Abstract

Multilingualism has become a significant feature of today’s linguistic scenario. When languages are in contact, one of them tends to dominate the others on account of the role given or due to the attitude of the language users. Language contact results in language-mixing. Language-mixing in formal, written, corpora in Pakistan manifests the spread of English. The present study aimed to investigate frequency of the occurrence of language-mixing in Urdu newspapers. Thus it investigated the extent to which language-mixing is prevalent and the processes it takes when it occurs in Pakistani Urdu newspapers. Two Urdu newspapers (published in morning) were taken (from February 1 to 15, 2014) as sample. The data from Urdu newspapers for linguistic analysis were analyzed through content analysis. The findings of the study showed that language-mixing is frequently done in Urdu newspapers and it takes various processes-insertion, hybridization and synthesis.

1. Introduction
Pakistan is a multilingual state where 6 major and 57 minor languages are spoken (Rehman, 2010). Urdu and English have coincided for many centuries thus they have influenced each other at various levels. The prevalence of the phenomena can be witnessed in different fields or domains of life and language use such as textbooks, newspapers, magazines as well as in media (Rasul, 2013). The role of English is pervasive in all the spheres of life and it plays a conspicuous role in the media (Sultana, 2009). The use of English in Urdu has led to the linguistic phenomena such as language-mixing and borrowing.

Language-mixing in spoken corpora is natural and can be unintentional. In fictional written corpus language-mixing is accepted as manifestation of creativity on the part of the author but code-mixing in written, non-fictional writings may reflect different trends. There is a need to see the occurrence of language-mixing in written, non-fictional writings, such as newspapers as they are read not only by educated community of the society, but for people without formal education as well newspapers are an important source of information. Previous studies in the local context have focused on the use of code-mixing and code-switching in spoken corpora but the present study will focus on written discourse. Print media includes newspapers, magazines, books, etc. Newspapers, in particular, are considered an important source of information as the masses have an easy access to them. The present study focused the following questions:

1) How frequent is the occurrence of language-mixing in Pakistani Urdu newspapers?

2) What types of language-mixing processes appear in Pakistani Urdu newspapers?

2. Literature review
When many languages exist together they impact each other in every aspect possible, be it phonology or lexicon (Mohideen, 2006). Language-mixing and borrowing are the results of language contact. It can take many forms (e.g. code-mixing, code-switching, borrowing, etc.) and many processes (e.g. insertion, hybridization, synthesis, etc.). Applied linguists and sociolinguists have always been concerned about the occurrence of language-switching and language-mixing in bilingual or multilingual societies. The existence of language-switching and language-mixing has been researched not only in many multilingual societies such as India, Hong Kong, Philippines, Malaysia, Indonesia, etc. (Li, 2000; Zakaria, 2011; Fong, 2011; Chantarothai, 2011; Erwin-Billones, 2012; Arumawanti, 2013; Adjei, 2014) but also in societies such as Chinese where bilingualism is still not common (Chun-xuan, 2010; Bi, 2011).

Myers-Scotton (1993) reported that initially many researchers had focused on reasons and motivations for language-switching or code-switching, which can be discussed under the interest of sociolinguists. With time structural interests also attracted attention of the researchers and the focus was more on grammatical aspect of language-switching.

