The prediction of morphemes distribution at surface and abstract level: the case of Embedded Language nonfinite verbs in Pashto-English Bilingual data

Arshad Ali Khan[1]
Nadeem Haider Bukhari[2]
Amina Khalid[3]
Syed Nasir Abbass[4]

Abstract

The present research explores the election of Embedded Language (EL) nonfinite verbs at surface level with respect to its activation at abstract level. It also investigates the argument that non-finite verbs are elected at lexical-conceptual level. The framework used in the present research is the 4-M model (Myers-Scotton, 2002) of morpheme classification for the prediction of distribution of the EL non-finite verbs at surface level in relation to abstract structure level. This research tries to answer the question: How the EL nonfinite verbs are elected at the abstract level in relation to surface level data? Approximately 8.49 hours bilingual data was collected from a sample of three groups is transcribed and only EL language verbs (nonfinite) are focused and discussed in the present research. According to the classification of 4-M model nonfinite verbs as content morphemes are salient only at lexical-conceptual level to its lemma specification of the speaker's intended semantic-pragmatic meanings. The nature of the EL nonfinite verbs also suggests its activation as fast and easy with low cost at the abstract level because its integration is made possible by the ML inflection morphemes for tense, aspect and outsider morpheme for subject-verb agreement. On the other hand the EL finite verbs are hard to find in the present data also suggest its complicated morphosyntactic nature. This is the first attempt in Pashto-English bilingual data that the 4-M model of morpheme classification is used to predict a relationship of the empirical data and its election at abstract level. The present study suggest that the model of morpheme classification could also be extended to the morpheme activation (L1 and L2) in second language acquisition.

Keywords: Embedded Language; nonfinite verbs; Abstract structure; 4-M model, Lexical-conceptual level

1. Introduction

The major goal of the present research work is to unfold the nature of entries of embedded language (EL) non-finite verbs in Pashto-English codeswitching data in relation to its representation at abstract level. In the corpus of CS the frequency of the different morphemes with a special focus on EL nonfinite verbs depend on the type of activation and election of the morphemes in the conceptual level or at the formulator

[1] Assistant Professor, Department of Linguistics and Communications, UMT, Lahore
[2] Professor, Department of English, University of AJ&K, Muzaffarabad
[3] Lecturer, Department of Linguistics and Communications, UMT, Lahore
[4] Associate Professor, LUWAMS, Lasbela

where the morphemes are structurally assigned (Myers-Scotton, 2017). For different languages the election of the morphemes at abstract level might be different due to the grammatical structure of a particular language. The study of verb in bilingual data is also useful as the verb is not only an integral part of argument structure but via morphological marking it is selected for tense, aspect and agreement. This research explores the claim of Myers-Scotton (2002, 2017) about the roles of 4-M model of morpheme classification and the link between the surface morphemes with the morphemes elected at abstract level, taking account of the abstract features that construct lexical elements and their morphological memberships.

The present study adapt a new proposal as presented in Myers-Scotton (2017) that the level at which morpheme types are "elected" in an abstract model of language production is a critical factor in predicting morpheme distribution across languages in bilingual data. The study explores the claim, that the nonfinite verbs do not project information about phi features (tense, aspect, agreement and modality) that is included in the abstract structure of finite verbs (Myers-Scotton, 2013). The very focus of the research in hand is only on the distribution and activation of EL nonfinite verbs in ML. In bilingual data the most prevalent EL switching in ML is noun and nonfinite verbs (Myers-Scotton, 2016, Khan, 2014, Backus, 1996). This study focuses on the nature of the integration of EL nonfinite verbs in light verb constructions as VP slots are responsible for selection of predicate-arguments structure, tense, aspect and subject-verb agreement. In order to sort out this relationship between actual data and it's realization at abstract level, this research tries to answer the question:

How the EL nonfinite verbs are elected at the abstract level in relation to surface level data?

The objectives of the study is to explore:

1. the distribution of EL nonfinite verbs at surface level and abstract level
2. the election of EL non-finite verbs and not that of EL finite verbs
3. the realization of EL non-finite verbs at lexical-conceptual level

In order to know this relationship between surface level data and abstract level data, 4-M model for the morpheme classification play a key role as the classification of morphemes is largely based on their surface distribution (Myers-Scotton, 2016). The following section highlights different types of morphemes.

This article is organized in different sections. Section (1.2) is about theoretical underpinning. Section (2) provides a short review of literature to EL nonfinite verbs in ML. The next section (3) is a review of the studies about the division of first and second language in cognition. It also highlights the abstract features and its activation.

1.1 The application of 4-M model

The theoretical framework used to discuss the present data is the 4-M model (Myers-Scotton, 2002). Before going to explain the 4-M model, a brief description to Matrix Language Frame (MLF) is crucial as it capture the asymmetry between the two languages in the bilingual speech. The Matrix Language Frame (MLF) model is based on content and system morphemes and focuses on the basic grammatical structure of a bilingual

clause (Myers-Scotton, 2016). In MLF one language function as Matrix Language (ML) and other as Embedded Language (EL). The EL contributes mostly the content elements "nouns" and "verbs" while the ML provides the grammatical frames. The ML provides a uniform structure to the bilingual clause and bilingual construction within the clause.

The 4-M model provides more explanation and classification to the morphemes types as in figure (1).

