The process of derivation is discussed in the sections to come. Gender marking of English loans actually concerns their adaptability in plural marking with native Urdu affixes.

5.2.2. Pluralisation of English Loans

As Winter (2008: 168) points out, Wegener (1999, 2002, 2003, 2004) argues that different plural forms e.g. those in Pennsylvanian German are integrated on a morphological level. The pluralisation of English loan nouns is the most frequent form of adaptation in Urdu, and many of its features are similar to those seen in Pennsylvania German (PG). For example, many English plurals have also entered Urdu, but not all are very commonly used. The plural morphology of English loan nouns shows changes more like those of native nouns. The loan stems are affixed with the plural morphemes, i.e. *-e* (m) *-ıjã*, *-ẽ* (f) (nominative forms) and *-õ̃* (oblique form). They are the general native Urdu plural markers. The following examples of masculine and feminine plurals in Tables (4a-4d) show that English loans are well integrated. Therefore, unlike Persian and Arabic, there is no morphological ground to differentiate between the pluralisation of the native nouns and that of the English loans.

There are two points to be noted in respect of oblique singular/plural forms. Like the singular oblique of the native masculine nouns, the singular oblique of English masculine nouns is the same as the nominative plural, i.e. it ends with *-e*. The oblique form of all other singular loan nouns is unchanged like their native counterparts. The oblique form of all loan plurals also follows the native oblique form and ends with -õ. Each of the four sets of masculine and feminine plurals contains one native Urdu example so that the morphological changes in the English loans can be compared and better understood:

Table 4a: Pluralisation of English Loans: Masculine Nouns (ending in *-a*)

Singular	Nominative Plural	Oblique Plural
1- det*a* 'data'	→ det**e** 'data'	/ det**õ̃**
2- sof*a* 'sofa'	→ sof**e** 'sofas'	/ sof**õ̃**

3- formul**a** 'formula' → formul**e** 'formulas' / formul**õ**

4- propagand**a** 'propaganda' → propagand**e** 'propagandas'/ propagand**õ**

5- pest**a** 'pesta' → pest**e** 'varieties of pasta' / pest**õ**

6. bet**a** 'son' → bet**e** 'sons' / bet**õ** (Urdu)

Table 4b: Pluralisation of English Loans: Masculine Nouns (all other endings)

Singular	Nominative Plural	Oblique Plural
1- daktər 'doctor'	→ daktər 'doctors'	/daktər**õ**
2- kəmɪʃən 'commission'	→ kəmɪʃn 'commissions'	/ kəmɪʃn**õ**
3- dʒəg 'jug'	→ dʒəg 'jugs'	/ dʒəg**õ**
4- pepər 'paper'	→ pæpər 'papers'	/ pæpər**õ**
5- mænɪdʒər 'manager'	→ menedʒər 'managers'	/ menedʒər**õ**
6- bhəŋg**i** 'toilet cleaner'	→ bhəŋg**i** 'toilet cleaners'	/ bhəŋg**iõ** (Urdu)

Many English loans treated as masculine (4a) end in -*a*. Their pluralisation is made by the alternation of the suffix -*a* with -*e*. The patterns of loan plurals, e.g. *formula* 'formula'→ *formule* 'formulas' are exactly the same as the native plural, e.g. *beta* 'son → *bete* 'sons'. In set (4b), comparing the loans, i.e. (1-5) with the native one (6), shows that many English loan nouns remain unchanged like their native counterparts in their plural forms. This means that masculine pluralisation by alternation of vowel is possible only if masculine loans end with -*a*. If they end with something other than -*a*, there are no changes for pluralisation.

The singular oblique patterns of the loans, following morphology of the native nouns, are the same as the nominative plurals. For example the nominative plural *formule* 'formulas', in set (4b), is also the oblique singular noun. There is no difference in the oblique plurals of the loan masculine nouns in the two sets (4a & 4b). The oblique plurals end with the suffix -**õ**. This

gives the picture that the loan oblique plurals, e.g. *formulõ* and native oblique plural, e.g. *betõ* (native) are also the same in forms. However, there is another possibility from the examples in set (4b). The oblique form of loan plural nouns, e.g. *kɪtʃɪn* 'kitchen' may also remain unchanged. The oblique plurals in general end with the suffix -õ, and there is no difference between masculine or feminine loan nouns. It is the same as the native nouns. The English feminine nouns (4c and 4d) also show this:

Table 4c: Pluralisation of English Loans: Feminine Nouns (ending in *-i/-ni*)

Singular	Nominative Plural	Oblique Plural
1- laɪbreri 'library'	→ laɪbrerɪj**ã** 'libraries'	/ laɪbrerɪj**õ**
2- ˌjunɪvəsɪti 'university'	→ junɪvəsɪtɪj**ã** 'universities'	/ junɪvəsɪtɪj**õ**
3- dɪgri 'degree'	→ dɪgrɪj**ã** 'degrees'	/ dɪgrɪj**õ**
4- edʒənsɪ 'agency'	→ edʒənsɪj**ã** 'agencies'	/ edʒənsɪj**õ**
5- daktərni 'lady doctor'	→ daktərnɪj**ã** 'doctors'	/daktərnɪj**õ**
6. bet**i** 'daughter'	→ bet**ij***ã* 'daughters'	/ betij**õ** (Urdu)

Table 4d: Pluralisation of English Loans: Feminine Nouns (all other endings)

Singular	Nominative Plural	Oblique Plural
1- plet 'plate'	→ plet**ẽ** 'plates'	/ plet**õ**
2- bɪldɪŋ 'building'	→ bɪldɪŋg**ẽ** 'buildings'	/ bɪldɪŋg**õ**
3- bʊk 'book'	→ bʊk**ẽ** 'books'	/ bʊk**õ**
4- məʃiːn 'machine'	→ məʃin**ẽ** 'machines'/ məʃin**õ**	
5- fradən 'cheater	→ fradən**ẽ** 'cheaters'	/ fradən**õ**
6. bhəŋgən 'toilet cleaners'	→ bəŋgən**ẽ** 'toilet cleaners' / bəŋgən**õ** (Urdu)	

In set (4c), English feminine nouns ending in *-i*, e.g. *laɪbrerɪ* 'library', are pluralized as *laɪbrerɪjã* 'libraries', the same as the native Urdu plural *betɪjã*. Further, the oblique plurals in every case continue to end with the suffix *-õ*. However, set (4d) shows a different picture. Like many native feminine nouns, e.g. *bhəŋgən* 'toilet cleaner', there are also English loan nouns, which are treated as feminine and do not end with *-i*, e.g. *bɪldɪŋ* 'building'. They are pluralized with the suffix *-ẽ* e.g. *bɪldɪŋgẽ* 'buildings' /*bəŋgənẽ* 'toilet cleaners'. All the loan examples in sets (4c and 4d) are evidence that there is a full integration of English loan nouns. The two sets show a general rule, namely, if the English loans end with *-i*, the plural ends with *-ɪjã*. If the loans end with something other than *-i*, the plurals end with *-ẽ*. Both the sets show that whether masculine or feminine, the oblique plurals end with *-õ*. Singular oblique remains unchanged in all nouns except for those given in set (4a), the same as the nominative plurals.

Forming plurals in Urdu is far more complex than in English. As seen in chapter 2, there are three types of plural patterns, i.e. native Urdu, Persian loan and Arabic loan patterns. In this scenario, English loan pluralisation seems to be according to native Urdu patterns. Therefore, applying Winter's (2008: 156) criteria of conformity, it can be observed that the English loan nouns are in conformity with TLS. There are some notable features.

Even though the pluralisation of English loans is based on native Urdu plural by suffixation, the suffixation itself is no more than an alternation of vowels. The alternation of vowels is the main feature in the gender and number morphology and the case morphology of the Urdu noun. Phonological alternations come in many shapes and sizes, and the processes behind them are equally varied, as are the kinds of factors which condition them (Davenport and Hannahs, 2005). The alternation forms nativised English plurals both with and without nasalization. With the nasalization, it makes changes at the suffix position, e.g. *bɔk*/ *bɔkẽ*

'books', as in sets (4b,4c and 4d) to form feminine plurals. To form masculine nominative plurals, the alternation of vowels at the suffix position does not require nasalisation, as in (4a). The oblique forms of plurals however require nasalization. The alternations are mainly caused by nasalised vowels and an allophone /~/of the alveolar nasal consonant /n/. There are a number of certain cases of nasalization in the alternation of vowels. One is grammatical (morphological), i.e. forming plurals, another is lexical, and still another is semantic. However, the lexical and semantic cases are not the issue here.

In addition to the alternation of vowels, all the plurals in the four sets (4a to 4d) also show the productivity of native Urdu plural suffixes and the adaptability of English bases. The plural suffixes are productive enough to attach to any of the loan nouns. Productivity is mainly said to be a function of the frequency of words, and the reason for the connection between frequency and productivity lies in the processing of words in the lexicon. As noted above, many English nouns are gradually becoming more frequent than the native ones. They are adaptable in native Urdu pluralisation. Dressler (2003: 32) postulates productivity as a primitive property of inflectional morphology (in parallel to grammatical productivity), and states, "if we conceive of productivity as a grammatical concept, then morphological productivity can still be defined…as: the possibility for language users to coin, unintentionally, a number of formations which are in principle uncountable." Discussing Wurzel's (1984) productivity rules, he mentions two important elements 1) the integration of loanwords with unfitting properties, which have to be fitted (accommodated) to the system adequacy of a borrowing language; 2) the integration of loanwords with the fitting criteria. The first, the integration of loanwords with unfitting properties is noted in derivation, as discussed in the relevant section ahead. For the integration of loan-words with already fitting properties, the integration must overcome only the obstacle of foreignness. Loanwords with

fitting properties here mean that they are adaptable and can be adjusted in the morphological system of borrowing language. One example, Dressler (2003) gives, is the entrance of German neuter *-o* into Slavic languages, where *-o* is the default ending of neuters (Sg.Nom. = Acc.) and the respective microclass the default for neuters: e.g. German *das Auto* > Polish *auto* (neuter), Slov. *auto* (masc.). German. *Auto* (neuter) becomes a neuter in Polish; but a masculine in Slovene demonstrates non-productivity of the Slovene neuter microclass in *-o*, but high productivity of its Polish correspondent. In the light of this, the pluralisation of English loans in Urdu can also be noted as showing the integration of loanwords with fitting properties. English loan nouns have fitting properties and are fully compatible in the Urdu plural system. They accommodate the Urdu gender system as well as the number system. Thus, the native suffixes are productive enough in these two aspects. They accommodate English bases without requiring any major changes, i.e. deletion or substitution of any other affixes. This shows that English bases can easily be extended by the addition of new items (the Urdu plural markers), and therefore seem to have been fully integrated into the Urdu pluralisation process with gender distinction, generally.

Gender distinction is visible in the derivation of words also. Derivational affixes possess different degrees of productivity, but this is only partially seen in the English loans. It is not even as much as the productivity of Urdu plural markers. However, the derivational patterns possess a variety of formations.

5.3. Adaptation of Loans by Derivational Means

The adaptation of English loans by derivation can be seen both by affixation and by compounding. This section discusses three major points 1) derivation by affixation 2) derivation by compounding 3) English loan affixes. Dressler (1993: 5028) considers that

morphological innovations (even if derived via the productivity derivation rule) always violate (at least a part of) lexical norms. This violation of norm is also evident in the derivation between Urdu and English morphological items. The major word-formation processes by English loans include affixation, compounding and acronymy, but reduplication, blending and clipping are also evident.

5.3.1. Derivation by Affixation

Urdu affixes are restricted in derivation as compared to gender and number marking. Some linguists hold the view that productive ways and means of word-formation are only those that can be used for the formation of an unlimited number of new words in a modern language. Bauer (2003) states, "a morphological process can be said to be more or less productive according to the number of new words which it is used to form". In light of this, Urdu derivational affixes are not as productive as gender and number affixes, which is nevertheless explicable in terms of grammatical marking in general: a noun in Urdu must be marked for gender and number. Generally, we don't see a large number of derivational patterns with English loan bases, although a variety of derivational changes can be seen. Moreover, Urdu borrows only one (or two) categories of a word. Other derivations of the same category can rarely be seen. For example the word *fraud* is very commonly used but its other categories i.e. *fraudulent* and *fraudulence* are unknown. Derivations from the loanwords are mostly in the native Urdu style, although some Persian and Arabic affixes are also productive.

