

Incorporating Verbs in Code-Switching: Insights from The Matrix Language Frame Model

Sami Saad Alghamdi, Ehab Saleh Alnuzaili¹, Hamdi Abdel Rasool Eljeam & Abdulbasit A. Alhaj
King Khalid University, Saudi Arabia

Nazir Ahmed Malik & Asad Ali
The University of Chenab, Gujrat, Pakistan

Abstract: This study investigates the incorporation of non-finite verbs in code-switching (CS), employing the Matrix Language Frame (MLF) model as its theoretical framework (Myers-Scotton & Jake, 2014; 2017). Drawing on both naturalistic and elicited data, this study examines how English non-finite verbs resist integration in the matrix language due to their complex selectional properties and morphosyntactic frames. Contrary to the 4-M model's classification of these verbs as content morphemes, this study identifies them as hybrid morphemes—linguistic items that possess $[\pm\theta]$ features and exhibit both lexical and functional behavior. Based on these findings, this study proposes a refined morpheme categorization within the MLF model, comprising three layers: content, system, and hybrid morphemes. In summary, the study concludes that not all non-finite verbs are equally incorporable and that the MLF model requires revision to accommodate hybrid morphemes. Furthermore, the findings of the study have pedagogical implications for teachers, emphasizing the need to recognize bilingual syntactic patterns, support authentic code-switching practices, and design curricula that reflect the natural linguistic behavior of multilingual learners. These insights offer a valuable understanding of bilingual competence and language development in multilingual communities.

Keywords: Incorporation, code-switching, content morphemes, $[\pm\theta]$ features, multilingual

Code-switching (CS), a dynamic phenomenon emerging from contact language, has garnered extensive scholarly attention across linguistics (e.g., Cedden et al. (2024), Muntendam et al. (2024), Rayo et al. (2024), and Zhong & Fan, (2023)), sociocultural (e.g., Munir et al. (2025) and Yim & Clément, (2021)), grammatical, (e.g. Ali et al. (2021a) and Jabbar et al. (2021)) and educational domains (e.g. Ellison & Si (2021) and Nguyen et al. (2022)). It has been examined through various angles, including borrowing (Treffers-Daller, 2025), code-mixing (López et al., 2017), and syntactic integration (Ali et al., 2020; 2021b; Alnuzaili et al. (2024); Chan, 2008; MacSwan, 2000; Sahib et al., 2021). Within multilingual contexts such as Pakistan, where Urdu

¹ Corresponding Author: an Assistant Professor, Dr. Ehab Saleh Alnuzaili, English Department, Applied College, King Khalid University, Saudi Arabia. E-mail; Ihabsaleh2023@gmail.com

and English coexist in formal and informal discourse, code-switching is not only a linguistic strategy but also a mirror image of sociopolitical identity and educational practices (Liu et al., 2024; Rababah, 2023; Soffer-Vital & Finkelstein, 2024; Vu et al., 2025). Despite its prevalence, the grammatical mechanisms underlying code-switching—typically verb incorporation—have remained underexplored.

The Matrix Language Frame (MLF) model (Myers-Scotton & Jake, 2014, 2017) facilitates a robust theoretical framework for analyzing how morphemes from diverse languages interact in bilingual utterances (Aburqayiq et al., 2025). According to this model, morphemes are conceptualized as packets of semantic and pragmatic meanings, distributed asymmetrically between the matrix language (ML) and embedded language (EL). Verbs, as the core of predicate argument structure (Myers-Scotton, 1997), play a pivotal role in constructing the grammatical integrity of code-switched sentences. Non-finite verbs—particularly infinitive, bare forms, and participials—are typically incorporated via direct inflectional morphology or through the do-construction strategy. However, empirical data from Urdu-English code-switching reveals derivations from established strategies, raising questions and challenging the applicability of the MLF model in this context.

Understanding non-finite verb syntax in Urdu-English code-switching is vital for shaping bilingual education, curriculum development, and language standardization policies (Mustafa et al., 2022; De Soete, 2025). In addition, it contributes to broader debates on language change, language shifting, and hybridization in postcolonial societies, where English largely functions as a language of power and prestige (Almasaeed, 2023; Morton, 2014).

Accordingly, this study aims to investigate the incorporation of English non-finite verbs in Urdu-English code-switching (CS), utilizing the MLF model as its theoretical framework (Myers-Scotton & Jake, 2014, 2017). By scrutinizing naturalistic and elicited data, the study seeks to evaluate the validity of the model and uncover the morphosyntactic properties that inhibit or facilitate verb incorporation (González-Vilbazo, & López, 2011, 2012).

Research Questions

RQ: 1. How are English non-finite verbs incorporated into Urdu-English code-switched utterances?

RQ: 2. To what extent do the selectional and morphosyntactic properties of Urdu-English non-finite verbs affect their incorporation into the matrix language?

RQ: 3. Does the Matrix Language Frame (MLF) model adequately account for the behavior of non-finite verbs in Urdu-English code-switching (CS)?

Literature Review

Spencer (1991) defined incorporation as “a word (typically a verb) forms a kind of compound with, say, its direct object, or adverbial modifiers, while retaining its original syntactic function” (p. 15). Baker (2009) opined that incorporation is a syntactic operation in which a lexical head—particularly a verb—moves within the clausal structure. Verb incorporation in code-switching involves morphological process related to the word formation and derivational morphology (Booij, 2010; Okeke & Okeke, 2022). Wohlgemuth (2009) also addressed typological aspects of borrowings. Unlike borrowing, code-switching is a language contact phenomenon that syntactically, semantically, and phonologically integrates a linguistic category from language X_L

into language Y_L often driven by social preferences. The asymmetric incorporation of two languages at the clausal level is considered a form of code-switching (Alnuzailli et al., 2025).

Code-switching (CS) refers to the juxtaposition or alternation of linguistic items from two distinct grammatical systems within a single sentence or discourse (see Baird et al., 2023; Balam et al., 2024; Gumperz, 1977; Kudjo Agama et al., 2023; Poplack, 1980; Poplack & Sankoff, 1981, for an overview). Scholars have conceptualized code-switching (CS) as a dynamic bilingual strategy, with Grosjean (1989) proposing bilinguals as “two monolinguals in one person,” and Weinreich (1953) emphasizing situational appropriateness in language switching. Building on these fundamental views, Myers-Scotton (1993; 2002; 2005) introduced the concept of *bilingual linguistic competence* (BLC) and later proposed the Matrix Language Frame (MLF) model (2000; 2014; 2017) which has been widely utilized to explain borrowing, code-switching, and code-mixing (Deuchar, 2020).

Linguists have investigated diverse grammatical and phonological aspects of code-switching (CS), including case-marking, phi-feature, word-order, coordination, and question-marking (henceforth q-marking) (Ashraf et al., 2021; 2025; Cable, 2007; Den Dikken, 2011; Ilyas et al., 2023). While extensive research exists on Spanish-English (Gosselin, 2022; MacSwan, 2000; 2020), Persian-English (Rahimi & Dabaghi, 2013), Pashtu-English (Khan et al., 2023), and Swahili-English (Myer-Scotton & Jake, 2014; 2017) CS, studies focusing on Urdu-English CS remain limited. Existing research studies (Alnuzailli et al., 2025; Ali et al., 2020; 2021b; 2023b; Jabbar et al., 2021; Malik, 2017; Munir et al., 2025) have not particularly addressed the role of non-finite verb forms, which are significant in constructing the abstract morphosyntactic frame that provides speakers with intent. However, this study aims to permeate that research gap by investigating the integration of non-finite verb forms in Urdu-English code-switching (CS) utilizing the Matrix Language Frame (MLF) model (Myer-Scotton & Jake, 2014; 2017).

According to Gallego (2010), “non-finite verbs had a hybrid categorical status. Even though they could not resort a technical apparatus at the time, they had the clear intuition that infinitives, gerunds, and past participles should be regarded as ‘verbal nouns’, ‘verbal adverbs’, and ‘verbal adjectives’ respectively” (p. 81). Later, Myers-Scotton and Jake (2017) asserted that matrix language (ML) finite verb frames in code-switching accommodate embedded language non-finite verbs, as these forms do not require interactions at the level of the formulator to activate their full salience, unlike finite verbs.

Assuming Gallego’s (2010) characterization of non-finite verbs as hybrid in nature, we posited that the lack of integration of non-finite verbs—such as those—into the ML frame necessitates a recalassification of morphemes within the 4-M model (Myers-Scotton & Jake, 2017). The categorization of morphemes has been conceptualized through two major theoretical frameworks: constraint-based theories, which distinguish between content and system morphemes (Belazi et al., 1994; Joshi, 1982; Maqsood et al., 2019; Myers-Scotton & Jake, 2014, 2017; Poplack, 1980; Poplack & Sankoff, 1981), and constraint-free theories, which differentiate between lexical and functional items (Gosselin, 2022; Mahootian & Santorini, 1996; MacSwan, 2000).

Myers-Scotton and Jake (2014) identified key content morphemes—non-finite, bare, infinitive, and participle—and investigated their incorporation in code-switching across diverse language pairs, including Turkish–Dutch (Backus, 1998), Finnish–English (Halmari, 1997), Xhosa–English (Myers-Scotton & Jake, 2017), Swahili–English (Myers-Scotton & Jake, 2000; 2014), Chiluba–French (Kamwangamalu, 1994), Ewe–French (Amuzu, 2011), Moroccan Arabic–

French (Bentahila & Davies, 1992), Persian-English (Mahootian & Santorini, 1996; Rahimi & Dabaghi, 2013), Arabic-English (Mustafa et al., 2022; Aburqayiq et al., 2025), Urdu-English (Alnuzailli et al., 2025; Malik, 2017; Munir et al., 2025; Si, 2011), Pashto-English (Khan et al., 2023), Spanish-English (Gosselin, 2022; González-Vilbazo & López, 2011; 2012; MacSwan, 2000; Pfaff, 1979), and Acholi-English (Myers-Scotton, 2005). These studies commonly employed the Matrix Language Frame (MLF) model (2000; 2014; 2017), utilizing role of *ML inflectional morphology* and *do-verb constructions* in syntactic integration. It is widely accepted that such verb forms impose minimal cognitive processing, facilitating their incorporation from embedded language (EL) into the matrix language (ML) frame (Pfaff, 1979). Do-verbs in the ML are theorized to select these content morphemes as complements, lubricating integration—an empirical intuition supported in Urdu borrowing (Maqsood et al., 2019). Indispensably, ML inflection must attach to do-verbs in the line with the System Morpheme Principle (SMP) and the Morpheme Order Principle (MOP), Urdu-English code-switching (CS) frequently exhibiting the syntactic applicability and functional significance of these forms (Myers-Scotton & Jake, 2017). Consider example (1a-1d).

