
The Dominant Language and Mixing Patterns

Dilkushi Senaratne
University of Kelaniya

ABSTRACT

This study investigates mixed discourse of Sri Lankan Sinhala-English bilinguals and attempts to describe the mixed patterns using Muysken's (2000) typology of code mixing (CM) and Myers-Scotton's (2002) Matrix Language Frame Model theory. The study treats insertion, alternation, and Congruent Lexicalization (CL) as manifestations of the same linguistic phenomenon.

This study will use the theories forwarded by both Muysken (2000) and Myers-Scotton (2002) to determine if there are differences in the manner of mixing. i.e. whether speakers alternate, insert or use CL patterns, and whether this is due to a dominant language. The MLF model will be used to analyze if a dominant language influences the patterns of mixing. The experiment contains samples of spontaneous speech productions of informants who are also interviewed.

The findings of the study will reveal whether there are mixed patterns in the Sinhala-English corpus and whether a dominant language influences mixed patterns. It will also reveal techniques that the bilingual employs when code-mixing which results in code mixes, hybrids, Sinhalizaciones and borrowings. The findings reveal that speakers resort to insertion, alternation and CL patterns when mixing codes and that one phenomenon may lead to another in the course of conversation. The findings of the study reveal that all types of mixing are present in the Sinhala-English corpus. In insertion, the presence of a dominant language can be determined whereas in alternation the participating languages share the structure. In CL, it is difficult to determine the dominant language as there is convergence of both grammars.

Key Words: Code - Switching, Bilingual, Sinhalization, Codemixing, Congruent Lexicalization

1. INTRODUCTION

Though phenomenally used by most speakers in Sri Lanka today, mixing Sinhala and English in informal conversations remains the subject of controversial debate in linguistic circles.

While some researchers use the term Code-Switching (CS) for alternating or mixing two languages in speech, the term CM will be used in this study to refer to all types of mixing, based on Muysken's (2000) typology. A typical example of a mixed utterance is given below:

- (1) call ekak diila aayet kiy?mu apitaa late venavaa nee.
[Call and tell him again if not we will get late.]

The example in (1) is typical of a Sinhala-English utterance. Observe the use of English lone lexical items such as 'call' and 'late'. The rest of the utterance is in Sinhala. Hence, the two words from English can be defined as 'nested' items in the 'host' language Sinhala.

Definitions of terms

Definitions of CM are ambiguous and vary. Even research students find it difficult to come to terms whether to use CM as an umbrella term that covers code-switching and borrowing. To obtain a general understanding of CM, a look at the usual definitions used by researchers in the field is necessary. In some instance, mixed data tends to be analyzed as code mixes or code switches. CS is generally defined as the use of two languages by a speaker in the same conversation. Definitions of CS are often linked to syntactic or morpho-syntactic constraints (Poplack 1980, Joshi 1985, Belazi et al 1994). It is important to keep in mind that the term CS is used when there is equal participation of two languages in the utterance. Since this paper attempts to describe mixed data using Muysken (2000) and Myers-Scotton (2002), it is important to take into account their views and definitions on CS. To Myers-Scotton (2002) CS is the use of two languages in the same clause³ which she calls intra-clausal switching. Inter-sentential switching, according to Myers-Scotton (2002), includes full sentences from both languages where each sentence will be a single clause. The same definition is elaborated by Poplack (1980) who views CS as comprising intra-sentential, inter-sentential, extra-sentential and tag switching. Hence, the distinction is made between each category of switching based on the speaker's proficiency of the languages concerned.

Musken (2000) defines CM as when lexical items and grammatical features from two languages appear in one sentence. The same view is held by Kachru (1978) who defines CM as a strategy used for the 'transferring' of linguistic units from one language to another. This transfer results in a 'restricted or not so restricted code of linguistic repertoire' which includes the mixing of either lexical items, full sentences or the embedding of idioms. Generally, CM is treated as a kind of 'transfer' of linguistic items, in most instances 'content words' or 'constituent insertions' from one language to another. In CM as opposed to CS, there is consensus that most often, the utterance (though bilingual) belongs to the structure of one language. Hence, in CM, it is observed that a dominant language can be identified. Furthermore, CM is structurally and sociolinguistically different to borrowing. Though CM shows affinity to its more celebrated counterpart (borrowing), there are significant differences between the two phenomena (Senaratne 2009). Where borrowings

are completely integrated into the host language, code-mixes are not¹. This observation is broadened in Muysken's (2000) typology of CM.

These definitions emphasize the influence of the proficiency or non-proficiency of languages that determine the type of mixing, employed by the speaker. It is evident from the results of the present study that there is influence of a dominant language at least in insertional CM. It is not so evident in other types of mixing patterns prevalent in the Sinhala-English corpus.