According to some researchers there are three types of language-switching (or code-switching)- intra-sentential switching, inter-sentential switching and tag switching (Poplack, 1978). Intra-sentential switching is insertion of a word or word phrase and it is often called language-mixing or code-mixing. Inter-sentential switching is switching at sentence and clause level. Muysken (2000) delineated the processes of language-mixing namely: “insertion”, “alternation” and “congruent lexicalization” (p.3). Insertion involves insertion of “lexical items” or “entire constituents”, the second process involves “alternation between structures from languages and the third process involves congruent lexicalization of material from different lexical inventories into a shared grammatical structure” (p.3).
Many researchers have focused on the occurrence of language-mixing in written and spoken discourse. Chun-xuan (2010) studied linguistic features of language-switching in Chinese sports news reports and established the point that in a bilingual country like China, code-switching in spoken language is not very common even its occurrence in “written discourses as news, novels and advertisement” are attracting interest of the researchers in China (p.167). The news writers do not employ language-switching as “accidental improvisation” rather it is “pre-planned and intentional” switching “with specific goals to accomplish in the minds of sports writers and the editors” (p. 171). The researcher also pointed out that “overuse” of language-switching in written discourse such as newspapers can put obstacle in the “standardization of Chinese writing system” and there is a need to ponder over the language policy regarding news writing in print media (p.173). “Newspapers are the potent source of information of any country. Its publication is issued periodically, to convey information and opinions on current events” (Sultana, 2009). Jalal (2010) has pointed out that extensive research has been on spoken language-mixing while very little research is in written language-mixing, while some have investigated language-mixing in fictional corpus ignoring non-fiction. She is of the view that written language-mixing is frequent in fiction than non-fiction; she associates it with “authentic” and “artificial”. McClure (as cited in Jalal, 2010) links the frequency of language-mixing in fiction and non-fiction with the “function of writing in non-fiction”. As written corpus is not fluent and casual we see few occurrences of code-mixing in non-fiction than oral communication. Writers or editors take account of minimizing the use of language-mixing in formal cases. Is it the case when we analyse Urdu newspapers where we see frequent occurrences of language-mixing? Callahan (as cited in Jalal, 2010) identifies that in non-fiction there is a deliberate attempt to minimize the frequency of code-mixing because of editing and proof reading and language-mixing “in written work involves more thought processing
and consideration” (p.45). Jalal (2010) opines that there is a “hierarchy” in the occurrence of language-mixing therefore we see more code switching in spoken or oral and informal writings than formal writings (p.46). Language-mixing “in official documents and non-fiction genres is much less and occurs only to perform specific functions” (Jalal, 2010, p.46).

Erwin-Billones (2012) examined the patterns of language-switching in formal writing such as newspapers and found out that language-switching patterns manifest them more than merely strategy employed by bilinguals; rather their prevalence is “creative linguistic process that reinforces a growing global language identity out of multiple language speakers in a world of shifting nationalities and boundaries”.

Kia, Cheng, Yee and Ling (2011) have found that the language of entertainment is different from other fields of life such as education, administration, law, etc., in stylistic features which justify for deviation from standard in language. It also leads to the perception that the news writers and reporters use such deviations to “attract the attention of the readers”.

Mareva and Mapako (2012) noted that despite the prevalence of language-switching and language-mixing in spoken conversation, they have permeated into the written language of Zimbabwe. Mohideen (2006) contends that English has “enriching” influence on the vocabulary of Bahasa Melayu. In Pakistan as well English and Urdu have been in contact for some centuries and influenced each other. Muhammad and Mahmood (2013) conducted a “corpus-based study” to explore “variation in Urdu language due to language contact”. They claim that “Urdu is a flexible language” and is “absorbing” words from other languages particularly English in such a way that its own structure is not “distorted”.

Rasul (2013) examined the instances of “borrowing and mixing in Pakistani children’s magazines”. The “trend” of using English vocabulary is found extensively in children’s Urdu
magazines which are expected to promote Urdu language. “In Pakistan children’s magazines are multilingual as are the people” (p.47). The question of the impact of language-mixing of English on the status of Urdu is an important one as Rasul (2013) “…how far this code-mixing adds to the richness of Urdu or affects is beauty; and what implications does it have with reference to the issues of globalization and Identity” ought to be given consideration in research (p.71).