- 1a. Some Englishmen traditional Indian women-ko **passand** karaten hain.
 Some Englishmen traditional Indian women-ACC like do are
 “Some Englishmen like traditional Indian women” (Myers-Scotton & Jake, 2017, p. 9)
- 1b. Some Englishmen traditional Indian women-ko **like** karaten hain.
 Some Englishmen traditional Indian women-ACC like do are
 “Some Englishmen like traditional Indian women” [Intended Expression]
- 1c. *Some Englishmen traditional Indian women-ko **pssand**-ten hain.
 Some Englishmen traditional Indian women-ACC like do are
 “Some Englishmen like traditional Indian women” [Intended Expression]
- 1d. *Some Englishmen traditional Indian women-ko **like**-ten hain.
 Some Englishmen traditional Indian women-ACC like do are
 “Some Englishmen like traditional Indian women” [Intended Expression]

The examples (1a-1d) highlight the *incorporation* of non-finite verbs into Hindi-English code-switching (CS). In (1a-1d), non-finite verbs (bare forms) are presented in boldface, *do-verbs* are underlined, and inflectional morphology is indicated in italicized. Example (1a), adapted from Myers-Scotton and Jake (2017), demonstrates the successful integration of the Hindi verb *pssand* (equivalent to English ‘like’) into the Hindi morphosyntactic frame, fulfilling both the System Morpheme Principle (SMP) and the Morpheme Order Principle (MOP). The substitution of *pssand* for *like* does not disorganize the syntactic integrity of the English verb phrase, thereby supporting the theoretical allegation that content morpheme—specifically non-finite forms—can be inserted with low cost and high syntactic compatibility into the matrix language (ML) frame.

In comparison, examples (1c-1d) as delineated by Maqsood et al. (2019) are ungrammatical due to the violation of the Grammatical Constraint (GC). Although light verbs play a significant role in sustaining grammaticality within code-switching (CS), this paradigm is not universally and consistently implementable across all forms of verbs. Specifically, bare forms of *passand* and *like* in examples (1c-1d) fail to incorporate directly into the ML inflectional structure (ten), resulting in syntactic dissonance. This discrepancy accentuates the variability in verb incorporation, which

is influenced by the *argument structure* of verbs (López et al., 2017). Different language theories have asserted their arguments on code-switching (CS).

Language Switching Theories

Language-switching theories are broadly categorized into constraint-based and constraint-free approaches (MacSwan, 2000). Constraint-based theories posit that language-specific constraints govern code-switching (CS), deriving syntactic well-formedness (see Belazi et al., 1994; Joshi, 1982; Maqsood et al., 2019; Poplack, 1980; Poplack & Sankoff, 1981, for an overview of constraint-based theories). Among these, the Free Morpheme Constraint (FMC) and Equivalence Constraint (EC) distinguish between free morphemes and bound morphemes, suggesting that code-switching (CS) is permissible only at points where the surface structure of both languages align. Joshi (1982) introduced the Close Class Constraint (CCC), classifying morphemes into open class (e.g. nouns, verbs, and adjectives) and closed class (e.g. inflectional morphemes, clitics, tense markers) categories. Belazi et al. (1994) further proposed the Functional Head Constraint (FHC) for code-switching (CS), dividing linguistic items into two categories: lexical and functional, each characterized by properties such as theta marking, case checking, and phi-features. Myers-Scotton and Jake (2000, 2014, 2017) advanced a “frequency-based model” of bilingual linguistic competence (BLC), later developed the 4-M model—an extension of the Matrix Language Frame (MLF) model (Jake & Myers-Scotton, 1997)—which categorizes morphemes into four types: content morphemes, bridge system morphemes, early system morphemes, and late outsider system morphemes.

However, the 4-M model focuses greater emphasis on system morphemes than content morphemes. Maqsood et al. (2019) refined this categorization by demarcating between lexical vs. fossil material, analogous to the lexical-functional dichotomy. They opined that lexical items necessitate fossil dummy material to map the grammatical structure across interface levels—phonological form (PF) and logical form (LF)—and proposed that both lexical and fossil elements contribute to the construction of the morphosyntactic frame. These theories (Belazi et al., 1994; Joshi, 1982; Maqsood et al., 2019; Myers-Scotton & Jake, 2014; 2017; Poplack, 1980; Poplack & Sankoff, 1981) apply to data on various languages. On the other hand, constraint-free theories, which differentiate between lexical and functional items (Gosselin, 2022; Mahootian & Santorini, 1996; MacSwan, 2000; Malik, 2017; Van Gelderen & MacSwan, 2008), assert that “no special mechanism” is essential to regulate code-switching (CS) or any contact language phenomena. The most prominent model of null theories is the minimalist model of MacSwan (2000), which efficiently exhibits bilingual linguistic competence.

Related Studies

Pursuing an examination of fundamental conceptualization in code-switching (CS) and non-finite verb incorporation, it is pertinent to critically examine prior research on verb integration within Urdu-English code-switching (CS). Maqsood et al. (2019) explored *borrowing* of nouns and verbs within the theoretical frame of the Minimalist Program (Chomsky, 2014). Their findings indicated that nouns could be freely borrowed into English through the do-construction verb strategy, while borrowing verbs was subject to grammatical constraint. This approach reflects a constraint-based theoretical stance, which arguably conflicts with the generative grammar

principle known as the third-factor design of language—an interface condition that promotes economy and efficacy in linguistic computation (Chomsky, 2005; Cai et al., 2024).

In contrast, Ali et al. (2021b) accommodated the third-factor language design in their evaluation of Urdu-English code-switching (CS), emphasizing the obligatory presence of the close-class item, *keh* (“that”) in Urdu-English code-switching (CS). They opined that the MLF model is redundant for accounting for Urdu-English code-switching (CS), particularly because *keh* (“that”) in Urdu and Hindi functions flexibly to restrict and mark the mood of the embedded clause.

According to Myers-Scotton and Jake (2014, 2017), morpheme external to maximal projection—such as *keh* (“that”) is classified as a late outsider system morpheme and must be supplied by the matrix language in code-switching (CS) (Sankoff, 1998; Subiyanto et al., 2024; Van Gelderen & MacSwan, 2008). These observations suggest that the current classification of morphemes within the MLF model permits reconsideration, as multiple empirical instances demand and challenge its exploratory adequacy (Chomsky, 2005; 2014).

Munir et al. (2025) conducted a study on Urdu-English code-switching (CS), particularly focusing on the frequency patterns and grammatical functions of inflected morphemes. Their analysis identified nine distinct inflectional morphemes that were prominently used within the Urdu-English bilingual corpus. These inflected morphemes served specific grammatical functions, primarily related to agreement. Two patterns of tense agreement were observed: short-distance agreement (SDA) and long-distance agreement (LDA). The LDA pattern was specifically influenced by the ergative marker, *-ne*, which inhibited nominative case agreement and facilitated cross-clausal alignment. Furthermore, the empathetic marker *-di* emerged as both pragmatically and syntactically salient, underscoring its relevance in bilingual discourse and grammatical construction (Fillmore, 1967).

Previous studies (e.g., Ali et al., 2020, 2021a; 2021b; Butt, 2014; Shim, 2016) have investigated various dimensions of Urdu-English code-switching (CS), offering a broad and diverse scope of inquiry. However, the integration of morphemes—particularly non-finite verb forms—into code-switching (CS) remains under-investigation within the Urdu-English bilingual context. While Myers-Scotton and Jake (2014; 2017), Maqsood et al. (2019), and Munir et al. (2025) have contributed valuable instantiations into Urdu-English code-switching (CS) patterns, they did not particularly address the integration of non-finite verb forms into bilingual constructions. This research gap exhibits the need of the hour for further research focused on the syntactic and morphological behavior of non-finite verbs in Urdu-English code-switching (CS) due to rich Urdu morphology (see Ashraf et al., 2025; Ali et al., 2025).

Matrix Language Frame (MLF) model (Myers-Scotton & Jake, 2000, 2014, and 2017)

In bilingual competence, two linguistic systems—referred to as X_L (Embedded Language) and Y_L (Matrix Language)—are integrated within bilingual speech. The Matrix Language Frame (MLF) model, developed by Myers-Scotton (1993), offers both descriptive and explanatory powers to account for how language is accessed and retrieved prior to its final surface realization (Myers-Scotton 1993, p. 45). The model articulates an asymmetrically uniform internal structure in bilingual utterances, distinguishing between the participating languages and categorizing morphemes into two primary types: content morphemes and system morphemes. Subsequently, Myers-Scotton & Jake (2017) refined the classification of morphemes into four subtypes: content morphemes, early system morphemes, bridge system morphemes, and late outsider system morphemes.

