Testing Matrix Language Framework Model On Urdu-English Online News Entity: A Creative Approach

Fatima Tu Zahra,Tahir Saleem, A Farhat Abdullah, Muhammad Arif Khan

Article Info	Abstract
	Codeswitching in the context of single or multiple conversations has been a myth for language experts. Matrix language framework (MLF) model proposed by Myers Scotton (1993) has become very popular for the analysis of language pairs, and is influential in determining matrix language in different language pairs. The aim of this study is to identify matrix and embedded language in Urdu-English data sets of health and science theme. MLF model is applied to an original article on Covid-19. Data sets include language pairs from a published article on the nature of coronavirus. A qualitative design was followed to arrange data sets, language pairs were identified, transcribed and coded carefully according to the Canonical Trilinear Representation. Three layers of data with the first layer of roman Urdu, the second layer of gloss and the third layer of English translation were further analyzed syntactically and morphosyntactically to show how they grammatically occur in the bilingual complementiser phrases. The findings of this study reveal that code-switching was permissible even when it led to structural asymmetry. English insertions received different Urdu markers of gender and number wherever required. Urdu adjectives played a significant role in realizing nouns. Some data sets allowed English insertions without Urdu markers. Moreover, the data supported matrix language frame, morpheme order principle and system morpheme principle and no counter example appeared against MLF model. Thus, the present study is a significant contribution in the related area.

Introduction

The alternation of two languages inside and outside of sentences –codeswitching (CS hereafter) has been studied mainly from two perspectives, i.e. Grammatical/syntactic and pragmatic/discourse (Romaine, 1995; Jake, 1994). Although these two aspects are related and a knowledge of one helps understand the other (Jake & Myers-Scotton, 1997; Myers-Scotton, 2002; Hebblethwaite, 2010). A number of researchers have proposed descriptive grammatical constraints in CS, rather than theoretical. Poplack's study on Spanish/English bilinguals (1980) was one of the most influential studies.Her study is specifically based on syntactically close languages, i.e. English / Spanish, butcounterexamples can be found in syntactically distant languages. In the 1980s some other models were proposed to define grammatical constraint, e.g. Chomskians' Government Binding model by Di Sciullo, Muysken and Singh (1986). Myers-Scotton and her associate Azuma developed a model based on speech production theory, i.e. Azuma's frame-content hypothesis (1993) and Myers-Scotton's Matrix Language Frame (MLF) model (1993). Myers-Scotton's MLF model is an abstract theoretical model, but it has been employed to examine contact phenomenon among a variety of languages since the 90s. Her MLF model was revised in Myers-Scotton, 1997 and extended sub-models were proposed (Jake & Myers-Scotton, 2002; Myers-Scotton & Jake, 2017; Kniaź & Zawrotna, 2020).

The concept of the MLF model is influenced by psycholinguistic theories. The most significant of the three are the differential activation of base language and guest language (Grosjean,1988 quoted in Myers-Scotton, 1993), the different retrieval process of closed class items and open items in Garret's speech error study (1975 quoted in Myers-Scotton, 1993), and lemmas in the mental lexicon linking conceptual information and grammatical function in Levelt's language production model (1989 cited in Myers-Scotton, 1993). Urdu-English code switching is evident in health and science related data sets. Recent covid-19 pandemic has proved that certain data sets are being produced every day. Frequent use of English content words in Urdu bilingual speech helps in realizing potentials and constraints of both languages. In a single conversation, there is always a mutual understanding between the speaker and interlocutors. The speaker is in the mode of communicating while considering semantic, pragmatic and socio-pragmatic aspects of interlocutors (Myers-Scotton 1996; Jake & Myers-Scotton, 2009). Both languages in connection have definite roles and MLF model explains these in interpretations.Myers-Scotton (1993a, 1997, p. 82) proposed System Morpheme Principle (SMP) along with Morpheme Order Principle (MOP) to present an argument on bilingual constituents of participating languages.