3. Methodology

Two Pakistani Urdu newspapers, published in the morning, were selected to investigate the frequency and the processes of language-mixing for which the population was Urdu newspapers which carried sound justification for its selection. Instead of choosing T.V programs as population the researcher took newspapers because they represent a form of language-mixing which is more formal (than spoken corpora) as it is written and published. Erwin-Billones (2012) argues the same:

...the formal CS [code-switching] is composed of more purpose with regard to language and meaning. When one writes for publication or performance, they create with forethought and intentionality. Their formal product carries with it a targeted meaning. [One such] case of formal CS [code-switching] in a published medium is the newspaper. (p. 41)

It was ensured that data analysis and findings do not bring harm to the reputation of the newspapers as the names of the newspapers were kept confidential. For linguistic analysis of newspapers, content analysis was done. Urdu newspapers have many sections- front page (includes captions and news from all the aspects namely political news, entertainment, sports, international and national news), sports news, education, health, city news, entertainment, business news. The
The present study took front page (page number 1) and international news page (page number 3) as they cover all the aspects of news briefly.

Table 3.2 Sample for newspapers

<table>
<thead>
<tr>
<th>Sample 1</th>
<th>February 1 to 15, 2014</th>
<th>Page 1 &amp; 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 2</td>
<td>February 1 to 15, 2014</td>
<td>Page 1 &amp; 3</td>
</tr>
</tbody>
</table>

To analyse data for frequency of language-mixing and for exploring the processes of language-mixing content analysis was done. Firstly, the frequency of English language-mixing items in newspapers was identified.

To understand the English influence, it is not enough to study the type of words that exist and the semantic categories they belong to. It is important to look at the actual frequency of the occurrence of words. A word has significance if used frequently. (Mohideen, 2006, p. 48)

Different researchers have devised frameworks to analyse processes of language-mixing. The present study uses framework used by Rasul (2006) for the “linguistic analysis”, as Rasul (2006) has used the term, of the instances of language-mixing. Rasul (2006) adapted Kachru’s (1978) framework by adding and subtracting categories which were not in Kachru’s (1978) work. She used many categories. The first category that she used was “insertion” such as “Single word insertion (noun insertion, adjective insertion, adverb insertion); Phrase insertion (noun phrase insertion, prepositional phrase insertion, verb phrase insertion); Clause insertion; Sentence insertion” (p. 115-116). The second category was “hybridization” such as hybridization within words and units within single words and compound word, etc. The third category used was
“synthesis” such as “Abbreviations and acronyms, modes and terms of address, names and titles, reduplication, inflection attachments, repetition, tags and fillers, greetings, blessings and interjections, connectors, miscellaneous” (p.115-116). Muysken’s (2000) processes of language-switching such as “insertion” and “congruent lexicalization” have been in line with the present study’s analysis but the present study uses the term “hybridization” for Muysken’s (2000) “congruent lexicalization”. From synthesis, the cases of abbreviations and acronyms were taken for analysis. The rest of the categories (by Rasul, 2006)-“ modes and terms of address, names and titles, reduplication, inflection attachments, repetition, tags and fillers, greetings, blessings and interjections, connectors, miscellaneous”- are found in spoken data and not in written data such as newspapers.

4. Findings

The study found that English words are frequently mixed in Urdu newspapers and takes many forms such as insertion, hybridization and synthesis.

Table 4.1 Frequency (f) of insertion of the language-mixing items in Sample 1 and Sample 2

<table>
<thead>
<tr>
<th></th>
<th>Insertion at word level</th>
<th>Sample 1</th>
<th>Sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Nouns</td>
<td>Workers, Warrant</td>
<td>327</td>
</tr>
<tr>
<td>1.2</td>
<td>Adjectives</td>
<td>Ideal, Liberal</td>
<td>13</td>
</tr>
<tr>
<td>1.3</td>
<td>Verbs</td>
<td>Appeal <em>ki</em></td>
<td>30</td>
</tr>
<tr>
<td>1.4</td>
<td>Gerunds</td>
<td>Funding, Training</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>Insertion at Phrase level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>Noun Phrases</td>
<td>Sui Southern gas company, Sunday magazine</td>
<td>377</td>
</tr>
</tbody>
</table>
Table 4.1 provides a brief overview of the result of the analysis of language-mixing items in insertion as per their syntactic categories in Sample 1 and Sample 2. The table shows that in Sample 1 total number of English noun phrases in Urdu newspapers were 377 which are the highest in the list. Total nouns were 327, 30 verbs, 13 adjectives, 10 gerunds, 6 phrasal verbs and 3 clauses were also in the list. The example of adjectival phrase is “good-natured”. There were total 766 English items in Sample 1 which occurred for 2319 times. Moreover, total number of English noun phrases in Urdu newspapers in Sample 2 is 391 which are the highest in the list. Total nouns were 336, 26 verbs, 20 adjectives, 14 gerunds, 19 phrasal verbs and 1 clause were also in the list. There were total 807 English items in Sample 2 which occurred for 2194 times. The occurrences show that English words are frequently used in Urdu newspapers and they belong to numerous syntactic categories, such as nouns, noun phrases, verbs, adjectives, gerunds and phrasal verbs.