The MLF model predicts that bilingual utterances universally conform to a uniform structural pattern governed by three core principles. First, the Uniform Structure Principle (USP) asserts that “a given constituent type in any language has a uniform abstract structure, and the requirements of well-formedness for this type must be observed whenever the constituent appears. In bilingual speech, the structures of the Matrix Language are always preferred. However, some embedded structures [i.e., Embedded Language islands as discussed below] are allowed “if Matrix Language clause structure is observed” (Myers-Scotton, 1993, p. 83). Second, the System Morpheme Principle (SMP) stipulates that in Matrix Language + Embedded Language constituents, all system Morphemes with grammatical relations external to their head constituent—those participating in the sentence’s thematic role grid—must originate from the Matrix language. Third, the Morpheme Order Principle (MOP) maintains that in mixed constituents containing singly occurring Embedded Language lexemes and any number of Matrix Language morphemes, surface morpheme order must reflect the syntactic structure of the Matrix Language (Myers-Scotton, 1993). The MLF model operates on the three core principles stated previously, central to understanding bilingual code-switching (CS). The dominant and most active language in a bilingual clause is referred to as the Matrix Language (ML), which provides a *morphosyntactic* frame for the bilingual clause (Jabbar et al., 2021). In contrast, the Embedded Language (EL) plays a structurally subordinate role, supplying only *content items*—such as nouns, non-finite verbs, infinitives, and participles—that fill syntactic slots specified by the ML.

Content morphemes are defined as those that assign or receive thematic roles (theta roles). In syntactic theory, these morphemes carry semantic and pragmatic features and are typically assigned by verbs within discourse. A verb projects its arguments within a constituent and selects Determiner Phrases (DPs) that function as *agent, theme, or patient* (Ali et al., 2023a; 2023b). These arguments are referred to as theta roles. Accordingly, nouns and adjectives receive theta roles, while verbs assign them, thus qualifying as content morphemes. In contrast, system morphemes do not assign and receive theta roles. All morphemes other than nouns, verbs, and adjectives—such as functional words, bound and free morphemes, plural markers, and derivational affixes—are categorized as system morphemes.

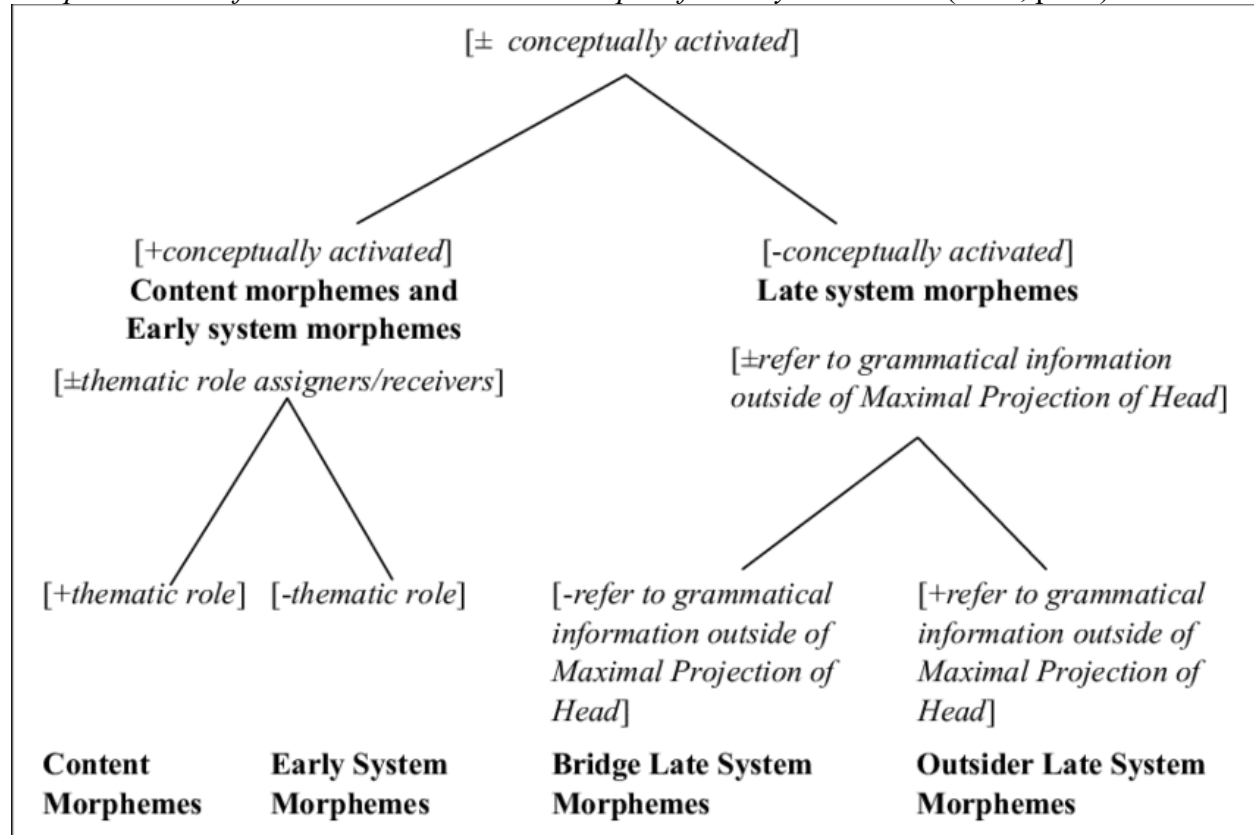
The 4-M model (Myers-Scotton, 2002; 2005; Myers-Scotton & Jake, 2000; 2014, 2017) extends the MLF model (Myers-Scotton, 1993; Jake & Myers-Scotton, 1997) by offering a more nuanced classification of morphemes based on their roles, incorporation, and accessibility in language production. It introduces the “Differential Access Hypothesis,” which posits that different types of morphemes are accessed at different levels of the production process. Particularly, content morphemes (CMs) are early system morphemes (SMs) are accessed at the level of the mental lexicon, whereas late SMs project only at the “level of the formulator” (Myers-Scotton, 2002, p.78).

Early SMs are conceptually activated at the mental lexicon but differ from content morphemes in that they do not assign and receive theta roles. These include bound morphemes and affixes such as determiners and plural markers. Although early SMs do not directly select content morphemes, they are structurally dependent on them, hence their classification as system morphemes. Bridge system morphemes facilitate the construction of syntactically larger constituents. They may project within their maximal projection, whereas outsider system morphemes rely on information external to the projection in which they structurally occur. Example bridge SMs include English *of* and possessive *’s* as well as Urdu *ka* and *k*, which help form larger phrases from content morphemes.

Late outsider system morphemes differ from bridge SMs in that their presence and form depend on information from other constituents or discourse elements. For instance, subject–verb agreement is realized through late outsider late SMs. According to the 4-M model (Myers-Scotton, 2002, 2005; Myers-Scotton & Jake, 2000, 2014, 2017), these morphemes are exclusively inserted by the ML, adhering to the model’s core principles. Nonetheless, the phenomenon “Embedded Language (EL) islands” may occur. These islands are maximal projections composed entirely of EL elements and are “not inflected with ML system morphemes, even though they occupy syntactic positions projected by the ML.

Figure 1

Morpheme Classification in the 4-M model adapted from Myers-Scotton (2002, p. 73).



Research Methodology

Research Design

This study adopts a *qualitative research design*, which is appropriate given the exploratory nature of the investigation into verb incorporation in Urdu-English code-switching (CS). This design aligns with previous studies in bilingualism, multilingualism, and sociolinguistics (Ali et al., 2020; Khan et al., 2023; Park, 2025), supporting the nuanced analysis necessary for understanding morphosyntactic patterns in bilingual constructions.

Participants

Participants were selected from 200 undergraduate students at the university level. Based on the study's inclusion criteria, only 100 balanced Urdu-English bilingual speakers were invited to participate in the study and were equally recruited based on gender (see Table 1). The Balanced Bilingual Scale (MacSwan, 2000; 2020) was employed to ensure that participants demonstrated fluency and proficiency in both languages. All participants had received education in elite-class institutions, indicating a high level of exposure to both Urdu and English in academic and social contexts. This demographic was determined to ensure linguistic competence and to reflect the bilingual speakers relevant to the current study's desiderata.

Table 1

Demographic of Participants

Sr. #	Number of participants	Sex	Age	First Language	Second Language	Level of Education
1	50	Male	21-26 years	Urdu	English	Undergraduate
2	50	Female	21-26 years	Urdu	English	Undergraduate
Total	100	-	-	-	-	-

Data Collection

Data were gathered through multiple sources to ensure rich and authentic manifestations of bilingual language use in natural settings (Poplack, 1980; Sankoff, 1998). These sources include naturalistic expressions, participants' observations, and social discourse (see Table 2). Naturalistic expressions were gathered using audio recordings as a research instrument from balanced bilinguals in natural, formal, and informal settings. Data were collected using participant observation as a research tool during classroom interactions (Aburqayiq et al., 2025). Furthermore, data were collected from academic peers using social discourse as a research instrument (Grosjean, 1989; Liu et al., 2024; Ninio, 2011). To ensure reliability and validity, all linguistic data used in analysis were counter-verified by two independent experts—professionally trained linguists and fluent Urdu-English bilinguals.

Table 2

Data of Urdu-English Bilingual Corpus

Sr. #	Type	Duration/Unit
1	Number of Recording Hours	4 Hours 38 minutes
2	Number of Urdu Expressions	970 Expressions
3	Number of English Expressions	670 Expressions
4	Number of code-switched Expressions	1648 Expressions
5	Number of verb code-switched Expressions	280 Expressions
6	Number of total groups	2 Groups
7	Number of participants in each Group	50 Members
8	Number of moderators	2 Moderators
9	Number of experts	2 Experts

In addition, elicited data were collected using the substitution method, wherein another replaced one morpheme from the same grammatical class but in a different language was substituted. This method permitted controlled manipulation of morphemes to examine code-switching (CS) behavior and verb incorporation patterns (Cedden et al., 2024; Nguyen et al., 2022). Participants were classified into two groups, each supported, guided, and supervised by a designated observer. These collected recordings were later transcribed into textual form for analysis. This data collection process was instructed and guided by the Matrix Language Frame (MLF) model (Myers-Scotton, 2002; 2005; Myers-Scotton & Jake, 2017), which supported theoretical foundations for identifying morphosyntactic frames and the behavior of non-finite verb forms in Urdu-English code-switching.