In this argument, Matrix language (ML) provides morphosyntactic frame, whereas lexical items such as nouns and verbs are inserted from Embedded language (EL). This analysis is focused on Urdu-English data sets. These data sets were converted into their English translation and then roman Urdu version. Data is available online atthe BBC Urdu website (https://www.bbc.com/urdu/science-54677744). The article was written to provide information about the nature of the virus and the very natural flow of the text made it highly suitable for this analysis. The lack of the "observer's paradox" (Labov, 1972, p. 79). Code mixing denotes a phenomenon where two languages are observed in a single sentence carrying lexical and grammatical features from both languages. However, there are set rules for the insertion of the items. These insertions are often unconscious, abrupt and unplanned but almost all productions seem to follow the same criteria. Muysken (2011) discusses code mixing with reference to language mixture. Urdu-English code switching can be intra-sentential or inter-sentential. CS can be structurally divided into inter-sentential CS and intra-sentential CS. The inter-sentential CS has been broadly studied in the sociolinguistic field, but grammatical constraints are not a major focus there. With the intra-sentential CS, grammatical constraints directly affect the behaviour of two, or more participating languages (Youssef, 2017). The MLF model is devised to explain intra-sentential CS as both Urdu and English are extensively used in formal education, public sector and industry. Though both languages are used in the expressions, code-switching confirms that Urdu is the only language that provides morphosyntactic frame of the clause in data sets as proposed by Myers Scotton (2005). Language production at the abstract level confirms bilingual constituents atthe surface level of data sets and this is discussed in MLF principles. This data set contains Urdu-English data with an intra-sentential code switching. Clyne (2003) and Muysken (2000) both argued that the cognitive aspects in language development and processing need careful scrutiny on the part of researchers.

2. Literature Review

Asymmetrical distribution of intra-sentential code-switching confirms presence of more dominant language as Matrix language and the language of inserting constituents as an embedded language (EL). ML can be frame language as this is the first language or the language that has a dominant position in the speech. Myers-Scotton's (1993) proposed that ML is responsible for the structure and provides a grammatical frameof the sentences were as ELs are inserted in the form of content words in speech. Myers-Scotton explained that CP (projection of Complementizer) lends itself to analysis more appropriately than 'sentences' as grammar of sentences can be isolated and not in contact. The position 1 of CPs is strong and it is used by many scholars (Myers-Scotton, 2002, p. 55). Two variations of CS are observed by Myers-Scotton and she proposed that the classical code switching happens when a speaker is efficient enough to produce ML grammatical structure. Besides the other form of code switching is likely to happen when grammatical structure is not provided from one dominant ML but both languages are taking part in the structure. (Myers-Scotton& Jake, 2017, p. 2). This composite CS has been detected when the speakers have vagueobjectives about the preferred ML. In this paper only classic CS is discussed. A bilingual CP can consist of 1) ML islands which have only ML morphemes 2) mixed constituents, including morphemes from both ML and EL and 3) EL islands consist of only EL morphemes. ML islands are made of ML morphemes and are under the control of ML grammar (Ziamari, 2007). On the other hand, EL islands are also well-formed by EL grammar, but they are introduced into an ML frame. Therefore, EL islands are under the constraint of ML grammar.

2.1. Content-system morpheme distinction

ML can be identified through careful analysis of content and system morphemes. Thematic roles are accepted or rejected by content morphemes such as nouns, verbs, adjectives and some prepositions. These content morphemes carry semantic and pragmatic properties of EL, crucial for sending or receiving messages in communications. On the other hand, system morphemes such as function words and inflections have nothing to do with thematic roles and they are responsible for relationships and agreements with content morphemes. Grammatical frames are produced through this collaboration. Bilingual CPs express system morphemes from ML and EL provides content morphemes. MLalso provides content morphemes. This distinction of content and system morphemes allow researchers to identify ML in bilingual CPs through two principles of Myers-Scotton.

The Morpheme-Order Principle: In ML+EL constituents consisting of singly-occurring EL lexemes and any number of ML morphemes, the surface morpheme order will be that of ML.

The System Morpheme Principle: In ML+EL constituents, all system morphemes which have grammatical relations, external to their head constituent will come from the ML ((Myers-Scotton, 1993, p. 83).