5.3.1.1. English Loanwords with Native Urdu and Persian/Arabic Derivational Affixes

This section discusses mainly native Urdu affixation, but some English loanwords taking native Urdu, Persian andArabic affixes are also highlighted. The following section then focuses on Persian and Arabic affixation separately. As stated above, most word formations

are the result of native Urdu affixation, which may be productive in just one meaning because that meaning is specially needed by the community at a particular time period historically. It can be non productive to coin scientific or technical terms. Only a few English loanwords, e.g. *park*/*parking*, and *administration*/*administrator* are borrowed with more than one category. Most loans are used in the original forms borrowed and the derivations, if any, are mostly made in the native Urdu manner, although native affixes are fewer than Persian and Arabic affixes. They do not only derive new words from English loans, rather they create new lexical items with loan bases. However, some Persian and Arabic affixes are productive enough to derive quite many words with English loans. This is illustrated in the following data:

Table 5: Some English Loans and their Adaptation in Urdu

a. frad 'fraud' (N)   →   1) fradɪj*a* 'one who gets benefits by treachery' (N) (m)

                               fradɪje 'cheaters' (plural and sing. obl) / fradɪj**õ** (pl.obl.)

            →   2) frad**i** 'cheater' (N/A)

                     frad**i** 'cheater' (plural and sing. obl) / fradɪj**õ** (pl. obl.)

b. bəm 'bomb' (N)   →   1) bəmb*a*r 'bomber' (N/A)   (m)   (used in native manner)

            →   2) bəmbar**i** 'bombardment' (N)   (f)

            →   bəmbar**ɪj*ã*** 'bombardments' / bəmbar**ɪj*õ*** (pl. obl)

c. fəlsəf*a* 'philosophy' → 1) fəlsəf**e** (sing. obl./pl) / fəlsəf**õ** (pl. obl)   (m)

                 2) fəlsəf**i** 'philosopher' (N/A)

                 3) fəlsəfɪj*ana* 'philosophical' (A)

                 4) *fəlsəfɪjət* 'status of a philosophy' (N)

d. dʒʊɡrafɪj*a* 'geography' (N) 1) dʒʊɡrafɪja*i* 'geographical' (A)

e. titʃər 'teacher'   →   titʃər**i** 'teaching' (N)

  d*a*ktər 'doctor'   →   d*a*ktər**i** 'doctorship' (N)

menedʒər 'manager' → menedʒəri 'management/managership' (N)

kəndəktər 'conductor' → kəndəktəri 'conductorship' (N)

f. dr*ama* 'dr*ama*' (N) → 1) dr*ama***i** 'dramatic' (A)

The loans undergo all the possible inflectional and derivational changes through which a native noun can undergo. Some examples of major changes, from Table 5, in comparison with Urdu counterparts are illustrated below:

(5a) frad**ɪj*a*** --------------------------frad**ɪje**----------------------frad**ɪjõ** (E)

kəbaɽ**ɪj*a*** 'wreckage buyer'----kəbaɽ**ɪje**--------------------kəbaɽ**ɪjõ** (U)

(5b) bəm**b*ar*** ------------------------bəmb*ar***i** --------------------bəmbar**ɪj*ã̃*** (E)

bʊrdb*ar* 'tolerant'--------------bʊrdb*ar***i** 'tolerance'**------**bʊrdbar**ɪj*ã̃*** (U)

(5c) fəlsəf**i** ----------------------------fəlsəf**ɪj*ana***----------------fəlsəf**ɪjət̪** (E)

vehʃ**i** 'barbarian'---------------- vehʃɪ**j*ana*** 'wild'----------vehʃɪ**jət̪** 'status of wilderness' (U)

vehʃ**ət̪** 'wilderness'

In the given examples, (5a) shows inflectional changes based on native morphology; (5b) also exhibits native based derivational and inflectional changes, and (5c) Persian and Arabic based derivational changes. Almost all loans in the table seem to be integrated, as they show various changes of categories and forms. The changes in the loan noun *frad* 'fraud' (5a) are in two different patterns, i.e. *frad**ɪja*** and *frad**i*** 'one who gets benefits by treachery' (both) which are new lexical creations. They are in conformity with the target language patterns only, and their semantic substitute does not exist in English, as a derivation of *frad*. This shows that native Urdu affixes, with English bases, can create new lexical items rather than merely deriving new words. Lexical creation is more important than the derivation of new words in that it introduces a new lexical item into a lexicon, while aderivation does nothing but changes the

category of already existed word in a lexicon. Both *fradɪja* and *fradi* are new introductions to Urdu lexicon. This shows the complete integration of some English loanwords in Urdu morphology. Both the forms are in concordance with those seen in the native patterns and are in common use. The first one, i.e. *fradɪja* is entirely parallel to the native changes, e.g. *dak* 'mail' (f) → *dakɪja* 'postman'. This change is from a (m/f) noun to form an agentive masculine noun showing the male specific job or profession. A noun undergoing morphological changes may be masculine or feminine, e.g. the English *frad* is masculine and the native *dak* is feminine, but the newly formed noun, i.e. the English *fradɪja* or the native *dakɪja* must be a masculine noun. In addition, it has no feminine form.

The second formation *frad* 'fraud' → *fradi* 'one who gets benefits by treachery' takes the native derivational affix -*i*, but the morphological changes in fact parallel the changes in Persian loanwords, e.g. *mehnət* 'hard work' → *mehnti* 'hard worker'. Both *fradi* and *mehnti* are used as a noun and an adjective, and there seems to be no distinction of gender. However, there is a possibility that the English noun takes the feminine form *fradən*, although it is not traditional. The newly derived noun may not necessarily be agentive.

The morphological changes in all the examples from (5b) onwards show that Urdu borrows one category words and derives, if it does, all other categories with mostly native affixes in a native manner. For example,. English noun *bəm* 'bomb' (5b) exhibits much morphological and phonological fusion. Morphologically, the noun is further converted into the noun/adjective *bəmbar* 'bomber', noun *bəmbari* 'bombardment' and a verb *bəmbari kərna* 'to bombard/to do bombing' (for which a dummy verb is required, which will be discussed later). Two of the nouns *bəm* and *bəmbar* are treated as masculine, but *bəmbari* is feminine. All these nouns take nominative and oblique plural morphology, just like the native Urdu

nouns do. This shows that the integration is the same as the integration of Persian-Arabic loans with various morphological changes. However, the number of such English loans is comparatively not large.

The integration of most English loans as in (5c-5d) is generally not so much with any affixes whether native or Persian/Arabic. However, like those in (5a) and (5b), these changes verify the perception that English loans have started undergoing some integration, similar to the ones seen in Persian/Arabic loans, although to a far smaller extent than the older ones. The loan *fəlsəfa* 'philosophy' (5c) can be pluralised as *fəlsəfe*. It can also take the suffix *-i* for the formation of an agentive noun/adjective *fəlsəfi* 'philosopher'. It shows morphological changes in the bases with even a higher degree of integration by taking Persian affix *-ıjana* and Arabic affix *-ıjət̪* to form an adjective *fəlsəfıjana* and an abstract noun *fəlsəfıjət̪* 'status of a philosophy'. It is thus unique in the sense that it takes the native suffix *-i*, it takes the Persian suffix *-ıjana*, and then it can also Arabic suffix *-ıjət̪*.

The changes in (5e) and (5f) are not various. They are from a noun to noun and noun to adjective. The nouns in (5e) show an agentive noun, e.g. *titʃər* 'teacher' changing into an abstract noun *titʃəri* 'teaching'. Like *fəlsəfa*, the noun *drama* in (5f) is unique in the sense that it can be affixed with native and Persian/Arabic derivational affixes. Affixing with the native *-i*, it is changed into the adjective *dramai*, but affixing with Persian and Arabic, it is changed into agentive nouns i.e. *dramanıgar* and *dramanəvis*. Thus, it shows more affixation of Persian morphemes than native morphemes.

All the morphological changes generally show the productivity of native derivational suffixes and the adaptability of English loans. The degree of productivity of affixations very much depends on the structural, lexico-grammatical and semantic nature of bases and the meaning

of an affix. The adaptability of English loans in pluralisation is widespread, but by derivational means it is not seen on a very large scale. However, various derivational patterns can be seen. There are changes in various loans, but each loan is restricted to one change in most cases. Those which show a high degree of integration and complete morphological adaptation are only some words. Complete morphological adaptation means a loan, e.g. *frad*, is able to take a maximum number of possible morphological forms, the same as shown by Persian and Arabic loans, and it looks like a native word. It should show the gender, number changes in the native manner; it should also show the derivational changes in the native manner. With this explanation, it can be said that the loans in (5a-5c) are completely integrated, although those in (5d-5f) are partially integrated.

The ability of the loans to take all forms of structural changes shows that some English loans are fully integrated, and the affixation or inflection is according to native (target language) patterns. The inflectional and derivational changes are both major and minor, and the word formations include all major categories of words i.e. nouns, adjectives and verbs. The new words derived from the English loans seem to be designed systematically, as they follow the native Urdu morphological rules exactly, to make the changes. Plag (1999) terms it the morphological productivity that can be defined as "the property of a given word formation process to be used to derive a new word in a systematic fashion". The newly derived words from English loans are the extended forms and the results of several applications of the word-formation rule in a systematic fashion, as the examples of *frad* 'fraud' shown ahead. The stems of the words form derivational relations in different degrees. What is meant by 'degree' is that the stem of a word, e.g. *art* takes various affixes, one after the other, to derive various formations, e.g. *artist*, *artistic artistically*. The English loan *frad* 'fraud' in base form shows the zero degree affixation. Then, by applying a suffix *-i* or *-ɪja* to the base, there is first degree

affixation converting an abstract noun into an agentive noun *fradɪ* or *fradɪja* 'cheater'. The application of second degree affixation to *fradɪja* changes it into nominative plural and the singular oblique form *fradɪje*. The noun takes plural oblique form as *fradɪjõ* in its third degree of affixation. Thus, it can be suggested that in the inflectional and derivational process the English loans are generally adaptable. They can be extended into a variety of forms in the Urdu morphological system, and some of them are completely integrated.

Generally, an integrated or non integrated loanword exists side by side with its native equivalent so that the substitution can be easily identified. It can be true for words like *bəmbari* 'bombing' but not for *fradɪja* 'cheater'. The native substitute, i.e. *golabari* 'bombing' for the nativised loan *bəm* 'bomb' does exist but is rarely used. It may be considered obsolete now, because the nativised loan *bəmbari* has replaced it. The native substitute for *fradɪja* may be *ḍhokabaz,* which is in fact used for 'cheater'. It is fully semantically compatible with *fradɪja*. The English loan is specific in meaning and implies cheating in terms of money, but the native noun refers to cheating in a general sense. This semantic production is also evidence that some English loans are fully integrated and are like native Urdu words.

The changes in (5a)-(5c) bring us to the conclusion that English loans have started taking the status of Persian and Arabic loans, which are so fully integrated (phonologically, morphologically) in Urdu that they seem to be original Urdu words. Their morphological changes are the same as those in native words and Persian-Arabic loans. Words like *frad* contain retroflex sound, which is the specific feature of native words, not found in Persian-Arabic loans. Therefore, the changes in *frad* like English loanwords seem to be more native than in Persian and Arabic loans. The structural changes and the convergence of the loans are

extreme, involving some types of interaction and mutual accommodation by English loanwords and native Urdu affixes.