Limitations of the Study

The study presents several limitations that should be acknowledged. First, the analysis of this study is restricted to a single language pair—Urdu-English—which limits the generalizability of the findings to other bilingual utterances. Second, the study exclusively engages with constraint-based models, such as the 4-M model (Myers-Scotton & Jake, 2014, 2017), while excluding constraint-free models that may offer alternate insights into code-switching behavior. Third, the participants were regionally confined, and in addition, the use of the 4-M model assumes an asymmetrical structure in language switching, where one language contributes as the matrix and the other is embedded, potentially overlooking more symmetrical bilingual interactions (Niaz & Ali, 2023).

Ethical Considerations

The study was approved by the Institutional Review Board (IRB) under reference number BoS78th. Prior to data gathering process, informed written consent was obtained from participants. To ensure confidentiality, all personal details were anonymized. During data collection, no participants were forced for participation, no gender discrimination was conducted while recruiting participants.

Data Analysis and Findings

Data analysis was conducted using the Matrix Language Frame (MLF) model (Myers-Scotton, 2002, 2005; Myers-Scotton & Jake, 2000, 2014, 2017), which offers both descriptive and exploratory power in understanding bilingual competence. The three core principles—Uniformity, Morpheme Order Principle (MOP), and System Morpheme Principle (SMP)—were utilized to analyze the structure and function of morphemes in Urdu-English code-switching utterances. The designed framework was based on the MLF model to classify morphemes into content and system morphemes. This framework was refined by the “Differential Access Hypothesis” (Myers-Scotton & Jake, 2017), which posited that different morphemes are accessed at different levels of the language production process. The substitution method was facilitated as a diagnostic tool to examine the incorporation of non-finite verb forms into the ML morphosyntactic frames and compatibility across languages.

7. Hum police-ko apna name **btaa-te** hain.
 We police-DAT our name tell-INF be. AUX. FUT
 “We tell the police our name” [Direct Inflectional Morphology]
8. Alia pizza **khaa-ti** hai.
 Alia pizza eat-INF be. AUX. PRS
 “Alia eats pizza” [Direct Inflectional Morphology]

In example (6), the Urdu non-finite verb **aa** (“come”) is incorporated, while example (7) features **btaa** (“tell”), also from the Urdu language in Urdu-English code-switching (CS). These examples (6-8) illustrate Urdu-English code-switching (CS), where the matrix and embedded languages interact in an irresistible and mutually reciprocal manner. The matrix language (Urdu) maintains structural dominance, while the embedded language supplies lexical items—functioning as “raw bricks” in constructing a unified morphosyntactic linearization, consistent with the System Morpheme Principle (SMP) and Morpheme Order Principle (MOP). In example (8), the Urdu verb--**khaa** (“eat”) is directly inflected with the matrix language morphology, *-ti* (“Habitual Aspect”), consolidating the structural uniformity of code-switching (CS). These instances indicate that the content morphemes are seamlessly inflected with the matrix language frame, preserving grammatical integrity and supporting cross-linguistic integration.

Integrating Non-finite Verbs into Urdu-English Code-Switching Through Do-verb Strategy

This section demonstrates the incorporation of verbs into Urdu-English code-switching (CS), particularly when they conform to the do-verb strategy proposed by Myer-Scotton & Jake (2014, 2017) under the Differential Accessibility Hypothesis. Consider instance in (9-14)

9. Maou-ka allmi day aaj **celebrate** ki-aa gi-aa.
 Mothers-ERG international day today celebrate LVB-INF be. AUX. FUT
 ‘International mother day was celebrated today.’ [Do-verb Strategy]
10. Alia Ali-ko **pasand**-kar-ti hai.
 Alia Ali-ACC like LVB-INF be. AUX. PRS
 “Alia likes Ali” [Do-verb Strategy]
11. Children launch **share** kar-tee hain.
 Children launch share LVB-INF be. AUX. PRS
 “Children share the launch.” [Do-verb Strategy]
12. Aqsa quran-e-pak **recite** kar-ti hai.
 Aqsa quran-e-pak recite LVB-INF be. AUX. PRS
 “Aqsa recites the holy quran.” [Do-verb Strategy]
13. People masjid-me **enter** ho-tee hain.
 People mosque-LOC enter LVB-INF be. AUX. PRS
 “People enter into the mosque.” [Do-verb Strategy]
14. Humra airplane ajj suba hi **arrive** ho-aa.
 Our airplane today morning also arrive LVB-INF.
 “Our airplane also arrived today’s morning.” [Do-verb Strategy]

The data presented in examples (9-14) were collected from naturalistic settings and reflect utterances produced by Urdu-English bilinguals. The non-finite verb forms—**celebrate** (9), **pasand** (10), **share** (11), **recite** (12), **enter** (13), and **arrive** (14)—employ the do-verb strategy, wherein inflection is partially affixed to the light verb but not with auxiliaries. *-kar* light verb exhibits agentive nature, and it introduces agent at the external argument positions (9-12). Conversely, *-ho* light verb is ergative, which does not potentially introduce agent arguments but partially licenses eventive arguments (13-14). Across examples (9-14), the same do-verb strategy is consistently noticed. In Urdu, two light verbs—*kar* and *ho*—are commonly employed to support verb incorporation (Bandi-Rao & Den Dikken, 2014; Ellison & Si, 2021; Shim, 2016; Si, 2011). This section also illustrates how verbs are incorporated into Urdu-English code-switching (CS) when they follow do-verb strategy sketched by Myers-Scotton and Jake (2014, 2017) under the Differential Accessibility Hypothesis. This strategy facilitates the integration of English content verb forms—infinitive, bare, participle, and non-finite—into the morphosyntactic frame of the matrix language (Urdu), maintaining grammaticality and structural coherence (Ali et al., 2021b).

Integrating Non-finite Verbs in Urdu-English Code-Switching Through Elicited Data

This section shows elicited data to evaluate the proposition of verb incorporation in Urdu-English code-switching (CS). As outlined in the methodology, a substitution method was applied in the elicited dataset. To tacitly test linguistic items, both naturalistic and elicited data were utilized. In the preceding section, the naturalistic dataset was analyzed using the Matrix Language Frame (MLF) model (Myers-Scotton & Jake, 2014; 2017), which demonstrated compatibility with Urdu-English code-switching (CS). Accordingly, the model should generate similar sentence types when applied to the elicited data, adhering to the same principles. To test the model's validity, only counter-language verbs were substituted, based on the pretence that non-finite verbs—such as infinitive, bare forms, and participles—are “free” and low-cost morphemes, similar to nouns (Myer-Scotton & Jake, 2014; 2017). These forms should integrate seamlessly into the morphosyntactic frame of the matrix language (Balam et al., 2023; Baird et al., 2023; Balam et al., 2024; Jabbar et al., 2021; Khan et al., 2023).

15a. Alia pasta **khaa-tii** hai.

Alia pasta eat-Inf be. Aux

“Alia eats pasta.”

[Natural Expression]

15b. *Alia pasta **eat-tii** hai.

Alia pasta eat-Inf be. Aux

“Alia eats pasta.”

[Elicited Expression]

16a. Chokidar room-me **soo-ta** hai.

Guard room-in sleep-Inf be. Aux

“Guard sleeps in the room.”

[Natural Expression]

16b. *Chokidar room-me **sleep-ta** hai.

Guard room-in sleep-Inf be. Aux

“Guard sleeps in the room.”

[Elicited Expression]

17a. Zoya letter **likh**-*ti* hai.

Zoya letter write-Inf be. Aux

“Zoya writes a letter.”

[Natural Expression]

17b. *Zoya letter **write**-*ti* hai.

Zoya letter write-Inf be. Aux

“Zoya writes a letter.”

[Elicited Expression]

Two patterns were used: the first example illustrates a natural expression, while the second substitutes only the verb with its English counterpart within the same sentence structure. This approach revealed two key instantiations. First, although argument structure is often considered “universal” and events embedded under control or raising predicates are assumed to be loosely connected to the matrix verb (Fillmore, 1967; Butt, 2014), the elicited example (15b) challenges this assumption. It shows that events are not governed by Urdu inflection, and the bare verb form diverges from its English equivalent. Second, example (16b) exhibits that the verb **sleep** is rarely observed in code-switching (CS) across language pairs. The Urdu verb in example (16a) and the English verb in example (16b) differ due to two factors: “argument structure” and “selectional properties”. Similarly, in example (17a), the Urdu verb **likh** (“write”) is directly inflected, whereas the English verb **write**, despite having a comparable argument structure, resists incorporation in example (17b). It fails to adopt either the do-construction strategy or direct matrix language inflection.

Examples (15a-17b) present paired constructions, where the first example in each pair reflects a natural expression and the second substitutes the content morpheme with an English non-finite verb form. The incorporation of substituted verbs in examples (15a), (16a), and (17a) demonstrates that content morphemes can be freely integrated into the morphosyntactic frame of the matrix language (Myer-Scotton & Jake, 2014; 2017). However, if all verbs were equally accessible, their incorporation should consistently yield grammatical structure. Instead, certain verbs resist integration, resulting in ungrammaticality.

To resolve this issue, Maqsood et al. (2019) proposed grammatical *constraints* to ensure the grammaticality of mixed sentences. Crucially, for successful incorporation, content morphemes must retain a functional supportive edge through a light verb. Examples (18a-18b) and (19a-19b) also follow a paired format, contrasting natural expression with an elicited one in which the verb is substituted with a counter-language equivalent. In these cases, grammaticality is preserved only when a light verb is present, aligning with the principles sketched by Myers-Scotton and Jake (2014, 2017) and Maqsood et al. (2019).