Phenomenon of code-switching (CS) has theoretical and empirical descriptions in the past studies. Woolford (1983) argued that sentences are produced from both languages and phase structure rules are used in this formation. Myers-Scotton (1983) views code switching as a phenomenon where marked and unmarked choices are available to the language users. Chan (1984) proposes that constituents are bound together while carrying semantic and syntactic rules that join single entries in a single language. Malmkjaer (1991, p. 62) proposed that language mixing can also be used to express emotion, close personal relations, harmony and to eliminate a third person from part of a discussion. A division is drawn between two types of linguistic collaboration as follows:

Code mixing is the use of elements, most typically nouns from one language in an utterance predominantly in another language and Code switching is a change from one language to another in the same utterance or conversation (Hamer & Blanc, 1989, p. 35 cited in Malmkjaer 1991, p. 62).

The arguments at which code switching can take place are between sentences, clauses, phrases and words. The exchanging is administered by different norms in different bilingual groups. It further concludes that code switching is more challenging when typologically diverse languages are elaborated than when the languages are typologically alike (Myers-Scotton & Jake, 2009, 2014;Namba, 2012).

MLF model was studied by various researchers both empirically and theoretically. Chun (2001) applied MLF model on Korean-Chinese language pairs.Muysken (2000, p. 3-10) proposes that the patterns of code mixing are often rather different from one another because there are several processes at work during code mixing. Macswan's (2000) and Chun's (2001) view that the MLF model is not a universal theory is the motivating factor in this study. Macswan (2014) asserts that a "language frame" is only necessary if the grammatical facts on code switching cannot be explained and so the model is not necessary. Chun found out that the MLF model does not adequately account for Korean-Chinese code switching. Another research on the application of MLF model data totally supports both the MOP and SMP (Hadei et al., 2017; Myers-Scotton & Jake, 2017). Hence, this study is initiated to examine Urdu-English data in the light of MLF model proposed by Myers-Scotton.

2.2. Objectives of the study

The current study tried to address the following research objectives:

1. To know whether the matrix language and embedded language are determined through Morpheme order principle in Urdu-English data sets.
2. To identify content morpheme and system morphemes through the scope of system morpheme principles.

3. Methodology

This analysis reveals code switching patterns of MLF model in Urdu-English. Our data set consists of an article from BBC Urdu science and health section. Corpus of science and health related sentences in the form of different interactions involving bilinguals' data provided several sets of CS patterns. Through MLF model data was used to realize the matrix language versus the embedded language, the system morpheme versus the content morpheme and the principles of the MLF model. The description of the three concepts in the following subsection is drawn from Myers-Scotton (1993). Data covers health and science topic with specific reference to the Corona Virus. In this article virus is well described using different words and contexts. On the other hand, health professionals and university professors are also quoted from beginning till the end. Data sets are composed by highly experienced authors. This article was selected because of its health and science related vocabulary. After careful organization, data were further divided into two major categories.

1. Usage of the word virus in different singular and plural forms and syntactic structures
2. References of professors

3.1. Data Collection

Data sets were selected and categorized into sentences and then to phrases and words, for further analysis. The datawere converted into Roman Urdu with online conversion tools. Some of the words were not picked by the software so the researcher manually converted them into Roman Urdu. All data have been transcribed in three layers. The first layer represents the data at the morphemic level. This data was taken from BBC article and then converted into Roman Urdu through an online tool. The second layer represents the data at translation level. Third level was also taken from the English version of the BBC Urdu website. Word with similar entries in same semantic positions were excluded from the final analysis. However, the same word with different morphosyntactic and semantic patterns were included in the analysis. In some places, only phrases are included in the discussion instead of completed clauses or sentences.

3.2. Data Coding

First option of the data coding was to use roman Urdu and English translation only. The example shows sentence in Roman Urdu and its English translation.

Text: Covid 19: Corona virus isqadar mohlik kiyun hai?
Translation: Covid-19: Why Corona virus is that much more hazardous?

4. Data Analysis

In this data set Urdu-English codeswitching is analyzed under MLF model proposed by Myers-Scotton. Single word entries along with phrases, clauses and sentences are analyzed in this research as quoted in Chun (2001, p. 4). These data sets are observed for both principles and roles of both languages are determined. Morpheme Order Principle and System Morpheme Principle were used to analyze whether bilingual constituents are coming from EL or ML. In the first place an Urdu word order (SOV) is followed in the data sets, instead of English word order (SVO).

When two or more languages are in contact, surface level words from one language are embedded in the receiver language. Both Urdu and English are asymmetrical in their structure and this asymmetry also involves

content words and grammatical elements. In Urdu-English data sets Urdu, one language are the "frame language" also called matrix language and the other one is embedded language or the guest language as proposed by Myers Scotton (2005).