The structural convergence is similar to the one Winford (2003: 11) notes in Northwest New Britain where languages of the Austronesian and non-Austronesian families and new contact languages i.e. pidgins, creoles, and bilingual mixed languages have become structurally isomorphic i.e. they are ancestrally different, though, their structures are similar. He considers that borrowing is sometimes "the incorporation of foreign features into a group's native language by speakers of that language…the borrowing language is maintained, though changed in various ways by the borrowed features, and that the agents of change are its native speakers." Based on this criterion, it can be said that the English loans are at the maintenance stage. They have converged into native form. However, it is just the convergence of some lexical items, despite the fact that a large number of English loans have become members of the Urdu lexicon. Like the examples in (5d-5f), there are many English loans which, though integrated with Urdu affixes, do not undergo all the derivational changes. They take some forms, but do not take others. Consider the following examples:

4a. dram*a* 'drama' (N) → dr*am*a**i** 'dramatic' (A)

  b dʒʊɡrafɪj**a** 'geography' (N) → dʒʊɡrafɪj*a***i** 'geographic' (A)

  c. edʒənda 'agenda' (N) → *edʒənda**i**

The two nouns (4a & 4b) ending with *-a* can be converted into adjectives by suffixing with *-i*, but the third noun (4c) is ungrammatical with the same suffix. Discussing the productivity of the English derivational suffix *-ist*, Katamba (1993: 68) states, "We do not have a carte blanche to use it with any noun. There are unexplainable gaps. For example, a follower of prophet Muhammed is not a *Muhammed**ist*** though the follower of Budha is a *Buddh**ist**.*" The

adaptation of English loans with Urdu affixes also shows the same unexplainable gaps and can be widely seen. Although the native Urdu derivational affixes, particularly *-i*, are productive and form new words with English loan bases, they do not always function as shown in (4c). There is no explanation as to why they do not form new words. The degree of productivity of affixation very much depends on the structural, lexico-grammatical and semantic nature of the bases and the meaning of the affixes. In general, it can be said that native Urdu affixes are productive and the English loans are adaptable, and so integration occurs. Some loans, e.g. *drama* in (4a), are adaptable more with Persian and Arabic affixes than with native Urdu affixes, as shown in the next section.

5.3.1.2. English Loanwords with Persian and Arabic Affixes

Some English loan nouns, e.g. *drama* (4a), are unique in their adaptation in that they take native Urdu, Persian and Arabic derivational affixes. Affixing with the native *-i*, the noun *drama* is changed into an adjective *dramai* 'dramatic', but affixing with the Persian *-nɪgar/-nəvis*, it is changed into an agentive nouns i.e. *dramanɪgar/ dramanəvis* 'dramatist', and affixing with Arabic *-ɪjət*, it is converted into an abstract noun *dramaɪjət* 'the state of being dramatic'. Although these are only a few examples, it is interesting to see some patterns of Persian/Arabic affixes attaching to English loans separately from native based changes:

Table 6: English Loanwords with Non Native Affixes

a. laɪn 'line'	(N) →	laɪnḍar	'(paper) having lines'	(A)
b. fəlsəfa 'philosophy'	(A) →	fəlsəfɪjana	'philosophical'	(A)
c. bor 'bore'	(A) →	1) borɪjət	'boredom'	(N)
	→	2) bor kərna	'to bore'	(V)
	→	3) bor hona	'to be bored'	(V)

d. drama	(N) →	1) drama**ɪjət**	'the state of being dramatic' (N)
	→	1) drama**nɪgar**	'dramatist' (N)
	→	2) drama**nəvis**	'dramatist' (N)

The English loan nouns *laɪn* 'line' (6a) and *fəlsəfa* 'philosophy' (6b) change into an adjective. Both take two distinctive Persian affixes. The noun *fəlsəfa* is unique in the sense that it may take native Urdu, Persian and Arabic affixes. The change in *borɪjət* (6c) is also unique but different from the first two in that a loan adjective takes an Arabic suffix *-ɪjət* for the formation of a noun, and it takes dummy verbs for the formation of a verb. The noun *drama* in (6d) is the most interesting among the examples in that it takes native Urdu, Persian and Arabic affixes. It changes into all major categories i.e. noun, adjective and verb.

Although English loans are generally adaptable, some of them don't accept Urdu affixes freely. Various affixes e.g. *-ɪja*, *ɪjət*, *-ɪjana* shown in (5a-5b) attach to some items e.g. *dʒʊgrafɪja*, *titʃəri*, *dramaɪjət* as in (5d-5f). However, the affixation of many others is ungrammatical with the same lexical item e.g. **dʒʊgrafɪjana*. In general, the attestation of all these affixes is not possible, and the integration of the English loans is generally limited with Persian and Arabic affixes, and mainly occurs with native Urdu affixes. Therefore, it can be said that, although English loans show their adaptability and some variations, they do not generally integrate with Persian and Arabic affixes. Those given are among a few examples. As Winford (2003: 24) points out, early scholars such as Müller (1875) and Jakobson (1938) were of the view that structural (linguistic) constraints were the primary determinants of contact-induced change. Arguing against them he emphasises that, "the wide body of evidence available to us now shows that practically any linguistic feature can be transferred from one language to another, if the circumstances are right. The reason is that extra linguistic

factors – the social ecology of the contact situation itself can override any purely structural resistance to change." In this scenario, we see that the modern vernacular language is based on the *Khaṛi Boli* dialect, of the Delhi region, spoken in 12th century, although the name *Ordu* was first used much later in 17th century. Since then, Persian and Arabic linguistic influence began with the arrivals of Afghan emperor Mehmood Gaznavi and then Mughals, as two languages got much importance. Therefore, the integration of Persian and Arabic lexical items and their affixes was convenient. However, the same Persian and Arabic affixes have not proved to be productive in their integration with English loans, which are adaptable enough with the native Urdu affixes. They adjust English loanwords very to some extent, as compared to native Urdu affixes. The social setting, the circumstances and the linguistic wellformedness do not allow the integration of English loans in every aspect of its morphology.

Analysing the English loans based on Wegner's criteria (cf Winter, 2008: 168)), it is observed that most loans, as in Table (2), are without any morphological changes, and therefore the zero degree inflection is according to a non-native pattern. The derivation by zero degree inflection is ascribed to simple loans whose stem is homonymous with a word-form and often with a root-morpheme. They are simple lexical items with no derivational changes. Secondly, there are also many English loan nouns, e.g. *titʃəri* 'teaching' (5e) whose derivational changes follow native rules. Therefore, the affixation of a morpheme is according to a native TL pattern. Thirdly, all the English loans taking Persian/Arabic derivational affixes cannot be counted as fully integrated, as they neither follow native (Urdu) nor non-native (English) morphological rules. Rather, they follow Persian/Arabic rules, and therefore such loanword integration is according to a peripheral native TL pattern, as defined by Winter (2008: 158).

Although such loanword integration and derivation by affixation, according to the peripheral native TL pattern, shows a few formations, the derivation by compounding is vast enough. The major categories of English loanwords combine with Urdu dummy verbs to derive verbs.

5.3.2. Derivation by Compounding

Derivation by compounding of English loans is seen in hybrid compounds. Most hybrid formations of English loans are not with the native Urdu words but with Persian/Arabic loans. These formations may be divided into two groups, i.e. hybrid compound nouns/adjectives, and compound verbs. Both compound groups show that English loanwords are adaptable with Persian and Arabic loanwords, already exist in Urdu, for the formation of compounds. The following subsection focuses on compound nouns and adjectives.

5.3.2.1. Hybrid Compounds: Noun and Adjective Formations

Hybrid compounds are the formation of two or more constituents from different languages. As pointed out earlier, Capuz (1997: 8) considers the terms *hybrid* or *loan blend* for them and states, in Urdu, there are a large number of hybrid compounds formed by one constituent borrowed from English and the second from Urdu, e.g. *dəbəl* 'double' (E) + *roti* 'chappati' (U) → *dəbəlroti* 'bread'. They are the third most frequent example of English loanword adaptation after pluralisation and compound verbs.

Regarding hybrid compounds in Pennsylvania German, Schach (1948: 121) cites Lambert's (1924) list of over a hundred hybrid compounds, in which the constituents from different languages may be wholly or partly integrated. For example, one loanword *fens* completely supplants the native words, but another loanword *hendəl* only partly replaces Pennsylvania German words, as in *griff*/ **hand***griff*. There is no partial integration of the Urdu/English constituents, however, the compounds sometimes reflect the assimilation of the constituents.

For example, *səbzimarkit* 'vegetable market' (8d), given below, is so commonly used in Urdu that the two constituents, i.e. *səbzi* 'vegetable' (Persian) and *markit* 'market' (English) seem to be assimilated with each other, so much so, that the compound seems to be a single entity. Sometimes, constituents may be synonyms and similar in meaning. Like those in Pennsylvania German, there may also be a division into two groups depending on those designating objects for which there are no Urdu words, and those which exist beside an Urdu doublet e.g. *phəl **frut*** 'a variety if fruits'. The Urdu constituent in the compound is semantically an Urdu doublet of the English constituent and both have the same meaning.

A hybrid combination of a native and a loanword is a general phenomenon. As there is a heavy influence of Persian and Arabic on Urdu, naturally there is a possibility that an English loan may blend with a Persian or Arabic word already present in Urdu. Classifying the hybrid formations of English loanwords with native Urdu, Persian and Arabic, it is interesting to see that most hybrid formations are those with Arabic loanwords and then Persian loanwords. In Kent's (1999) terms such hybrid formations as pseudo loanwords. The examples in the tables (8)–(10) show that most hybrid formations are pseudo type of loan formations. However, the combinations with native Urdu words as well as original English compounds are also seen. Most compounds are nouns, and refer to food, administration and institutions etc.

Table7. Hybrid Compounds: NN formations

a – b   (NATIVE)—NN/NV

a. phəl 'fruit' (N) + frut 'fruit' (N) → phəl frut '(a variety of) fruits'   (N)

b. ræl 'rail' (N) + gaɽi 'vehicle' (N) → ræl gaɽi 'train'   (N)

c. maḏər 'mother' (N) + tʃoḏ 'fuck' (V) → maḏər tʃoḏ 'abusive remarks for somebody' A/N

c – d (PERSIAN)

c. tʃɪkən 'chicken' (N) + kəbab 'roast meat' (N) → tʃɪkən kəbab '(a dish of) roasted chicken'

 d. səbzi[13] 'vegetable' (N) + markit 'market' (N) → səbzimarkit 'vegetable market'   (N)

e – f   (ARABIC)

e. sɪti 'city' (N) + nazɪm 'administrator' (N) → sɪti nazɪm 'city mayor'   (N)

f. səfəha 'page' (N) + nəmbər 'number' (N) → səfəha nəmbər 'page number'   (N)

Table 8. Hybrid Compounds: AN formations

a - (PERSIAN)

a. ɪmt̪ehani 'examination' (A) + farm 'form' (N) → ɪmt̪ehani farm 'examination form'   (N)

b – d   (ARABIC)

b. behri 'naval' (A) + kəmando 'commando' (N) → behri commando 'naval commando' (N)

c. tʃif 'chief' (A) + sahɪb 'Mr.' (N) → tʃif sahɪb 'a title formerly given to a politician'   (N)

d. qomi/subai 'national/provincial' (A) + əssembli 'assembly' (N) → qomi/subai əssembli

   'national/provincial assembly'   (N)

Table 9: Hybrid Compounds: NA

a – b   (ARABIC)—NA

a. sekretəry 'secretary' (N) + xardʒa 'foreign' (A) → sekretəry xardʒa 'secretary of foreign

   affairs' (N): Similarly: sekretəry xəzarna 'secretary of finance', sekretəry d̪axla 'secretary

   of internal affairs'   (N)

b. rʊpe 'rupees' (N) + fi 'per' + kɪlo 'kilo' (A) → rʊpe fi kɪlo 'rupees per kilo'   (N)

---

[13]. Although used in many south Asian languages, *səbzi* is originally a persian word for vegetable (Persian online dictionary: *Aria Dic*-- http://api.ariadic.com/?word=&lcode=20)

Most combinations are NN and AN, but there are some examples of NA and NV as well. Classifying the hybrids within these categories, it can be seen that the NN class has English loanword combinations with native Urdu words (7a–7b), Persian loanwords (7c–7d) and Arabic loanwords (7e–7f). Most of these compounds refer to food items, transport, and administration. The AN class is mostly formed with Arabic loanwords (8b-8d) and then Persian loanwords (8a). Semantically, these compounds generally refer to political institutions. The NA and NV classes may be formed with native Urdu and Persian loans, but most structures are formed with Arabic loans. They mainly refer to government administrative machinery (9a). One formation (9b) refers to an interesting fact that a compound may also be a trilateral structure, i.e. a formation of native Urdu, (Arabi/Persian) and English loanword. Here *rʊpe* is plural of Sanskrit based word *rʊpeja'*, *fi* is Persian and *kɪlo* is an English loan.