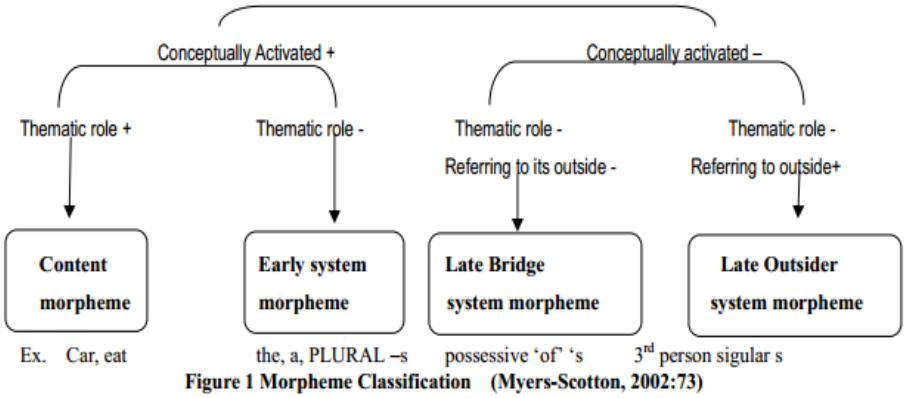

Figure 1 Morpheme Classification (Myers-Scotton, 2002:73)

The major focus of 4-M model is the classification of morphemes at abstract level, between conceptually activated, and structurally assigned morphemes (Myers-Scotton, 2000). Actually it differentiate between different types of morphemes election at abstract level such as at mental lexicon only content and early system morphemes (SMs) are elected, but late SMs do not become salient until the level of the formulator (Myers-Scotton, 2002a, p. 78).

1.2 Conceptually activated Morphemes: Content Morphemes and Early SMs

According to the 4-M model, these types of system morphemes (content) are activated in mental lexicon because they are activated with the speakers' pre-linguistic intensions and convey most of the semantic and pragmatic meaning (Myers-Scotton, 2017). Nouns are the most prevalent content morphemes which receive the thematic role of a clause and verbs. Due to this level of activation they are considered as conceptually activated morphemes. The other type of conceptually activated morphemes are early system morphemes. In English language theses morphemes are definite article (the), and the indefinite articles (a, an), determiners and plural marking (Myers-Scotton, 2002).

In example (1), that clause of Pashto as ML whereas the EL noun 'youth' is inserted in the subject position and at the object position the EL noun 'study' is inserted. Pashto provides the morphosyntactic frame and the embedded elements in italics follow Pashto rules. The morpheme *-i* is suffixed with the transitive auxiliary *kaw* for the 3rd person singular subject agreement with the verb *kaw-i*.

1. Pashto-English (Author, 2014, p. 107)
   che youth hpala study na kaw -i

…that.CP youth own study not do.PRS.IPFV -3SG
'… that youth is not doing their studies.'

Apart from nouns and verbs other lexical categories are also function as content morphemes and they are assigned thematic role such as in case of prepositions. In (2), *before* assigns a temporal thematic role to "tomorrow evening" (Myers-Scotton, 2017, p. 4).

2. Swahili–English (Myers-Scotton, 1993a, p. 124)
   Labda, […] **bring it at my home**. U-let-e **before** kesho jioni.
   Perhaps well, bring it to my home. 2s-bring-subjunct before tomorrow evening
   "Perhaps […] bring it at my home. You should bring it before tomorrow evening."

1.3 Late System Morphemes

Late system morphemes (LSM) are classified into two types the bridge and outsiders system morphemes. They are responsible for the main architecture of the clause and is responsible for the relationship of VP and NPs within the clause. They are elected later at the formulator and they are structurally-assigned.

To develop larger constituents within the phrases the bridge system morphemes play a crucial role as it keeps these constituents in well-formed in the relevant language. In the possessive construction "of" is the best example of the Bridge system morphemes. It joins the phrases, and sometimes the clauses as in the case of 'that' type (Myers-Scotton, 2009).

Outsider late system morphemes determines agreement morphology and tries to retain the co-indexing such as between verb and its arguments. Some of the pronominal clitics co-indexing subject and object is an example of outsider system morphemes such as in Spanish and other Roman languages (Myers-Scotton, 2016). Other aspect of outsider system morpheme is case marking as it clearly indicates predicate-argument structure. Example (3) shows the integration of Dutch NP in the grammatical frame of bilingual clause inflected with Turkish instrumental suffix.

3. Turkish-Dutch (Backus, 1996, P. 105)
   ondan sonar **lauw water**-<u>nan</u> yıkayınca …
   then after lukewarm water-with wash.while …
   "and then, while you're washing [it] with lukewarm water …"

2. EL nonfinite Verbs in the ML frames

The insertion of the EL verbs in ML frame are all nonfinite forms of the verb (Myers-Scotton, 2017, p.13, Backus, 1996, p.213). The distribution of EL nonfinite verbs in different CS data could be classified into three categories: (1) in the light verb construction it function as complement to the light verb in ML especially in "do" construction; (2) in the second type the EL nonfinite verb is inserted with the help of ML verbalizer element and it could be a clitic or an affix; (3) in third type the nonfinite verbs are fully inflected in ML frames.

In various works (Backus, 1996, Myers-Scotton, 2013) on code switching, LVCs are defined as a conjugation of a light or helping verb, usually translated as "do" or "make", and a lexical item which gives the semantic content of the construction. "do" construction is the most frequent of all EL nonfinite verb construction and is called nativizing auxiliaries (Backus, 1996, p. 213). The "do" construction is responsible to provide predicate-argument of the clause as it receives all the necessary inflection of the ML (Myers-Scotton, 2017, Backus, 1996, p. 212). The complement of this verb is EL nonfinite verb and especially infinitive form that carries the intended meaning of the speaker's semantic- pragmatic intention. In the Turkish-Dutch bilingual corpus the Dutch infinitive occur very frequently (Myers-Scotton, 2016, p. 14) as in (4).

4. *Turkish–Dutch* (Backus, 1996, p. 238)
   Ja maar toch, millet **kijken** yapıyor
   yeah but still people **watch**-INF do-PROG-3S
   "Yeah, but still, everybody is watching you."