18a Alia Ali-ko **pasand** kar-*ti* hai.

Alia Ali-ACC like LVB-INF be. AUX. PRS

“Alia likes Ali”

[Natural Expression]

18b. Alia Ali-ko **like** kar-*ti* hai.

Alia Ali-ACC like LVB-INF be. AUX. PRS

“Alia likes Ali”

[Elicited Expression]

19a. People masjid-me **enter** ho-*tee* hein.

People mosque-LOCin enter LVB-INF be. AUX. PRS

“People enter into the mosque.”

[Natural Expression]

- 19b. People masjid-me **dakhil** ho-tee hain.
People mosque-LOCin enter LVB-INF be. AUX. PRS
“People enter into the mosque.” [Elicited Expression]

While grammatical constraints contribute significantly, this study posits that the selectional properties of verbs are a more decisive factor in determining the choice of verb. In examples (18a-18b) and (19a-19b), both the matrix language and embedded languages dynamically interact within mixed constituents, governed by the System Morpheme Principle (SMP) and the Morpheme Order Principle (MOP). These examples also satisfy the universal principle of constituent Uniformity. As Butt (2014) notices, “events embedded under a control or raising predicate are less tightly connected to the matrix verb.” This phenomenon is supported through examples (20a, 20b, and 20c) and (21a, 21b, and 21c). The pair-wise examples are given below.

- 20a. Students trip-pe **jaa-na** **chaha-tee** hain.
Students trip-LOC go-INF want-INF be. AUX. PRS
“Students want to go to trip.” [Natural Expression]

- 20b. *Students trip-pe **go-na** **chaha-tee** hain.
Students trip-LOC goINF want-INF be. AUX. PRS
“Students want to go to trip.” [Elicited Expression]

- 20c. *Students trip-pe **jaa-na** **want-tee** hain.
Students trip-LOC go-INF want-INF be. AUX. PRS
“Students want to go to trip.” [Elicited Expression]

- 21a. Larkian novel **paar-na** **chaha-tien** hain.
Girls novel read-INF want-INF be. AUX. PRS
“Girls want to read the novel.” [Natural Expression]

- 21b. *Larkian novel **read-na** **chaha-tien** hain.
Girls novel read-INF want-INF be. AUX. PRS
“Girls want to read the novel.” [Elicited Expression]

- 21c. *Larkian novel **paar-na** **want-tien** hain.
Girls novel read-INF want-INF be. AUX. PRS
“Girls want to read the novel.” [Elicited Expression]

Instances in (20a-21c) indicate the complex verb constructions, as articulated by Butt (2014), involving two verbs: a lexical verb and a raising verb such as *chaha* (“want”) in Urdu. The English verb *want* shares similar syntactic properties with *chaha*. Each examples are depicted in three forms: (a) a natural expression, (b) a construction in which the non-finite verb form (e.g. *jaa-na* “go” and *paar-na* “read”) is substituted with its English counterpart (*go*, and *read*), and (c) a version where the raising verb *chaha* is replaced with the English verb *want*. The elicited examples (20b-20c) and 21b-21c) are ungrammatical, exhibiting that neither the content morpheme nor the

raising verb is freely inserted and incorporated into the matrix language frame. Moreover, do-verb strategies fail to preserve grammaticality in Urdu-English code-switching (CS).

Based on naturalistic and elicited data, this study hypothesizes that the failure of verb incorporation in Urdu-English code-switching (CS) is due to two factors. First, the hybrid nature of certain verbs prevents them from conforming to grammatical constraints proposed for Urdu-English mixed data (Maqsood et al., 2019). Second, the argument nature of verbs varies across languages, and the complexity of these structures inhibits their integration into the matrix language frame. Given these findings, the current study proposes a revision of morpheme classification to more precisely account for the syntactic behavior of verbs in Urdu-English code-switching (CS).

Discussion

The incorporation of non-finite verb forms into Urdu-English intrasentential code-switching has been widely researched (Ali et al., 2021b; Belazi et al., 1994; Gonzalez-Vilbazo & Lopez, 2012; Gosselin, 2022; Grimstad et al., 2017; Jabbar et al., 2021; Lopez et al., 2017; Malik, 2017; Maqsood et al., 2019; Myers-Scotton & Jake, 2014, 2017; Poplack, 1980; Rahimi & Dabaghi, 2013; Sankoff, 1998; van Gelderen & MacSwan, 2008). However, these studies have not systematically categorized content morphemes into distinct, layered classifications. If bilingual speakers select morphemes at the *conceptual level* to construct mixed utterances—technically called an intra-CP structure—it can be assumed that multiple lexicons interact to produce hybrid morphemes.

The Matrix Language Frame (MLF) model (Myers-Scotton, 2002, 2005; Myers-Scotton & Jake, 2000, 2014) provides a unified framework for analyzing and accounting for code-switching, code-mixing, and borrowing across diverse language pairs. Its 2017 extension further classifies system morphemes, *Early and Late SMs*. Further, the *Late System* is categorized into *Bride System* and *Outsider System* categories, with *Late morphemes* subdivided into Bridge and Outsider types. However, this study shifts focus on content morphemes (non-finite verb forms—infinitive, bare, and participle), which received limited theoretical attention in prior research studies (Ali et al., 2021b; Beckett & Kobayashi, 2020; Hashmi et al., 2024; Khan et al., 2023; Rahimi & Dabaghi, 2013; Saram et al., 2023).

Addressing RQ1: How are English non-finite verbs incorporated into Urdu-English code-switched utterances?

This study hypothesized that not all non-finite verb forms—whether bare, infinitive, or participial—are freely incorporated in the morphosyntactic frame of the matrix language. English and Urdu non-finite verb forms often resist integration via either direct matrix language (ML) inflectional strategy or do-construction insertion strategies (Myers-Scotton & Jake, 2014). Upon insertion, these verbs show distinct properties related to argument structure and selectional constraints, which influence the incorporation.

Addressing RQ2: To what extent do the selectional and morphosyntactic properties of Urdu-English non-finite verbs affect their incorporation into the matrix language?

This study found that non-finite verb forms are not uniformly “low-cost” morphemes (Myers-Scotton & Jake, 2014). In lieu, they demonstrate complex selectional and argument

structures that thwart their incorporation. These verbs are language-specific and lexically encoded in a cogent form, with predicates and events tenaciously bound to their syntactic behavior. Consequently, some non-finite verbs resist both direct ML inflection and do-verb strategies, proposing a hybrid nature that diverges from previously established constraints (Poplack, 1980; Sankoff, 1998; Belazi et al., 1994; Maqsood et al., 2019).

Addressing RQ3: Does the Matrix Language Frame (MLF) model adequately account for the behavior of non-finite verbs in Urdu-English code-switching?

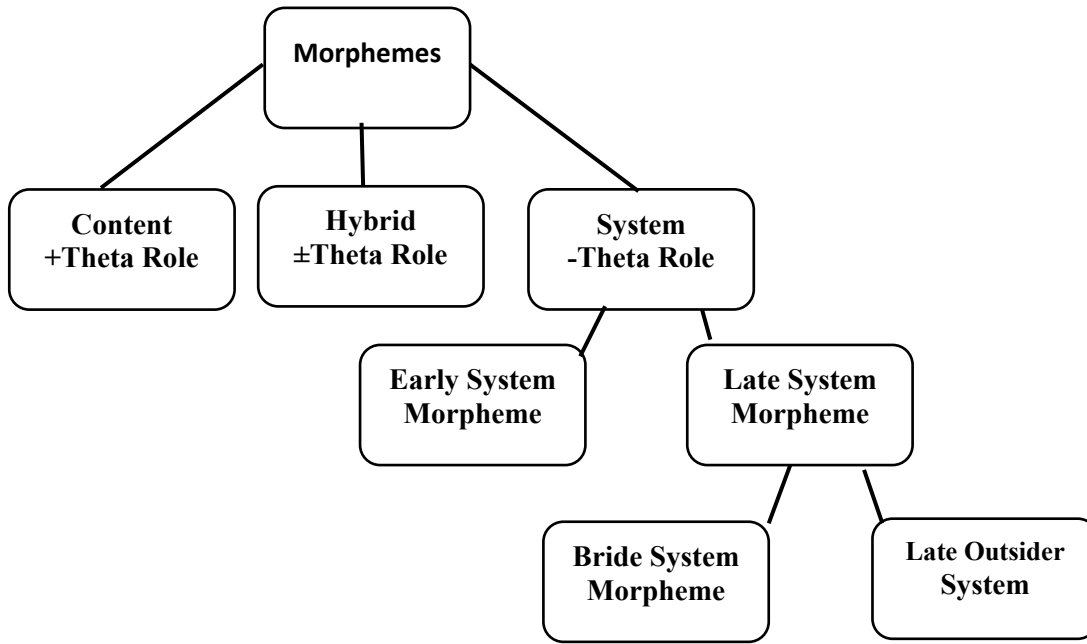
While the MLF model provides a robust framework for system morphemes, it does not entirely account for the behavior of non-finite verbs in Urdu-English code-switching (CS). The present study proposes a revision to the model by introducing a third category of morphemes—hybrid morphemes—which demonstrate properties of both content and system morphemes. These morphemes challenge the binary classification of lexical and functional items (Maqsood et al., 2019) and require a more nuanced approach to accommodate structurally complex and free-word-order languages.

Theoretically, Myer-Scotton and Jake (2014; 2017) describe morphemes processing across three levels: (a) the *conceptual level*, where intention and morpheme selection occur; (b) the *functional level* at level, where late system morphemes are incorporated; and (c) the *positional Level*, where phonological realizations take place. This study posits that not all content morphemes directly select bridge system morphemes at the conceptual level. Some attach at the functional level, while others—classified as hybrid morphemes—optionally obtain functional material during derivation (Ali et al., 2021b; Ali et al., 2025).