Not only single word insertions (workers, election) but phrase insertion (Election tribunal, life guards) were found in the data from Sample 1 and Sample 2. There were also clause insertions (3 in Sample 1 and 1 in Sample 2) in the data from newspapers but they were names of some book or movie such as “Welcome to London” (film), “I am Malala” (book). Phrasal verbs such as “mark-up”, “shutter-down”, and “kick off” have been found which do not have equivalents in Urdu and when used in English give required meanings.

Table 4.2 Frequency of hybridized items (within single words) in Sample 1 (S1) and Sample 2(S2)
It was also found that the total number of hybridized items (hybridization within single words), in Sample 1 are 42 and their total occurrence is 164, in Sample 2 are 56 and their total occurrence is 144 which means they have appeared frequently even in a single page. For words such as “university” (jaamia) and “seat” (nashist), there are Urdu equivalents but hybridized forms “universityun” and “seaton” suggest linguistic innovation but it can also be seen as distortion of language.

Table 4.3 Frequency of hybridized items (within units) in Sample 1 (S1) and Sample 2(S2)
It was also found that the total number of hybridized items (hybridization within units), in Sample 1 are 135 and their total occurrence is 339, in Sample 2 are 166 and their total occurrence is 287 which means they have appeared frequently even in a single page.

Table 4.4 Frequency of abbreviations and acronyms in Sample 1 (S1) and Sample 2(S2)

<table>
<thead>
<tr>
<th>Abbreviations and acronyms</th>
<th>Examples</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>S1</td>
</tr>
<tr>
<td>CNG, DPO, ISPR, PIA, IDPs, IS, NADRA, NATO</td>
<td>95</td>
<td>94</td>
</tr>
<tr>
<td>Total number of occurrences</td>
<td>350</td>
<td>252</td>
</tr>
</tbody>
</table>

It was found that total number of English acronyms and abbreviations, in Sample 1, are 95 and they have occurred for 350 times. Number of English acronyms and abbreviations, in Sample 2, are 94 and they have occurred for 252 times. They include designations (SHO, DSP), names of institutions (NADRA, ISPR), organizations (UN), parties (MQM, PP), and movements (IS). The use of acronyms and abbreviations in place of Urdu words can be either for ease of writing. Their frequent use suggests that the news writers assume that the newspaper readers will understand them even if the full form of those acronyms and abbreviations are not given. Some words are known to public in their abbreviated form only and common readers do not understand their Urdu names, such as NADRA, ISPR, SHO, etc.

Overall frequency of language-mixing items shows that English words are frequently mixed and used in Urdu newspapers.

4.5 Frequency of language-mixing items
Table 4.5 shows that in Sample 1 overall frequency of language-mixing items was 3,172 words and 2,887 in S2. They include English words inserted, hybridized or synthesized into Urdu lexicon and structure.

5. Discussion

According to Kamwangamalu (1989), the questions important in linguistic approach to language-mixing are about differences between language-mixing and language-switching, types of language-mixing, patterns, and rules that govern this phenomenon. Muysken (2000) illustrates that “structural analysis” of code-switching is as important as “sociolinguistic analysis”. The findings of the study revealed three processes of language-mixing - insertion (single word insertion, phrase insertion, clause insertion), hybridization and synthesis (abbreviations and acronyms). These processes have also been found in Rasul (2006) with changes as she used spoken data to analyse the linguistic phenomena while this study has taken written data.