In Swedish/Persian bilingual data (Lotfabbadi, 2002) the dominant pattern of LVCs is the Swedish bare infinitive verbs integrated by the Persian auxiliary *Kardan* 'do'. The 'do' construction receives all the inflections morphology from the ML language and is responsible for predicate-argument structure. The complement of this bilingual verb phrase is EL *besviken* 'disappointed' which carries the semantics of the speaker's intended message.

5. Swedish/Persian (Lotfabbadi 2002, p. 111)
   man-o *b*esviken kard –i
   me-ra disappointed did-2SG
   You made me disappointed.

Muysken (2000) has provided examples from Sranan/Dutch/English mixed verbs with a structure of EL lexical items and ML light verb *kare*. The ML *kare* "do' is inflected for tense, aspect at formulator for the predicate-argument structure. This is common in the Indic languages, such as in the followings example (6):

6. Sranan-Dutch (cited in Muysken, 2000, p 185)

a.	*onti* kare	'to hunt'	Sranan
b.	*train* kre	'to train'	English
c.	*bewijis* kare	'to prove'	Dutch

In example (7), EL nonfinite verb (English gerund) 'teaching' appears in conjugation with the light verb *kaw* "do". The English gerund teaching is a good example of nominalized verb. This integration is made possible by the Pashto light verb which is marked for tense, aspect and subject verb agreement.

7. Pashto-English (Khan, 2014, 2018)
   Mong hapal [V teaching kaw -o]
   1PL own teaching do.PRS.IPFV -IPL
   'We do our own teaching.'

In some languages verbs are integrated with the nativizing affixes such as in German, loan verbs are incorporated with *ier-en* (verbalizer-INF), as in case of "to arrange" become *arrangieren* (Myers-Scotton, 2017, p. 14). In example (10) the English verb 'pretend' is inflicted for tense and aspect by the ML verbal morphology *–at-* (Halmari, 1997, p. 111).

8. Finish-English (Halmari, 1997; cited in Myers-Scotton, 2016, p. 14).
   Me **pretend**at-in olevan elefantteja
   We pretend.vm-pass/past. to.be elephants
   "We pretended to be elephants"

Backus (1996. p) has provided a detailed description of EL non-finite verbs integration in ML frame. Backus (1996) argued that in some languages there are special auxiliaries which nativize the EL verbs and some auxiliaries simply insert ML frame. He focused on two major patterns; Pattern 1: Auxiliaries and Pattern 2: Morphology

Pattern 1 auxiliaries has already been discussed as "do/ make" construction and his second type of auxiliary pattern 2 is also discussed in example 8. The following examples show the morphological integration of EL non-finite verbs in ML frame.

In example (10a – 10b) a French stem of infinitive verb is preceded by Lingala prefixes and in turn the French stem of infinitive precedes the tense, aspect and suffices. The following examples show a very different process of EL non-finite verb incorporation in ML.

9. (Bokamba, 1998; cited in Backus, 19996, p. 223)
   (a) Ko- ***compared*** *–re* ("to understand)
   (b) Ko- ***compared*** *–re* ("to ask")

On the integration of the non-finite verbs in ML frame all the previous studies (Backus, 1996; Muysken, 2000; Myers-Scotton, 1993, 2017 & Khan, 2014, 2015) are agree that only non-finite verbs are very common type and EL finite verbs are rare in the CS literature (Backus, 1996).

In pattern (2) the morphological integration of the EL nonfinite verbs has been discussed in Backus (1996). While incorporating the foreign (French) stems in German a derivative suffix *–(i)er-* is added such as *kummuniseera* in Swedish, and *cummuniceren* in Dutch and *kommunizieren* in German.

In Spanish-English (cited in Backus, 1996, p. 220) most of the English verbs (stems) have been nativized with the Spanish verbalizer *–ear* such as in example (10):

10. (cited in Backus, 1996; p. 220)
    watch *–ear*
    work – ear
    mop –ear
    swim -ear

The Michif language has nativized the Europeans verbs stems with a verbalizer *–er* such as in example (11) the English stems are incorporated with the French verbalizer *–er.* The stem and the verbalizer are both from European origin such as the article *le* is also from French.

11. (cited in Backus, 1996, p. 219)

a. ka- *le*- ***charge***-*er*-t ("that it charges")
b. ni-ki-*le*-**gag**-*er*-n ("I bet")
c. *le*-***celebrate***-*er*-*w*-ak ("to celebrate")

2.1 Classification of the Abstract Features and their Levels of Election

The researchers interested in the language representation in the mind is heavily influenced by Garrett's (1975) claim that a word's syntactic and morphophonological specifications are stored separately from each other in the mind.

The focus of most researchers (Myers-Scotton, 2017, De Bot, 1992, & Wei, 2000) in the language production model is on the activation of L1 and L2 lexical items and it's representation in the abstract level. While using two languages in a single clause could be leveled with its representation and activation in a bilingual mental lexicon. The priming techniques is used to know the mental associations between words of two languages and the possible channels of activation; whereas activation shows the availability of representations at different levels of processing (Costa, 2005: 309). In the light of the studies (Francis (2005), Jiang (2000) & Kroll and Tokowicz (2005) conducted on the representation of the language at abstract level, it could be concluded that mental connections can exist between words from different languages at each level of representation (i.e. lexemes, lemmas and concepts) and also across levels of representation (e.g. a lexeme from Language A can be mentally associated with a lemma from Language B). Wei (2002) assumes that each lemma is marked for a specific language and each lemmas in the bilingual mental lexicon activate and supports the realization of an actual lexeme and specific sets of morphosyntactic procedures in the formulator.