Drawing on Urdu-English code-switching (CS), this study proposes a revised MFL model that classifies morphemes into three broader layers: (i) content, (ii) system, and (iii) hybrid. Although the previous studies (Khan et al., 2023; Maqsood et al., 2019; Smit et al., 2023) have supported the MLF model's assumptions and stipulations, none have reconstituted its framework to address the core challenges posed by non-finite verb incorporation. The updated model aims to fill this theoretical gap by accounting for the complex nature of verbs and their diversified incorporated behavior in bilingual utterances.

In response to the predetermined research questions, this study proposes a revised Matrix Language Frame (MLF) model (Figure 2) based on two core hypotheses: First, not all non-finite verbs—whether bare, infinitive, or participial—are successfully incorporated into the morphosyntactic frame of the matrix language. Second, these verbs do not consistently behave as either content morphemes or system morphemes. Instead, they exhibit distinct properties that secure them functioning as “free” morphemes within the matrix language structure. Their incorporation is subject to grammatical constraints, as introduced by Maqsood et al. (2019), though this study does not introduce additional constraints. Instead, it reclassifies morpheme types to reflect the syntactic behavior of non-finite verbs better. The morphosyntactic structure of both natural and mixed sentences is fundamentally shaped by the verb's argument structure and selectional properties. These properties vary significantly across types: transitive verbs typically select one determiner phrase (DP) as complement, intransitive verbs may project a prepositional phrase (PP), and ditransitive verbs require two DPs, often realized through do-verb constructions. This diversity contributes to the complexity of verb incorporation in code-switched utterances.

Figure 2.
Re-classification of Morpheme Types



Accordingly, this study reformulates the Matrix Language Frame (MLF) model's classification of morphemes by introducing a more foundational categorization based on argument structure and selectional properties. Content morphemes, considered the lexical raw material of language, typically bear [+θ] features, indicating their capacity to assign thematic roles (Myer-Scotton & Jake, 2014; 2017). Conversely, system morphemes carry [-θ] features, signifying that they do not assign a thematic role and instead serve functional or grammatical purposes. This study introduces a third category--*hybrid morpheme*—which possesses dual properties: they may both assign and receive theta role, encoded as [±θ]. Due to this dual nature, hybrid morphemes perform functional operations that extend beyond the binary classification of content and system morphemes. The revised framework continues to operate within three core principles of the MLF model as posited by Myer-Scotton and Jake (2014; 2017), while offering a more nuanced account of morpheme behavior in bilingual constructions.

Conclusion

In summary, this study examines the incorporation of non-finite verbs in Urdu-English code-switching (CS), collecting both naturalistic and elicited data. While the 4-M model classified such verbs as content morphemes, findings reveal that certain non-finite forms resist integration due to their complex argument structure and selectional properties. These verbs do not behave like typical “free” content morphemes and fail to confirm either direct matrix inflection or do-verb insertion strategies. To address this, the study proposes a revised classification of morphemes into three categories: (i) content, (ii) system, and (iii) hybrid. Hybrid morphemes possess [±θ] features, enabling them to both assign and receive thematic roles and operate with functional flexibility. Additionally, in response to RQ1 and RQ2, this study summarizes that not all non-finite verbs are equally incorporable. RQ3 is addressed by suggesting that the Matrix Language Frame (MLF)

model necessitates amelioration to account for hybrid morphemes in Urdu-English code-switching. The study offers implications for teachers and educators, enhancing awareness of how bilingualism manifests in classroom discourse and enabling educators to understand better the syntactic patterns students create (Dar et al., 2024; Ali et al., 2025). The findings also confirm curriculum design by exhibiting the need to accommodate multilingual learners' natural code-switching tendencies. Pedagogical strategies should reflect authentic bilingual usage rather than impose rigid monolingual norms, allowing students to express themselves more fluently and accurately.

Acknowledgement

The authors extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through a small Research Group under grant number RGP.1/61/46

References

- Ali, A., Jabbar, Q., & Kiani, H. (2021a). Clausal-internal scrambling in the Urdu language: A derivation by phases. *REiLA: Journal of Research and Innovation in Language*, 3(1), 52-60. <https://doi.org/10.31849/reila.v3i1.5968>
- Ali, A., Dar, N. K., & Ashraf, J. (2025). On Agreement of Urdu Relative Clauses. *International Journal of Advanced Social Studies*, 5(2), 76-87. <https://doi.org/10.70843/ijass.2025.05209>
- Ali, A., Jabbar, Q., & Malik, N. A. (2020). No functional restriction and no fusion linearization on intrasentential codeswitching; a minimalist explanation. *Ijee. org*, 9(4), 130-145.
- Ali, A., Jabbar, Q., Malik, N. A., Kiani, H. B., Noreen, Z., & Toan, L. N. (2021b). Clausal-internal switching in Urdu-English: An evaluation of the Matrix Language Frame Model. *REiLA: Journal of Research and Innovation in Language*, 3(3), 159–169. <https://doi.org/10.31849/reila.v3i3.6774>
- Ali, A., Malik, N. A., & Zain. (2023a). Split tense projection in Urdu: An illusion. *Pakistan Journal of Language Studies*, 7(1), 16–31. Retrieved from [//pjls.gcuf.edu.pk/index.php/pjls/article/view/222](http://pjls.gcuf.edu.pk/index.php/pjls/article/view/222)
- Ali, A., Younis, A., Jabbar, Q., & Niaz, S. (2023b). Morphosyntactic study of Urdu ESL learners: A derivation by interface. *Journal of Studies in Language, Culture and Society*, 6(2), 36–43. <https://asjp.cerist.dz/en/article/239075>
- Almasaeed, N. (2023). Effects of Social Power and Distance on the Realization of Requests in Jordanian Bedouin Arabic. *Journal of Ethnic and Cultural Studies*, 10(5), 29–44. <https://doi.org/10.29333/ejecs/1728>
- Alnuzaili, E. S., Alghamdi, S. S., Ali, A., Almadani, Mohammed. A., Alhaj, A. A., & Malik, N. A. (2025). Code-switching beyond phases. *Cogent Arts & Humanities*, 12(1). <https://doi.org/10.1080/23311983.2025.2564881>
- Alnuzaili, E. S., Waqar Amin, M., Saad Alghamdi, S., Ahmed Malik, N., A. Alhaj, A., & Ali, A. (2024). Emojis as graphic equivalents of prosodic features in natural speech: Evidence from computer-mediated discourse of WhatsApp and Facebook. *Cogent Arts & Humanities*, 11(1). <https://doi.org/10.1080/23311983.2024.2391646>

- Ashraf, J., Mehmood, N., Ali, A., & Jabbar, Q. (2021). Possessor in Urdu nominal phrases. *Educational Research (IJM CER)*, 3(6), 30–37. https://www.ijmcer.com/wp-content/uploads/2023/07/IJM CER_E03603037.pdf
- Ashraf, J., Munir, Z., & Ali, A. (2025). Nominal licensing in Urdu-Hindi applicative construction. *Journal of Arts and Linguistics Studies*, 3(1), 193–211.
- Amuzu, E. (2011). *Ewe contact database*. Unpublished corpus, University of Ghana, Legon.
- Backus, A. M. (1998). *Two in one: Bilingual speech of Turkish immigrants in the Netherlands* (Doctoral dissertation, Tilburg University). Available at: <https://research.tilburguniversity.edu/en/publications/two-in-one-bilingual-speech-of-turkisch-immigrants-in-the-netherl>
- Baird, B., Balam, O., & Couto, M. C. P. (2023). On the linguistic analysis of Central American Spanish in monolingual and bilingual/multilingual contexts. In B. Barid, O. Balam, & M.C.P. Couto (Eds.), *Linguistic advances in Central American Spanish* (pp. 1–16). Brill.
- Balam, O., Pérez Leutza, L., Michalski, I., & Parafita Couto, M. del C. (2024). A semantic analysis of bilingual compound verbs in two contact Spanish communities. *Research in Corpus Linguistics*, 12(1), 140–170. <https://doi.org/10.32714/ricl.12.01.06>
- Balam, O., Stadthagen-González, H., Rodríguez-González, E., & Parafita Couto, M. D. C. (2023). On the grammaticality of passivization in bilingual compound verbs. *International Journal of Bilingualism*, 27(4), 415–431. <https://doi.org/10.1177/13670069221097772>
- Baker, M. C. (2009). Is head movement still needed for noun incorporation?. *Lingua*, 119(2), 148–165. <https://doi.org/10.1016/j.lingua.2007.10.010>
- Bandi-Rao, S., & Den Dikken, M. (2014). Light switches: On v as a pivot in codeswitching, and the nature of the ban on word-internal switches. <https://doi.org/10.7551/mitpress/8338.003.0010>
- Beckett, G. H., & Kobayashi, M. (2020). A Meta-study of an Ethnographic Research in a Multicultural and Multilingual Community: Negotiations, Resources, and Dilemmas. *American Journal of Qualitative Research*, 4(1), 85–106. <https://doi.org/10.29333/ajqr/8267>
- Belazi, H.M., Rubin, E.J., & Toribio, A. J. (1994). Code switching and X-Bar Theory: The functional head constraint. *Linguistic Inquiry*, 25(2), 221–237. <http://www.jstor.org/stable/4178859>
- Bentahila, A., & Davies, E. E. (1992). Code-switching and language dominance. In *Advances in Psychology* (Vol. 83, pp. 443–458). North-Holland. [https://doi.org/10.1016/S0166-4115\(08\)61510-1](https://doi.org/10.1016/S0166-4115(08)61510-1)
- Booij, G. (2010). *Construction morphology*. Oxford University Press
- Butt, M. (2014). Control vs. complex predication. *Natural Language & Linguistic Theory*, 32(1), 165–190. <https://doi.org/10.1007/s11049-013-9217-5>
- Cable, S. (2007). *The grammar of Q: Q-particles and the nature of Wh-fronting, as revealed by the Wh-questions of Tlingit* (Doctoral dissertation, Massachusetts Institute of Technology).
- Cai, Q., Lin, Y., & Yu, Z. (2024). Factors influencing learner attitudes towards ChatGPT-assisted language learning in higher education. *International Journal of Human–Computer Interaction*, 40(22), 7112–7126. <https://doi.org/10.1080/10447318.2023.2261725>
- Cedden, G., Meyer, P., Özkara, B., & von Stutterheim, C. (2024). The “code-switching issue”: transition from (socio)linguistic to cognitive research. *Bilingualism: Language and Cognition*, 1–14. doi:10.1017/S1366728924000737
- Chan, B. H. S. (2008). Code-switching, word order and the lexical/functional category distinction. *Lingua*, 118(6), 777–809. <https://doi.org/10.1016/j.lingua.2007.05.004>