Like general Urdu compounds, most compounds with English constituents are from the endocentric class. The other classes of compounds, the copulative compounds *sɪti nazɪm* 'mayor', are few in number. The patterns also reveal that the English loanword may appear both as a head and a modifier disregarding whether the compound is right headed or left headed. For example, *qomi əssembli* 'national assembly' (8d) is right headed, but *sekretəry xardʒa* 'secretary of foreign affairs' (9a) is left headed. The English loan constituent is the head and the native one is the modifier in both compounds.

There are also some compounds, e.g. *ræl gaɽi* 'train' in which the English constituent may function as a modifier and the Urdu constituent as the head of the hybrid compound. The two compounds, i.e. *ræl gaɽi* and *sɪti nazɪm* show that whether the head constituent is the native *gaɽi* or a Persian/Arabic loan *nazɪm*, the appearance of an English constituent as modifier also forms a variety of compounds. These examples are similar to the ones Schach (1948: 123)

observed in hybrid compounds in Pennsylvania German. For example, he considered the English loanword in compounds, e.g. *ke spaɪ* 'cheese pie' as the final or basic element, but others, e.g. *paɪgruʃt* 'pie crust', *paɪ* is the descriptive or definitive element.

AN or NA compounds in Tables 8 and 9 need some explanations that they are compounds and not just an adjective modifying a noun syntactically. Hüning (2008) considers that the sharp distinction between syntactic structure of an adjective modifying a noun and morphological word formation by an adjective and a noun, found in many grammatical theories, fails to account for the functional similarities between phrases and A+N compounds. However, there is a division of functions between syntax and word formation. Some of the functional differences, he mentions, are as follow: 1) the naming function is typical for the compounds. It is the task of compounds with an adjectival first element to characterize or set apart something as a special category. In contrast, phrases have the task not to name, but to describe entities. Some of the phrases in Dutch lose their inflectional marking at the end of the adjective when used as names. 2) Compounds are stressed on the first constituent, phrases on the second e.g. *bláckbird* vs. *black bírd*. 3) Compounds are written as one word, phrase as two. Thus it is often difficult to define the distinction between compounds and phrases, but there are certain elements e.g. naming, stress pattern and orthography. Looking at Tables 8 and 9, the examples given only seem to show names of some animate or inanimate objects, although the other criteria Hüning (2008) defines, i.e. stress patterns and orthography, may be seen in other compounds. Unlike English, stress pattern plays no role. There may be some formations orthographically seem to be single words, but there is non among the given ones. One point of difference is that the adjectival constituent is not the first element, whether it is an English or Urdu constituent. The compounds formed may be left branching e.g. *qomi*

əssembli 'national assembly' or right branching e.g. *sekretəry xardʒa* 'secretary of foreign affairs' disregarding the position of English constituent.

Hybrid compounds are sometimes formed due to the absence of a convenient general term in Urdu and an overabundance of specific terms in English loans for the same purpose. The formation thus becomes a necessity. This is because Urdu has no adequate vocabulary in some aspects of life, particularly in science and law. For example, there is no native substitute for the English loan constituent *nəmbər* 'number' in the compound (7f). Urdu vocabulary for much of the government machinery has also long been discarded and thought to be obsolete now. Therefore, the word *sekretəry* in *sekretəry xəzana* 'secretary of finance' (9a) and the like do not have Urdu counterpart. Thus, the English loans are convenient to use in the hybrids.

The compounds formed about government machinery are the most among all compounds, which cover various aspects of life. The formations are generally related also to some food items, court and transport, and show various semantic features similar to those seen in hybrid compounds in Pennsylvania German. For example, *phəl frut* '(a variety of) fruits' (7a) shows that they sometimes exhibit semantic duplications in the two constituents (one Urdu and the other English), which are synonyms. However, the compound refers to 'a variety of fruit'.

In short, the hybrid compounds of nouns and adjectives have various formations belonging to various aspects of life. Most of them are from the endocentric class, but there are a certain number of other examples, e.g. copulative. The English constituents may function both as head and modifier without any distinction whether a compound is right headed or left headed. Most of these compounds are on a peripheral target language pattern. They are formed by English loans with the combination of Persian and Arabic loans that already exist in Urdu.

This shows that English loans are not adaptable with Persian and Arabic affixes, but that they conveniently adapt with Persian and Arabic loanwords in compound forms.

This adaptability is not restricted to the formations of nouns and adjectives only. The dominant majority of these hybrid compounds by English and Perso-Arabic constituents is in verb formations. Winter (2008: 168) observes that it is doubtful whether it is possible to have borrowed verbs without any morphological integration at all. Borrowed verbs are normally either inflected according to a productive pattern in the TL or used in periphrastic forms, which avoid an inflection of the borrowed item proper. For example, an English loan verb *kritisaize* 'criticise' is used as *kritisaise kərna* 'to finish', because a loanword for expanding verbal repertoire must be facilitated by a device called a dummy verb.

5.3.2.2. Hybrid Compounds: Verb Formations

The use of verbs without any morphological integration in the native morphology is not possible. No verbs enter with their original structure, probably due to Urdu's morphologically complex verbal system. As seen with Persian and Arabic loan verbs, the verbal structure is formed by compounding the English adjective, noun, verbal noun or even a base verb with an Urdu dummy verb, e.g. *kərna* 'to do' (for active) and *hona* 'to be' (for passive). As noted in Chapter 4, the two are restructured forms of the Persian auxiliaries *kərḍən* and *buḍən/ʃŏḍən* respectively. Thus, the integration in compounding is on a peripheral TL pattern, and Persian auxiliary verbs used as dummy verbs in Urdu play a key role. It can be said that English loanwords combining with Urdu dummy verbs are the most possible and the most frequent forms of their adaptation with Persian-(Arabic) loanwords. Consider the following examples:

Table 10: Hybrid Verbs with English Loans

a. kritisaize 'criticise'→ kritisaize kərna 'to criticise'/ kritisaize hona 'to be criticised'

b. ərændʒ 'arrange' → ərændʒ kərna 'to arrange' / ərændʒ hona 'to be arranged'

c. rɛnəveɪt 'renovate' → rɛnəveɪt kərna 'to renovate' / rɛnəveɪt hona 'to be renovated'

d. eksɛpt 'accept' → eksɛpt kərna 'to accept' / eksɛpt hona 'to be accepted'

e. əpruv 'approve' → əpruv kərna 'to approve' / əpruv hon*a* 'to be approved'

f. beg → beg kərna 'to beg' / beg hona 'to be begged'

g. responsəbɪlti → rɪsponsəbɪləti ḍena/ lena 'to assign/take responsibility'

h. rɪspons 'response' → rɪspons ḍena 'to respond'/rɪspons lena 'to take response'

i hæpi 'happy' → hæpi kərna 'to make happy' / hæpi hona 'to be happy'

j. æŋgri 'angry' → æŋgri kərna 'to make angry' / æŋgri hona 'to be angry'

The examples (10a-10f), e.g. *kritisaize* (10a), show that English base verbs are the bases with Urdu dummy verbs to derive verbs. The dummy verbs are in active and passive forms, but those in (10g-10h) are only in active forms. The latter ones i.e. *ḍena* 'to give' and *lena* 'to take' are also different in that they show handing over and taking over something e.g. a certain job. Unlike Persian and Arabic base verbs that may undergo modification to attach to a dummy verb, English base verbs, as in (10a-10f), do not undergo any morphological changes. An English base verb can only be used as a verbal noun, and the function of verb is adopted by a dummy verb. In general terms, verb borrowing is supposed to be the most complicated in the derivation of new words. Reed (1948: 243) notes regarding English loans in Pennsylvania German morphology that the verb conjugation is the same in the case of loanwords (English) as for native words (Pennsylvania German), and phonetic rules are observed accordingly; if the English verb ends with [d/t], it is conjugated just like German verbs with root-endings in

[d]. This shows morphological changes in the loan constituents. In contrast, there is no alternation in the English loan base words, in their integration with Urdu dummy verbs.

Moreover, in contrast to Persian/Arabic bases which may be nouns, adjectives or verbal nouns, the English bases may also include basic forms of verb. As mentioned earlier, discussing degrees of grammatical productivity, Dressler (2003: 35) also points out the integration of loanwords with unfitting properties in which foreign words which are not adaptable in the normal circumstances can be integrated into the system by a given derivational rule. For this, a rule must have maximum productivity in order to overcome the two obstacles of foreignness and unfitting properties. For example, English (and German) verbs have no thematic vowels, thus a thematic vowel must be added when adapting an English loan-verb into a Romance language: E. *to dribble* > It. *dribbl-a-re*, Sp. *dribl-a-r*. The English base verbs in Urdu i.e. *kritisaize* (10a), *rɛnəveɪt* (10c) and *beg* (10f) with unfitting morphological properties can be integrated into the Urdu verbal system by obligatory derivational verbal constituents, e.g. *kərna* 'to do' (dummy verb) to form verbs as *kritisaize kərna* 'to criticise'. This shows that although Persian and Arabic affixes aren't productive and derive few words in combination with English loans, the dummy verbs are productive. They can accommodate not only English loan nouns, adjectives and verbal nouns but also base verbs, which cannot be integrated into other forms and with other categories of words for verbal capacity. Matras (2009: 182) argues that languages tend to adopt morphological adaptation strategies to accommodate complex loan verbs, which can be interpreted from full acceptance of the "verbness" to the denial and consequent need to reconstruct the verbness. Therefore, it can be said that the Urdu system of dummy verbs functions perfectly for the acceptance of English verbs which are otherwise not attested in Urdu morphology. There seem no alternations in English loan bases in the compound verbs.

The discussion of hybrid compounds, whether they are nouns, adjectives or verbs, shows that the integration of English loanwords is based on the pattern of Persian and Arabic loanword integrations with native Urdu words. However, this recent loanword integration may be far less (in amount) than the older ones.

Spread across a long period of time, the integration of Persian and Arabic words with native Urdu words caused an integration of the loan affixes in the Urdu morphological structures. The time period of English loanword integration is far shorter than that of Persian and Arabic. Therefore, it is noteworthy that English loanword integration has not yet caused any integration of English loan affixes in Urdu morphology. However, certain examples show that there is some usage of English loan affixes, mostly in an informal language.

5.4. English Loan Affixes in Urdu

For morphologically rich languages, like Urdu, a lot of grammatical information is expressed via affixes. Its rich variety in the affixation process is due to its nature of borrowing and absorbing a large number of loanwords and their affixes. Therefore, it can be expected that English affixes may also find some place in Urdu morphology. However, in contrast, English affixes have not yet been able to get a place in its morphological structures. They are used in certain positions but mainly in the informal language or in the media by the elite. Rubba (2004: 2) thinks that English has three inflectional categories. They are correlated with nouns in number, verbs in Tense/Aspect, adjectives (and adverbs) in comparison. The continuous use of English loanwords and code switching has caused these three types of loan inflections to enter Urdu. There may seem plural markers, adjectival markers and progressive markers. Although most are restricted to informal use, some may be coming up parallel to some of the present Urdu affixes whether native or Persian and Arabic loans, as shown below:

5a. Plural markers: *-ies* and *-s*

   ʃərab**ies** 'drinkers' / sərd̪**ɑrs** 'tribal lords'

   b. Adjectival markers (with city names):   *-ite*, *-ean*, *-i*

   Karachi**ite** / Lahore**an** 'a citizen of Karachi/Lahore':

   c. Progressive marker (used as an inflectional suffix):   *-ing*

   kh*a* 'eat' ₊ ing → kh*a***ing** 'eating'

   d. Nominal markers: (emphasis)   *-ation*

   gəɽbəɽ 'disorder' + -*a*tion → gəɽbəɽ*ation* 'disorder'

The use of the same word in the increasing exposure to English (Smeaton 1973) can lead to the introduction of the patterns of such lending languages into the receiving language. This can be observed in English plural patterns, which are gradually getting attraction from the speakers of Urdu. The loan plural markers *-s* and *-ies* seem to make a place in Urdu morphology, if persistently used in the media. They are often used with English loan plurals, but they are also used with Urdu plurals, e.g. *mullahs* 'religious scholar' and *sərd̪ars* 'tribal lords' in informal language. Although Urdu has gender marking, with this pluralisation by English markers, there is no distinction of gender, e.g. *tʃəpaṯ̪ies* 'chapatis' (f) and *nəmaz***ies** 'people praying' (m). It is only by assumption that a noun is differentiated in its gender.