The distinction between ML and EL in respect to the grammatical function is vital as the morphosyntactic procedures is not equally govern by the two languages. One language has greater input in the resulting string in bilingual speech production (Myers-Scotton & Jake, 1995; Wei, 2001). In the speech production of a bilingual speaker the language-specific semantic/pragmatic feature bundles are activated at conceptual level as preverbal message/intention. In the mental lexicon this preverbal message at lemma level is mapped with lemma. The activated lemma for grammatical encoding and morphological encoding send information to the formulator at functional and positional level (Wei, 2002, P.694).

Wei (2002) argues that the switch between the EL and ML is possible if the lemma of the two languages (EL & ML) are in congruence and inappropriate congruence at structure restrict the choices of the switch and that's why EL fixed expression as in Muysken (2000) and the switch as EL Islands in Myers-Scotton (1993) do not follow ML morpheme order but the EL morpheme order.

Chinese and English shares the same word order V-O, and the Chinese-English bilinguals can easily switch the EL verbs, verb phrases, EL noun and noun phrases in the same slots because of the congruence in EL and ML structure (Wei, 2002). In example (12) the EL infinitive verb "drive" is used in the slot of object to the ML main verb "xue" *learn*. The Chinese finite verb "xue" is used the same order as in English.

12. Chinese-English (Wei, 2002, p. 701)
    ta gang dao, ta dei xue *drive.*
    he just arrive he must learn drive
    ''He just arrived and must learn how to drive.''

Wei (2002) has presented a distinction in activation between content and system morphemes. In the production model the prototypical content morphemes are nouns, verbs, descriptive adjective, most prepositions, and free standing pronouns. Prototypical system morphemes are quantifiers, specifiers, and inflectional affixes.

The abstract level contain semantics, syntactic, pragmatic, and morphological and sociolinguistics information which is critical to the speaker's intention of the selection of these information (Myers-Scotton & Jake, 2013; Wei, 2002). The abstract level model can be classified in three categories on the basis of its function which includes lexical-conceptual structure, predicate-argument structure and the level of morphological realization patterns (Myers-Scotton, 2013).

In example (13) the activation of the EL islands verb phrase "call me" is accessed at lexical-conceptual level for the speaker's semantic- pragmatic intended meaning. As the counter part of EL island "call me" is easily accessed than its Chinese counterpart "da dianhau gei wo" whereas its literal meaning 'make phone to me'.

13. Chinese-English (Wei, 2002, p. 702).
    name ni mingtian *call me.*
    then you tomorrow call me
    ''Then you call me tomorrow.''

In code-switching data EL nouns are the most frequent embedded elements in the ML because they only convey a speaker's intended message at the abstract lexical conceptual level; that is, an EL noun need not match an ML counterpart at the levels of predicate–argument structure and morphological realization patterns for tense, mood and aspect (TMA). Unlike EL nouns, EL verbs occur only infrequently in some language pairs, although very often in other pairs, as noted above. This makes their switching something of a puzzle. The analysis here offers at least a partial answer, including the conclusion that EL verbs in naturally occurring CS imply a sense in which the bilingual cognitive system is open to modifications. In this paper, we only consider classic CS in which all the critical grammatical elements in bilingual constituents come from one language, the ML.

Figure 2, illustrates how the bilingual language production model adapted in this study is classified into three parts; mental lexicon, lexical conceptual level and formulator.

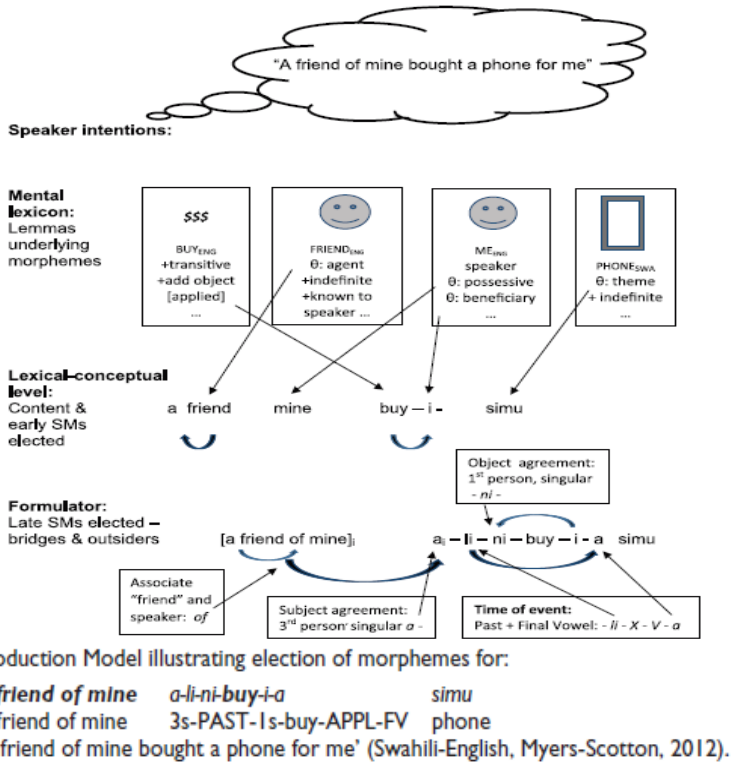

Figure 2. Morpheme activation in the bilingual mental lexicon
(Adapted from Myers- Scotton & Jake, 2017, p. 12)

Figure 2 provides ample explanation to the direct and indirect activation of the early SMs and content morphemes at the conceptual level and how late SMs, bridges and outsiders, are activated at the level of the formulator. Wei (2002) argued that in the case of intrasentential CS the speakers ignores the ML cues and switch to an EL item when no ML lexical item is available to the speaker's intended meaning. The bilingual speakers are making lexical choices in the preverbal message and that is why the conceptual information and the language cue work together to activate the most appropriate language items in the mental lexicon. Myers-Scotton (2017, p. 9) has presented three levels to the abstract lexical structure underlying the abstract level model. The first level, lexical-conceptual level refers to semantic and pragmatic meanings. The second level is that of predicate-argument structure and the third level is that of morphological realization of patterns. These three levels will also be discussed in analyzing the EL nonfinite verbs in Pashto-English bilingual data.