Table 1

Distribution of English single insertions

Type	Number
Nouns	29
Verbs	2
Adjectives	0
Adverb	0

This study confirms English insertions in the form of nouns and verbs. Here nouns are switched with highest frequency from EL to ML frame and it confirms what is proposed by Myers-Scotton (2002). This data is aligned with many other studies mentioning nouns as the most common form of the codeswitched constituent in a sentence (see Haugen, 1969; Health, 1981; Myers-Scotton, 1993; Naseh, 2002). Target data sets carry unique adjective-noun combinations where Urdu adjectives are used with English nouns and this is prevalent for the most used element "VIRUS".

English Nouns in Urdu Frame

In Urdu-English target data set English nouns were found to occur with or without Urdu markers of plurality and gender. Urdu case markers and light verbs were also modified to adjust English nouns. The English single noun virus was found to occur with and without Urdu markers. There is only one case in which it is marked with Urdu plural marker virus-on and doctor-on.

    (1) Covid-19
    (2) Virus

Example 1 shows that the term is used universally as it is. Example 2 target word Virus was frequently used in different forms. It has the highest frequency of 29 out of 77 English insertions. In some places the English insertion of the virus is adjusted through Urdu case markers, plural markers, gender markers and even light verbs and adverbs. Urdu adjectives are also used to modify English noun insertions. This data also confirms both principles of MLF model. Lexical words such as nouns and verbs are prototypical content morphemes whereas function words with affixes appear with other words of agreement. These words are good examples of system morphemes such as determiners and case words or postpositions in Urdu.

Table 2

English Insertions in Urdu frame

Word	Category	Number
Covid -19	Noun	8
Virus	Noun	29
Corona virus	Noun	9
Flu	Noun	6
Chemicals	Noun	3
Infection	Noun	4
Viral factory	Noun	1
November	Noun	1
October	Noun	1
Bi bi si	Noun	1
Interferon	Noun	1
Laboratory	Noun	1
Joker card	Noun	1
Hospital	Noun	1
Driver	Noun	1
Sars cov-2	Noun	1
Europeans ka America	Noun	1
Inside health	Noun	1
Career	Noun	1
Cellular doorway	Noun with plural marker	1
Ace 2 receptor	Noun with plural marker	1
Viral infection	Adjective	1

Metabolic		1
Protein	Verb	1
Virus-on	English words	1
Doctor-on	English words	1
Infected		1
Update		1
Test		1
Total		87

Table2 shows 77 different English insertions in the article with only two verbs. Nouns were in the ML islands with Urdu adjectives and case markers.

(a) aik mamooli se virus
   a minor virus

(b) kamal ka virus
   wonderful virus

(c) naye virus
   new virus

(d) fraib dainay wala munfarid virus
   Unique deceptive virus

(e) bairooni virus
   external virus

(f) bilkul mukhtalif virus
   totally different virus

(g) chaar insani corona virus
   four human corona virus

(h) yeh virus
   this virus

(i) cheechak ke virus
   small pox virus

(j) faraar ho jane wala virus
   virus that escapes

(k) qaatil virus
   Killer virus

All given examples (a to k) have the word virus in common. This English insertion in Urdu frame carries Urdu modifiers and structure. In data sets, Covid-19 is used as it is and only at one place English transcript is used to communicate the name of H1N1.

1: *The Matrix vs the Embedded Language*

Matrix language determines the morphosyntactic framefor code switching data sets. According to MLF model ML have more morphemes and determines tense, aspect and the agreement of the datasets. Following example illustrates both ML and EL languages.

(1) aap laboratory mein bemaar khulion ko dekhen ge to yeh aap ko bemaar yani infected nahi nazar ayen gay
   If you look at diseased cells in the laboratory, they will not look sick, that means infected

In (1) Urdu is the ML because it gives the sentence, tense and pronoun "aap" along with plurality markers -khul-yon, daikh-ain and "gay" at the end. Urdu case marker 'ko' also confirmed that Urdu is the ML of the data sets. Two morphemes, Laboratory and infected further confirm that English is the embedded language here.