The second important affixation comes with the adjective markers *-ite* and *-ean* that are affixed with city names. However, they are mainly restricted to print media. The two affixes *-ite* and *-ean* were known after they were first used in English papers, and it was gradually increased in talk shows in the electronic media. Otherwise, they are not generally used. The general public would rather use *-i* with any city name, e.g. *Hyderabad**i*** or even *Lahor**i***. This marker seems more adaptable in the context of its use with country names, e.g. *Pakistan**i*** or

Israeli. Despite its common use internationally with many other countries, it is no less than a native adjectival marker with Pakistani city names e.g. *Mirpuri*, *Sejalkoti* and *Islamabadi* etc. The extent of its adoption cannot be claimed, though it looks like a native adjectival marker.

Other possible affixes are the English progressive marker *-ing* and the nominal marker *-tion*, used only rarely in the informal language. People may say *kha* 'eat' *+-ing → khaing* (mẽ 'I' + *a:m* 'msango'+ *khaing* 'eating' → mẽ *a:*m *khaing*) rather than use a progressive from Urdu. However, it is not in accordance with the Urdu syntactic structure for progressive sentences, and does not seem to integrate in Urdu morphology. Although the nominal marker *-tion* has a higher chance of fusing into the system, it is equally uncommon. A little semantic change is seen in the use of *-ation* e.g. *gəɽbəɽation* 'disorder'. The noun with its use is more emphatic.

Apart from plural markers, other affixes have fewer chances of integration. An interesting observation is the use of some affixes by a specific class of people. For example, the adjectival markers for city names, i.e. *-ite* and *-ean*, are mostly used by the elitist or educated people. This is also because they are frequently used in English newspapers e.g. Daily Dawn, the largest English newspaper. The progressive and the nominal markers i.e. *-ing* and *-ation* on the other hand are used by the common people, not very educated. Unlike the adjectival markers *-ite* and *-ean*, the progressive and the nominal markers *-ing* and *-ation* are restricted to informal language. Both do not seem to be integrated. Only English plural markers among all affixes have some possibilities of integration due to their use without any distinction.

The discussion on the English affixes and on the adaptation of English loans is summarised in the following section with some results drawn.

5.5. Summary of the Chapter and the Results Drawn

English loanwords are commonly used in various forms, and the nativisation shows morphological changes in gender marking, pluralisation, derivation, hybrid compounds, reduplication and the affixation of English morphemes. However, the discussion was restricted to a few most important points.

The native Urdu gender and number marking of English loans, e.g. *daktərnɪjã* 'doctors' (f), is in common use. The adaptability of loans is restricted in derivation by affixation, although they exhibit various affixation types from native Urdu, Persian and Arabic sources. The evidence of lexical creation, e.g. *fradɪja/fradi*, by native Urdu affixation shows that English loans are adaptable in Urdu morphology and may take the status of Persian and Arabic loans. The affixation of Persian and Arabic morphemes, e.g. *fəlsəfɪjana* 'philosophical' and *borɪjət̪* 'boredom', can also be seen with some loanwords. The derivation by compounding with Persian and Arabic constituents, e.g. *fɪnɪʃ kərna* 'to finish' and *sɪti nazɪm* 'city mayor', is the most widespread adaptation of English loans after pluralisation. However, English loan constituents show no morphological changes. This restriction is clearly evident also in the nativisation of English affixes except for some instances e.g. the plural markers *-s* and *-ies*.

In sum, English loans are adaptable and native Urdu affixes are productive, in the process of inflection and derivation by affixation. Derivation by compounding of English loans is more frequent with Persian and Arabic loans. There are some results drawn from the discussion on the whole adaptation process.

5.5.1. Results Drawn from the English Loanword Adaptation

 The description of the morphological adaptation of English loans in light of the three sources of morphological structures draws some findings, as shown below.

1. The productivity of the native Urdu affixes and the adaptability of the English loanwords show that Urdu affixes whether the gender and number markers (Section: 5.2.) or derivational affixes *-i/-ɪja* (Section: 5.3.) are productive and English loans are adaptable. Therefore, the integration takes place with the native Urdu affixes only.

2. Although the Persian and Arabic loan affixes function in Urdu, and may also be productive in the original languages as well, they are not generally productive with the English loanwords (Section: 5.3.) except for a few dummy verbs (modified from Persian auxiliaries), Persian adjectival marker *-ana* and Arabic nominal marker *-ɪjət̪*. In addition to these few, there can rarely be any integration of English loanwords with the Persian and Arabic loan affixes. This is despite the fact that English loanwords are adaptable in Urdu, but their adaptability does not extend to the Persian and Arabic loan affixes.

3. When Urdu borrowed a word from Persian and Arabic, it also borrowed the derivational patterns (Chapters 3 and 4), and so various other derivational categories of the same word. In contrast, when Urdu borrows from English, it does not generally borrow the other categories of the same word. Rather, it derives from the loanwords according to its own derivational system. Words like *bəmbarɪ* 'bombing' and *fəlsəfi* 'philosopher' (Section 5.3.) are good examples of the adaptation.

4. As derivational changes of English loanwords by affixation are far fewer than derivational changes of Persian and Arabic loanwords, it is clearly seen that the formation of new words by the combination of native Urdu affixes and English loanwords is limited. However, as the examples show (Section 5.3.), it is interesting to see a variety of new words derived. A process of lexical creation, e.g. *fradɪa/fradi*, with relevance to native Urdu morphology, is

also evident and even more interesting. The new words are a lexical creation rather than simply derivation, as there is no such derivation in the English lexicon for their meaning.

5. More frequent derivation of words is achieved however by hybrid compounding (Section 5.3.2.1.), not by affixation. The hybrid compounds derived by English loanwords with Persian and Arabic loanwords are greater in number than those with native Urdu words. It is noteworthy that the adaptability of English loans does not take place with Persian and Arabic affixes, but the formation of compounds with the older loans is not prohibited.

6. Although the patterns of English loanwords adaptation are the same as the patterns of Persian and Arabic loanword adaptations, the extent of English loanword adaptations is far less than Persian and Arabic loanword adaptation.

7. English loans generally show no structural changes with any affixes from the three sources i.e. native Urdu, Persian and Arabic. This is particularly seen in the English loan constituents in hybrid compounds (Section: 5.3.2.1) and is unlike the older Persian and Arabic loans whose constituents were often incorporated, as in Urdu copulatives (Chapter 4).

8. Section (5.4.) shows that English affixes have not yet found any proper place in Urdu morphology, except in some instances of plural markers -*s* and –*ie*. On rare occasions, they are used in informal language e.g. *vəderas* feudal lords'. Moreoves, they are used as nominative plurals only, and there is no use as oblique forms.

These results drawn from the discussion on the use of English loan affixes and English loanwords adaptation require some general conclusion to be made in a separate chapter.

****

Chapter 6

General Conclusion from the Discussion

6.1. Introduction

Persian, Arabic and English have been very influential languages in the world. A large number of the world's languages have borrowed from each of them. Urdu is unique in that it is a recipient language from them all with a large number of loanwords. However, the manner of borrowing is different. In the case of Persian and Arabic, the borrowing is historical and very old. The influence began from the early development of Urdu after the 12th century, and became even stronger in the Mughal era in the Indian subcontinent. The influence of English loanwords is more recent, and seen particularly in the last decade mainly after the introduction of electronic media.

There are various aspects of the linguistic influence from the three great languages, but unfortunately no research has been done into any of these aspects relative to borrowing into Urdu, whether lexical, phonological, morphological or syntactic. The aim of this thesis is to see how Persian, Arabic and English loanwords are morphologically adapted. Any research into the issue becomes even more interesting when one considers that two prominent languages, Persian and Arabic, have had a historical influence on Urdu with various morphological changes. The loanwords from them have played a key role in its early lexical and morphological development. Thus, Urdu morphological structures come from the three sources i.e. native Urdu, Persian and Arabic. Considering the morphological structures from these three sources as three aspects of Urdu morphology, the study looks at how the loanwords from English are adapted. English loawords are more recent and so treated as foreign. The hypothesis at the beginning of the study was that the affixation, whether

inflectional or derivational, might be on native Urdu patterns but that the compounding of English loans is more with the older loans (Persian and Arabic).

The present study is descriptive due to the non availability of relevant literature and non availability of any theory dealing with the morphology of loans. The bigger part of the study deals with Urdu morphological structures with reference to the three sources of lexical items. The next sections summarises the Urdu morphological structures first and then the adaptation of English loans in the light the three sources. They are followed by some discussion.

6.2. Summary of the Discussion on Urdu Morphology: Native Urdu, Persian and Arabic

The discussion begins in the first chapter with an introduction into the study and an overview of the sociolinguistic background of Urdu. Persian and Arabic loanwords play a part in the morphology of Urdu in particular, and south Asian languages in general. English loanword influence is much more recent, and so a brief comparison is made between the morphology of English loanwords in five south Asian languages i.e. Urdu, Punjabi, Sindhi, Pashto and Marwarhi. The study focuses in particular on gender and number morphology and the formation of compounds.

Chapters 2, 3 and 4 discuss various aspects of Urdu morphological features. Morphology is a comprehensive field of study. All of its aspects cannot be covered in a short work. Therefore, only some selected points, i.e. the gender, number and case morphology of nouns, derivational affixes and some types of compounds are elaborated on.

The native Urdu gender, number or case morphology is based on the alternation of vowels, which also shows nasalisation in feminine plurals. The native gender markers *-a* (m) and *-i* (f) are heavily dominant. Masculine is the basic gender and feminine is only a derived form.

Some masculine nouns also end with -*i*, and in this case the feminine nouns must end with -*ən*. A large number of nouns do not have any specific masculine marking, and their feminine forms must end with -*ni*. These markers in general express gender rule. However, some native nouns deviate from this rule. The rule becomes further irrelevant in loanword morphology. Some Persian and Arabic loans take the Arabic feminine marker -*a*, which is a homophone of the native masculine gender -*a*. Persian in fact expresses no gender marking, and its loans in Urdu adopt native gender mostly.

There are also deviations in pluralisation. Native masculine nouns generally end with -*e*, if singular nouns end with -*a*, otherwise they remain unchanged. Feminine plurals end with -*ɪjã*, if their singular form takes -*i*. All other feminine nouns take the plural marker -*ẽ*. Although most loans take native plural markers, some e.g. -*an*, and *aṯ* (both masculine) from Persian and -*in* (m) and -*aṯ* (f) from Arabic have also entered Urdu morphology and function like native plural markers. The most interesting factor in pluralisation is the non Arabic loans' formation of broken plurals, a feature that is only associated with Arabic loanwords. It shows the nativisation of the Arabic patterns of broken plurals, but it also shows deviation from the native Urdu and the Arabic pluralisation rules.

