Strategies for incorporating nouns and verbs in code-mixing: the case of Pashto-English bilingual speech

Arshad Ali Khan and Pieter Muysken
Strategies for incorporating nouns and verbs in code-mixing: the case of Pashto-English bilingual speech

Arshad ALI KHAN
University of Management and Technology, Lahore, Pakistan

Pieter MUYSKEN
Radboud University Nijmegen and Stellenbosch University

Abstract

A recurring issue in the analysis of code-mixed speech concerns the strategies for incorporating nouns and verbs. How can nouns and verbs from one language be felicitously incorporated into sentences from another language? This paper analyses this question with a case study of English-Pashto bilingual speech. English nouns appear in determiner phrases marked with Pashto case endings, and English verbs have an associated light verb with Pashto inflections. Code-mixing and lexical borrowing have led to a large amount of indigenized English elements in Pashto. Code-mixing is a device of indigenization whereby Pashto speakers of Khyber Pakhtunkhwa (KP) have adopted and nativized the English elements in their local use.

Key words: Indigenization, code-mixing, insertion, borrowing, loanwords, case marking, bilingual compound verb, monolingual complex predicate, light verb
1. Introduction

A recurring issue in the analysis of code-mixed speech concerns the strategies for incorporating nouns and verbs. How can nouns and verbs from one language be felicitously incorporated into sentences from another language? Both nouns and verbs need embedding when taken from one language to another. Verbs are often linked to functional categories expressed through inflection, including marking for Tense, Aspect and Mood, and for person and number. Nouns are the core of determiner phrases, and associated with functional categories expressing case, quantity, and definiteness. In different bilingual language pairs, different strategies are found to create these links and associations, since the functional categories involved are generally taken from the matrix language (Myers-Scotton, 1993).

This paper analyses this question with a case study of English-Pashto code-mixed bilingual speech. English nouns appear in determiner phrases marked with Pashto case endings, and English verbs have to carry Pashto inflections. Code-mixing and lexical borrowing have led to a large amount to indigenized English elements in Pashto. Code-mixing is a device of indigenization whereby Pashto speakers of Khyber Pakhtunkhwa (KP) have adopted and nativized the English elements in their local use.

2. Background on the bilingual community

In the bilingual speech community of Khyber Pakhtunkhwa (KP) (Pakistan), in the region on the border with Afghanistan, the use of English and Pashto can be characterized in terms of diglossia, where English has an official status and is used in formal settings such as in the office, in writing applications, in teaching, and in courts. It can be said that function of English is mostly restricted to writing1. On the other hand, Pashto is the dominant language of the speech community where it is used in the market, in the play grounds, in the family setting, with friends, in the café, in an office, and in informal educational settings.

It is not only the post-colonial setting but also the prestigious nature of English and its instrumental quality in getting jobs and social status which has made it so attractive in Pakistan and across the globe. The real source of English-Pashto bilingualism in KP are the electronic media, the education system and ESL classrooms. In KP, especially in the Mardan division, the Peshawar division and the Malakand division, code mixing (hereafter CM) is the driving force to introduce borrowings in the Pashto speech community. The following are the two extracts taken from Khan (2011), where he reports on an interview about the use of English CM in Pashto. In the extract, the first respondent show solidarity with the Pashto speech community; using English lexical items is not a threat to his identity as a Pakhtun.

---

1. The present article is a shortened and partly elaborated version of some of the chapters of the PhD thesis defended by Arshad Ali Khan at the University of Azad Jamu & Kashmir (Pakistan) and supervised by Pieter Muysken while Khan was on a visit to Nijmegen (Netherlands), funded by a Pakistan Government grant. The work of Pieter Muysken was supported by the ERC Advanced Grant Traces of Contact.
I am very proud of being a Pakhtun. I am conscious of my nationality. I am pretty concerned about my own language Pashto. I love to use Pashto while speaking to any Pakhtun but as an educated person I don’t mind to use English when it is necessary. (Khan, 2011: 122)

In response to the first question: why do you use English words in Pashto, the respondent states:

I use English while talking to other Pashto speakers for the better interpretation of my feelings and thoughts. I feel easy and comfortable in conversation by mixing two languages. (Khan, 2011: 122)

The above two extracts show the attitude of the Pashto speakers, where they show solidarity and identity with the Pashto speech community but along with that they do not overrule the use of English in their conversation. It is used to facilitate the interlocutors in face to face conversation. It is used as a tool to achieve the better end in discussion. Its use encourages the users in educated speech community and gives them an identity of an educated person. English code mixing is a prevalent and an established pattern of communication in the Pashto speech community of Yousafzi dialect in KP.

Code mixing (CM) is the innovative force of indigenization in the Pashto speech community of Afghanistan and Pakistan. In order to explore the process of indigenization, Penzel (1961) has focused on English, French, and German loanwords in Pashto of Afghanistan. Penzel has revealed the areas where Western ways, cultures, and technology have influenced the country. In the Pashto spoken in Afghanistan, contacts with the intellectual and elite class are the decisive factor in the adaptation of the loanwords (Penzel, 1961). Penzel’s study focused on geographical terms, measures, months, clothing, objects, food, administrative and political terms, scientific and technical vocabulary, motor transport and educational, medical, cultural and social terms. In his study, he has focused on the short vowels (/i/, /a/, // , /u/) of Pashto and on the phonological adoption of the Western loanwords. He has also studied plural formation of the Western loanwords in Pashto. This study has used three sources for the analysis of the loanwords in Pashto. The first source of the study is the study conducted by Penzel (1961) on Western loanwords in Pashto. The second source is the transcribed data of English-Pashto CM used in the present research.

Before going to analyze the CM data, it is required to give a brief introduction to the difficulties in splitting CM and loanwords. In (1) and (2), taken from the present English-Pashto data set, the two verbs ‘start’ and ‘play’ in theBCVS are bare verbs conjugated with the light verb kaw ‘do’ or ‘make’, which may carry person and aspectual markers.

(01) che [[da de format] ba sanga [start kaw-o]]

\[
\text{COMP GEN DM.PRX format CL.FUT how start do.PRS.IPFV-1PL} \\
\text{‘… that how we will start this format.’}
\]
(02) *media [[negative role] hu [play kaw-i]]
'Media indeed play a negative role.'

In (1), the word ‘start’ is embedded as verb in the bilingual verbal complex and is combined with the transitive light verb *kaw*. The embedded verb ‘start’ is in its bare form without inflection or morphological integration, and inserted into the bilingual verbal complex. As such, it is a strong example of code switching. Its Pashto counterpart *shuroo* has the same function. The use of the embedded verb ‘start’ is very frequent and is used in multiple contexts and with multiple concepts:

(03) **G L O S S**

start kawal  to start
motor start kawol  to start a motor
par'ifa start kawal  to start a paper
dukan start kawal  to start shop keeping
k:arobar start kawal  to start a business
rote start kawal  to take a meal
laba start kawal  to start a game

The use of ‘start’ in multiple contexts makes it more akin to borrowing than code switching. Following Myers-Scotton (2002), however, it does not make too much sense to separate code switching from borrowing. The same holds in (2), where the English verb ‘play’ is used in conjunction with the light verb *kaw*. The EL verb ‘play’ demonstrates numerous functions with multiple concepts in Pashto.

(04) **G L O S S**

play kawal  to play
Sports:  shot play kawal  to play shot
Politics:  role play kawal  to play a role
Drama:  kirdar play kawal  to play a role
Music:  gana play kawal  to play a song
cassette play kawal  to play a cassette

In (3) and (4), the speaker can flexibly use these verbs ‘start’ and ‘play’ with different concepts and in different contexts.

The bare embedded elements in the host language constitute an analytical problem. In Pashto, nouns can be case marked (especially in the oblique case) and pluralized, as we will see below. The nature of the singular embedded element is always problematic: do we count it as borrowing or code mixing?

(05) **Pashto**  **G L O S S**

yaw cup  a cup
In (5), a few singular nouns are listed which are very frequently used in the Pashto speech community. These nouns in their singular form do not show any phonological or morphological change but if a four year old child is shown the pictures of the words in (5), s/he will pronounce only the English loanwords, and not their counterpart in Pashto.

The morphological change of an alien word into a host language varies from language to language. The examples (1) to (5) show that the marking of an embedded element into a matrix language does not provide ample proof to determine the embedded word as a code mix or a loanword.

Another type of adoption involves a conceptual innovation, as found in the embedded word ‘trout’. Trout fish are found in the Swat valley of KP. When analyzing a documentary from AVT Khyber (see below), it was found that the word ‘trout’ was symbolically used for a hunter (agent or a person working on commission). Those agents convince patients to visit a doctor in order to get money out of this bargain, and are known as ‘trout’. The word ‘trout’ was very common among people who were working in that era. The word ‘trout’ as single noun does not show any sign of phonological and morphological integration but it has been used in different contexts. The same word is also used for fish and most of the people are marking the same word with Pashto plural marker -α:n as trout-α:n.

<table>
<thead>
<tr>
<th>(06) Trout</th>
<th>SINGULAR</th>
<th>PLURAL</th>
<th>OBlique</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>fish</td>
<td>trout</td>
<td>trout-α:n /trouts</td>
<td>trou:n-o</td>
<td>trout (fish)</td>
</tr>
<tr>
<td>agent</td>
<td>trout</td>
<td>trout-α:n</td>
<td>trout-α:n-o</td>
<td>trout (agent)</td>
</tr>
</tbody>
</table>

Myers-Scotton (1993, 2002) has defined two different types of borrowings: cultural and core borrowings. Cultural borrowings constitute words which express concepts that do not exist in the lexicon of the recipient language. These are the words which do not have counterparts in the host language. Many of them are lexical elements related to technology and science and in some cases, they can even constitute new words for the donor language as well. However, cultural borrowings are not always words referring to science but also words that describe items of clothing or food which, as it was aforementioned, do not have a lexical equivalent in the lexicon of recipient language (Myers-Scotton, 2002). On the other hand, core borrowings are not new words but words that are already expressed by an equivalent lexical item in the recipient language. Although there seems to be no reason for the existence of this kind of borrowing, it does not only take place, but according to the data collected for the present study, it constitutes a large amount of overall borrowing.

3. Data collection

In order to collect data for our different research questions different methods were adopted: questionnaires and transcriptions of recorded speech.
The CM data concerned natural and spontaneous conversations. The collected data covers television shows of a Pashto language TV channel AVT Khyber, i.e. Khyber News, Khyber Watch and Khyber club. The data covers a wide range of topics i.e. political, social, cultural, pedagogical, and women's rights. All of the participants are bilinguals and do not consider code switching as stigmatized behavior. The only dialect used throughout these shows is the Yousafzi dialect. Two different formats have been used in these clips; some are group discussions about a certain topic and some of the clips are based on face to face interviews with participants.

Most of the participants are highly educated and they shared the same cultural norms of the Pashto speech community. The data collected from YouTube were categorized as in Table 1.