(2) agar aap apna test karwaen ge
   But if you do your test

The above example shows the presence of Urdu light verbs along with English main verb 'test'. Matrix language has given a masculine identification to the verb through 'Apna'. Besides, there are other words of Urdu, for example, aap, karwa-ain and 'ge' are all revealing Urdu agreement.

(3) unhein infection hai, aur yeh sirf aik' joker card' hai jo yeh virus khail sakta hai .
   They have an infectionand it's just a 'joker card' that this virus can play can play.

Finally, masculine markedness 'sak-ta' at the end of the sentence further confirms the morphosyntactic structure of Urdu. EL noun forms, infection, joker card and virus are not modified and used as it is.

Mayer Scotton frequency criteria of ML are also applicable here.

(4) naya virus un viruson hi se bohat had tak milta julta tha
The new virus was very similar to the viruses.

(5) ab tak mazeed chaar insani corona virus aaye hain

So far, four more human corona viruses have been reported

(6) hamaray jism mein virus ki tadaad

The number of viruses in our body.

Example (4), (5) and (6) have ML islands, naya, milta julta tha, insaani, hamaray and EL islands in the form of virus and corona. There is an ML + EL constituents, 'naya virus' Un virus-on hi say', 'chaar insani corona virus aaye hain'and 'virus ki tadaad'. Example 4 supports both principles as well. The Urdu word order is followed here. In addition, although content words or main subject (virus) is from English, subject agreement and system morphemes are from Urdu. This data support Swahili as the ML in this example. English virus is receiving Urdu plurality marker-on.

2. The System Morpheme Vs Content Morpheme

Content morphemes are either thematic role receivers or assigners and most of the system morphemes belong to the functional category, such as inflectional morphemes. Quantification also differentiates between content and system morphemes. System morphemes show quantification as well.

(7) University of manchester ki professor Tracy Hassel kehti hain ke' yeh bilkul naya hai, is liye hamein nahi maloom ke is ke khilaaf mudafat mojood hai ya nahi. '

"It's brand new, so we don't know if it's immune," said Tracy Hassel, a professor at the University of Manchester.

Example (7) shows that content morpheme (University of Manchester) is working well with the system morpheme (ki, kehti hain) of the matrix language, Urdu. ML has kept its system morphemes while embedding content morphemes from EL.

3: The Principles of the MLF Model

THE ML HYPOTHESIS

Morphosyntactic frame of data sets is based on ML grammar. Myers-Scotton explained this principle through **morpheme order principle** and **system morpheme principle**.

(8) Covid-19: Why is the corona virus so deadly?

Covid-19: corona virus is qader mohlik kyun hai?

Example (8) shows the morpheme order as Urdu is closing sentence with *hay*as compared to English, deadly. Example (4) shows that the plural marker of Urdu virus-on is used in one sentence and it gives strength to the idea that Urdu is working as ML here.

(9) Pehlay se jism ke mdafati nizaam ki tayari ke fuqdaan ko Europeans ke America ke braazmon mein cheechak ke virus ko liye jane ke waqeye se mawazna kya ja sakta hai

The lack of pre-existing immune systems can be compared to the incidence of the smallpox virus Europeans' continents of America.

Example (9) shows Urdu morpheme order and system morphemes in all structures. In the first place 'Europeans *ke* America *ke*' and then in the second place 'cheechak *ke* virus *ko*',case markers balance content morphemes and express strong position of Urdu system morphemes, ki-a, ja sak-ta, hai, etc.

(10)Ye bohat he kamal ka virus ha

'This is an amazing virus'

Example 10 confirms that one of the participating languages is ML and it is supplying morpheme order and subject-verb agreement in these constituents. Urdu has given 'virus' a masculine marker 'ka' in the above example. On the other hand, morpheme order confirms that Urdu is an ML here. Urdu light verb "ha" is at the end instead of English word 'virus'. Example (10) shows blocking of EL system morpheme. ML has blocked EL content morpheme so that it is not realized as system morpheme of the ML.An obligatory EL island is also expected through accidental usage of EL morpheme. Example of this phenomenon was not found in this article, however absence or presence of a plurality marker in the example (4), (5) and (6) shows strong influence of ML that determines -plurality marker of the morphemes. Urdu is identified as ML.

(11)Ye aik joker card ha jo ye virus khel sakta ha

'This is a joker card which this virus can play.'