There are also deviations in the derivation by affixation. Urdu contains a variety of native and loan derivational affixes, discussed in Chapter 3. There are more prefixes than suffixes, but most derivations are made by suffixes. Native affixes are few and based on vowels mainly, but they are the most productive of all. Persian affixes are mainly prefixes and though largest in number, they only derive a few words. However, some affixes, e.g. *na* 'no/not' and -*ana* are quite productive and may attach to non Persian words very often. There are not many Arabic affixes, but some, e.g. *la* 'no/not' and *ɣær* 'not/without' (both negative markers) also

show productivity. The dominant among loan derivational affixes are negative markers, e.g. *na*, *be* from Persian and *la*, *bɪla* and *ɣer* from Arabic. All loan affixes express semantics. Some of them have interchange ability with or without semantic changes in the derivations. This shows variations in the Urdu derivation. They seem more often in the various patterns of Arabic derivations by modification of the base, as the patterns are nativised by taking affixes from native Urdu, Persian and even from Arabic. For example, the Arabic tri consonantal root *zlm* derives the noun *zɒlm* 'tyranny' and adjective *zalɪm* 'tyrant'. But the native *-i* and Persian *-ana*, make further derivations of adjectives, *zɒlmi* 'cruel man' *zalɪmana* 'oppressive'. Such derivations are frequent, although there are some irregularities in their formations.

Arabic loan derivational patterns by modification of bases also show an interaction between morphology and phonology. Many morphological changes are caused by phonological changes. Therefore, a discussion on some important aspects of Arabic loanword phonology was necessary. Although most Arabic phonemes lose their phonetic value and are replaced by native Urdu phonemes, some are still part of the Urdu phonetic inventory. Fricatives /x/, /ɣ/ and the uvular stop /q/ were not only diachronically substituted for native phonemes in native words but also created new lexical items. Thus, they are now part of the Urdu phonetic inventory and function like native phonemes. However, as recently noted, in a shift back the borrowed velar fricatives /x/, /ɣ/ and the uvular stop /q/ are now substituted by the velar stops /kʰ/, /g/ and /k/. The data confirms Wohlin's (2010: 38) observation of the reversal of sounds regarding Indo European and Germanic languages, but the features in Urdu are different.

Phonology affects not only derivation by affixation but also derivation by compounding. In Urdu, compounding of constituents from two different sources is a frequent function. These hybrid compounds are more frequent than normal, i.e. native + native, combinations.

Therefore, a discussion with respect to the source languages in this aspect seemed irrelevant in chapter 4, which focuses rather on the various types of compounds. By types of classification, most compounds are endocentric and verbal compounds. Then, they include exocentric, copulative and postpositional compounds.

Urdu compounds may be left branching (modifiers that come before head) or right branching (modifiers that come after head). Both constituents are often synonyms or near synonyms, although they may be from two different sources e.g. in *zɵlm-o-sɪt̪əm* 'tyrannies' N_1 is Arabic and N_2 is Persian. Urdu compounds are generally formed by distinctive compounding words, each of which may form several combinations; they may also have infixes that structurally link both constituents. There are three infixes *-e-*, *-o-* and *-b-* that perform grammatical, semantic and phonological functions. They are mostly found in endocentric and copulative compounds. No verbal compounds contain any infixes. The key constituents in verbal compounds are dummy verbs, the modified forms of Persian auxiliaries.

Dummy verbs are a major tool for verb formation in Urdu, particularly by compounding with loanwords. This phenomenon of hybrid verb formation is also very frequent in English loanword adaptation, which is discussed in in Chapter 5.

6.3. English Loanword Morphology in the Light of Persian and Arabic Morphology

The central hypothesis is shown to be true, that the affixation of English loanwords, whether inflectional or derivational, tends to be on native Urdu patterns. The compounding of English loans is more like that of the older loans (Persian and Arabic), which form a bigger part of the Urdu Lexicon and function like native Urdu affixes. The study has also found that English loans are integrated more with native Urdu affixes. The adaptation of English loans is not as complete as that of the two older loan influences, and so Urdu does not generally borrow any

derivations of a loan, as it did with Persian and Arabic with or without any morphological changes. The related supplementary questions regarding the structural patterns, the manner of loanword modification, the elements affected and the English loan affixes support this fact. Urdu does not generally borrow English affixes, but some, e.g. plural markers, are in the informal language.

Examining the questions, in detail, the study has found that English loans are adaptable and their integration takes place with native Urdu affixes, which are productive and conveniently attach to loanwords. There are both inflectional and derivational changes. Inflectional changes are extreme and only based on native Urdu patterns. English loan nouns take the native gender markers *-a* and *-i* and the plural markers *-e* (m) and *-ɪjã̃/-ẽ* (f).

In derivation by affixation, English loans are generally more adaptable with native affixes, which are productive enough. Words like *fradɪja,* 'one who commits fraud' even show lexical creation. Although Persian and Arabic loan affixes function like native affixes in Urdu, and may be productive in the original languages as well, they are not generally productive with English loanwords. Except for a few dummy verbs, such as the Persian adjectival marker *-ana* and the Arabic nominal marker *-ɪjət̪*, no other Perso-Arabic element is integrated with English loans. This is despite the fact that the latter is adaptable, but the adaptability does not extend to the former and so the integration cannot occur freely.

However, frequent derivation is made mainly by the verb formation of English loanwords with Urdu dummy verbs. Although the pattern of verb formation is native, the major dummy verbs are the restructured forms of Persian auxiliaries. This makes compound verbs according to the peripheral target language system. These hybrid formations are greater in number than those with any other words from the three sources. It is interesting to see that English loans

are not adaptable with Persian and Arabic affixes, but the formation of compounds with the older loans is not prohibited.

The adaptability of the recent loans is limited irrespective of the morphological patterns they follow. It is far less than the adaptation seen in Persian and Arabic loans. There are two factors, 1) the influence spread across centuries' long period and 2) Urdu's own developmental stage differentiate the capacity of the past and present influences. English loanword influence is spread over barely a century and a half. Its extreme form is seen during the last few decades, particularly with the introduction of electronic media. Urdu is already a well established language and the huge language contact has caused only some morphological changes. In contrast, the Persian and Arabic influence began with the early development of Urdu, nearly a thousand years back.. Therefore, it is natural to see far more morphological features in the two older languages than in the latter one.

Another point is the morphological systems of the lending languages, Perso-Arabic and English in this case, that may affect borrowing languages, as Urdu in the present case. Al-Jabri (2010: 3) argues that English has a morphology which is not as productive in its overall derivation as some other morphological systems such as those found in Semitic languages (e.g. Arabic). With lexical borrowings from Persian and Arabic, Urdu also borrowed derivational patterns. Therefore, a number of other derivational categories of the same word can be seen. These derivational patterns are now part of Urdu morphology, as the affixes in the loan derivational patterns may attach to the bases other than from their source languages. Thus, they function like native affixes. However, this is not true with English loanwords. Urdu does not normally borrow from English the other categories of the same word. Rather, it derives words according to its own derivational system. Words like *bəmbari* 'bombing' are

good forms of the adaptation. This shows that English derivational (and inflectional) morphemes have not been integrated into Urdu. This is why English loanwords depend on native Urdu morphemes or in some cases Persian and Arabic morphemes already present. The only English loan affixes, sometimes used mainly with English loanwords, are the plural markers *-s* and *-ies*. On rare occasions, their affixation with Urdu bases, e.g. *mullahs* 'priests' also occur in informal language. However, the proposition here is that there is some lexical conditioning for it. The nouns ending in *-a* and *-i* e.g. *mulla*(*h*) 'priest', *vədera* 'feudal lord' and *vahabi* 'a religious sect' may be the convenient bases for English loan plural markers. The nouns are similar to English nouns e.g. *visa*, *baby*, which are more easily adaptable than others, possibly because of the harmony in the endings of such nouns in the two languages. This harmony in the endings suggests a possibility of English plural markers entering into Urdu. Presently, as mentioned above, their occurrence is very limited to a specific class of people. The adjectival markers *-ite* and *-ean* are in even more limited use.

Thus, English loanwords have brought into Urdu no inflectional or derivational patterns, unlike Persian and Arabic morphological patterns. Therefore, the creation of new lexical items and new forms also cannot be expected. It is possible when some patterns have been adopted, which can be noted in the case of Arabic loanwords, e.g. a broken plural form *xəvaṯin* and a formation of a noun from a noun e.g. *xəbərɪjəṯ* 'status of news'. However, there are some examples of neologism, e.g. *fradɪa*/*fradi*. The new words are a lexical creation rather than simple derivation, as there is no derivation parallel to these in the English lexicon for their meaning. Al-Jabri (2010: 8) argues that, in general, current approaches to lexicalisation tend to treat concepts as related entities; words, on the other hand, are viewed as isolated items that just happen to be attached to concepts. Both the creations *fradɪa* and *fradi* refer to a person who gets personal gain by treacherous means. They fulfil the lexical need to represent

a person of treacherous acts. Although the Persian loan *ḏəɣ**baz*** 'cheater' or native Urdu derivation *ḏhoke**baz*** are its synonyms and already present in Urdu, there is a key semantic difference between the older forms and the new creations. The older forms are used in the generic sense, while the new creations are used in a specific sense for the one who gets some financial benefits by treacherous means. This shows that English loans have the ability of lexicalisation for a specific semantic concept.

Nevertheless, this explanation is only partly true. More importantly, many English loanwords have retroflex sounds which make them distinct from Persian and Arabic loans. It is also necessary to note that retroflex sounds are seen in South Asian English, not in British or American English. The lexical creations are not due to loanwords' phonemic influence, as seen in the creations by Arabic sounds. Along with lexical borrowings from Arabic, some Arabic phonemes, e.g. the velar fricatives /x/ and /ɣ/ also entered Urdu. In contrast, no English phonemes entered Urdu. However, there is a phonological harmony between many (South Asian) English loanwords and native Urdu words, which have the retroflex sounds /t/ and /d/. This common feature increases the chances of neologism by English loanwords in general. This is verified by many other similar loans, when they conveniently take less frequent feminine markers, e.g. *-ni* and **-ən** in *daktərni* 'lady doctor' and *ʃradən*. Due to the gender distinction in Urdu, the original loans are treated as masculine, the basic gender, and they only need to be feminised. Thus, it can be claimed that any English loan with retroflex sounds will probably be more conveniently adapted and lexicalised than other loans. This is expected also due to the fact that the loans with a retroflex sound, e.g. *totəl* and *daɪve* more conveniently enter the Urdu lexicon and substitute the Urdu counterparts, i.e. *kɒl* and *tʃhəlaŋ*, which are nearly obsolete now.

This further confirms the adaptability of English loanwords however it may be limited when compared to that of Persian and Arabic loanwords. Lee (2004: 1) considers that change in a language occurs either quickly or slowly; the rate of change is unpredictable. There are only some morphological changes in English loanwords, but there is a process of change and English loanwords adapt in Urdu morphology.

However, there are no morphological changes in the English constituents in the hybrid compound verbs when combined with Urdu dummy verbs. Even if there are English loan verbs, they cease to function as a verb, and the verbal capacity is adopted by the dummy verbs. Urdu adaptation morphology is functional in establishing the verbness of loan verbs, but without any morphological changes in them. This is unlike the older Persian loan auxiliaries which function as dummy verbs in their restructured forms and have diachronically helped in establishing a system of compound verb morphology in Urdu in the absence of a base verb. Matras (2009: 182) argues that languages adopt morphological adaptation strategies to accommodate complex loan verbs. English loan verbs alone cannot function with their verbal capacity. Therefore, the Urdu system of dummy verbs functions perfectly to accommodate them, which are otherwise not adaptable. The dummy verbs function the same way with bases from native Urdu, Arabic or English. As required, they show tense/aspect changes, but no changes are seen in the English loan constituents. Therefore, the adaptation of English loans in compound verbs seems no different from the older loan bases. Like older loans, they only follow the set patterns established by Persian auxiliaries. However, certain examples, e.g. Persian noun *xəriḍ* 'purchase', show that older loans may adopt the status of base verbs without requiring a dummy verb. It is not possible with English loans, and so their adaptation is far less than that of Persian and Arabic loans.