Some findings of the data are in line with Rasul (2006) who took spoken data to analyse the processes of “language hybridization”. Rasul (2006) has discussed synthesis as consisting of multiple sections such as titles, abbreviations, connectors, fillers, etc. The categories such as “tags and fillers”, “reduplication”, “repetitions”, “connectors”, “interjections”, which Rasul (2006) has discussed, are specific to spoken discourse which she focused. These are not found in non-interactive written discourse or writings such as newspapers. Although the occurrence of
hybridized items, abbreviations and acronyms are very high in newspapers which have been focused in this study.

A very important form of linguistic innovation is “language hybridization” (for e.g. vison (visas), partyaan (parties/jammat), maliyaati fund. In the present study many hybrid items were found in Sample 1 and Sample 2. They included terms such as “partyan” (parties), “buson” (buses) which are neither English nor Urdu. Such hybridized items manifest patterns of language variation and linguistic innovations. Kachru (2005) has called “hybridization” a consequence of “Englishization” where one of the parts is from English. There were words in the samples which can be considered a consequence of “Englishization” (e.g. “Najkaari commission”, “jet tayarey”, “Halal-i-ahmar society”, “mega mansuubey”, “parlimaani committee”). These hybrid items have been further categorized by Kachru (1986) into two: “open set items” and “close set items”. The first type “open set items” does not abide by grammatical constraint while “close set items” have some constrains. Rasul (2006) has used the second type “close set items” as “inflection attachments” (such as police walla). Rasul (2006) opines that these “hybrid forms produced through this process have given birth to code-mixing” and led to “Englishes” in South Asian context (Rasul, 2006, p.310). There are many hybridized names for parties, organizations, etc. “Muslim League Nuun”, “Awami National party”, “Mutahidda Qaumi Movement”.

Ehsan and Aziz (2014) have pointed out as mostly people are negligent about the equivalent Urdu words. Not only this but words whose equivalents are common, are used in Urdu news despite their availability. The present study found that words such as “party” (jamaat), “law and order” (aman-o-aman) have their equivalents but have been frequently used in Urdu newspapers. Ehsan and Aziz (2014) found out that Urdu has lexical gap in terms of words related to technology, administrative terms, etc. They found that words related to colour (such as navy blue, white) are
used as well as words as “presiding officer”, “mobile market”, “megawatt” are without equivalents in Urdu. Some borrowed words such as ‘traffic’ are used and acronyms such as “CCTV”, “CNG”, “ISPR” are used as their Urdu equivalents are “technical and more difficult to pronounce and understand” (Ehsan & Aziz, 2014, p.167).

English words are flexible and used as nouns, adjectives and verbs, for example “record” which function as a noun as well as a verb. The highest number of insertion was at phrasal level, noun phrase which shows that English is widely used in Urdu not only at word level but at phrase level as well. The English words used at phrase level are examples of code-mixed items rather than borrowed ones.

Some English words which are mixed in Urdu are borrowed from English to fill lexical gap. Borrowing is limited to borrowing of some parts of speech as nouns, adjectives and some other, but code-mixing is not limited to some types or categories. Borrowing is limited in innovations as compared to code-mixing which involves extensive set of vocabulary and is not limited in innovation while mixing and crossing boundaries of the other language. Many English words are used in Urdu due to the lexical gap which leads to the borrowing of English words. English words are frequently borrowed in economics, technology, education, military terms, legal terms, etc. Rasul (2013) states that code-mixing and borrowing in written text such as magazines can be due to various reasons or it might be a result of “an oversight” and “lack of awareness” about language-mixing (p.71).