<table>
<thead>
<tr>
<th>Clip #</th>
<th>Khyber Watch/ Khyber Topic</th>
<th>Participants (M = male; F = female)</th>
<th>Duration in minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Khyber club EP# 1 Computer</td>
<td>M: 3; F: 1</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>Khyber club EP# 11 Electronic/ print media and its influence on our life</td>
<td>M: 3; F: 1</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>Khyber club EP# 16 The use of tuition in education</td>
<td>M: 3; F: 1</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>Khyber club EP# 16 The use of tuition in education</td>
<td>M: 3; F: 1</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>Khyber club EP# 14 Peace, prosperity and development</td>
<td>M: 4</td>
<td>14</td>
</tr>
<tr>
<td>6</td>
<td>Khyber club EP# 11 Electronic media</td>
<td>M: 3; F: 1</td>
<td>14</td>
</tr>
<tr>
<td>7</td>
<td>Khyber club EP# 13 Electronic media</td>
<td>M: 3; F: 1</td>
<td>14</td>
</tr>
<tr>
<td>8</td>
<td>Khyber News My right (women issues)</td>
<td>F: 2</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>Khyber News My right (women issues)</td>
<td>F: 3</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>Khyber News My right (women issues)</td>
<td>F: 3</td>
<td>14</td>
</tr>
<tr>
<td>11</td>
<td>Khyber News Ismail (Singing)</td>
<td>M: 2; F: 1</td>
<td>19</td>
</tr>
<tr>
<td>12</td>
<td>Khyber News Personal political views</td>
<td>M: 2; F: 1</td>
<td>14</td>
</tr>
<tr>
<td>13</td>
<td>Khyber News Personal political views</td>
<td>M: 2; F: 1</td>
<td>14</td>
</tr>
<tr>
<td>14</td>
<td>Khyber News Personal political views</td>
<td>M: 2; F: 1</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>Nun Sahar (Education and progress) VC Sarhad University</td>
<td>M: 2</td>
<td>52</td>
</tr>
<tr>
<td>16</td>
<td>Nun Sahar Self and Reality</td>
<td>M: 1; F: 1</td>
<td>14</td>
</tr>
<tr>
<td>17</td>
<td>Nun Sahar Self and Reality</td>
<td>M: 1; F: 1</td>
<td>14</td>
</tr>
<tr>
<td>18</td>
<td>Khyber Club Politics: Care taker Prime minister</td>
<td>M: 4</td>
<td>14</td>
</tr>
<tr>
<td>19</td>
<td>Nun Sahar Unemployment</td>
<td>M: 4</td>
<td>9</td>
</tr>
<tr>
<td>20</td>
<td>Nun Sahar Unemployment</td>
<td>M: 4</td>
<td>14</td>
</tr>
</tbody>
</table>
Table 1: Topic, gender, and time of TV fragments

<table>
<thead>
<tr>
<th></th>
<th>Topic</th>
<th>Gender</th>
<th>Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>Nun Sahar Classics</td>
<td>M: 1; F: 1</td>
<td>39</td>
</tr>
<tr>
<td>22</td>
<td>Nun Sahar Classics</td>
<td>M: 1; F: 1</td>
<td>27</td>
</tr>
<tr>
<td>23</td>
<td>Khpal Etwar</td>
<td>M: 1; F: 2</td>
<td>27</td>
</tr>
<tr>
<td>24</td>
<td>Khpal Etwar</td>
<td>M: 1; F: 2</td>
<td>21</td>
</tr>
<tr>
<td>25</td>
<td>Khpal Etwar</td>
<td>F: 3</td>
<td>14</td>
</tr>
</tbody>
</table>

Approximately 12 hrs of 35 clips were downloaded from the online source (YouTube) for the analysis of the CM data. Out of 35 clips only 25 clips, a total of approximately eight/8 hrs were selected. In order to select reliable and appropriate data the clips were selected by six native speakers of Pashto.

The first criteria of selection of the different shows were their language and participants. Most of the participants were educated and they were using English words/phrases with ease and spontaneously. The second criterions of selection of the online data were their topics. Their topics were about the culture, politics, unemployment, education and women rights in Pashto speech community of KP. Here something which is really good for this research is the notion of “observer’s paradox” (Labov, 1972), because all of the talk shows which we have downloaded are held in a very natural setting. There was no such risk that the participants in shows felt observed for the research purposes, and their speech style remains unchanged throughout the debate held in the shows. Through this way we managed to collect valid data for our CM studies.

In order to transcribe English-Pashto CM data the recordings were transcribed into Roman English. In order to give it proper glossing and morphemic identification, the data was transcribed with the help of Tool Box. All data have been transcribed in three layers. The first layer represents the data at the morphemic level. The second layer represents the data at the gloss level, and the third the translation. It is found that the data were quite homogeneous. Each transcribed sentence has a reference number according to the participant gender, function and name of the recording.

The English-Pashto bilingual transcribed data is discussed and analyzed on the basis of the previous approaches to the bilingual data. The data is also analyzed on the basis of the literature on monolingual Pashto (Butt, 1995; Akhtar, 2003; Babrakzai, 1999; Robson and Tegey, 1996; Roberts, 2000).

4. Nominal constituents

The present study does not include the entire loanword corpus but has only presented the most frequent examples. Before turning to the syntactic properties of English nouns in Pashto, their phonological and morphological integration as loans will be briefly discussed, as well as the most prominent semantic domains in which we find Pashto nouns. The pho-
nological and morphological integration will be discussed in relation to plural markers and oblique case markers.

In Table 2, the loanwords have been put under the three headings: loanwords, plural, and oblique case. Under the different major categories which have been found in Penzel (1961), the AVT Khyber news documentaries and from the English-Pashto bilingual data in the present study. The loanwords may be marked with the plural markers -e, -una and the oblique marker -o. Most of the monosyllabic loanwords in Pashto retain their phonological form. The loanword wikat / wlkAt / shows more phonological integration than its English counterpart ‘wicket’ /wlk/ll. In most of the following English loanwords, a drastic phonological change has taken place as marked in the first column. /WαskAt/ is pronounced in English as /wets t anomalies/.plural and oblique markers and show major morphological changes as shown in the table. A major phonological integration may be seen as the English loanword /da kiar/ ‘doctor’ is pronounced in English as /dαkt/. and the loanword /hαspə:λ/ in English is pronounced as /hαs.pl.t/. One important point that has been discussed is that most of Pashto native speakers have problems with the English vowels, so, the major phonological change might be because of L1 interference in L2. The native speakers of Pashto pronounce the word ‘break’ as /br 1 k/, not as English /breik/.

Table 2 also illustrates some of the key semantic domains in which English nouns occur in Pashto, and also shows the amount of phonological variation related to modifications.

<table>
<thead>
<tr>
<th>Loanwords</th>
<th>Plural</th>
<th>Oblique</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sports</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bet</td>
<td>bet-una</td>
<td>bet-una-o</td>
<td>bat</td>
</tr>
<tr>
<td>ball</td>
<td>ball-una</td>
<td>ball-una-o</td>
<td>ball</td>
</tr>
<tr>
<td>wikat</td>
<td>wikat-e</td>
<td>wikat-o</td>
<td>wicket</td>
</tr>
<tr>
<td><strong>Measures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gram</td>
<td>gram-una</td>
<td>gram-una-o</td>
<td>gram</td>
</tr>
<tr>
<td>kilogram</td>
<td>kilogram-una</td>
<td>kilogram-una-o</td>
<td>kilogram</td>
</tr>
<tr>
<td>inch</td>
<td>inch-una</td>
<td>inch-una-o</td>
<td>inch</td>
</tr>
<tr>
<td><strong>Clothing and objects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>teibal</td>
<td>teibal-e</td>
<td>teibal-o</td>
<td>table</td>
</tr>
<tr>
<td>kot</td>
<td>kot-una</td>
<td>kot-una-o</td>
<td>coat</td>
</tr>
<tr>
<td>waskat</td>
<td>waskat-e</td>
<td>waskat-o</td>
<td>waistcoat</td>
</tr>
<tr>
<td><strong>Scientific and Technical Vocabulary</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>haspatə:λ</td>
<td>haspatə:λ-una</td>
<td>haspatə:λ-una-o</td>
<td>hospital</td>
</tr>
<tr>
<td>telefon</td>
<td>telefon-una</td>
<td>telefon-una-o</td>
<td>telephone</td>
</tr>
<tr>
<td>tlktar</td>
<td>tlktar-e</td>
<td>tlktar-o</td>
<td>tractor</td>
</tr>
</tbody>
</table>
Table 2: Loan words from a number of central semantic domains

<table>
<thead>
<tr>
<th>Category</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inanimate</td>
<td>film, shopping, telephone</td>
</tr>
<tr>
<td>Feminine</td>
<td>wicket, waistcoat, tractor, puncture</td>
</tr>
<tr>
<td>Animate, Masculine</td>
<td>artist, minister, student, lecturer</td>
</tr>
</tbody>
</table>

Table 3: Conditions on the affixation of the various Pashto plural suffixes for English word

The oblique suffix -o is simply added in all contexts, sometimes replacing a final vowel.

4.1. Single elements

Table 4 shows the distribution of inserted single elements, both in absolute figures and in percentages. As Table 4 shows, nouns are the most frequent embedded items in Pashto. Out of 140 single embedded elements, 80 times a noun, 40 times a verb and 20 times a modifier were found in the data. In the entire data-set only one noun, college-una is suffixed
with the-una plural marker. In the present data the morpho-syntactic frame of Pashto is very important in determining the distribution of the embedded elements. The different morpho-syntactic constructions are very helpful in determining the role of the different morphemes in the process of CM. In English-Finish code switching (Poplack, 1988) the English elements are affixed with the Finish marker for its morphological integration. These elements are described as nonce borrowings (Poplack et al., 1988). In Swedish/Persian (Lotfabbadi, 2002) 87 receive Persian inflectional morphology, while in Pashto the English single nouns function as bare nouns or in different constructions as given in Tables 4 and 5. In the present study not a single embedded noun is preceded by an embedded language (English) determiner. The embedded noun may be marked by the possessive marker da.

<table>
<thead>
<tr>
<th>Type</th>
<th>#</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nouns in Bare DP constructions</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>Noun in Determiner complex construction</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Noun in Determiner Phrase construction</td>
<td>23</td>
<td>16</td>
</tr>
<tr>
<td>Noun in Prepositional Phrases</td>
<td>24</td>
<td>17</td>
</tr>
<tr>
<td>Noun in (da) possessive construction</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>Adjectives</td>
<td>16</td>
<td>11</td>
</tr>
<tr>
<td>Adverbs</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>V+ Light verb construction V+ particle + Light verb</td>
<td>40</td>
<td>29</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>140</strong></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: English single elements in different constructions of Pashto

*English bare nouns.* In the following sentences the English bare nouns in italics have been inserted in the frame of Pashto. In English-Pashto CM 14% of insertion of bare nouns has been recorded.

In (7) the bare noun ‘youth’ does not take Pashto inflections but the sentence is well-formed. The insertion of the English verbal element in mixed verbal complexes will be explained in detail below. The verb is inflected for the Pashto tense, aspect, and agreement. The inserted word ‘youth’ is in congruence with Pashto structure and its counterpart in that language would be zwanan.