Morphological changes in English loans follow the morphological patterns formed by native Urdu, Persian and Arabic. This seems to be a similar situation to the one seen in Turkish loanwords in Aivaliot (a dialect of Greek). Ralli (2011: 185) argues that Turkish loans are adapted to Aivaliot morphology, following specific requirements, which are imposed by the most productive derivational and inflectional processes of Greek, but they are also constrained by features innate to the donor language, i.e. Turkish. There are constraints in English morphological rules functioning in Urdu, as the loan affixes are not active. English loanwords only follow some requirements imposed by derivational and inflectional patterns of the native Urdu, Persian and Arabic. The changes in them are restricted to pluralisation mainly, and therefore they have not yet played any role in Urdu development.

Baik and Shim (1998: 276) advise, "One must not forget that language change is the result of dynamic interaction between internal linguistic development and the external socio-psycho-econo-political factors in language contact situations". Both the factors have played a role in the language change in Urdu as well. Some major factors are given in the following section.

6.4. Factors Influencing the Degree of Adaptation

Internal Linguistic development is an ongoing process, which is also evident in Urdu. One major source of linguistic development is language contact. Persian and Arabic language contact with Urdu contributed to its linguistic development from the earliest times. In contrast, the role of English is very limited due to the shorter period of language contact. Although English has lent no morphological patterns, it does play a role in the enrichment of the Urdu lexicon. The use of the same word in the increasing exposure to English with time and frequency (Smeaton 1973) in the long run, can lead to the introduction of the patterns of such a lending language into the receiving language (Urdu in the present case). This can be

seen in the original plural patterns of English loans, which are gradually increasing in their uses by the speakers of Urdu.

Contact between the two languages correlates the linguistic nature of both, which is also a major factor influencing the adaptation process. The mechanism of native Urdu word formation is important. The productivity of its affixes associates with the adaptability of English loanwords to adopt a morphological pattern. English loans are adaptable as much as native affixes are productive. They do not associate with Persian/Arabic affixes, which though productive in Urdu and in their original languages are unproductive with English loans in Urdu. However, their compound formation with the older loans is not prohibited. Thus the productivity and the adaptability function in due course for their integration.

Ralli (2011: 185) argues that the morphological adaptation of loans is mainly constrained by three factors: (i) the degree of intensity of socio-cultural contact in a language-contact situation (Thomason and Kaufman 1988); (ii) the recipient-language word-formation mechanisms, and (iii) the productivity of morphological patterns. These three factors in fact help in understanding Urdu morphological patterns. The second and third have just been seen, and so the focus is on the first one now. Language contact between two languages depends mainly on external socio-psycho-econo-political factors, which are thus even more important generally and even more extreme in Urdu. The first among them is the use of electronic media. Consistent use of English affixes, e.g. plural markers *-s* and *-ies* by the elite in the electronic media may cause the formation of an English morphological pattern in Urdu. As noted in the first chapter, English is considered to have a high status and prestige as a global language. The use of English loanwords in daily Urdu is not just a status symbol it has become a social habit. The use of plural markers is rare, but it is the new phase of the social

habit. The use may vary and may thus influence the degree of the adaptation process. It is in the highest degree with plural marking and the hybrid compounds but not with the derivation by affixation. This is mainly because of the attitude of the speakers for whom the use of nativised English plurals and the hybrid compounds is convenient.

The increasing education and literacy rate is also a factor for this. Although the huge majority of people are still uneducated, ordinary people continue to get education. English is the medium of instruction in all private educational establishments and from higher secondary classes and onwards in all public institutes. This makes the use of English loans – or at least exposure to them – increase day by day. The dominant role of electronic media and the presence of a large number of English language centres have made it even more rapid. The rapid increase in their use causes loanwords to undergo more nativisation and makes it more possible for their adaptation.

6.5. Implications of the Study

There is a scarcity of linguistic research in Pakistan except for a small amount of work done by a few individuals. This conclusion can be reached from the fact that there is just one Department of (theoretical) Linguistics established by Azad Jammu and Kashmir University a few years back, and one Journal of Linguistics published by the same university. Urdu is no exemption from this scenario. Thus, this study is the first attempt in the research on Urdu morphology. It therefore forms the basis for further research in various aspects of this field.

The first area of further research is the adoption of a theoretical approach. It is now convenient in terms of data analysis, as a description is already present. It is suggested that the lexeme based approach is suitable for future study. Words as the bases are present whether

they are native words or loanwords, and the affixes of one source attach to the words of the other source.

The study of loanword morphology in Urdu within a lexeme based model can help the development of a theory of loanword morphology in general. The absence of such a theory was strongly felt at the beginning of this study, and so it became the reason for the descriptive work done. There are two major points that can specifically be dealt with by the development of a theory of loanword morphology. First, it can help in the evaluation of the morphological adaptation phenomena by loanwords in general. Secondly, it can provide a tool to compare and discuss the linguistics issues involved in the case of a language which is the recipient language from more than one language. The data description of Urdu provides a good base.

There is also a possibility that an OT based approach may be adopted to see the constraints in the morphological adaptations of loans in Urdu. Loanwords from various sources present an interesting base for the comparative analysis of the constraints.

OT can be particularly helpful in the morphology-phonology interface, which is also observed in this work, e.g. affixation of the Arabic loan definite article *ʊl/əl* and assimilation of the lateral. This can be a rich area of linguistic research, for Urdu, as only some surface elements were highlighted. There are other important aspects, for example the role of a non neutral affix, e.g. *-gi* that causes the phonological changes in the stems. There must be other non neutral affixes, which need to be seen.

Due to this morphology-phonology interface, Arabic loanword phonology functioning in Urdu had to be highlighted. This shows that some Arabic loan phonemes not only substitute for native phonemes in native Urdu words but also function in combination with native

retroflex phonemes in lexical creations. The introduction of Arabic phonemes in Urdu was a diachronic process and provides us with another rich area to study the process and its present function as compared to the native Urdu phonemes. Moreover, as many phonemes are now being substituted by native phonemes even in Arabic loanwords, it shows the reversal of Arabic sounds from Urdu. This presents another idea which can broaden the discussion introduced, on the rise and fall of Arabic phonemes in Urdu.

****

References

Abdallah, A. S. 2006. *Teaching Word Roots and the Broken Plural.* London: A Project by CETL School of Oriental and African Studies, University of London.

Abbas, S. (2002*).* "Language of the Armies, Urdu: A Derivative of Persian and Avestan" In *Persian Language and Literaterature*. Iran Chamber Society. Retrieved: 02/2011 http://www.iranchamber.com/literature/articles/language_of_armies.php.

Ahmed, R. 2008. "Unpacking indexicality: Urdu in India". *Texas Linguistic Forum*, 52: 1-9.

Ahmed, T. 2008b. "Instrument as Path: Reconciling the Semantic Usages of Urdu *se* and Pashto *pa*." Paper presented at *Universität Konstanz*, Germany. Retrieved. 9/2010. http:www.ling.uni-konstanz.de/pages/home/tafseer/instrument.pdf.

Aronoff, M. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.

Aronoff, M. 1996. *Morphology by Itself*. Cambridge, MA: MIT Press.

Azim, A. (1975) "Urdu - A Victim of Cultural Genocide?" In Zafar Imam (ed.) *Muslims in India*. New Delhi: Orient Longman, pp. 256-272.

Baik, M. J. and Shim, R. J. 1998. "Martin Jonghak Baik and Rosa Jinyoung Shim: Reply" *World Englishes*, 17(2): 273-279.

Bauer, L. 2003. *Introducing Linguistic Morphology*. Edinburgh: Edinburgh University Press.

Bauer, L. 2005. "Les Composés Exocentriques de L'anglais". In D. Amiot (ed). *La Composition Dans une Perspective Typologique*. Arras: Artois Presses Université. pp 35-47.

Baugh, A. and Cable, T. 1978. *A History of the English Language* (3rd Edition). Englewood Cliffs: Prentice-Hall Inc.

Baumgardner, Robert J. 1998. "Word-Formation in Pakistani English" *English World-Wide*, Volume. 19, Number 2, John Benjamins Publishing Company. pp. 205-246(42).

Baumgardner, Robert J. 1990. "The Indigenisation of English in Pakistan". *English Today*.6: Cambridge University Press. pp 59-65.

Benczes, R. 2004. "On the Analysability of English Exocentric Compounds". *Jezekoslovija*, 5.1-2: pp 1-21.

Bogel, T., Butt, M., Hautli A., and Sulger, S. 2007. "Developing a Finite-State Morphological Analyzer for Urdu and Hindi". In *Proceedings of the Sixth International Workshop on Finite-State Methods and Natural Language Processing*. Potsdam.

Booij, G. 1992. "Compounding in Dutch". *Rivista di Linguistica*, Volume 4(1) pp. 37-59.

Booij, G. 2005a. *The Grammar of Words: An Introduction to Linguistic Morphology*. Oxford: Oxford University Press.

Booij, G. 2005b. "The Interface between Morphology and Phonology: A Comment on Ellen Kaisse: Word Formation and Phonology". *Skase Journal of Theoretical Linguistics*, pp. 246-99.

Bukhari, N. H. 2009. *The Syntax of Serial Verbs in Gojri*. Ph.D Thesis, Newcastle University, United Kingdom.

Butt, M. 1995. *The Structure of Complex Predicates in Urdu*. Stanford, California: CSLI Publications.

Capuz, J. G. 1997. "Towards a Typological Classification of Linguistic Borrowing. *Revista Alicantina de Estudios Ingleses*, 10: 81-94.

Carlisle, J. F. 1987. "The Use of Morphological Knowledge in Spelling Derived Forms by Learning-Disabled and Normal Students". *Annals of Dyslexia*, (volume 27): 90-108.

Carstensen, B. 1986. "Euro-English". In D. Kastovsky and A. Szwedek (eds). *Descriptive, Contrastive and Applied Linguistics* Berlin, New York: Mouton de Gruyter, 827-835.

Crawford, C. (2008). "An Evolutionary Account of Loanword-Induced Sound Change in Japanese". *University of Pennsylvania Working Papers in Linguistics*. Vol. 14. Iss:1. Proceedings of the 31st Annual Penn. Linguistics Coloquium.

David, A., Maxwell, M., Browne, E. and Lynn, N. 2009. *Urdu Morphology*. Maryland: Centre for Advanced Study of Language. University of Maryland.

Davenport, M. and Hannahs, S.J. 2005. *Introducing Phonetics and Phonology*. London, New York: Hodder Arnold.

Dressler, W. U. 1993. "Word-Formation: Poetic License". *The Encyclopaedia of Language and Linguistics*, 9: 5028-5029.

Dressler, W. U. 2003. "Degrees of Grammatical Productivity in Inflectional Morphology". *Italian Journal of Linguistics/Rivista di Linguistica,* 15(1): 31-62.

Durani, N. 2007. *Typology of Word and Automatic Word Segmentation in Urdu Text Corpus*. M.S Thesis, National University of Computer and Emerging Sciences. Lahore, Pakistan.

Grimes, B. 2000. "Pakistan". In Ethnologue. *Languages of the World. 14th ed*. Dallas, Texas Summer Institute of Linguistics.

Gumperz, J. J. and Wilson, R. 1971. "Convergence and Creolisation: a Case from the Indo Aryan/Dravidian Border in India". In D. Hymes (ed.). *Pidginization and Creolization*

of Languages. Cambridge: Cambridge University Press, 151–167.

Hafez, O. 1996. "Phonological and Morphological Integration of Loanwords into Egyptian Arabic". Cairo. *Monde Arabe*, *Première Série*, 27: 383-410.

Hardie, A. 2003. *The Computational Analysis of Morpho-Syntactic Categories in Urdu*, PhD Thesis, Lancaster University.

Hardie, A. (2003). "Developing a Tagset for Automated Part-of-Speech Tagging in Urdu". In Archer, D, Rayson, P, Wilson, A, and McEnery, T (eds.) *Proceedings of the Corpus Linguistics*. pp 298-307

Haugen, E. 1950. "The Analysis of Linguistic Borrowing". *Language*, 26: 210-231.

Hoffer, B. 1980. "English loanwords in Japanese: some cultural implications". *Language Sciences,* 12 (1): 1-21.