- Chomsky, N. (2005). Three factors in language design. *Linguistic inquiry*, 36(1), 1-22. <https://doi.org/10.1162/0024389052993655>
- Chomsky, N. (2014). *The minimalist program*. MIT Press
- Dar, N. K., Khan, M.S., Naz, R., & Ali, A. (2024). Assessing semantic perception, morphological awareness, reading comprehension and delay time processing in autistic children. *Journal of Arts and Linguistics Studies*, 2(3), 1737-1760. <https://jals.miard.org/index.php/jals/article/view/182>
- De Soete, A. (2025). ‘For the Dutch-speaking students, I will translate it’: code-switching as discretionary transformation of EMI language policy. *Current Issues in Language Planning*, 1–24. <https://doi.org/10.1080/14664208.2025.2471158>
- Den Dikken, M. (2011). Phi-feature inflection and agreement: An introduction. *Natural Language & Linguistic Theory*, 29, 857-874. <https://doi.org/10.1007/s11049-011-9156-y>
- Deuchar, M. (2020). Code-switching in linguistics: A position paper. *Languages*, 5(2), 22. <https://doi.org/10.3390/languages5020022>
- Ellison, T. M., & Si, A. (2021). A quantitative analysis of age-related differences in Hindi–English code-switching. *International Journal of Bilingualism*, 25(6), 1510–1528. <https://doi.org/10.1177/13670069211028311>
- Fillmore, C. J. (1967). The case for case. <https://linguistics.berkeley.edu/~syntax-circle/syntax-group/spr08/fillmore.pdf>
- Gallego, Á.J. (2010). On the prepositional nature of non-finite verbs. *Catalan Journal of Linguistics*, 9(1), 79–102. <https://doi.org/10.5565/rev/catjl.95>
- González-Vilbazo, K., & López, L. (2011). Some properties of light verbs in code-switching. *Lingua*, 121(5), 832–850. <https://doi.org/10.1016/j.lingua.2010.11.011>
- González-Vilbazo, K., & López, L. (2012). Little v and parametric variation. *Natural Language & Linguistic Theory*, 30, 33–77. <https://doi.org/10.1007/s11049-011-9141-5>
- Gosselin, L. (2022). Bilinguals have a single computational system but two compartmentalized phonological grammars: Evidence from code-switching. *Glossa: A Journal of General Linguistics*, 6(1), 147. <https://doi.org/10.16995/glossa.5800>
- Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and language*, 36(1), 3-15. [https://doi.org/10.1016/0093-934X\(89\)90048-5](https://doi.org/10.1016/0093-934X(89)90048-5)
- Hashmi, U. M., Alghamdi, S. S., Shahzad, M., Malik, M. Y., Alnuzaili, E. S., Hamad, M. M., & Omar, A. (2024). Intercultural communication, faith and impression management: former Muslims’ engagement with social media. *Language and Intercultural Communication*, 25(2), 226–241. <https://doi.org/10.1080/14708477.2024.2411266>
- Halmari, H. (1997). *Government and codeswitching*. John Benjamins Publishing Company.
- Ilyas, Y., Noureen, H., & Ali, A. (2023). Syntactic layer of coordination and conjuncts agreement: Evidence from Pakistani English newspapers. *Journal of Education and Social Studies*, 4(3), 683–691. <https://doi.org/10.52223/jess.2023.4328>
- Jabbar, Q., Ali, A., Malik, N. A., Mahmood, N., & Wasif, M. (2021). Morphosyntactic sub-categorization of lexical verbs. *Webology*, 18(6), 4145-4165.
- Jake, J. L., & Myers-Scotton, C. (1997). Codeswitching and compromise strategies: Implications for lexical structure. *International Journal of Bilingualism*, 1(1), 25–39. <https://doi.org/10.1177/136700699700100103>
- Joshi, A. (1982). Processing of sentences with intra-sentential code-switching. In J. Horecky (Ed.), *Coling 1982: Proceedings of the Ninth International Conference on Computational*

- Linguistics* (pp.145–150). North Holland Publishing Company. <https://aclanthology.org/C82-1023.pdf>
- Kamwangamalu, N. M. (1994). siSwati-English code-switching: The matrix language principle and linguistic constraints. *South African Journal of African Languages*, 14(2), 70–77. <https://doi.org/10.1080/02572117.1994.10587033>
- Khan, A. A., Khalid, A., & Saleem, T. (2023). The evidence of Embedded Language islands: The case of Pashto-English codeswitching. *International Journal of Multilingualism*, 21(4), 2073–2096. <https://doi.org/10.1080/14790718.2023.2212907>
- Kudjo Agama, F., Hebbard, M., & López-Fitzsimmons, B. M. (2023). Negotiating the Messiness of Teaching Linguistic Justice Online: Reflections of Multilingual Writing Instructors During COVID. *American Journal of Qualitative Research*, 7(1), 128–148. <https://doi.org/10.29333/ajqr/12872>
- Liu, Y., Liu, F., Wang, Z., & Mei, Y. (2024). Multilingualism, language choice, and identity construction: Diasporic Ukrainians in Shanghai. *Journal of Sociolinguistics*, 28(2), 42–60. <https://doi.org/10.1111/josl.12652>
- López, L., Alexiadou, A., & Veenstra, T. (2017). Code-switching by phase. *Languages*, 2(3), 9. <https://doi.org/10.3390/languages2030009>
- MacSwan, J. (2000). The architecture of the bilingual language faculty: Evidence from intrasentential code switching. *Bilingualism: language and cognition*, 3(1), 37–54. <https://doi.org/10.1017/S1366728900000122>
- MacSwan, J. (2020). Theoretical approaches to the grammar of code-switching 1. In E. Adamou & Y. Matras (Eds.), *The Routledge handbook of language contact* (pp. 88–109). Routledge.
- Mahootian, S., & Beatrice Santorini, B. (1996). Code switching and the complement/adjunct distinction. *Linguistic Inquiry*, 27(3), 464–479. <http://www.jstor.org/stable/4178946>
- Malik, N. A. (2017). No mixed grammars, no phonological disjunction: A new perspective on intra-sentential code-switching. *Lingua*, 194, 51–66. <https://doi.org/10.1016/j.lingua.2017.05.002>
- Maqsood, B., Saleem, T., Aziz, A., & Azam, S. (2019). Grammatical constraints on the borrowing of nouns and verbs in Urdu and English. *SAGE Open*, 9(2). <https://doi.org/10.1177/2158244019853469>
- Mohammad Hussein Aburqayiq, A., Altakhaineh, A. R. M., & Alsariera, A. H. (2025). Code-mixing between Arabic and English among Jordanians on social media. *Cogent Social Sciences*, 11(1), 2491705. <https://doi.org/10.1080/23311886.2025.2491705>
- Morton, J. M. (2014). Cultural code-switching: Straddling the achievement gap. *Journal of Political Philosophy*, 22(3), 259–281. <https://doi.org/10.1111/jopp.12019>
- Munir, Z., Saddique, A., & Ashraf, J. (2025). Inflectional morpheme and frequency patterns in Urdu-English code switching: A corpus-based study. *Journal of Arts and Linguistics Studies*, 3(3), 5013–5032. <https://doi.org/10.71281/jals.v3i3.452>
- Muntendam, A., & Parafita Couto, M. C. (2024). Code-Switching at the Interfaces. *Languages*, 9(8), 258. <https://doi.org/10.3390/languages9080258>
- Mustafa, M. B., Yusoof, M. A., Khalaf, H. K., Rahman Mahmoud Abushariah, A. A., Kiah, M. L. M., Ting, H. N., & Muthaiyah, S. (2022). Code-Switching in Automatic Speech Recognition: The Issues and Future Directions. *Applied Sciences*, 12(19), 9541. <https://doi.org/10.3390/app12199541>