(07) Electronic media 0008

\[
\text{pa de ke monga [youth] [target kaw-u]}
\]

\[
at \text{DM.FR} \text{ OBL} \text{ PRN.1PL youth target do.PRS.IPFV-1PL}
\]

‘In this we target youth.’
In (8) the same English bare noun ‘youth’ in the subject position is preceded by the Pashto complementizer che and in object position ‘study’ is preceded by the reciprocal pronoun hpala. Pashto is the language which provides the morphosyntactic frame and the embedded elements in italics follow Pashto rules. The transitive verb in the present tense is in agreement with the subject but in the past tense the same agreement would be marked on the direct object. The morpheme -i is suffixed with the transitive auxiliary kaw for the 3rd person singular subject agreement with the verb kaw-i. In (8) Pashto clearly is the ML.

(08) che youth [hpala study] na kaw-i

\[\text{COMP youth RECP study not do.PRS.PFV-3SG}\]

‘… that youth is not doing their studies.’

In (9) the embedded element ‘violence’ in subject position occurs as a bare noun. The morphosyntactic frame is from Pashto. The possessive construction and the preposition phrase in the oblique case show Pashto grammar, as seen in də in the possessive construction and the preposition ke in the postpositional phrase, following the morpheme order of Pashto. Subject-verb agreement on the intransitive auxiliary keg is co-indexed by the 3rd person marker-i.

(09) Khyber Club 1: media role 0010-M2

che violence [də dunia pa hr corner ke] keg-i

\[\text{COMP violence GEN world LOC every corner Obl become.PRS.PFV-3SG}\]

‘That violence is happening in every corner of the world.’

Pashto as the ML provides the morpho-syntactic frame for the bare nouns ‘answer’ in (10) and ‘education’ in (11). Subject-verb agreement is via the subject 2pl marker-o suffixed on the verb ‘balance sh’ in (10):

(10) Electronic media 0015

[[[staso answer] [balance sh-o]]

\[\text{POSS.2PL answer balance COP.PST.PFV-2PL}\]

‘Your answer had become balanced.’

In (11) the 1pl marker-o marker is suffixed on the verb ‘start kr’ in (11), while there is a complex prepositional phrase. The preposition ‘through’ seems to be attached to the Pashto locative pa.

(11) Khyber Club 1: media role 0005-M3

pa kal-u ke mong [da TV pa through] education [start

\[\text{LOC village-PL Obl PRN.1PL GEN TV LOC through education start}\]

kar-o do.PRS.PFV-1PL

‘In villages we can start education with the help of TV.’

*Complex Determiner Phrase constructions.* In English-Pashto code mixing data complex determiner constructions (Muysken, 2000) are found only in 3% of the cases.
In English complex determiner constructions a demonstrative cannot precede the indefinite determiner, while in the Pashto demonstrative pronouns can precede the indefinite pronoun *yaw*. In the bilingual NP the English head follows the string of Pashto, and satisfies the Pashto pattern.

In (12) the embedded noun ‘image’ in italic is preceded by a Pashto determiner complex structure, and follows Pashto grammatical morpheme order, preceded by the demonstrative *daa*, the indefinite article *yaw* and the intensifier *dair* along with the Pashto modifier *h*. Subject-verb agreement for the demonstrative pronominal subject is marked with the suffix *i* on the verb. The verb is inflected for the Pashto tense, aspect, and number and is coindexed with *daa* in the NP.

(12) nu aya [[daa yaw dair ha image] [ha-yi]]
then if DM.PRX one.M,SG very good image show.PRS.IPFV-3SG
‘If this shows a very good image, …’

**Determiner Phrases.** In the following examples the focus is on the embedded element in the Pashto determiner construction. The EL single elements inside Pashto in determiner constructions constitute 16% of the total English-Pashto CM data. The determiners which precede the embedded elements are the demonstrative, possessive adjective, quantifiers and indefinite *yaw*. The demonstrative, indefinite, and the quantifier determiners are suffixed according to head gender and number.

In the complementizer that clause (with morpheme *che*) in (13), the embedded element ‘issue’ in subject position is preceded by the Pashto *daa* functioning as determiner. The example retains Pashto word order.

(13) Electronic media 0016
Che [[daa issue] [daira sensitive]] da]
COMP DM.PRX issue very sensitive COP.PRS.IPFV.F.3SG
‘… that this issue is very sensitive.’

In (14) the embedded element ‘future’ in the NP is preceded by a possessive adjective *zamong* that functions as a determiner. The morpheme distribution inside the NP phrase is that of Pashto congruent with its English counterpart ‘our future’. The main verb *taba* is inflected by the intransitive COP sh for tense and aspect and subject-verb agreement is observed by the pronominal marker-*o* on the COP sh coindexed with the 1st person Plural subject.

(14) The use of tuition in Education 031-M2
[zmonga future] [taba sh-o]
POSS.1PL future destroy COP.PRS.PFV.1PL
Our future has been destroyed.’

In (15) there are English embedded elements in positions in the intransitive *sh* clause. The form ‘collapse’ is the main verb from English inflected by the intransitive COP sh for Pashto
tense and aspect, and subject-verb agreement is marked on the intransitive sh by the third person pronominal marker i. The other embedded element in the subject position is preceded by Pashto quantifier toul.

(15) Peace, prosperity and development 0001
    Nu [[toul system] [collapse sh-i]]
    then all system collapse COP.PRS.PFV-3SG
‘Then the complete system collapsed.’

In examples (16) and (17) Pashto is the matrix language providing the morpho-syntactic frame to the embedded elements. In (16) the embedded noun target is preceded by Pashto indefinite morpheme yaw, following the same morpheme order as that of Pashto.

(16) The use of tuition in Education 035-Caller
    Medical aw engineering mong [yaw target] [jor kre medical and engineering PRN.1PL one.M.SG target make do day]]
    COP.PRS.IPFV.M.3SG
‘We have made medical studies and engineering a single target.’

In (17) the embedded element time is preceded by the Pashto DST demonstrative pronoun agha, following the morpheme order of Pashto. Subject-verb agreement in the two examples is marked by the pronominal marker -i on the transitive auxiliary ka-i.

(17) The use of tuition in Education 012-F1
    k aghwi [agha time] tapoos [na-ka-i]
    if DM.PL.DST DM.DST time ask NOT-PRS.PFV-3PL
‘… if they do not ask questions at that time.’

Nouns in Prepositional Phrases. In Pashto syntax pre- and postpositions are combined, with the oblique case at the end. The oblique case in the pre-postposition always functions in the indirect object position. In the relationship of Pashto NPs constituents with its verbal predicate is determined by the case marking system of Babrakzai (1999:14). Sometimes, the NP in the preposition is encoded as oblique as in the possessive construction with the possessive preposition da in the locative construction with the pre or postposition pa/ke or bande and in the dative construction with the postposition ta Babrakzai (1999). About 17 % of the total single embedded elements in the English-Pashto were recorded.

In (18) the focus is on ‘forum’, preceded by the Pashto adverb dase in the PP construction, working as indirect object to the transitive auxiliary kaw. Subject-verb agreement is on the transitive auxiliary kaw by the 3rd person singular marker i.

(18) Khyber Club 1: media role 0004-Caller1
    da k dase sez-una [pa dase forum ban-de] [discuss kaw-i]
    DM.PRS if such thing-PL LOC such forum in-OBL discuss do.PRS.IPFV-3SG
‘If such things are discussed on such forums.’
In (19) the focus is on the EL element 'show' in the locative construction working as indirect object to the past tense transitive auxiliary ku. In order to establish a core argument with the verb ku in the verbal position the noun 'show' in pre-postposition is marked as oblique case. Pashto has a split tense agreement and in the past tense the agreement is triggered by the direct object which is identical in the form with nominative. The auxiliary k along with the perfect marker wa for it is marked by the 3rd person pronominal u for subject-verb agreement.

(19) The use of tuition in Education 033-M1
\[
\text{taso mong ta } [[\text{pa } \text{show ke} ] [\text{call} ] [\text{wa-k-u}]]
\]
\[
2\text{PL.ERG PRN.1PL to LOC show OBL call PFV-3SG-PST.3SG}
\]
‘You called us into the show.’

In (20) the embedded noun ‘class’ in the intransitive construction is triggered by the morpheme ta in the Pashto morpho-syntactic frame. The word order inside the bilingual constituent ‘class ta’ is [N P], unlike its English counterpart. The word order constraint for Pashto as verb final is observed. The agreement between the argument agha and the verb razi is marked by the 3rd person pronominal morpheme -i. The distribution of the morphemes is from Pashto.

(20) The use of tuition in Education 034-M2
\[
\text{k agha } [[\text{class } \text{ta} ] [\text{regular} ] [\text{na raz-i}]]
\]
\[
\text{DM.DST class to regular not come.PRS.IP-FV-3SG}
\]
‘… if he is not coming to the class regularly.’

In examples (21) and (22) the word order is from Pashto as Pashto is a rigidly verb final language. Inside the bilingual phrase as in (21) it is [N P] and (22) also has the same [N P] following Pashto word order inside a prepositional phrase. In (21) the agreement between the argument aghwi and the COP wi is marked by the morpheme i.

(21) Khyber Club 1 032-M2
\[
\text{da aghwi } [[\text{education } \text{s ra} ] [\text{interest} ] \text{ziath } [\text{wi}]]
\]
\[
\text{GEN DM.PL.DST education with interest many COP.PRS.IP-FV.3PL}
\]
‘Those mothers who are uneducated have a great interest in education.’

In (22) the agreement between the subject agha and the COP day is marked by the morpheme y. In the above two examples the embedded language elements are maximally controlled by the Pashto morpho-syntactic frame.

(22) Khyber Club 1 003-M1
\[
\text{agha zma } [[\text{hee side ta} ] [\text{nast day}]]
\]
\[
\text{DM.DST POSS.1SG right side to sit COP.PRS.IP-FV.M.3SG}
\]
‘He is seated at my right side.’
Possessive constructions (də). In Pashto the possessive relation is marked by the preposition də. The possessive preposition də can work as conjunction with the other postposition as example də Ahmad sra, ‘with Ahmad’ and də Ahmad dapara, ‘for Ahmad’ (Babrakzai, 1999).

In (23) there are three bilingual constituents. In the object position the noun ‘leadership’ is preceded by the distal demonstrative, and accompanied by the oblique preposition ke; the second insertion ‘youth’ is preceded by the morpheme də in the possessive construction, and the third insertion a noun preceded by the Pashto morpheme sa. One thing which should be noted is that when the experiencer is marked with the genitive marker then the copula agrees with the non-oblique NP (Babrakzai, 1999). ‘sa role’ functions as nominative subject and agrees with the copula day. The functional morphemes are from Pashto, which provides the morphosyntactic frame.

(23) aw [[agha leadership] ke] [[də youth]] [sa role] [day]]
and DM.DST leadership in GEN youth what role COP.PRS.IPFV.M.3SG
‘And what will be the role of youth in that leadership.’

In (24), when the experiencer is marked by the possessive də, then the subject-verb agreement is marked on the non-oblique NP. Subject-verb agreement is between the unmarked NP galti and the copula da, satisfying the word order and functional category constraints. Pashto provides all functional morphemes.

(24) Electronic media 0010
[[daa də youth] galti] da
DM.PRX GEN youth mistake COP.PRS.IPFV.F.3SG
‘It is a mistake of youth.’