Hüning, Matthias (2010). "Adjective + Noun Constructions between Syntax and Word Formation in Dutch and German". In Sascha Michel and Michael Onysko (eds.), *Cognitive Perspectives on Word Formation*. Berlin: Mouton de Gruyter. pp 195-215.

Hussain, S. 2004. "Letter-to-Sound Conversion for Urdu Text-to-Speech System". In the *Proceedings of Workshop on Computational Approaches to Arabic Script-based Languages*, COLING, Geneva.

Ijaz, F. and Moin, M. R. 2003. Gender and number generation of Urdu nouns. Islamabad. *CRULP*. pp 52-56.

Jakobson, R. 1938. "Sur la theorie des affinities phonologiquesentre des lang" ues. *Actes du Quatrieme Congres Internationalede Linguistes*, pp 48-59.

Junior, G. R. (2005) "The Languages of the World" In *Ethnologue* 15th Edition Dallas Tx SIL International Online Version. http://www.ethnologue.com/. Retrieved (04/2008).

Kahnemuyipour, A. 2000. "On the Derivationality of Some Inflectional Affixes in Persian" Chicago. Paper presented at the Linguistic Society of America Annual Meeting, *LSA*.

Kaisse, E. 2005. "The Interface between Morphology and Phonology". In: P. Štekauer and R. Lieber. *Handbook of English Word-Formation*. Dordrecht: Springer, 25-47.

Kashif, D. 2004. *Jadeed Dersee Urdu*. Karachi: Qamer Kitab Gher 'Qamer Book House'.

Katamba, F. 1993. *Morphology*. London: Macmillan Press Limited.

Kent, D. B. 1999. "Speaking in Tongues: Chinglish, Japlish and Konglish". *KOTESOL* Proceedings of PAC 2. (The Second Pan Asian Conference*). Seoul.

King, R. E. 2000. *The Lexical Basis of Grammatical Borrowing: A Prince Edward Island*

French Case Study. Amsterdam, Philadelphia: John Benjamins.

Keane, E. 2001) *Echo Words in Tamil*. PhD Thesis, Trinity College, Oxford.

Koul, O. N. 2008. *Modern Hindi Grammar*. Springfield, USA: Dunwoody Press.

Lambert, M. B. 1924. *A Dictionary of the Non-English Words of the Pennsylvania German Dialect*. Lancaster: PA

Lee, J. J. 2004. "The Impact of English on the Post-1945 South Korean vernacular. In *Karen's Linguistic Issues.* Retrieved: http://www3.telus.net/linguisticsissues/koreanvernacular.

Lieber, R. and Scalise, S. 2006. "The Lexical Integrity Hypothesis in a New Theoretical Universe". *Lingue e Linguaggio*,1: pp 7-32.

Martindale, J. 2009. *Languages of India.* Sydney: Martindale's Language and Translation Centre.

Matras, Y. 2009. "Formal and Functional Aspects of Verb Borrowing". In *Morphologies in Contact*. An International Symposium Financed by the Volkswagen-Foundation. Universitat Bremen.

McGregor, R.S. 1972. *Outline of Hindi Grammar*. Oxford: Clarendon Press.

Megerdoomian, K. 2000. "Unification-Based Persian Morphology". In *Proceedings of CIC Ling.* Alexander Gelbukh, ed. Centro de Investigación en Computación IPN, Mexico.

Mohanan, T. 1994. *Argument Structure in Hindi.* PhD Thesis, Stanford University, CA.

Moizuddin, M. 1989. *Word Forms in Urdu*. Islamabad: National Language Authority.

Müller, M. 1875. *Lectures on the Science of Language*. New York: Scribner.

Nawaz, S. 2002. "Deletion Rules in Urdu Language". *CRULP*, pp 220-222.

Naim, C.M. 1999. *Introductory Urdu.* Cambridge: University of Chicago.

Olsen, S, 2000. "Copulative Compounds: A Closer Look at the Interface between Morphology and Syntax". In G. Booij and J. van Marle (eds.). *Yearbook of Morphology.* Dordrecht: Kluwer. pp 279-320.

Olsen, S. 2000. "Composition". In G. Booij, C. Lehmann and J. Mugdan (eds), *Morphologie/Morphology* 1. Berlin, New York: Mouton de Gruyter, pp 897-916.

Paradis, C. and LaCharité, D. 2001. "Guttural Deletion in Loanwords". *Phonology*, 18.2: pp 255-300.

Plag, I. 1999. *Morphological Productivity: Structural Constraints in English Derivation*,

Berlin: Mouton de Gruyter.

Plag, I. 2003. *Word-formation in English*. Cambridge: Cambridge University Press.

Poplack, S., Wheeler, S. and Westwood, A. 1989. "Distinguishing Language Contact Phenomena: Evidence from Finnish-English Bilingualism". *World Englishes*, pp 389-406.

Qureshi, A. H. and Akram, M. 2008. "Hyberdization in Urdu and Inter-Cultural Perspective". Jamshoro, Pakistan. Sine nomine (no named publisher). pp -1-12.Retrieved. 06/2011 http://www.victimsofterror.info/wp-content/uploads/2011/03/HYBRIDIZATION-IN-URDU-Jamshoro.pdf.

Radcliffe, R. 1990. "Arabic Broken Plurals". In M. Eid and J. McCanhy (eds). *Perspectives on Arabic Linguistics II*. Amsterdam: John Benjamins, 94-119.

Rai, A. 2000. *Hindi Nationalism.* New Dehli: Orient Longman Private Limited.

Ralli, A. 2011. "Morphology in Language Contact: Verbal Loan-Blend Formation in Asia Minor Greek (Aivaliot)". In Th. Stolz, M. Vanhove, H. Otsuka and A. Urdze (eds.), *Morphologies in Contact (Studia Typologica 12).* Berlin: Akademie Verlag, 185-201.

Reed, C. E. 1948. "The Adaptation of English to Pennsylvania German Morphology". *American Speech*, 23: pp 239-246.

Rehman, T. 1998. "Linguistics in Pakistan [Canada]: A Country Report". In Rajindra Singh *The Yearbook of South Asian Languages and Linguistics* (ed) (Delhi: Sage Publications,). pp 184-196.

Rehman, T. 2004. "Language Policy and Localization in Pakistan: Proposal for a Paradigmatic Shift, Crossing the Digital Divide". *SCALLA Conference on Computational Linguistics* 5th to 7th January, Katmandu.

Rehman, T. 2005. "The Muslim Response to English in South Asia: With Special Reference to Inequality, Intolerance and Militancy in Pakistan". *Journal of Language Identity and Education*. Volume 4, Number 2. Lawrence Erlbaum Associates Inc. pp 119-135.

Riaz, K. 2007. "Challenges in Urdu Stemming (A Progress Report)", In *Proceeding of BCS IRSG Symposium: Future Directions in Information Access (FDIA).*

Rizvi, S. M. J. 2007. *Development of Algorithms and Computational Grammar for Urdu*. PhD. Thesis, Pakistan Institute of Engineering and Applied Sciences.

Rubba, J. 2004. *An Overview of the English Morphological System*. San Luis Obispo: California. Polytechnic State University.

Ryding, K. C. 2005. *A Reference grammar of modern standard Arabic.* London: Cambridge University Press.

Sankoff, G. 2001. "Linguistic Outcomes of Language Contact". In P. Trudgill, J. Chambers and N. Schilling-Estes, eds. *Handbook of Sociolinguistics*. Oxford: Blackwell, pp 638-668.

Sapir, E. 1921. *Language: An Introduction to the Study of Speech* (Chapter IX. P. 2.). New York: Harcourt Brace, New York: www.bartleby.com (Online Books Publisher) (Date accessed: January 2009).

Schach, P. 1948. "Hybrid Compounds in Pennsylvania German". .*American Speech*: *A Quarterly of Linguistic Usage*, 23: pp 121-134.

Schmidt, R. L. 1999. *Urdu: An Essential Grammar.* London. Rutledge.

Seidenberg, M.S. and Gonnerman, L. M. 2000. "Explaining Derivational Morphology as the Convergence of Codes". *Trends in Cognitive Sciences,* 4 (9): pp 353–361.

Shahid, S. R. 2002. "A Study on Glottal Stops in Urdu". Islamabad. *CRULP Annual Report*.

Shah, S. M. 2002. "Urdu Nasal Consonant and Phonological Behaviour". Islamabad: *Akhbar-e –Urdu.* Muqtadra Qaumi Zuban. 'Urdu Language Authority'.

Shapiro, M. C. 2003. "Hindi". In C. George and D. Jain (eds.). *The Indo-Aryan Languages.* Routledge. pp 250-285*.*

Smeaton B. H. 1973. *Lexical Expansion due to Technical Change*, *as Illustrated by the Arabic of Al Hasa*, *Saudi Arabia*. Bloomington: Indiana University Publications.

Sogaard, A. 2004. "Compounding Theories and Linguistic Diversity". In: Frajzyn-Gier, Zygmunt, Hodges, Adam & Rood, David S. (eds.), *Linguistic Diversity and Language Theories.*Amsterdam: John Benjamins, pp 319-337.

Takashi, K. (1990). "A Sociolinguistic Analysis of English Borrowings in Japanese Advertising Texts", *World Englshes*, Volume 9, 327-341.

Tatsioka, Z. 2008. "Code-switching and English Loanwords in the Greek Media". *Newcastle Working Papers,* 14: 129-144.

Thakur, A. K. 1997. *Genitive Construction in Hindi*. M. Phil Thesis, University of Delhi.

Thomason, S. G. 2001. *Language Contact*. Washington, DC: Georgetown University Press.

Thomason, S. and Kaufman, T. 2008. *Language Contact*, *Creolization*, *and Genetic Linguistics.* Berkeley: University of California Press.

Vajpeyi, K. 1958. *Hindi Shabdanushasan.* Kashi: Nagri Pracharni Sabha.

Verdaguer, I. 1996. "Making Sense of Neologisms". *English Teaching Forum*, 34 (3): 98-106.

Versteegh, K. 2001. "Linguistic contacts between Arabic and other languages". In *Arabica, Linguistique Arabe: Sociolinguistique et Histoire de la Langue*. pp 470-508.

Waqar, A. and Waqar, S. 2002. "Identification of Diphthongs in Urdu and Their Acoustic Properties". *CRULP*, pp 16-26.

Watson, J. C. E. 2002. *The Phonology and Morphology of Arabic.* London: Oxford University Press.

Wegener, H. 1999. "Die Pluralbildung im Deutschen – ein Versuch im Rahmen der Optimalitäts theorie". Linguistik Online 4, 3/99. <http://www.linguistikonline.de>

Wegener, H. (2002). "Zur Funktion des s-Plurals im heutigen Deutsch". In Wiesinger, P.(ed.), Akten des X. Internationalen Germanistenkongresses Wien 2000 "Zeitenwende– Die Germanistik auf dem Weg vom 20. ins 21. Jahrhundert". Vol. 2: Entwicklungstendenzen der deutschen Gegenwartssprache. *Lexikologie und Lexikographie*, Bern et al.: Lang. pp 109-116.

Wegener, H. 2003. "Normprobleme bei der Pluralbildung fremder und nativer Substantive". *Linguistik Online* 16, 4/03. <http://www.linguistik-online.de>.

Wegener, H. 2004. "Pizzas und Pizzen – die Pluralformen (un)assimilierter Fremdwörter im Deutschen". *Zeitschrift für Sprachwissenschaft*, 23: pp 47-112.

Winford, D. 2003. An *Introduction to Contact Linguistics*. London: Blackwell.

Winter, F. E. 2008. "Studying Loanwords and Loanword Integration: Two Criteria of Conformity". *Newcastle Working Papers*, 14: 156-176.

Wohlin, A. 2010. "A 5-Dimensional Analysis of Linguistic Data"- Part II'. Retrieved on 22 February 2011, from: www.u5d.net/Linguistics/Part%20II-45p.pdf

Wurzel, W. U. 1984. *Flexionsmorphologie und Natuerlichkeit.* Berlin: Akademie-Verlag.