3. Research Methodology

3.1 Sample description

Twenty-four fluent English-Pashto bilinguals were selected in the present research work. All participant were educated and Pashto-English codeswitching was a prevalent style of

the sample population. They were sharing the same regional background and culture of Pashto speech community. They learned Pashto as their first language and English as their second language in English language classroom. Their ages ranged from 19 years to 50 years. Some of the participants were students but most of them were busy in their jobs and businesses. They were divided into mixed gropes irrespective of their age, education and nature of their jobs. The purpose of the mixed groups was to collect data with different patterns. The following tables (1, 2 and 3) show the demographics of these groups.

Table 3.1 Demographic information of group 1

Participant	Ah	Zu	Ra	Ma
Age	35	24	22	26
First language	Pashto	Pashto	Pashto	Pashto
Bilinguals	Pashto/English/ Urdu	Pashto/English/ Urdu	Pashto/English/ Urdu	Pashto/English/ Urdu
Medium of instructions in School	Pashto/English/ Urdu	Pashto/English/ Urdu	Pashto/English/ Urdu	Pashto/English/ Urdu
Medium of instructions university	Urdu/English	Urdu/English	Urdu/English	Urdu/English
Highest qualification	MS	MA	BS	MA
Socio-economic status	Middle class	Middle class	Lover middle class	Lower class
Occupation	Lecturer	Businessman	Student	Nurse in Hospital

Table 3.2 Demographic information of group II

Participant	Ba	Mu	Ra	Mush
Age	40	35	27	29
First language	Pashto	Pashto	Pashto	Pashto
Bilinguals	Pashto/English/ Urdu	Pashto/English/ Urdu	Pashto/English/ Urdu	Pashto/English/ Urdu
Medium of instructions Schooling	Pashto/English/ Urdu	Pashto/English/ Urdu	Pashto/English/Urdu	Pashto/English/ Urdu
Medium of instructions university	Urdu/English	Urdu/English	Urdu/English	Urdu/English
Highest qualification	PhD	MA	MS	BS
Socio-economic status	Upper-middle class	Middle class	Lover middle class	Lower class
Occupation	HR manager	Coordinator	Student	School teacher

Table 3.3 Demographic information of group III

Participant	Ah	Zu	Ra	Ma
Age	29	24	25	48
First language	Pashto	Pashto	Pashto	Pashto
Bilinguals	Pashto/English/ Urdu	Pashto/English/ Urdu	Pashto/English/ Urdu	Pashto/English/ Urdu
Medium of Schooling	Pashto/English/ Urdu	Pashto/English/ Urdu	Pashto/English/ Urdu	Pashto/English/ Urdu
Medium of university	Urdu/English	Urdu/English	Urdu/English	Urdu/English
Highest qualification	MS	MA	MBBS	BS
Socio-economic status	Middle class	Middle class	Middle class	Lowe-middle class
Occupation	Research fellow	Agriculture	Doctor	Policeman

3.2 Data collection Procedure

In each group the data was recorded in different three settings. In each settings only one topic was administered for discussion. The researcher worked as moderator and it worked really well as the researcher was a part of the same speech community. Table (4) demonstrates the distribution of the topics and the duration of the discussion with the groups.

Table 3.4 Topics and duration of discussion

Groups	Topic 1	Topic II	Topic III	Time
I	Female Education	Unemployment	Poverty	2: 45 hours
II	Present Government	Human rights	Electronic media	3: 21 hours
III	Unemployment	Human rights	Female Education	2: 33 hours
Total Time				8:49 hours

3.3 Data Coding

Out of 8 hours and 49 minutes data only 154 bilingual clauses were transcribed in roman English. Only the data with EL verbs in ML were placed into three steps. The first step represents the data at the morphemic level. The second step represents the data at the gloss level, and translation at the third level.

3.3.1 Structural Analysis of EL nonfinite verbs in bilingual Light Verbs Construction (LVC)

In the entire bilingual data, 39 instances of EL elements especially nonfinite verbs in ML light verb constructions (LVCs) were found. The most dominant pattern is the light verb construction *kaw* 'do' or 'make' which is inflected for tense, aspect and gender with EL nonfinite verb. The second dominant pattern is the intransitive light verb *keg* 'become'. The EL nonfinite verbs (table, 1) could be fitted in the four categories of verbs; phrasal verbs, gerunds, participles, and infinitive. The EL nonfinite verb; present participle and past participle are observed only three times entire data.

Table 3.5 Distribution of the EL verbs in ML light Verb Constructions (LVCs)

ML Construction	EL Verb (infinitive)	EL Participle	EL Gerund	EL Phrasal verb
Kaw (do/make)	22		1	
Keg (become)	9			2
Copula (be)	2	2	1	

As argued (Myers-Scotton, 2017, p. 13) that the EL verbs are embedded in the ML finite verbs construction and the EL verbs are nonfinite verbs. In Pashto-English CS data (Author, 2014 & 2015) the morphosyntactic frame is provided by Pashto language and EL lexical items are integrated in different constructions of Pashto language. The EL nonfinite verbs are inserted in ML frame because the EL nonfinite verbs are not activated at the formulator as the ML finite verbs do. The most remarkable function of the EL nonfinite verbs in Pashto-English CS data is the role of the Pashto light verbs. The Pashto-Light verbs are marked for tensed, aspect and gender in LVCs and the EL nonfinite verbs only carry the speaker's intended meaning (Khan, 2018). In the Abstract level model the EL nonfinite verbs are not salient at the predicate-argument structure as it is activated only to carry the speaker's intended semantic-pragmatic meaning at the lexical-conceptual level (Myers-Scotton, 2017, p. 13). Wei (2002, p. 691) states that the speaker's choice at the preverbal level of lexical-conceptual level play a crucial rule in selecting the lemmas of the specific language in the bilingual mood and then morphosyntactically it is realized at the functional level of predicate-argument structure.