- Myers-Scotton, C. (1993). Common and uncommon ground: Social and structural factors in codeswitching. *Language in Society*, 22(4), 475–503. <https://doi.org/10.1017/S0047404500017449>
- Myers-Scotton, C. (2002). *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford University Press, USA.
- Myers-Scotton, C. (2005). Embedded Language elements in Acholi/English codeswitching: What's going on?. *Language Matters: Studies in the Languages of Southern Africa*, 36(1), 3–18. <https://hdl.handle.net/10520/EJC59724>
- 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. <https://doi.org/10.1177/1367006915626588>
- Myers-Scotton, C., & Jake, J. L. (2000). Testing the 4-M model: An introduction. *International journal of bilingualism*, 4(1), 1–8. <https://doi.org/10.1177/13670069000040010101>
- Myers-Scotton, C., & Jake, J. L. (2014). Non-finite verbs and negotiating bilingualism in codeswitching: Implications for a language production model. *Bilingualism: Language and cognition*, 17(3), 511–525. <https://doi.org/10.1017/S1366728913000758>
- Nguyen, L., Yuan, Z., & Seed, G. (2022). Building educational technologies for code-switching: Current practices, difficulties and future directions. *Languages*, 7(3), 220. <https://doi.org/10.3390/languages7030220>
- Niaz, S., & Ali, A. (2023). Explicit learning triggers sensory motor competence: An experimental study of Pakistani ESL learners. *Journal of Studies in Language, Culture and Society*, 6(1), 36–42. <https://asjp.cerist.dz/en/article/229872>
- Ninio, A. (2011). *Syntactic development, its input and output*. Oxford University Press.
- Okeke, G. T., & Okeke, C. O. (2022). On the semantic-pragmatic interface of Igbo verbs of perception. *Cogent Arts & Humanities*, 9(1). <https://doi.org/10.1080/23311983.2022.2025991>
- Park, S. M. (2025). Qualitative Research Methods in Ethnic Communities: A Framework for Studying Language and Cultural Preservation. *Journal of Ethnic and Cultural Studies*, 12(3), 34–52. <https://doi.org/10.29333/ejecs/2185>
- Pfaff, C. W. (1979). Constraints on Language Mixing: Intrasentential Code-Switching and Borrowing in Spanish/English. *Language*, 55(2), 291–318. <https://doi.org/10.2307/412586>
- Poplack, S. (1980). Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: toward a typology of code-switching1. *Linguistics*, 18(7-8), 581–618. <https://doi.org/10.1515/ling.1980.18.7-8.581>
- Rababah, L. (2023). Examining Speech Acts in Jordanian Advertising: Pragmatic Functions, Linguistic Features, and Rhetorical Devices. *Journal of Ethnic and Cultural Studies*, 10(5), 212–223. <https://doi.org/10.29333/ejecs/1722>
- Rahimi, M., & Dabaghi, A. (2013). Persian–English codeswitching: A test of the Matrix Language Frame (MLF) model. *System*, 41(2), 322–351. <https://doi.org/10.1016/j.system.2013.01.023>
- Rayo, W., Barrita, A. M., Cabrera Martinez, L., & Carbajal, I. (2024). The complexity in bilingual code-switching research: A systematic review. *Languages*, 9(6), 217. <https://doi.org/10.3390/languages9060217>
- Sahib, H., Hanafiah, W., Aswad, M., Yassi, A. H., & Mashhadi, F. (2021). Syntactic configuration of code-switching between Indonesian and English: Another perspective on code-

- switching phenomena. *Education Research International*, 2021(1), 3402485. <https://doi.org/10.1155/2021/3402485>
- Sankoff, D., & Poplack, S. (1981). A formal grammar for code-switching. *Paper in Linguistics*, 14(1), 3–45. <https://doi.org/10.1080/08351818109370523>
- Sankoff, D. (1998). A formal production-based explanation of the facts of code-switching. *Bilingualism: language and cognition*, 1(1), 39–50. <https://doi.org/10.1017/S136672899800011X>
- Saram, M., Ali, A., Mahmood, A., & Naz, R. (2023). Neural trigger of speaking skills in autistic children: An intervention-based study. *Journal of Education and Social Studies*, 4(3), 424–430. <https://doi.org/10.52223/jess.2023.4302>
- Shim, J. Y. (2016). Mixed verbs in code-switching: The syntax of light verbs. *Languages*, 1(1), 8. <https://doi.org/10.3390/languages1010008>
- Si, A. (2011). A diachronic investigation of Hindi–English code-switching, using Bollywood film scripts. *International Journal of Bilingualism*, 15(4), 388–407. <https://doi.org/10.1177/1367006910379300>
- Smit, J., Chisari, L. B., Kouns, M., Øyehaug, A. B., Savelsbergh, E., & Hajer, M. (2023). Inclusive STEM Teaching from a Language Perspective: Teacher Learning in a Professional Development Program. *European Journal of STEM Education*, 8(1), 07. <https://doi.org/10.20897/ejsteme/13643>
- Soffer-Vital, S., & Finkelstein, I. (2024). Empowering Teachers: Multicultural Social and Emotional Learning (MSEL) Among Arab Minority Teachers. *Journal of Ethnic and Cultural Studies*, 11(1), 39–57. <https://doi.org/10.29333/ejecs/1638>
- Spencer, A. (1995). Incorporation in Chukchi. *Language*, 71(3), 439–489. <https://doi.org/10.2307/416217>
- Subiyanto, A., Nurhayati, N., Suryadi, M., & Suwarno, P. (2024). Motion verbs related to Javanese traditional fishing activities: a natural semantic metalanguage approach. *Cogent Arts & Humanities*, 11(1). <https://doi.org/10.1080/23311983.2024.2338979>
- Treffers-Daller, J. (2025). The simple view of borrowing and code-switching. *International Journal of Bilingualism*, 29(2), 347–370. <https://doi.org/10.1177/13670069231168535>
- Van Gelderen, E., & MacSwan, J. (2008). Interface conditions and code-switching: Pronouns, lexical DPs, and checking theory. *Lingua*, 118(6), 765–776. <https://doi.org/10.1016/j.lingua.2007.05.003>
- Vu, N. T., Chi, D. N., & Nguyen, H. T. M. (2025). Exploring Students’ Intercultural Communicative Competence in the Light of English-Medium Instruction (EMI) in Higher Education in Vietnam. *Journal of Ethnic and Cultural Studies*, 12(3), 152–175. <https://doi.org/10.29333/ejecs/2168>
- Weinreich, U. (1953). The Russification of Soviet minority languages. *Probs. Communism*, 2, 46. <https://heinonline.org/HOL/LandingPage?handle=hein.journals/probscmu2&div=58&id=&page=>
- Wohlgemuth, J. (2009). *A typology of verbal borrowings*. Mouton de Gruyter.
- Yim, O., & Clément, R. (2021). Acculturation and attitudes toward code-switching: A bidimensional framework. *International Journal of Bilingualism*, 25(5), 1369–1388. <https://doi.org/10.1177/13670069211019466>
- Zhong, Z., & Fan, L. (2023). Worldwide trend analysis of psycholinguistic research on code switching using Bibliometrix R-tool. *Sage Open*, 13(4), 21582440231211657.

Notes on Contributors

Sami Saad Alghamdi is working as an Assistant Professor in General Course Department, Applied College, King Khalid University, Saudi Arabia. He has published his research in national and international journals. His areas of interest are applied linguistics, ELT, Structural Equation Modeling, and language teaching. He has guided many undergraduate and graduate students in their research. He will be called through email ID sasaalghamdi@kku.edu.sa.

Nazir Ahmed Malik is chairing the Department of Languages, the University of Chenab, Gujrat, Punjab, Pakistan. He has been guided many MS and PhD students in Pakistan. His research includes syntax, sociolinguistics, bilingualism and linguistics theory. He has developed a ground-breaking theory of No mixed grammar. He is also the reviewer of SAGE journals, Lingua and linguistics approaches on bilingualism. He has published his research on national and international platforms also. He actively participated in countless international conferences as keynote speaker. He will be contacted through email ID malik.nazir@gmail.com.

Ehab Saleh Alnuzaili is working as an Assistant Professor in English Department, King Khalid University, Abha, Saudi Arabia. His research areas include applied linguistics, ELT, and visual discourse. He has published his research in well-regarded national and international journals. He will be corresponded through email ID ihabsaleh2023@gmail.com.

Hamdi Abdel Rasool Eljeam is working as an Assistant Professor in English Department, Applied College, King Khalid University, Saudi Arabia. He has a keen interest in language teaching, semantics, Autism and applied linguistics. He has published his research in many national and international journals. He will be called via email ID heljeam@kku.edu.sa.

Abdulbasit A. Alhaj is working as an Assistant Professor in English Department, Applied College, King Khalid University, Saudi Arabia. He has a keen interest in language teaching, pragmatics, ELT and applied linguistics. He has published his research in many national and international journals. He will be contacted via email ID bastalnour@gmail.com.

Asad Ali is working as an Assistant Professor in Department of English at the University of Chenab, Gujrat, Punjab, Pakistan. His areas of interest are matrix syntax, neurolinguistics, morphology and computational linguistics. He has been teaching and researching at undergraduate as well as graduate level in the University of Lahore, Pakistan. He has published his research in national and international well-reputed journals. He is a reviewer of many linguistic journals including Sage open, Francis and Taylor. He has an overall index of 8 in 2024 in Pakistan. He will be corresponded via email ID masad7721@gmail.com.

ORCID

Sami Saad Alghamdi, <https://orcid.org/0000-0003-2899-429X>

Nazir Ahmed Malik, <https://orcid.org/0000-0003-0017-2833>

Ehab Saleh Alnuzaili, <https://orcid.org/0000-0002-0600-4470>

Hamdi Abdel Rasool Eljeam, <https://orcid.org/0000-0001-5516-5046>

Abdulbasit A. Alhaj, <https://orcid.org/0000-0001-6782-2314>

Asad Ali, <https://orcid.org/0000-0001-7398-9583>

List of Abbreviations:

1=first person
2=second person
3=third person
ACC=accusative
AUX=auxiliary
BLC=bilingual linguistic competence
COMP=complementizer
CM=code mixing
COP= copula
CS=code switching
LVB=light verb
DAT=dative
DET=determiner
EL=embedded language
FEM=feminine
FoL=faculty of human language
FUT= future
HAB= habitual
INF =inflection
INFI=infinitive
LIs=lexical items
LOC=locative
MS= masculine
ML=matrix language
MOP=morpheme order principle
NEG= negative
PST=past
PRS=present
PL=plural
SG= singular
SMP=system morpheme principle
USP=uniform structure principle
VP= verb phrase
VBL, VBZ=verbalizer
H=hybrid, m refers to morphemes and – and + are the featural values of the morphemes based on which it enters the derivation. Normally, morphemes only bear one feature, either + or – on the binary distribution, but the case of a hybrid is totally different from normal morphemes.