(12)aap ke naak ko aik viral factory me badal deta ha.

'Which can turn your nose into a viral factory'

Only one of the participating languages supplies a certain type of system morpheme. Example 11 and 12 shows that content morphemes (virus, joker card, viral factory) are from EL whereas all system morphemes are from Urdu.

Some other nouns are also there.

Viral factory

Blood-forming chemicals

University of Cambridge

(13)ye bohat he mukhsoos qism ka infection ha

'This is a very special kind of infection'.

Example 14 further gives Urdu case marker "ka" to the English noun, infection.

(14) Cellular doorway ki wajah

Because of cellular doorway

English noun cellular doorway is used in Urdu structure. This example shows "ki" after English insertion.

English verbs in Urdu frames:

(15)   update ki gayi

updated

(16)   Agar aap apna test karwain gay

If you want to get yourself tested.

Example (15) show modification of English verb. Urdu frame and system morphemes are used to adjust El and English insertion is adjusted. English verb 'test' is used in Urdu frame here. Urdu possessive apn-a gives gender marker to the English verb "test".

Table 3

Usage of the Word Virus and its Different Forms

Singular	Plural
Corona virus	Corona viruses
Virus	Number of viruses
A minor virus	Similar to the viruses
dangers of a virus epidemic	Four more human corona viruses
A new virus or flu	
This corona virus	
An external virus	
Joker card	
Killer virus	
Attacking virus	
Different virus	
New virus	
The smallpox virus	
Hostage to the virus	

Table 3 shows target the word virus in different morphological and morpho-syntactic patterns of bilingual code switching in Urdu-English data sets. English translation of the data sets shows that English insertion has not affected the structure of the frame language and in all examples that structure from only one source is the preferred option. In bilingual speech, the structures of the Matrix Language are always preferred as proposed in the Uniform Structure Principle (Myers-Scotton, 2002). Joshi (1982) was one of the first to refer to the frame-building language as the Matrix Language and to the other participating language as the Embedded Language. Three premises of MLF model are identified in the data sets. First premises, confirm that both languages do not participate equally and this also aligns uniform structure principle. Example 24 reveal the matrix language dominance in the data. The second premise is about the percentage of morpheme types. These premises limit the Embedded Language to specific types of participation and example (24) shows a modification of English nouns in Urdu frame. We can observe high frequency of content words from EL, mostly nouns. The third premise proposes that Matrix language is more active, but both languages are "on" when a speaker is speaking or writing through codeswitching. First two premises are supported by MLF with empirical evidence. In Urdu data sets both languages seem to be active and Urdu is more in use as compared to El (English).

Table 4

Names of the teachers

	Teacher's references
1	Says Paul Lehner, a professor at the University of Cambridge in the UK
2	Professor Guyaka says
3	Said Tracy Hassel, a professor at the University of Manchester
4	Mauro Giaca, a professor at King's College London, says
5	Beverly Hunt, a professor at King's College London says
6	Sir Steven O'Reilly, a professor at the University of Cambridge, says

Table 4 shows the names of professors in Urdu frames.

There are some other very interesting examples of insertion in data sets. These examples provide further evidence for the support of MLF model.

(17) Protein
(18) Metabolic nizam
(19) Flu
(20) Career
(21) Inside health

The above examples are common English nouns. In one place example (18) English word metabolic is accompanied with the Urdu wordNizam.

(22) Itnay saaray mareezon ke group
    A group of so many patients

In this example Urdu casemarker 'ke' is used to mark plurality of the word mareezon.

(23) Virus tehelta -hwa barhta chala jata hai
    The virus travels and grows

Example (23) shows Urdu frame for the word virus and how the words are modifying English insertion.

(24) A: hamaray jism mein virus ki tadaad
b: Virus ko is baat ki koi parwah nahi
c: Corona virus ke khilaaf
    d:Corona virus se mil kar
    e: Flu ke baad
    f: Makhsoos qisam ka infection
    g: Virus ki waba ke khatraat
    h:Naya virus un viruson hi se bohat had tak milta jalta
    i:Kayi doctoron ne apni kahaniyan bayan ki hain

Urdu postpositions follow nouns or pronouns and mark grammatical functions, location, movement or extendedin time and space. Common postpositions are ka, ke, ki, KO, ne, se, main, par etc. (Schmidt 1999). Some verbs mark objects with –ka, whereas –KO is used when the indirect object of a verb is marked as in (b). The indirect object preceded the direct object. Example (24) he shows possession and –KO is marking subject here.