According to 4-M model, verbs are content morphemes and is responsible to assign thematic role to arguments and that's why it carries meaning (Myers-Scotton, 2017, p. 13). According to Myers-Scotton, (ibid) if the EL nonfinite and nouns both carry meanings then why EL nouns are prevalent in CS data then the EL verbs. On the other hand the finite verbs have a different procedure than that of nouns and nonfinite verbs, while activation in the bilingual mental lexicon. The finite verbs undergo at different level; such as for semantic-pragmatic it is partially elected at the lexical-conceptual level and for the tense, aspect and gender features it is elected at formulator in the abstract level model. It is also elected for the maximal projection to produce larger constituents such as clause, for the agreement features and case assigning and that's how the election of the finite verbs are more complex in nature (Myers-Scotton, 2017).

The finite verb is elected at two different levels in the language production model and it could be more costly and complex and that's why only the EL verbs could be found in the bilingual LVCs. Table 1, provides different construction of the ML where the EL verbs are nonfinite or in to infinitive form.

In (14), the EL nonfinite verb "play" is functioning as content morpheme and is inserted with the help of ML transitive light verb *kaw* "do" which is inflected for tense, aspect and subject-verb agreement. According to 4-M model the content morphemes; the EL nonfinite verb "play" is conceptually activated for semantic pragmatic meaning. In Pashto-English bilingual clause the EL, nouns 'media' in subject position and 'role' in object position is incorporated in Pashto morphosyntactic frame.

The motivation behind this insertion is the nature of the EL nonfinite verb “play” as it is easy to be incorporated as infinitive as it has also replaced its Pashto counterpart *adda* in nonfinite slot. In order to satisfy the predicate-argument structure the ‘do’ construction receives all the inflections from the ML. According to the argument presented in image (1) there is an obvious relationship between different abstract levels with the empirical data.

14. Pashto-English codeswitching
    Masharan hmisha *positive role* play kaw -i
    Elders.NOM always positive role play do.PRS.IPFV -3SG
    ‘The elders always play a positive role.’

In order to satisfy the predicate-argument structure the Pashto light verb kr ‘do’ is marked for tense, aspect and subject-verb agreement at the formulator as the phi features for tense, and aspect is elected at formulator. According to the 4-M model the nonfinite verbs such as ‘divided’ is a content morpheme which is salient at the lexical-conceptual level to fulfill the speaker’s semantic-pragmatic intentions. Myers-Scotton (2014) while discussing the integration of EL nonfinite verbs in ML has provided few examples from Swahili-English (Myers-Scotton, 1993a, p.120) CS data. The insertion of the English verbs *spoil* and *select* is made possible when it receives the inflection morphemes from Swahili language.

15. Pashto-English codeswitching
    …Che dwa *groups* ke mong *divide* kr -o
    …that.CP two groups in.OBL 1PL.NOM divide do.PRS.PFV -1PL
    ‘That we have divided in two groups’

The EL infinitive “target” is inflected for the pie with the ML light verb *kr* “do” for tense, aspect, and subject-verb agreement. The election of the EL nonfinite verb ‘target’ is more convenient than its counterpart *pə-naha* of Pashto language. The predicate-argument structure confirms the thematic role assigned to the two arguments *mong* ‘we’ and ‘youth’ but it is the Pashto light verb *kaw* that is elected at the formulator according to the classification of 4-M model. The lemma specification at the preverbal level play a crucial role in the selection of the EL nonfinite verb at lexical-conceptual level.

16. Pashto-English codeswitching
    pə de ke mong *youth* *target* kaw -u
    at this in.OBL 1PL.NOM youth target do.PRS.IPFV -1PL
    ‘In this we target the youth.’

In example (17), the EL nonfinite verb ‘teaching’ (present participle) appears in ML light verb construction *kaw* “do”. The EL nonfinite verb as content morpheme is elected at lexical-conceptual level for the semantic-pragmatic intention of the speaker and

contributes to the core semantics of the construction. The selection of the nonfinite verbs in ML finite verbs slot is always easy to retrieve because the production cost in the activation of the EL nonfinite verbs is lower than that of ML finite verbs (Myers-Scotton, 2017, p. 16) as it is salient at one level and for inflection it is the light verb *kaw* which is elected at formulator for predicate-argument structure, tense, aspect and modality. The IPFV in the following example and the English present participle shows that it satisfy the speaker's intention to convey incomplete aspect.

17. Pashto-English codeswitching
    Taso dair ha *teaching* kaw -y
    2PL very good *teaching* do.PRS.IPFV -2PL
    'You are doing great teaching'.

Example (18) is cited in Myers-Scotton (2014) The EL present participle is not inflected directly but is inserted in the ML frame by Hungarian verbalizer only to realize aspect. It satisfy the speaker's intention to convey incomplete aspect as the incorporation of past participle will show complete aspect in such Hungarian-English CS.

18. *Hungarian–English* (Bolonyai, 2005, p. 320)
    Nem lenne jobb ott-marad-ni egy kicsi-t **socializing**-ol-ni
    not be/COND better there-stay-INF a little socializing-VBZ-INF
    "Wouldn't it be better to stay and socialize a little?"