Using Muysken's (2000) typology of CM, this study views CM as incorporating insertion, alternation, and congruent lexicalization (CL) patterns of mixing. These mixing patterns are prevalent in the Sinhala-English corpus. Thus, CM will be used as a cover term to refer to all types of bilingual speech where empirical evidence reveals the use of two languages in an utterance: clause or sentence of a bilingual.

2. OBJECTIVES

The objective of this study is to explain the distributions or patterns in English-Sinhala mixing using mainly Muysken's (2000) typology of CM and Myers-Scotton's (2002) MLF model. Muysken's (2000) typology of CM will reveal the workings of all types of mixing patterns (insertion, alternation and CL) in the Sri Lankan bilingual's informal speech. To determine the dominant language, this study uses Myers-Scotton's (2002) Matrix Language Frame Model (MLF) theory. The present study will only focus on the English-Sinhala corpus with emphasis on the dominant language of speakers who code-mix between English and Sinhala.

The purpose of this study is to reveal that all types of mixing patterns exist in the speech of Sinhala-English bilinguals. Using Muysken's (2000) theory of CM, this study reveals that Sinhala-English bilinguals resort to insertional, alternational and CL patterns of mixing which in turn result in mixed types. This study will also reveal that defining the dominant language in insertional CM is easier than in alternation and CL.

3. LITERATURE REVIEW

Traditionally, research on language change mostly focused on internally motivated changes rather than external ones. Historical linguistics concentrated on internally motivated factors that change the internal character of a language, mainly in syntax (e.g. change in word order in English from SOV to SVO) and phonology (e.g. the Great Vowel Shift in English). Internally motivated changes locate processes either in the language system or 'native speaker creativity'². It is apparent that internal factors (or mostly monolingual

¹See Senaratne (2009: 235)

²Jones, M. C. and Singh, I. (2005: 4)

factors) focused on what happens within a speaker rather than the external forces that influence a speaker to change his/her way of using language in a given context. Studies in Contact Linguistics have shifted the focus of research in its entirety. Bilingualism has gained ground and there are more external reasons for languages to change than internal ones. The bilingual's use of two languages is now viewed as 'socially significant' (Gumperz 1982: 72), emphasizing the speaker as a skilled performer using both languages at his/her disposal as a tool in society. The bilingual's use of the two languages depends on the domain, topic and interlocutor. It is also governed by social and individual norms.

CM is now the focal point of research in contact language phenomena. Are these mixed utterances governed by one grammar or two grammars, which would suggest that mixing takes place at specific turns or points? On the other hand, are they unconstrained making available a number of possibilities to the bilingual to switch and mix at any point or turn, which would account for the simplicity and the fluidity with which it is employed? What are the social constraints that govern language mixing in bilingual societies? What are the norms in bilingual societies and how are they different to monolingual societies? How does language processing take place in the bilingual brain when producing mixed utterances? These questions reflect the many challenges that lie ahead for researchers engaged in this field.

At present, scholars have agreed on many issues. They are at a consensus that mixed utterances are constrained and that these constraints are structural, social and individual. There is also agreement that bilingual norms are different to monolingual norms, and hence bilingualism cannot be judged by monolingual standards. To an extent researchers agree that the two lexicons of the bilingual contribute to a code-mixing grammar, and that in language processing in the bilingual, the mixed grammar provides clues to the fluency of the two languages in the bilingual.

Researchers also focus on a universal approach towards CM. Such a universal approach needs to include issues such as the bilingual context, the period of contact between languages, the status, the typology, the competency of speakers, and the motivations for activation of certain languages over others. In post-colonial countries where CM with English is a relatively widespread phenomenon, other issues such as the status of the imperial language after colonization, the amount of fostering and nurturing it has received by the local governments after independence, and attitudes towards languages need to be considered. In this context, both Muysken's typology of CM and Myers-Scotton's theory of the MLF model plays a significant role.

3.1 Muysken (2000) - A typology of CM

Muysken (2000) in his theory of CM suggests that there are 'asymmetrical insertion, symmetrical alternation and CL models', and identifies three types of CM strategies in intra-sentential CM namely insertional CM, alternational CM and CL. Observing CM patterns in bilingual data from typologically different languages as well as similar ones, he explains that these three strategies are governed by different structural, psycholinguistic and social constraints.

The dynamic nature of contact situations, giving rise to a variety of mixing phenomena, is best explained in Muysken's theory where he proposes that the three strategies are related to each other. Accordingly, single and multi-word mixes are categorized into insertion, alternation and CL patterns, based on their structural elements, positioning within the utterance, and syntactic relation to the rest of the sentence.