In the bilingual PP example (25) the embedded noun ‘environment’ is marked by the possessive morpheme də. The embedded noun ‘environment’ is preceded by the Pashto morpheme də, following Pashto word order.

(25) Electronic media 0019
de də pehawar də environment-na ha habar da-y
3SG.PRX GEN Peshawar GEN environment-LOC good knowledge COP.PRS.IPFV.M-3SG
‘He is very much acquainted with the environment of Peshawar.’

In (26) there are three bilingual constituents. The embedded noun ‘student’ in the bilingual NP is preceded by agha ‘that’; the AP ‘part time’ as embedded island modifies ‘learning’ in the bilingual PP and the noun ‘learning’ in the PP is preceded by possessive morpheme də. In the bilingual PP the embedded noun follows Pashto word order. In English the possessive structure would be ‘for his own learning’ where as in Pashto it is ‘of own learning for’. The verb kai is coindexed on the subject for the subject-verb agreement by the 3rd person singular pronominal marker i.
(26) The use of tuition in Education 010-Caller 1
[agha student] [part time] [do hapal learning] [dapara] ka-i
DM. DST student part time GEN own learning for do.PRS.IPFV.3SG
‘That student is doing that part time for his own learning.’

In (27) Pashto is the matrix language and provides the morpho-syntactic frame to the embedded elements preceded in the bilingual PP by the possessive d/da. The word order constraint is observed in the bilingual phrase as embedded nouns follow Pashto word order as may be seen in the bracket [P N P N], where the embedded elements are preceded by the possessive marker do. In Pashto the bilingual constituent order is of tuition of concept as its English counterpart is ‘the concept of tuition’ [DET N P N]. Subject-verb agreement is marked on the copula yum, which agrees with the 1st person pronominal marker m.

(27) The use of tuition in Education 028-M2
Za [[da tuition da concept]] helaf zeka [yu-m]]
1SG GEN tuition GEN concept against because COP.PRS.IPFV.1SG
‘That is why I am against the concept of tuition.’

EL Adjectives. English adjectives were inserted in 41 instances, where 16 were found as single elements and 25 as a phrase (EL Island).

In (28) the embedded adjective ‘important’ modifies the Pashto noun habr-e. In the mixed constituent the English modifier is preceded by the Pashto determiner string, and follows Pashto word order. The functional category constraint is observed as the Pashto present tense COP is inflected with the 3rd person plural pronominal morpheme-e for person and number agreement.

(28) Khyber Club 1: media role 0020-M3
[yaw so dair-e important habr-e] di
one.M.SG few very-PL important talk-PL COP.PRS.IPFV.3PL
‘I have a few very important words.’

In the entire data set predicative adjectives were used only two times. In attributive position English adjectives were very often preceded by Pashto intensifiers and the indefinite article. In the entire data set there is no single case where the adjective is affixed with any kind of morpheme. In Lotfabbadi (2002) 37% of the EL Swedish adjectives were affixed with Persian bound morphemes. In the present data, according to the definition of code switching and borrowing in Muysken (2000), the English elements in Pashto matrix language behave as CM elements.

In (29) Pashto is the matrix language and the two wellformedness constraints (word order and functional category) are satisfied. The directive pronoun ra functions as subject, and is marked with the 1st person plural pronominal marker-u on the verb kaw. Example (29) satisfies word order and functional category constraints.
In (30) the embedded adjective ‘standard’ in a bilingual NP is preceded by Pashto indefinite yaw and followed by the NP head tareek-a, marked by a gender suffix-a. The bilingual constituent in the determiner complex structure is different from the English structure. The bilingual constituent in the following example is [DET DET A N] whereas in English the same construction will be ‘a standard technique’, [DET A N]. The copula da is co-indexed with tareek-a. Pashto provides the morpho-syntactic frame.

(30) Khyber Club 1: media role 0016-M3
[(da yaw standard tareek-a] da]
dm.prx one.msg standard technique-f cop.prs.ippv.f.3sg
‘It is a standard technique.’

In (31) the embedded adjective ‘regular’ is in attributive position modifying the head mulaqat ‘meeting’. The morpheme order of the bilingual NP is in congruence with the English morpheme order with attributive adjective. The agreement pattern of subject and verb follows Pashto subject-verb agreement. The 3rd person pronominal marker-i on the intransitive auxiliary is suffixed for subject-verb agreement.

(31) Khyber Club 1 004-M2
os ba zamong taso sra [regular mulaqath] [keg-i]
now cl.fut poss.1pl 2pl with regular meeting become.prs.ippv.3pl
‘Now we will have regular meetings with you.’

In (32) the English adjective ‘regular’ in the ‘if’ clause is inserted in predicate position, modifying agha in the subject position. In English-Pashto CM data only 2 instances with the predicate adjective were found. In Pashto the embedded adjective in attributive or predicate position is not affixed for the number, person or gender of the head noun. In the example below the predicate structure of the embedded adjective is congruence with the English. In the following example the directive pronoun ra is prefixed to the intransitive verb z1, to show the movement of the subject agha. The directive pronoun is coindexed with the subject. Subject-verb agreement is marked on the 3rd person nominal marker i suffixed with the intransitive verb stem ‘z’.

(32) The use of tuition in Education 034-M2
k agha [class ta] regular [na ra-z-i]
if dm.dst class to regular not cop.prs.ippv.3sg
‘If he is not coming to the class regularly.’
4.2. English multi-word strings Pashto CM sentences

In this part the focus is on multi-word strings forming embedded island (Myers-Scotton, 1993). These are phrases within a bilingual clause where the words show a structural dependency relationship which makes them well-formed in the embedded language. They often are collocations. It is different from single element code mixing because in a bilingual clause the island retains the word order and functional categories of the embedded language. In Table 5 the types of multi-word strings in the data set are listed.

<table>
<thead>
<tr>
<th>Type</th>
<th>#</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>EL NPs Island in ML</td>
<td>19</td>
<td>24</td>
</tr>
<tr>
<td>In ML PPs</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>In the Determiner construction</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>In the possessive construction</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td>Embedded Island as PPs</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Adjective in EL Island</td>
<td>27</td>
<td>34</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>80</strong></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: English multi-word strings Pashto CM sentences

*Bare Embedded NPs.* In the following sentences English island have been embedded in larger units following Pashto grammar, often collocations or formulaic string of words. In English-Pashto embedded islands. In the English-Pashto code mixing data 24% of such types are recorded.

In (33) the constituent ‘check and balance’ is found in the Pashto frame. It is well-formed and follows the English word order. It is a common formulaic phrase in English, although in English itself it is plural rather than singular. Subject (covert) verb agreement is marked on the verb wasat –u by the 2nd plural marker-u.

(33) Peace, prosperity and development 0003

Che pa de bande [\textit{check and balance}] [wasat-u]

COP at CL on check and balance keep.prs.ippf-1pl

‘… that we should keep checks and balances on it’

In example (34) ‘live show’ as such is well-formed in English, but overall grammar is Pashto. Subject verb agreement is marked on the cop da for the 3rd person subject ‘live show’.

(34) The use of tuition in Education 001-M1

da [\textit{live show}] da

dm.prs live show cop.prs.ippf.p.3sg

‘This is a live show.’
Strategies for incorporating nouns and verbs in code-mixing: the case of Pashto-English bilingual speech

In (35) the constituent ‘negative role’ has been found in a Pashto frame. Subject verb agreement is marked on the transitive auxiliary kaw inflected by the 3rd person singular marker i, co-indexed with the subject media. Interestingly, many content words in the sentence come from English.

(35) Khyber Club 1: media role 0001-Caller1
    media [negative role] hu [play kaw-i]
    media negative role indeed play do.PRS.IPFV-3SG
    ‘The media indeed play a negative role.’

In (36) the constituent ‘political analyst’ is well-formed following English rules. The example shows a common English collocation phrase in a Pashto frame. Subject verb agreement is marked on the cop yu, co-indexed with the 1st person singular subject. The tense, aspect, and the agreement on the cop yu suggest that Pashto is the ML.

(36) Electronic media 0025
    monga sə [political analyst] ya expert na yu
    PRN.1PL like political analyst or expert not COP.PRS.IPFV.1PL
    ‘We are not political analysts or expert like.’

Adpositional phrases. In the PP constructions in the following examples we find EL islands preceded by a pre or/ pre-postposition (except for the possessive construction with de). The nouns in pre-postposition are marked oblique.

In (37) ‘live show’ is embedded in an oblique construction. The PPs constituents in the pre-postposition with ta is a clear indication that the ML controls the larger constituents. The subject verb agreement is marked on the transitive auxiliary k inflected by the 1st person pronominal marker u. Pashto has split ergativity and agreement in the past tense is marked on the direct object. Tense, aspect, and the agreement marker on the auxiliary k suggest that Pashto is the ML.

(37) Khyber Club 1: media role 0012-M1
    che taa ma [ta live show ke] [call wa-k-u]
    COMP 2SG.ERG 1SG.OBJ to live show OBL call IPFV-do.PST-1SG
    ‘… that you called me in a live show’

In (38) the plural noun ‘colleges’ is embedded in a Pashto oblique construction frame. The form ‘colleges’ contains the morpheme -s for plural. Subject verb agreement is marked on the intransitive auxiliary keg inflected by the 3rd person pronominal marker i. The morpheme i is coindexed with the reduplicated subject NP [concept muncept]. The word order and the functional category constraints determine that Pashto is the ML. Tense, aspect, and the agreement marker on the auxiliary keg suggest that Pashto is the ML.

In the embedded NP ‘concept’ has been reduplicated with meaningless muncept to mark indefinite reference. Many embedded words can be duplicated in the same way: school
mul, college malej, university munawrsty, etc. This strategy is also used in Urdu.

(38) The use of tuition in Education 036-Caller
[pa college-s ke] [concept muncept] [na clear keg-i]
LOC college-PL OBL concept.REDUP not clear become.PRS.IPFV-3SG
‘In colleges the concept and so on will not be going to become clear.’

In (39) ‘current era’ in the PP is oblique but is well-formed in English. The subject verb agreement is marked on the cop yu by the 1st person plural pronominal marker u. The morpheme u is coindexed with the subject mong. The other morphemes, the complementizer ch and the directive pronoun –ra function to cement the morpho-syntactic frame of Pashto.

(39) Khyber Club 1 037-M1
che mong [pa current era ke] [ra-wan yu]]
COMP 1PL LOC current era OBL 1PL-go COP.PRS.IPFV.1PL
‘Now as we are going into the current era.…’

Determiner constructions. In the following examples the focus is on Pashto determiner constructions. In the English-Pashto code mixing data 14% of the constituents in the determiner construction were recorded, with a Pashto morpho-syntactic frame. The EL single elements inside Pashto in determiner construction were 16% of the total English-Pashto CM data. The determiners which precede the embedded islands are the Pashto demonstrative, possessive adjective, quantifiers and indefinite yaw.

In (40) ‘golden rule’ is placed in the morpho-syntactic frame of Pashto, but follows the internal structural dependency relations of English and is internally well-formed, except that the plural suffix is missing. Subject verb agreement is marked on the cop wi, inflected by the 3rd person pronominal marker i. The tense, aspect, and the agreement marker on cop wi are evidence that Pashto is the ML.