In example (19) the distribution of the nonfinite verb "release" in the surface data is an obvious example of content morpheme. The argument presented in Myers-Scotton (2017, p. 9) about the nature of the lemma underlying a morpheme depends on the semantic-pragmatic features or grammatical features is crucial. The empirical evidence strongly suggest that the English nonfinite verb "release" is elected at lexical-conceptual level for the speaker's intention (semantic-pragmatic) and even the thematic role is assigned at the same level as it has taken only a single argument 'cassette' in the intransitive light verb *shw* "become" construction.

19. Pashto-English codeswitching
    Halaq way -i Che c*assette* *release* shw -a
    People say.PRS.PFV -3PL That.CP cassette release become.PRS.PFV -3SG
    'The people say that the cassette has been released.'

The EL nonfinite verb 'destroy' is inflected for tense, aspect and agreement by the Pashto intransitive light verb *sh –i* (become). In example (20) the EL nonfinite verb "destroy" is classified to be activated in the lexical-conceptual level and has taken a single argument 'system' in the intransitive light verb *sh* 'become' construction. According to the language production model the finite verb is elected both at lexical-conceptual level as well as at the level of formulator for tense, aspect, mood and agreement and indicate high cost production as compared with the election of EL nonfinite verbs at lexical-conceptual level and indicate low cost production.

20. Pashto-English codeswitching

Bia	toul	*system*	*destroy*	sh-i	
Then	complete	system	destroy	become.PRS.PFV	-3SG

"Then the complete system is destroyed."

In present data only two instances of phrasal verb were found, in example (21), the integration of EL phrasal verb 'brought up' with ML light verb is innovative. The counter part of EL 'brought up' in Pashto language is *pervarish,* actually in Pashto there no such structure as phrasal verb and that is why it is an innovation into the Pashto language. At the preverbal level for the lemma specification EL phrasal verb is more convenient to the speaker and its selection at the lexical-conceptual level in the production model is a strong indicator of the speaker's semantic pragmatic intention. According to 4-M model of morpheme classification the EL phrasal verb is the early system morpheme and is activated at preverbal position in the language production (Myers-Scotton, 2016).

21. Pashto-English code switching

…Che	zamong	hə	*brought*	*up* wa	-sh	- i
…That.CP	IPL	good	brought up	PFV	-become.PRS.-3PL	

'…That we should have best brought up.'

In (22), the English embedded verb 'apply' is inflected with the Pashto light verb *kaw* "do" for tense, aspect and agreement. The EL verb is nonfinite as it supply only the speaker's intended semantic and pragmatic meaning and according to the classification of the 4-M model the EL infinitive 'apply' is a content system morpheme. According to different lexical abstract level the projection of Pashto light verb for tense, aspect is at level of predicate-argument structure (Wei, 2002, p. 704). It also indicates the incomplete aspect of the speaker's intended meaning.

22. Pashto-English codeswitching

Za	hapal	toul wisdom	*apply*	kaw	-m
ISG	own	all wisdom	apply	do.PRS.IPFV	.ISG

"I apply all my wisdom."

It is not only the LVC that allows the EL nonfinite verbs to be inserted in the ML but the ML copula (be) also provide space to the EL nonfinite verbs to be selected at the lexical-conceptual level. The English present participle nonfinite verb 'missing' is a good example of adjectival verb. It's counterpart in Pashto is *mojud na* 'present' 'not'. The activation of the present participle "missing" is more convenient than its Pashto counterpart. The activation of the Pashto *mojud –na* is activated at two steps. The insertion of the EL noun "episodes" also demonstrates the speaker's semantic-pragmatic intension as according to 4-M model the plural markers from EL is a type of early system morphemes and is activated at conceptual level and pose no threat to the ML morphosyntactic frame.

23. Pashto-English codeswitching
...Che Sudais da dwa -o episodes na missing wu
…that.CP Sudais of two -ERG episodes from missing COP.PST.IPFV.3SG
'That Sudais was missing from two episodes'.

In Pashto language, the use of EL participle is very common such as "justified" "tested" but, only two instances have been recorded in the present data. In example (24), the EL past participle "connected" is embedded as complement to Pashto present imperfective copula *yum* 'be'. The selection of the EL verb is a strong example of the speaker's intended meanings as its counterpart *munsalik* in Pashto language do not carry the same semantic-pragmatic meaning. In the bilingual mood the lemma is specified at preverbal level and for the actual realization of the speaker's semantic-pragmatic intention then it is elected at lexical-conceptual level in the abstract model. The bilingual VP is making a passive construction with the EL past participle. Actually such insertion of past participle in the ML COP construction is making a passive structure. In active construction it function as infinitive such as *connect kaw -am* "connect do –ISG". The present data also show that most of the EL present participle as in example (24) or past participle in example (25) are embedded in the ML COP construction. The EL infinitive forms are embedded within the "do" construction of ML.

24. Pashto-English codeswitching
Mong tasu sra *connected* ye
IPL 2PL with connected COP.PRS.IPFV.2PL
'We are connected with you'.

4. Conclusion

The present study demonstrates that Pashto as ML provides the morphosyntactic frame and the EL nonfinite verbs are integrated in the light verb construction of ML. In some contact situation (Myers-Scotton, 2013) the EL nonfinite verbs are integrated with the help of an extension or verbalizer and that extension overtly mark the verb transitive but in the present data the integration or insertion is made possible by the Pashto light verbs. The study supports the stance that EL nonfinite verbs are elected at lexical-conceptual level to carry the semantics-pragmatics intention of the speaker. It also supports the stance (Myers-Scotton, 2013, P. 522) that the EL nonfinite verbs carry the speaker's intended meanings and is not elected at formulator. Nonfinite verbs function differently from finite verbs whereas to realize the argument-predicate structure and morphological patterns that are always available to participate in ML frame but these level are not salient to nonfinite verbs and that's why the argument-predicate structure to the nonfinite verbs is elected at the conceptual-lexical level.