Above examples carry system morphemes of Urdu with English content words. Example 4 h and 24 i show Urdu plural marker "-on" attached to English insertions. Example 24 "he" shows marked adjective "nay-aa" of Urdu for English noun. This adjective shows the gender and the number of the noun "virus". Marked adjective change to agree with nouns in number and gender. Certain verbs require objects marked with –se as in (24) d. Here –se is modifying verb to complete sentence.

(25) a. Sardi walay flu
    b. Faraar ho jane wala virus hai

In the above example the suffix –wala and –walay is following adjectives and preceding English nouns. This suffix is agreeing with the noun it qualifies.

(26) a. bilkul mukhtalif virus hai.

Urdu adjectives are modifying English noun in the example (26). The Urdu word order is followed here and embedded language English is providing content morpheme only.

(27) a. kamikaze ka ikhraj
    b. kemikalz ke ikhraj

Both forms of the phrases confirm English insertions in different Urdu contexts. Both English insertions are same, but Urdu case makers confirm successful code switching between two languages.Urdu clitics (ne, ko, mẽ, se, pər, t̪ək and kɑ) are tagged to the nouns and thus adjustments to the English insertions are made through clitics. These words adjust plurality and gender markers of Urdu.

5. Conclusion

Urdu-English data sets confirm many aspects of MLF model. Urdu genitives (ka/ki/ke) largely depend on gender and number of the head noun. Urdu-English code switching constraints are revealed through the Matrix Language Frame Model (MLF) proposed by Myers Scotton. Urdu-English bilingual data empirically support MLF model, and in the entire data set, there is not observed even a single counterexample. Selected data of Urdu-English was asymmetrical in a structure where Urdu served as the matrix language providing the morphosyntactic framefor well-formed sentences. It further confirmed that frame language determined morpheme order of the bilingual constituents and subject-verb agreement was also maintained by Urdu acting as matrix language. The study also reveals that late system morphemes are activated to maintain structure at the

threshold level. Three fundamental premises of MLF model are also observed (1) contributing languages do not play identical roles in the bilingual item, (2) in bilingual elements within this clause, not all morpheme forms can come similarly from the ML and EL, (3) both languages appear to be dynamic, though matrix language is more in use than EL. Surface words and morpheme order of Urdu along with system morphemes confirm Urdu as matrix language. System morphemes of Urdu seem to adjust verbs and adjectives in order to work out the gender and plurality of English insertions in Urdu frames. Intra-sentential codeswitching further separates content morphemes of English in almost all data sets (Myers-Scotton 1993, 1997).

5.1. Limitations

This analysis is limited to one article only. The theme of this article was science and health. This article was uploaded on the BBC Urdu website on 25th October and then updated on 10th November 2020.

References

Al-Bataineh, H., & Abdelhady, S. (2019). Cree-English intra-sentential code-switching: Testing the morphosyntactic constraints of the Matrix Language Frame model. *Open Linguistics*, *5*(1), 706- 728.

Azuma, S. (1993). The frame-content hypothesis in speech production: Evidence from intrasentential code switching. *Linguistics*, *31*(6), 1071-1094.

Hadei, M., & Ramakrishna, R. A. (2017). English single content morphemes in Persian structure: Applying the Matrix Language Frame Model (MLF) and 4M Model. *International Journal of Bilingualism*, 21(4), 433-453.

Hebblethwaite, B. (2010). Adverb code-switching among Miami's Haitian Creole–English second generation. *Bilingualism: Language and Cognition*, *13*(4), 409-428.

Hole, Y., & Snehal, P. (2019). Challenges and solutions to the development of the tourism and hospitality industry in India. African Journal of Hospitality, Tourism and Leisure. 8 (3), 1-11

Hole Y., Hole S.P., & Wagh.V. (2019). Omni channel retailing: an opportunity and challenges in the Indian market. Journal of Physics: Conference Series, 1362 (2019), 1-12.

Hole, Y., & Snehal, P. & Bhaskar, M. (2018). Service marketing and quality strategies. Periodicals of engineering and natural sciences,6 (1), 182-196.