(40) Peace, prosperity and development 0007
da Islam [[toul golden rule] ba halta [apply w-i]]
GEN Islam all golden rule CL.FUT DM.DST apply COP.PRS.IPFV-3SG
‘All golden rules of the Islam will be applied there.’

In (41) the embedded island ‘second pause’ is preceded by the Pashto indefinite determiner yaw. Subject-verb agreement in the ergative is marked on the direct object. The agreement on the verb is between the 3rd person pronominal marker e and the bilingual NP [yaw second pause] in object position. The morpheme e is co-indexed with the object the bilingual NP [toul golden rule] for information about its form and makes an maximal projection.

(41) Electronic media 0014
agha [yaw second pause] [wahast-e]
DM.DST ONE.M.SG second pause take.PST.PFV-M.3SG
‘He had taken a second pause.’
In (42) the embedded island ‘type trainings’ is preceded by the Pashto demonstrative determiner agha ‘that’. The head noun in the bilingual NP is suffixed by plural -s. ‘type trainings’ is odd in English because of its plural. Subject-verb agreement is marked by the 1st person pronominal marker u, coindexed with the subject mong.

(42) My right (women Issues) 0029
bia mong de la [agha type trainings] [arrange
then 1PL.NOM DM.PRX.F for DM.DST type trainings arrange
kaw-u]
do.PR.SPPFV-1PL
'Then we arrange for them that type of training.'

Possessive constructions. Possessive da has the same function with the constituent as it has with single embedded nouns in the English-Pashto CM data. In the present data 15% of the constituents in the possessive construction were recorded. In the following examples the focus is only on the function of the constituent in the morpho-syntactic frame of Pashto.

In (43) the embedded island ‘social events’ is embedded in the possessive construction of Pashto. In the possessive construction the embedded island is preceded by the Pashto demonstrative de but the head noun in the bilingual NP is suffixed by plural -s and follows the internal structural dependency relations of English. Subject-verb agreement is marked on the verb by the second person pronominal marker-e.

(43) Khyber Club 1: media role 0003-Caller 1
k taso zamong [[da [de social events]] pa
if2 PL.NOM POSS.1PL GEN DM.PRX social events LOC
bara ke [wa-gur-e]
about OBL IPPV-look.PRS-2PL
‘… if you look about our social events.’

In (44) below the embedded elements are a good example of Pashto word order, and it is reflected in the CM example (45):

(44) [da jwand [buniadi haq-una]
gen life basic right.PL
‘the basic rights of life’

Subject-verb agreement is marked on the verb with the 3rd person singular morpheme marker -i. The two functional category and word order constraints justify that Pashto is the ML. The English element is split up, and follows Pashto order in part.

(45) Peace, prosperity and development 0002
Agha ta [[da life] [basic rights] [namelaweg-i]]
DM.DSTO GEN life basics rights notmeet.PRS.IPFV-3SG
‘He does not meet the basic rights of life.’
In (46) the embedded island ‘previous caller’ is preceded by the possessive marker *da* as in Pashto. The constituent follows the morpheme order of English and follows the internal structural dependency relations of the EL as the adjective ‘previous’ in attributive position modifies the head ‘caller’. ML frame is realized by the functional category constraint as the subject-verb agreement is marked on the verb by the second person pronominal marker-*am* on the subject *za*. In the possessive construction the constituents NP is preceded by the marker *d* in order to satisfy the functional category constraint of the matrix language. The two proposed constraints word order and functional category have maintained the Uniform Structure Principle and have confirmed that Pashto is the matrix language.

(46) Khyber Club 1: media role 0009-F1
[[*da  previous  caller*] habara] za [viewer sra] [share kaw-am]
of previous caller talk 1SG viewer with share PRS.PFV-1SG
‘I will share the point of view of the previous caller with the audience.’

**Adpositional phrases.** Many embedded island are adjuncts when functioning as adverbial phrases of time or place. Their structure is formulaic, and that is why they function in chunks, composite units, or composite expression (Backus, 1999; Myers-Scotton, 2002). In present data about 5% of the constituents PPs are like this.

In (47) there are two embedded elements. The gerund ‘blackmailing’ is used as a noun and the embedded PP ‘in a sense’ is used as an idiom functioning as adjunct. In the passive construction the subject verb agreement is marked on the infinitive transitive Pashto verb *way-*al-e inflected by the 2nd person pronominal marker *e*. The morpheme *e* is coindexed with the covert 2nd person plural subject for information about its form and maximal projection. The word order and the functional category constraints support that Pashto is the ML and has qualified that the constituent is constrained by the morphosyntactic frame of Pashto. The tense, aspect, and the agreement marker on the verb suggest that Pashto is the ML.

(47) The use of tuition in Education 030-M1

**Blackmailing** hu [in a sense] war-ta na-shay way-al-e
blackmailing indeed ina sense 2SG-TD not- PRS.PFV say-2PL
‘In one sense indeed it cannot be said blackmailing.’

In (48) the embedded element ‘as a profession’ is a modifying the demonstrative *daa* refer to the topic in the discourse. The embedded PP ‘as a profession’ is a modifier and function as adjunct in the bilingual clause but follows the internal structural dependency relations of the EL. Pashto is a verb final language and the word order of the example is verb final. The morpheme *i* is coindexed with the covert subject anaphoric referenced by the demonstrative *daa* for information about its form and maximal projection. The word order and the functional category constraints support that Pashto is the ML and has qualified that the constituent is constrained by the morphosyntactic frame of Pashto. The tense, aspect, and the agreement marker on the COP suggest that Pashto is the ML.
The use of tuition in Education 011-F1

che [as a profession] də cha [yi]

‘... if as a profession someone has it.’

Adjective + Noun combinations. In English-Pashto CM data total 32% of the embedded island has been modified by the attributive adjectives. In the single element insertion data 16% of the adjectives recorded in the morpho-syntactic frame of Pashto. The following examples 53 to 55 have already been discussed in 40, 45 and 50. In the English-Pashto CM data the EL adjectives were found only in attributive position. In the current data not a single adjective is affixed by the ML morphemes but is preceded by the Pashto demonstrative, intensifier and indefinite article.

In (49), repeated from (2), the constituent ‘negative role’ is placed in the morpho-syntactic frame of Pashto. It follows English morpheme order and the internal structural dependency relations of English.

In (50) the constituent is placed in the possessive construction following the morpho-syntactic frame of Pashto. The possessive marker də precedes the constituent ‘previous caller’, where the attributive adjective ‘previous’ modifies the head noun ‘caller’.

5. English verbs in KP Bilingual Compound Verbs

5.1. The data

In English-Pashto code-mixing data 40 instances of the combination of the embedded element noun, adjective and verb were recorded, the so-called Bilingual Compound Verbs (BCVs). These constructions need to be seen in the context of the Light Verb Constructions (LVCs) that are very common in the region. In the literature different researchers have treated the LCVs differently. It has been named complex predicate, compound verb, composite verb, and Light Verb Constructions (Cattell, 1984; Grimshaw and Mester, 1998; Butt, 1995; Muysken, 2000; etc.). The major focus here is on how the grammaticality is observed in bilingual sentences. We begin with some basic examples.
In (51), repeated from (15), the English elements in two positions appear in a Pashto Present perfective tense frame, marked with the \textit{cop} \textit{sh}. The English root ‘collapse’ is joined together by verbal inflection \textit{sh} for Pashto tense and aspect. In the verbal complex phrase the subject verb agreement is marked with the third person \textit{i}.

\begin{verbatim}
(51) Peace, prosperity and development
nu [toul system] [collapse sh-i] then all system collapse \textit{cop.prs.pfv-3sg}
‘Then the complete system collapsed.’
\end{verbatim}

In (52) the embedded verb ‘use’ in the verbal complex is preceded by the Pashto transitive auxiliary marked for tense and aspect. The English verb ‘use’ carries the semantic load in the sentence and functions as the content lexical item. Auxiliary \textit{kaw} is suffixed by the 3\textsuperscript{rd} person singular pronominal marker \textit{i} for the subject-verb agreement with the subject \textit{da}.

\begin{verbatim}
(52) Khyber Club 1 031-M3
da hr sook [use kaw-i] & A D V. & P R X.
& every who use & do.prs.ipfv-3sg
‘It is used by every person.’
\end{verbatim}

In (53) ‘connected’ the English root in the bilingual verbal complex is followed by the verbal inflection \textit{yu}. The English past participle verb ‘connected’ carries the semantic load in the sentence and functions as content lexical item. The \textit{cop} \textit{yu} carries subject-verb agreement with the 1\textsuperscript{st} person plural subject \textit{monga}.

\begin{verbatim}
(53) Khyber Club 1 035-M3
monga agha halaq-o sra [connected yu] & P R N. & 1.PL & D M. & D S T.
& people-OBL with connected & \textit{cop.prs.ipfv.1pl}
‘We are connected with those people.’
\end{verbatim}

Example (54), despite the English lexical elements embedded in bilingual verbal complex, is well formed. The example supports the morpheme order constraints because the bilingual VP and everything else in the sentence follows Pashto order. In bilingual verbal complex [\textit{na attach kaw-am}] in the embedded verb root is inflected for tense and aspect by Pashto transitive auxiliary \textit{kaw}.

\begin{verbatim}
(54) Khyber Club 1 036-M2
agha za [computer sra] [na attach kaw-am] & D M. & P R N. & 1.SG & D M. & P R N. & 1.SG
& computer with not attach & do.prs.ipfv-1sg
‘I am not going to attach that with the computer.’
\end{verbatim}

In (55) the switch includes the English-Pashto bilingual LVC [\textit{support kar-i}]. The lexi-
cal English main verb 'support' is inflected by the auxiliary kar for the Pashto tense and aspect. The example supports the morpheme order constraints because the bilingual VP and everything else in the sentence follows Pashto order. The 3rd person plural pronominal marker i suffixed on the auxiliary kar and is coindexed with the subject moor aw plar.

(55) The use of tuition in Education 039-M1

\[\text{paker de che moor aw plar ye [pa aghe needed ke] [support kar-i]}\]

'Obl support do.prs.pfv-3pl

'It is needed that mother and father should support him in it.'

In (56) the English root 'conclude' is embedded in the bilingual VP, and it is inflected by Pashto auxiliary kaw for the tense and aspect. The example supports the word order constraint because the bilingual VP and everything else in the sentence follows Pashto order, because only one language supplies the morpheme order. The transitive auxiliary kaw is suffixed by the Pashto morpheme am to establish subject-verb agreement with the 1st person singular subject za.

(56) The use of tuition in Education 018-M3

\[\text{za ba da habare [conclude kaw-am]}\]

'I will conclude this discussion.'

In (57) the embedded phrasal verb 'brought up' in the bilingual VP is followed by the Pashto tense and aspect marker wash, suffixed by the Pashto morpheme i to establish subject-verb agreement with the 3rd person plural subject aghwi. The particle 'up' plays no independent grammatical role.