In discussing the borrowed words Kachru’s (2006) hypotheses ought to be discussed which provide explanation for borrowing of English words in Urdu. Kachru (2006) has elaborated two important hypotheses of borrowings: “dominance hypotheses” and “deficit hypotheses”. Kachru
(2006) defines “deficit hypothesis” as a result of the absence of equivalent word in borrowing language and it borrows to bridge this gulf and repair “linguistic deficit”. Mohideen (2006) discusses some interesting examples of the languages which have borrowed words from English. In the samples words such as “visa”, “bus”, “cricket”, “hockey” are borrowed from English. Some words are different in phonology only, “bomb” (bam), “college” (kaalij), “award” (awaard).

Kachru (2005) has also pointed out that some language-mixing items are used with “nativized phonology”. He stated:

"Englishization at lexical level involves processes such as the following: 1, loan words (nativized in phonology), the intrusion of lexical items is found in every domain, however the registers of science, technology, fashion, television, cinema, and advertising have particularly high frequency of such items. (Kachru, 2005, p.104)."

The other hypothesis by Kachru (2006) is “dominance hypotheses” for borrowing which entails the borrowing of words from the dominant language. For example, English is a global language and its influence is immense in all the domains of language use and all the fields of life, be it education, science, technology, food industry, fashion, entertainment, literature, etc. It is due to this dominance that words are borrowed from English in Urdu.

Different words become part of everyday use and English vocabulary is prevalent in all the domains of language use and so is in Urdu newspapers. Kachru (2005) has used the word “Englishization” for the widespread use of English in different fields of life and domains of language use. In the samples there were borrowed words in sports, entertainment, economics, etc. Kachru (2005) has also stated that the number of items borrowed in particular areas such as “science, technology, fashion, television, cinema, and advertising” are more in number than other
areas (p.104). English has also borrowed words from Latin, French, Spanish, Greek and German in its days of evolution (Kachru, 2005). The role of English as a borrowing language has been turned into a lending one as it is adding the treasure of vocabulary of different languages. Kachru’s (2005) “dominance hypotheses” for borrowing also accounts for the reason of borrowing from English because of its power and the prestige attached to it.

5. Conclusion

The findings of the data showed that English words are frequently used in Urdu newspapers. It identified three processes of language-mixing- insertion of English words, hybridization and synthesis.

The study found out that there might be multiple implications for the frequent use of English in Urdu written text. English words are used in Urdu newspapers because of lexical gap thus establishing “deficit hypotheses” of Kachru (2006). Another reason for the use of English words is the dominance and prestige that English enjoys in Pakistani society which leads to insertion of English words in Urdu newspapers. Another reason could be sheer negligence of the newspaper writers or editors who use English words despite the fact that there are Urdu equivalents available in Urdu lexicon.

English words are inserted at word, phrase and clause level. They appear in different syntactic categories such as nouns, noun phrases, adjectives, gerunds, phrasal verbs and clause, which shows that English words can easily be inserted in the structure of Urdu without distorting it. This study will contribute in the future research in the field of language variation, hybridization and language change. Previous studies in the local context have focused on the use of language-mixing and language-switching in spoken corpora but the present study has focused on written discourse,
formal and published form of written text such as newspapers. The present study will be of great importance not only to the applied linguists and sociolinguists but for policy makers as well, as it will provide an insight into the need for language policy in print media. Newspapers, in particular, are considered an important source of information as the masses have an easy access to them. Therefore they can be used or be helpful in determining trend for using language or language use.

There is also a need to determine pedagogical implications using the findings of studies on language mixing so as to further explore the use of English words in Urdu textbooks as well. Language mixing in spoken discourse receives great deal of acceptance by scholars, educationalists, general public, teachers and students but in written discourse language mixing faces resistance. There is dearth of research studies available in this area and this remains less explored in Pakistan.

References


Chantarothai, T. (2011). *The effect of code-mixing (Thai-English) in print advertisements on product knowledge and understanding, attitudes and purchase intention: Consumer’s educational background as a moderator* (Master’s thesis), Thammasat University, Thailand, Bangkok.