Jake, J. (1994). Intrasentential code-switching and pronouns: On the categorial status of function elements. *Linguistics, 32*, 271–298.

Jake, J. L., & Myers-Scotton, C. (1997). Code-switching and compromise strategies: Implications for lexical structure. *International Journal of Bilingualism, 1(1)*, 25–39.

Jake, J. L., & Myers-Scotton, C. (2002). Second generation shifts in sociopragmatic orientation and codeswitching patterns. In A. Rouchdy (Ed.), *Language contact and language conflict in Arabic. Variations on a sociolinguistic theme* (pp. 317–330). Routledge.

Jake, J. L., & Myers-Scotton, C. (2009). Which language? Participation potentials across lexical categories in code-switching. In L. Isurin, D. Winford, & K. de Bot (Eds.), *Multidisciplinary approaches to code switching* (pp. 207–242). John Benjamins Publishing

Joshi, A. (1982). Processing of sentences with intra-sentential code-switching. In *Coling 1982: Proceedings of the Ninth International Conference on Computational Linguistics*.

Kniaź, M., & Zawrotna, M. (2020). Embedded English verbs in Arabic-English code-switching in Egypt. *International Journal of Bilingualism*, 1367006920976909.

Labov, W. (1972). *Sociolinguistic patterns* (No. 4). University of Pennsylvania Press.

MacSwan, J. (2014). *A minimalist approach to intrasentential code switching*. Routledge.

Malmkaejar, K. (1991) The Linguistic Encyclopaedia. London and New York: Routledge.

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

Muysken, P. (2011). *Codeswitching*. In R. Mesthrie (Ed.), The Cambridge handbook of sociolinguistics (pp. 301–314). New York: Cambridge University Press.

Myers-Scotton, C. (1993). Common and uncommon ground: Social and structural factors in codeswitching. *Language in society*, 475-503.

Myers-Scotton, C. (1997). *Duelling languages: Grammatical structure in codeswitching*. Oxford University Press.

Myers-Scotton, C. (2005). *Multiple voices*. Blackwell publishing.

Myers-Scotton, C. M., & Jake, J. L. (2017). Revisiting the 4-M model: Codeswitching and morpheme election at the abstract level. *International Journal of Bilingualism*, *21*(3), 340-366.

Myers-Scotton, C., & Jake, J. (2009). A universal model of code-switching and bilingual language processing and production. In B. E. Bullock & A. J. Toribio (Eds.), *The Cambridge handbook of linguistic codeswitching* (pp. 336–357). Cambridge University Press.

Myers-Scotton, C., & Jake, J. L. (2014). Nonfinite verbs and negotiating bilingualism in code- switching: Implications for a language production model. *Bilingualism: Language and Cognition*, 17(3), 511–525.

Namba, K. (2012). Non-insertional code-switching in English–Japanese bilingual children: alternation   and congruent lexicalisation. *International journal of bilingual education and   bilingualism*, *15*(4), 455-473.

Poplack, S., Wheeler, S., & Westwood, A. (1989). Distinguishing language contact phenomena: evidence from Finnish-English bilingualism. *World Englishes*, *8*(3), 389-406.

Schmidt, R. L. (1999). *Urdu, an essential grammar*. Psychology Press.

Singh, R. (1989). Bilingualism, Language Contact, and Morphological Theory. *Language and Society: Steps Towards an Integrated Theory*, *27*, 72.

Youssef, I. (2016). English-Cairene Arabic classroom code switching: An interactional-sociolinguistic   approach. *International Journal of Arabic-English Studies*, 16, 7–28.

Ziamari, K. (2007). Development and linguistic change in Moroccan Arabic-French code-switching. In   C. Miller, E. Al-Wer, D. Caubet, & J. C. E. Watson (Eds.), *Arabic in the city: Issues in dialect contact and language variation* (pp. 275–290). Routledge.

Author Information

Fatima Tu Zahra	**Farhat Abdullah**
M.Phil Scholar, Department of English, University of Central Punjab Lahore Pakistan	Lecturer in English, Department of English, University of Central Punjab Lahore Pakistan
Dr Tahir Saleem	**Dr Muhammad Arif Khan**
Assistant Professor, Department of English, University of Central Punjab Lahore Pakistan	Assistant Professor