(57) The use of tuition in Education 017-M1

\[\text{che da aghwi h [brought up wash-i]}\]

'Comp gen dm.pl.dst good brought up become.prs.pfv-3pl

'That they should have best brought up.'

5.2. Monolingual complex predicates in Pashto

The monolingual complex predicate has researched in a wide range of languages, where the complex predicate comprises a light verb and a lexical element (noun, verb, adjective) (CatteII, 1984; Grimshaw & Mester, 1998; Butt, 1995; etc.). In the Indo-Iranian languages such as Urdu, Hindi and Panjabi, the first verb is called the main verb and the one at the right side expressing aspectual properties is called light verb. The light verb in the complex predicate is studied in detail for its aspectual role (Akhtar, 2003). Butt (1995) claims that the light verb construction (LVC) forms a single constituent in the complex predicate. To determine the semantic role of the light verb in complex VP is a challenge. The verbs are not entirely devoid of semantic predicative power either as there is a clear difference between take a bath and give a bath. The verbs thus seem
to be neither at their full semantic power nor at a completely depleted stage. Rather, they appear to be semantically light in the sense that they are contributing something to the joint predication. However, it is relatively difficult to characterize this component (Butt, 1995).

These light verbs have corresponding main verbs which can function as a predicate in a clause. When they are used as light verb, they are more auxiliary in nature than the main verb and their role changes to aspectual or operator ‘verb’, frequently translated as ‘do’ or ‘make’. In the monolingual or bilingual predicate the light verbs do not take their role as main verb but they mainly contribute aspectual information which appears in conjunction with a lexical element; this contributes to the core semantic content of the construction (Akhtar, 2003). These verbs are semantically bleached and usually follow a collocation pattern with a special class of verbs as the intransitive light verb collocates with the intransitive main verb.

In examples (58) and (59), Akhtar (2003) has differentiated the aspectual light verb from the main verb in the complex predicate. In (58), the verb *giaa* ‘go’ itself refers to an event and because of this semantic nature, Akhtar called it main verb. In (59), it is a light verb because it only contributes aspectual information to the clause.

(58) 
\[
\begin{array}{ccc}
\text{PRN.3SG.M} & \text{daakh} & \text{de} & \text{giaa} \\
\text{Post} & \text{give} & \text{go.pst} \\
\end{array}
\]
‘He delivered the mail and went away.’

(59) 
\[
\begin{array}{ccc}
\text{PRN.3SG.M} & \text{nalka} & \text{jeor} & \text{giaa} \\
\text{Hand.pump repair} & \text{go.pst} \\
\end{array}
\]
‘He fixed the hand-pump and went away.’ (Akhtar, 2003: 100)

A considerable amount of work has been done on the light verb in monolingual complex predicates of South Asian languages, but no satisfactory agreement has been proposed which can account for the licensing of the light verb (Butt, 1995; Akhtar, 2003). The only satisfactory account that has been proposed is the productive nature of the light verb in complex predicate of making new verbal category at the grammatical level. However, in the present study, the focus is on the bilingual complex predicate and the role of the light verb as emerging new verbal category. In order to understand the nature of the light verb construction in English-Pashto bilingual data, it is important to highlight the role of the light verb in the monolingual complex predicate of Pashto (Robson and Tegey, 1996; Babrakzai, 1999).

Robson and Tegey (1996) have classified the verbs as belonging to three different classes. They distinguished the three classes, such as simple verb, derivative verb and doubly irregular verb, and described the verbs according to perfective and imperfective aspects. In Babrakzai (1999), the verb has been categorized as transitive and intransitive nature according to its function and thematic role in the sentence. He has discussed in detail how the nominal elements (adjectives, nouns) and verbal elements (verbs) form compound verbs with the transitive and intransitive auxiliaries. Roberts (2000), in his study, has focused on the verb structure according to aspects, stem variation, and its function within sentence.
Babrakzai (1999) has further differentiated between intransitive ‘inchoative verb’ and ‘light verb’. The inchoative verbs are derived from stative or adjectival stems where the subject is affected by the event. The aspect plays a role to show the change in the subject with the help of the intransitive auxiliaries. In a schematic summary:

\begin{verbatim}
<table>
<thead>
<tr>
<th>Verb Type</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inchoative verb</td>
<td>light verb</td>
</tr>
<tr>
<td>Light verb</td>
<td></td>
</tr>
<tr>
<td>Light verb</td>
<td></td>
</tr>
</tbody>
</table>
\end{verbatim}

5.3. Intransitive compound verb constructions with the inchoative light verb

A light inchoative verb is made of a verbal element or nominal element. Babrakzai (1999) has differentiated two types of light verb. In the first group, the nominal element functions as ‘subject’ of the auxiliary, and in the second group it functions as object to the intransitive auxiliary. If the predicate has triggered another argument, then it would be in the oblique form:

\begin{verbatim}
(61) dltha da cricket lube kegi
     adv:Prx gen cricket play become.prs.ippv.3sg

‘Here cricket is played.’
\end{verbatim}

In the second group of the light inchoative verbs, the verbal element and the intransitive auxiliary make a compound verb. In this group, it functions as single predicate and takes another noun which functions as subject:

\begin{verbatim}
(62) Juwar karale ked-al
     maize sow become.pst.ippv.m.3pl

‘The maize was being sowed.’
\end{verbatim}

In (62), the verb agreement is triggered on the subject juwar ‘maize’; the other two elements karale and kedal function as a compound verb.

The first group where the nominal elements function as subject includes the following verbs (63):

\begin{verbatim}
(63) Verbal/Nominal elements Intransitive auxiliary Gloss
    koshish [try]    kedl [become]    ‘try’
    pekar [thought]  kedəl [become]    ‘think’
    lobe [play]     kedəl [become]    ‘play’
\end{verbatim}

The verb in the second group (64) is made of a verbal element that is why it takes subject to develop an agreement with the compound VP. The second group where the verbal elements function as object includes the following verbs:
5.4. Transitive compound verb constructions with the light verb ‘kawəl’

In the transitive construction, most of the compound or complex predicates are made of the nominal element and the transitive auxiliary kawəl ‘do’ and occasionally other verbs:

(Babrakzai, 1999: 139)

In the complex predicate, the nominal functions as a direct object with the transitive auxiliary or another verb, which needs an agent subject. Such compound verbs follow a transitive pattern: nominative-accusative in the present tense and ergative-absolutive in the past:

(66) mina  pekar  kaw-i ch nan baran raz –i
Meena  thought  do.PRS.IPFV-3SG  comp  today  rain  come.3SG
‘Meena thinks that it is going to be rain today.’

(67) mine  pekar  kaw–ə  a ch war  ye  raz –i
Meena.ERG  thought  do.PRS.IPFV-M.3SG  COM  brother  CL.F.3SG  come.3SG
‘Meena was thinking that her brother was coming today.’ (Babrakzai, 1999: 141)

In (66) and (67), the constituents in the complex predicate are nominal elements. If the constituent in the compound verb is a verbal element, instead of a nominal one, then in such a construction the verb requires an argument as its direct object (Babrakzai, 1999).

In (68), the verbal element shuro ‘start’ and the transitive auxiliary kawəl ‘do’ take another argument, majlis ‘meeting’.
Strategies for incorporating nouns and verbs in code-mixing: the case of Pashto-English bilingual speech

(68) ta majlis po pinzo baj-o shuro kar
you.ERG meeting LOC five o'clock-OBL start do.PST.M.3SG
‘You started the meeting at five o’clock.’

5.5. Pashto complex verbs and Aspect driven asymmetries

Roberts (2000) has underlined the role of aspect in compound verbs which is responsible for the syntactic and morphological constituency in Pashto. In the compound verb construction in perfective aspect, the verb behaves differently than the verb in imperfective aspect. Compound verbs in perfective aspects behave as two units. According to Roberts (2000), the two elements of an imperfective compound verb function as single word at phonological level. In contrast, the two elements of the verbs in the perfective aspect behave as two words at the phonological level.

Merger and clitic placement. The morphological process of a compound verb merging into a single word in imperfective aspect may involve a noun or an adjective ending with a consonant sound. Then the initial [k] of the transitive auxiliary kaw is dropped and the rest of the auxiliary is added to the noun or adjective to form a single word (Robson and Tegey, 1996: 109). In Roberts (2000), the focus is on Pashto clitics and their placement in the compound verb, while here it lies on the role of the aspects in the bilingual compound verb (BCV). It will be of great interest to see whether the two auxiliaries in the two aspects behave in the same way in the BCV as they behave in the monolingual complex predicate.

(69) zgal aw-am dee
    run do.PRS.IPFV-1SG cl.2SG
‘I am making you run.’

(70) khaaysta kaw-am dee
    beautiful do.PRS.IPFV-1SG cl.2SG
‘I am making you beautiful.’ (Roberts, 2000: 37)

In present imperfective (69) when the adjective zgal ends in the consonant [l], the [k] of the auxiliary has been dropped, but in (70), the adjective ends in a vowel sound, so there is no need to drop the [k] from kaw.

(71) jubal k-am dee
    injured do.PRS.PFV-1SG cl.2SG
‘I injure you.’

(72) Khaaysta kr-am dee
    beautiful do.PRS.PFV-1SG cl.2SG
‘I make you beautiful.’ (Roberts, 2000: 38)

In (71) and (72), it is evident that the [k] of transitive kr in the perfective aspect has not been dropped. The compound verb in the imperfective aspect works as single unit and
in the perfective aspect, it works as two units. This can be tested with the clitics merging in compound verb of the perfective and imperfective aspects.

**Clitic placement in Compound Verbs.** Clitic placement may be seen in the following Present Imperfective and Present perfective sentences in (73):

(73) zgal aw-am dee  run  do.PRS.IPFV-1SG  CL.2SG  
'I am making you run.'

(74)*zgal dee aw-am  run  CL.2SG  do.PRS.IPFV-1SG  
'I am making you run.'

(75) khaaysta kaw-am dee  beautiful  do.PRS.IPFV-1SG  CL.2SG  
'I am making you beautiful.' (Roberts, 2000)

(76) *khaaysta dee kaw  -am  beautiful  CL.2SG  do.PRS.IPFV-1SG  
'I am making you beautiful.'

(77) khaaysta k-am dee  Beautiful  do.PRS.PFV-1SG  CL.2SG  
'I make you beautiful.'

(78) khaaysta dee k-am  beautiful  CL.2SG  do.PRS.PFV-1SG  
'I make you beautiful.' (Roberts, 2000: 37-39)

The [k] has been dropped in (73) and has not been dropped in (75), but still there was no room for the 2SG clitic to be placed inside the imperfective compound verb. In present perfective compound verb, (77) and (78) the 2SG clitic has been shifted in spite of the auxiliary [k] that has not been dropped. From the discussion, it can be concluded that the possibility for clitic placement inside the compound verb is determined by aspect.
5.6. Bilingual Compound Verbs in the code switching literature

Romaine (1995) has described the BCVs as a feature of bilingual verities. In various works on code switching, BCVs are defined as a conjunction of a light or helping verb, usually translated as “do” or “make”, and a lexical item which gives the semantic content of the construction. The BCVs constructions have several names in different literature. Muysken (2000) has used the cover term ‘helping verb’, Myers-Scotton (2002: 134) refers to them as the ‘do-construction’, while Wohlgemuth (2009: 104) refers to them as the Light Verb Strategy and argues that it is “preferable to use the broader term Light Verb Strategy as the cover term for all of these constructions (bilingual VP)”.