The study show that at the abstract level the EL nonfinite verbs do not go for the congruence checking within the framework of ML as claimed in Wei (2000) but at preverbal level the lemma is specified for the anticipated language in the bilingual mood. The thematic role of the EL nonfinite verbs are checked at the level of conceptualizer so that to satisfy the speaker semantic-pragmatic intentions. The most dominant pattern of EL nonfinite verbs insertion is in the "do" construction of ML. The "do" construction in

the ML is realized for tense, aspect and it marked case in the predicate-argument combination. The late system morphemes are activated at the level of formulator, and that it plays a pivotal role in the construction of the Pashto ML frame. The study validate the claim that the production cost of the EL nonfinite verb is low as it is not checked for tense, aspect and agreement at formulator but is activated only at lexical level. The study also highlights the flexibility of the cognitive system during the bilingual mood.

References

Annamalai, E. (1971). Lexical insertion in a mixed language. *Papers from the Seventh Regional Meeting. Chicago Linguistic Society,* 20-27. Chicago: University of Chicago.

Babrakzai, Farooq (1999). Topics in Pashto syntax. Manoa, HI: University of Hawai'i at Manoa dissertation.

Backus, Ad M. (1999) Evidence for lexical chunks in insertional codeswitching. In *Language encounters across time and space*, E.L. Brendemoen & E. Ryen (Eds.), 93-109. Oslo: Novus Press.

Costa, A. (2005) Lexical access in bilingual production. In J. Kroll and A.M.B. De Groot (eds) Handbook of Bilingualism: Psycholinguistics Approaches (pp. 308_325). Oxford:Oxford University Press.

De Bot, K. (2004). The multilingual lexicon: Modelling selection and control. The International Journal of Multilingualism 1, 17_32.

Francis, W. (2005). Bilingual semantic and conceptual representation. In J. Kroll and A.M.B. De Groot (eds) Handbook of Bilingualism: Psycholinguistics Approaches (pp. 251-267). Oxford: Oxford University Press.

Garrett, M. (1975). The analysis of sentence production. In G. Bower (ed.). The Psychology of Learning and Motivation (Vol. 9; pp. 133_177). New York: Academic Press.

Jiang, N. (2000) Lexical representation and development in a second language. Applied Linguistics 21, 47_77.

Author, (2018).

Author, (2018).

Author, (2014).

Author, (2015).

Kroll, J. and Tokowicz, N. (2005) Models of bilingual representation and processing: Looking back and to the future. In J. Kroll and A.M.B. De Groot (eds) Handbook of Bilingualism: Psycholinguistic Approaches (pp. 531_553). Oxford: Oxford University Press.

Lotfabbadi, Leyla Naseh (2002). Disagreement in Agreement - a Study of Grammatical Aspects of Codeswitching in Swedish-Persian Bilingual Speech. Doctoral dissertation, Stockholm University.

Muysken, Pieter (2000). *Bilingual speech: A typology of code-mixing*. Cambridge University Press.

Myers-Scotton, C, and Jake, J. L. (2017). Revisiting the 4-M model: Code-switching and morpheme election at the abstract level. *International Journal of Bilingualism* 21: 340–66.

Myers-Scotton, Carol (1993). *Duelling languages: Grammatical structure in Code-switching.* Oxford: Oxford University Press.

Myers-Scotton Carol (2002) *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press.

Myer-Scotton, C. (1992a). Constructing the frame in intrasentential code-switching; *Multilingua, 11*, 101-27.

Myers-Scotton, C. (1993a). *Duelling Languages: Grammatical structure in Code-switching.* New York: Oxford University Press.

Myers-Scotton. C. and Jake, J. L. (2000). Four types of morpheme: Evidence from aphasia, codeswitching and second language acquisition. *Linguistics* 38:1053–100.

Myers-Scotton. C. and Jake, J. L. (2009). A universal model of code-switching and bilingual language processing and production. In Barbra E. Bullock and Alemida Jacqueline Toribio (Eds.), *The Cambridge Handbook of Linguistic Code-switching*, New York, Cambridge University Press, pp. 336-357.

Myers-Scotton, C. (2006). *Multiple Voices*: An introduction to bilingualism, Oxford and Malden, MA: Blackwell Publishing.

Richards, J.C. (1976) The role of vocabulary teaching. TESOL Quarterly 10, 77_89.

Ringbom, H. (1987) The Role of the First Language in Foreign Language Learning. Clevedon: Multilingual Matters. Lexical Transfer 123

Siegel, Jeffrey (1987) *Language contact in a plantation environment*. Oxford: Oxford University Press.

Treffers-Daller, J. (1994). *Mixing Two Languages: French-Dutch Contact in a Comparative perspective.* Berlin: Mouton de Gruyter.

Wei, L. (2001). Lemma congruence checking between languages as an organizing principle in intrasentential codeswitching. *International Journal of Bilingualism,* 5, 153–173.

Wei, L. (2002). The Bilingual Mental Lexicon and Speech Production Process. Brain and Language 81, 691-707

Appendix A. List of Abbreviations

1, 2, 3	first, etc. person	f	feminine	PFV	perfective
		FUT	future	PL	plural
ADV	adverb	GEN	genitive	POSS	possessive
BCV	bilingual complex verb	IPFV	imperfective	PRN	pronoun
		KP	Khyber Pakhtunkhwa	PRS	present
CL	clitic	LOC	locative	PRX	proximate
CS	code switching	LVC	light verb construction	PST	past
CP	complementizer			RECP	reciprocal
COP	copula	M	masculine	REDUP	reduplication
DM	demonstrative	OBJ	object	SG	singular
DST	distal	OBL	oblique		
ERG	ergative				