Muysken (2000) studied different corpora. In the chapter on “Bilingual Verbs”, Muysken (2000) has declared code mixing as innovative, productive and leading to a structure not present in either of the languages in contact. He has proposed the following four main types in the bilingual corpus (Muysken, 2000: 184):

| The verb is inserted into a position corresponding to a native verb, in adapted form or not |
| The new verb is adjoined to a helping verb |
| The new verb is a nominalized complement to a causative helping verb in a compound |
| The new verb is an infinitive and the complement of a native auxiliary |

Muysken (2000) argues in (79) that there is a lexical structure of the type (V kare) where the helping verb such as ‘make’ or ‘do’ is in conjunction with the left-most alien verb, which carries the semantic meanings. This is common in the Indic languages, such as in the following examples of Sranan/Dutch/English mixed verbs in Suriname Hindustani, Sarnami:

(79) a. ontí kare to hunt’ Sranan
b. train kre ‘to train’ English
c. bewijis kare ‘to prove’ Dutch (Muysken, 2000: 185)

In (79), the process is completely productive and does not entail phonological or semantic integration into the host language (Muysken, 2000). The same pattern of insertion is observed in Tamil/English data (Annamalai, 1971):

(80) avan enne confuse pannitten
he me confuse did
‘He confused me.’ (Annamalai, 1971; cited in Muysken, 2000: 215)

In BCVs, the bilinguals exploit the resources of the languages in contact while code mixing. In different languages, different morphological strategies have been developed in the bilingual VP.

In some languages, a bare verb is conjugated with a light verb or helping verb. In other languages, stems are affixed with the native markers. In the third types of code mixing contact,
stems are adapted as in the following examples, where the French verbs can only be introduced into Dutch language when the stem is affixed with -er:

(81) bless-er-en 'hurt' (Fr blesser)
condemn-er-en 'condemn' (Fr condemner)
concurrence-er-en 'complete' (Fr concurrence)

(Treffers-Daller, 1994; cited in Muysken, 2000: 191)

In many language contact settings, the languages that have a verbalizer can easily native an embedded language noun or verb into the matrix of the host language.

In Siegel (1987) English/Hindi code mixing data, it was noted that the left-most lexical item in the bilingual VP is not a verb stem. In (82), the mixing involves very common words which refer to daily activities and do not represent cultural borrowings.

(82) sain kar sign
wait kar wait
marit kar marry (in a civil ceremony)

(English/Hindi: Siegel, 1987; cited in Muysken, 2000: 208)

In the Lotfabbadi (2002) Swedish/Persian bilingual data, a Swedish bare infinitive verb occurs to the left of the Persian auxiliary, whereas in Persian counterpart, this element is noun. The same structural pattern in monolingual Panjabi VP is also observed (Romaine, 1995). The dominant pattern of BCVS is the Swedish bare infinitive verbs integrated by the Persian auxiliary kardan ‘do’.

(83) man-o besviken kard –i
me-obl disappointed did-2sg
‘You made me disappointed.’ (Lotfabbadi, 2002: 111)

5.7. English-Pashto Bilingual Compound verbs

The total set of mixed BCVS found in the present data is 43, where the bilingual pattern can be fitted into five categories of verbs: participles, gerunds, phrasal verbs and nouns as shown in Table 6. The most frequent combination of the English verb is with the transitive light verb kaw ‘do’ or ‘make’, and the intransitive light verb keg ‘become’. In entire data set, two examples of the English past participle and English gerund are observed. Only a single example of the intransitive light verb ra- zam ‘com-ing’ is observed. In other languages like Hindi and Urdu other elements such as daina ‘gave’, gya ‘go’, lia ‘take’ are the light verbs (Akh tar, 2003; Butt, 1995). The table also shows 2 examples of the verb particle (brought up) in the English-Pashto data.
Strategies for incorporating nouns and verbs in code-mixing: the case of Pashto-English bilingual speech

<table>
<thead>
<tr>
<th>English</th>
<th>kaw</th>
<th>keg</th>
<th>copula</th>
<th>others</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical elements</td>
<td>(do/make)</td>
<td>(become)</td>
<td>(be)</td>
<td>ra-zam</td>
</tr>
<tr>
<td>Verb</td>
<td>22</td>
<td>9</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Noun</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participle</td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Gerund</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Verb particle</td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

Table 6: Distribution of the Embedded Lexical items in the BCVS patterns in English-Pashto CM

It is clear that the most dominant pattern of BCVS is the combination of two light verbs, transitive kaw ‘do’ or ‘make’ and intransitive keg ‘become’, and the English verb. In Punjabi compound monolingual data, the left-most element cannot be a verb but can only be a noun or adjective (Muysken, 2000). In Pashto monolingual compound, the left-most element can be a nominal element and verbal element in the intransitive keg ‘become’ construction. In the transitive construction, most of the compound or complex predicate aree structured around the nominal element and the transitive auxiliary kawal ‘do’ and occasionally other verbs (Babrakzai, 1999). The monolingual data discussed in Babrakzai (1999) shows the same pattern of combination of N+V and Adj+V.

In the English-Pashto bilingual data, there are three patterns of BCV, following the monolingual compound verb and simple verb. In (84), the copula represents the helping verb ‘be’ in Pashto. The light verb in the proposed structure represents the two auxiliary verbs transitive kaw ‘do’ or ‘make’, intransitive keg ‘become’, and the light verb razam ‘coming’. The light verb razam is marked for tense, person, and agreement. The English elements in the bilingual VP must be in conjunction with the Pashto helping verbs where the helping verb is used as cover term for Pashto cop ‘be’ light verb kaw, keg and razam.

The proposed BCVS structure equivalents to Pashto Monolingual structure are as follows:

(84) Monolingual VP
   a. [native verb + cop]
   b. [native verb + light verb]
   c. [native verb]

   Bilingual VP
   a. [alien verb + cop]
   b. [alien verb + light verb]
   c. [alien verb + light verb]

Thus, the above three structural patterns have been observed in the English-Pashto CM data. The most striking pattern of replacement in the monolingual VP is the last one (84c). Here the Pashto native verb is a single verb inflected for tense, aspect and agreement marker and replaced by [V+ light verb] structure in the bilingual VP. The most dominant pattern of change is (84b), where the [native verb + light verb] is replaced by the [alien verb + light verb (kaw, keg)]. All the above three bilingual VP patterns are productive, and not a single example entails phonological or semantic integration into Pashto morphosyntax.
In English-Pashto bilingual data, not a single verb like ‘go’, ‘come’, and ‘keep’ is observed. The only possible explanation is that in Pashto, its counterpart is treated as a light verb and the proposed pattern of BCV is [English element + light verb or copula]. It is not possible that two light verbs function in the same BCV structure. We now provide some illustrations.

**English verbs with the light verb kaw.** In (85), repeated from (2), the BCV structure is [play kaw-i], where the alien verb ‘play’ is conjugated with the aspect marked light verb kaw-i. The English verb ‘play’ can be replaced by its Pashto counterpart ada. It is a typical example of replacing the Pashto native verb in monolingual compound verb construction. In the entire English-Pashto bilingual data set, not a single case of English verbal integration can be found. This is a very productive process, as noted earlier in this paper. On the other hand, the Pashto verb ada is very specific in taking the noun ‘role’ in its argument. The use of English verb ‘play’ is really a productive entry into Pashto. The English elements into Pashto can only be introduced when there is a helping verb in the bilingual VP.

(85). Media [negative role] hu [play kaw-i]
media negative role indeed play do.PRS.IPFV-3SG

‘Media indeed play a negative role.’

In (86), the helping verb kr-o in the BCV is multimorphemic and is marked for tense, aspect and agreement. The English verb ‘divide’ appears in conjunction with the light verb kaw-o. The pattern of the BCV is equivalent to the Pashto monolingual structure. The English verb has replaced the Pashto counterpart taqsem ‘divide’ and is embedded in the native verb slot. The issue of borrowing is really complicated with the English verb in Pashto. As the EL verb ‘play’ in (20) does not entail any morphological and phonological integration in Pashto. One of the reasons with borrowing in Pashto is the multimorphemic nature of Pashto helping verbs which have reduced the possibility of integration of the English verb in Pashto structure.

(86) Che [[dwa groups ke] mong [divide kr-o]]
COMP two groups in.OBL 1PL.NOM divide do.PRS.IPFV-1PL

‘That we have divided in two groups.’

In (87), repeated from (1), the embedded verb ‘start’ appears in conjunction with the Pashto transitive light verb kaw-o. It is a clear example of replacement of the Pashto monolingual counterpart verb shuro. The process in BCVS is completely productive as the English alien verb ‘start’ does not imply phonological or semantic integration into Pashto morphosyntax.

(87) che [[də de format] ba sanga [start kaw-o]]
COMP GEN DM.PRX format CL.FUT how start do.PRS.IPFV-1PL

‘You should say to the dear audience that how we will start this format.’

The BCV construction is innovative, especially when its Pashto counterpart is a simple word marked by a transitive/intransitive marker. In (88), the counterpart to the bilingual VP [use kaw-i] in Pashto is istimal- aw-i. In the monolingual VP the stem istemal is suffixed by the transitive marker aw to mark the tense, aspect and subject-verb agreement with the pronomin
marker –i. In Pashto, the stem takes a direct suffix marker for tense and aspect, but when the same imperfective structure is used in a BCV, it changes the entire structure. In (88), the BCV [use kaw –i] functions as two parts, but its Pashto counterpart istimal-aw-i function as a single part.

(88) da hr sook [use kaw-i]
DM.PRX each person use do.PRS.IPfv.3SG
‘This is used by every person.’

Nouns. The embedded verb ‘target’ appears in conjunction with the Pashto transitive light verb kaw-u as shown in (89), repeated from (7). The Pashto light verb kaw-u is a multi-morphemic word and is used as marker of tense, aspect, and agreement. It is a clear example of replacement of the Pashto monolingual counterpart verb pәnaha ‘at target’ but in Pashto, it is a compound verb combination of the preposition pә ‘at’ and the noun naha ‘target’. The process in the BCV is completely productive as the English alien verb ‘target’ does not entail phonological or semantic integration into Pashto morphosyntax.

(89) pә de ke monga youth [target kaw-u]
LOC DM.PRX OBL PRN.1PL youth target do.PRS.IPfv-1PL
‘In this we target the youth.’

Gerunds. In the BCV example (90), the English gerund ‘teaching’ appears in conjunction with the light verb kaw-o. The English gerund ‘teaching’ is a good example of a nominalized verb. The English gerund in Pashto is always used as nominalizer and functions in the start of a sentence. The English nominalized verb contributes to the core semantics of the construction.

The role of the light verb kaw-o is to mark the tense aspect and agreement of the construction. In the BCV, the English gerund does not entail phonological and morphological integration in Pashto. The present BCV pattern is highly productive and innovative in Pashto.

(90) Mong [[hapal teaching] kaw-u]
PRN.1PL own teaching do.PRS.IPfv-1PL
‘We do our own teaching.’

English verbs with the intransitive light verb keg. In (91), the light verb shw-a ‘become’ is marked for tense and aspect, and further inflected by the 3rd person pronominal marker-a for subject-verb agreement. In the BCV, the left-most lexical element, the bare infinitive verb ‘release’, carries the meaning. The insertion of the English bare infinitive is an innovation because its counterpart ragl-a ‘came’ is an intransitive verb having a stem inflected for tense, aspect and person. Semantically the two stems ‘release’ and ragl-a are by no means equivalent but here the English verb ‘release’ carries the same meanings in the following construct.

(91) Halaq wayi che [[cassette] [release shw-a]]
people say.PRS.IPfv COMP cassette release become.PRS.IPfv-3SG
‘People say that the cassette has been released.’
In (92), the BCV structure [alien verb + light verb], the embedded verb ‘use’ is used as a bare infinitive form without any direct inflection or affixation. The other remarkable feature beside the transitivity is that the verb derived from adjective stem is marked perfective by the stem and an auxiliary verb. On the other hand, the imperfective aspect is marked with a stem suffixed by a nominal marker for agreement (Babrakzai, 1999).

(92) da computer [[all over the world] [use keg-i]]
DM.PRX computer all over the world use become.PRS.PFV-3SG
‘The computer is used all over the world.’

In English-Urdu bilingual verbal complexes, the lexical item is combined with a Urdu light verb such as use huta hain and use krtha hain. In English-Pashto bilingual verbal complexes this process of code mixing is highly innovative as it produces an entirely new structure.

(93) Pashto English-Pashto (BCV) Urdu English-Urdu (BCV)
Istimalegi use keg –i Istemal huta hain use huta hain
Istamalwai use kaw –i Istemal krtha hain use krta hain

The BCV [collapse sh-i] in (94) has two parts where the English verb is in conjunction with the Pashto intransitive light verb sh-i ‘become’. The Pashto light verb sh-i is multimorphemic and is marked for tense, aspect and agreement.

(94) Nu [[pura system] [collapse sh-i]]
thencomplete system collapse become.PRS.PFV-3SG
‘Then the complete system collapsed.’

Nouns. In (95), the English noun ‘balance’ is used as verb in the bilingual verbal complex with the Pashto intransitive light verb sh-o. The English verb ‘balance’ is compatible with the argument NP ‘answer’. On the other hand, its Pashto counterpart, barab or masawi would not be not compatible with ‘answer’. The English noun ‘balance’ is an example of a nominalized verb in the BCV.

(95) staso answer [balance sh-o] 
PSS.2PL answer balance become.PST.PFV-2PL
‘Your answer had become balanced.’

Phrasal verbs. In (96), the insertion of the English phrasal verb ‘brought up’ in the light verb construction is very productive and innovative. In Pashto, the counter part of ‘brought up’ is pervarish but in Pashto, there is no such structure as a phrasal verb and that is why it is an innovation into Pashto.
(96) Che da aghwi hə [brought up wa-sh-i]
COMP GEN DM.PL.DST good brought up become.PRS.PFV-3PL
'That they should have best brought up.'

**English lexical elements with a Pashto copula.** In (97), repeated from (40), the English embedded verb ‘apply’ is conjoined with the Pashto copula wi.

(97) da Islam [toul golden rule] ba halta [apply wi]
GEN Islam all golden rule CL.FUT DM.DST apply COP.PRS.IPFV.3PL
'All golden rules of Islam will be applied there.'

**English gerunds with the Pashto copula.** In the Pashto speech community, the use of the English gerund is very common, but in the present data, only few examples were observed, with the Pashto transitive light verb kaw-o and the copula. In the BCV example (98), the English gerund ‘missing’ appears in conjunction with the Pashto imperfective past tense copula wo. The Pashto copula by no means is equal with the English past tense ‘was’ because Pashto copula has more a multi-morphemic function than the English ‘was’. The English gerund in Pashto has many more functions as it can be used as noun, verb, and as adjective. The English gerund ‘missing’ is a good example of an adjective verb. This type of construction is highly innovative and productive in Pashto structure. The Pashto monolingual counterpart of the bilingual VP is shamil na wo. In the Pashto monolingual equivalent, there is a negation marker in the VP slot, while in bilingual VP, the English adjective has turned into a verb.

(98) che Ajab Khan [da yaw dwa-o program-un-o na]
COMP Ajab Khan GEN one.M.SG two-OBL program-PL-OBL from
[missing wo] missing COP.PST.IPFV.3SG
'That Ajab Khan was missing from one and two programs.'

**Participles.** In (99), the English past participle appears in conjunction with the Pashto present imperfective copula yu ‘be’. In Pashto, the use of the participle is very common but in the present data, only two instances have been observed. The English participle is a strong example of code switching, and the process is innovative and productive, not showing any phonological and semantic integration into the Pashto morphosyntactic frame.

(99) monga agha halaq-o sra [connected yu]
PRN.1PL DM.DST people-OBL with connected COP.PRS.IPFV.1PL
'We are connected with those people.'

**English lexical elements with the light verb ra-zam.** In entire data set, only such two examples are observed. In Pashto syntax, the light verb ra- zam is not a typical type of auxiliary verb like the transitive auxiliary kaw ‘do/make’ or intransitive keg ‘become’, but it is more like the Urdu light verbs jaa, ‘go’, gya ‘went’, etc. In the BCV, the participle ‘prepared’ is used as a past stem conjugated with Pashto light verb ra- zam. The counterpart of the English participle
‘prepared’ in Pashto is the adjective *tyar* ‘ready’. The Pashto verb *tyar* is used in the verbal slot next to the light verbs in the bare from and does not take a tense or aspectual marker. The English verb ‘prepared’ in the Pashto structure is an innovation as it deviates from the Pashto monolingual VP structure.

(100) za [class ta] [prepared ra-zam]

P R N
PRN.1SG class to prepare 1SG-come.PRS.PFV.1SG

‘I come to class prepared.’

6. Comparing nouns and verbs

The study shows that in language contact phenomena, code mixing is a driving force of indigenization of the donor words in the host language. The results and discussion suggest that there are no exact rules to differentiate between borrowing and loanwords. In uninflected embedded words, it is difficult to determine that the word as a code mixing, borrowing or loanword. In the present data, the two types of borrowing, cultural and core loanword elements are very frequent.

However, here there is a potential distinction between the nouns and the verbs, as to their mode of incorporation into KP bilingual speech. Nouns and verbs differ, at first glance, in two ways, morpho-phonologically and grammatically.

*Morpho-phonology.* The study shows that majority of the noun loanwords may be morphologically and phonologically integrated into Pashto by plural, reduplication, and oblique marking.

In contrast, Pashto does not mark the embedded elements inside the *BVC*. The *BVCs* in the Pashto-English data show no evidence of Pashto verb morphology directly attached to English verbs.

*Grammar.* English nouns are directly inserted into positions reserved for Pashto nouns according to the rules of the Pashto matrix grammar. There are no special grammatical adaptations or special constructions.

In contrast, the *BVC* is the most striking and innovative pattern of insertion of verbs in English-Pashto bilingual speech. The bilingual pattern can involve the five categories of verbs, participles, gerunds, phrasal verbs and nouns. The most frequent conjugation of the English verb is with the transitive light verb *kaw* ‘do or make’ and the intransitive light verb *keg* ‘become’. The light verbs (e.g. *kaw, keg*) are suffixed with Pashto aspect, tense, and agreement markers. The data is compatible with the Myers-Scotton ML model (1993, 2002), with the proviso that in Pashto itself the construction does not involve two verbs.

However, there is also a significant parallel between nouns and verbs. The English plural marking – *s* is sometimes lacking and sometimes redundantly present, where English itself would not have plural. It thus seems to be a primarily lexical and not grammatical feature of mixed Pashto-English bilingual speech. Apart from that, there are no other functional nominal
elements from English present in the utterances cited, except in a few fixed expressions. Demonstratives, possessives, adpositions, indefinite articles, and quantifiers are consistently from Pashto. In the same way, there are no functional elements from the verbal domain in Pashto.

The **BCV** constitute a privileged construction, in which English verbs can be productively incorporated into Pashto. Nonetheless, these verbs do not occupy the position they would need to productively be affixed by Pashto morphological markers. This remains limited to the light verb.

**List of abbreviations**

<table>
<thead>
<tr>
<th>1, 2, 3</th>
<th>ADV</th>
<th>COMP</th>
<th>COP</th>
<th>DM</th>
<th>DST</th>
<th>ERG</th>
</tr>
</thead>
<tbody>
<tr>
<td>first, etc. person</td>
<td>adverb</td>
<td>complementizer</td>
<td>copula</td>
<td>demonstrative</td>
<td>distal</td>
<td>ergative</td>
</tr>
<tr>
<td><strong>f</strong></td>
<td><strong>ADV</strong></td>
<td><strong>COMP</strong></td>
<td><strong>COP</strong></td>
<td><strong>DM</strong></td>
<td><strong>DST</strong></td>
<td><strong>ERG</strong></td>
</tr>
<tr>
<td>feminine</td>
<td>future</td>
<td>light verb construction</td>
<td>masculine</td>
<td>object</td>
<td>oblique</td>
<td></td>
</tr>
<tr>
<td><strong>PPFV</strong></td>
<td><strong>FUT</strong></td>
<td><strong>IPFV</strong></td>
<td><strong>M</strong></td>
<td><strong>OBJ</strong></td>
<td><strong>OBL</strong></td>
<td></td>
</tr>
<tr>
<td>perfective</td>
<td>future</td>
<td>imperfective</td>
<td>masculine</td>
<td>object</td>
<td>oblique</td>
<td></td>
</tr>
<tr>
<td><strong>PL</strong></td>
<td><strong>GEN</strong></td>
<td><strong>PRN</strong></td>
<td><strong>PRS</strong></td>
<td><strong>PRX</strong></td>
<td><strong>PSG</strong></td>
<td></td>
</tr>
<tr>
<td>plural</td>
<td>genitive</td>
<td>pronoun</td>
<td>present</td>
<td>proximate</td>
<td>singular</td>
<td></td>
</tr>
<tr>
<td><strong>POSS</strong></td>
<td><strong>IPFV</strong></td>
<td><strong>PRX</strong></td>
<td><strong>PSG</strong></td>
<td><strong>RECP</strong></td>
<td><strong>REDUP</strong></td>
<td></td>
</tr>
<tr>
<td>possessive</td>
<td>imperfective</td>
<td>proximate</td>
<td>singular</td>
<td>reciprocal</td>
<td>reduplication</td>
<td></td>
</tr>
<tr>
<td><strong>PRS</strong></td>
<td><strong>LOC</strong></td>
<td><strong>PST</strong></td>
<td><strong>PG</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>present</td>
<td>locative</td>
<td>past</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>KP</strong></td>
<td><strong>IPFV</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Khyber Pakhtunkhwa</td>
<td>imperfective</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>LVC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>light verb construction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>M</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>masculine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OBJ</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>object</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OBL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>oblique</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
References


Strategies for incorporating nouns and verbs in code-mixing: the case of Pashto-English bilingual speech