Primarily, the most important aspect in Muysken's interpretation of mixing is the distinction he makes between borrowing and CM phenomena. Muysken's (2000) theory observes that patterns of borrowing exist in all the mixing strategies analyzed. Hence, borrowing exists in insertion, alternation and CL. Muysken suggests that many single word mixes that have the appearance of 'nonce borrowings'³ are in fact N or NP insertions.

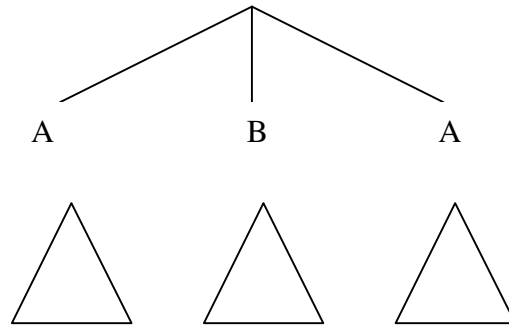
Furthermore, his analysis proposes that counter examples for the 'nonce borrowing theory' are results of insertional CM strategies and not alternational CM strategies where mid-word mixing does not exist. Most significant is the fact that the counter examples forwarded for the 'nonce borrowing theory' comes from languages that are distant to each other and which display asymmetrical mixing patterns (as in the case of Sinhala and English). Counter-arguing the criteria proposed for the analysis of 'nonce borrowings', Muysken suggests that (a) the absence of pronouns, (b) the absence of determiners, demonstratives and quantifiers accompanying the English noun and (c) the presence of case marking with the English noun do not necessarily indicate the mixing process as either borrowing or insertion. Basing his analysis on Tamil-English CM (from Sankoff, Poplack and Vanniarajan 1990 as quoted by Muysken 2000: 78), Muysken proposes that:

Both the absence of English determiners and the presence of Tamil case marking are due to the incompatibility of the English determiner system involving pronominal separate elements and the Tamil system which crucially involves case affixes.

³ Poplack (1982)

This observation is crucial for the analysis of borrowings and insertions in the Sinhala-English corpus. According to Muysken (2000), insertion occurs when entire constituents or lexical items are inserted into a structure from another language. He proposes the following pattern for insertion:

(2)



Insertional CM occurs when single, content words (such as nouns and adjectives) are morphologically integrated in the bilingual utterance. Observe the following example:

- (3) a. Chay-ta *las dos de la* noche-ta chaya-mu-yku.
 That- AC the two of the night- AC arrive-CIS-1pl
 [There at two in the morning we arrive] (Muysken 2000: 63)
- b. mokakhari individual kenek
 whatever individual NM.IND.PRO about
kīyāṇə word *ek-ak.*
 say.RL word NM-IND
 [A word **used for an** individual.] (Senaratne 2009: 146)

The word order of the utterances displays a nested a b a structure. There is grammatical relationship before and after the switch. This means that the elements preceding and following the insertion are related. Hence, the characteristics of insertion according to Muysken (2000) are as follows: they are single, selected content words. Apart from these, there are also dummy word insertions and telegraphic insertions. In other words, the structure of one language dominates the utterance.

Muysken describes alternation as a strategy of mixing where the two languages remain separate in the bilingual utterance as A...B unlike in insertion. He reaffirms theoretical views by Poplack (1980) that when there is linear word order equivalence between the two languages, alternation takes place also suggesting that in alternation the symmetrical

relations of the languages concerned play a key role. Discourse particles and adverbs are analyzed as alternations. In addition, alternation is observed to display a non-nested A...B...A structure, which means that the elements preceding and following the 'switched string' are not 'structurally' related. In other words, the participating languages play an equal role in the mixed utterance.

Muysken (2000) views alternation as a strategy akin to Poplack's inter-sentential CS. The striking structural characteristic is that alternation displays the 'juxtaposition' of L1 and L2. It is observed that in alternation, a sentence begins with language A and ends in language B. However, structural features of alternation are not that simple to identify.

In alternation, it seems that halfway through the sentence, one language is replaced by another (Muysken 2000). Alternation has broadened to include, tags, exclamations, interjections, self-corrections, doubling, flagging, conjunctions and repetitions. In alternation, several constituents are mixed in an order that displays linear equivalence. The typical context in which alternation occurs requires the symmetrical involvement of languages. There is a true switch from one language to the other involving both grammar and lexicon. Muysken (2000) proposes the following pattern for alternation where A and B represent languages and argues that when the switched string is preceded and followed by elements, not structurally related, this instance can provide a case for alternation.

(4)



The example in (5) by Treffers-Daller (1994) is cited by Muysken (2000: 97) as his example (4) as a case for alternation:

- (5) a. A.....B non-nested form
 Bij mijn broer y a un ascenseur en alles
 ' At my brother's place/there is an elevator/ and everything'.
 (Treffers-Daller 1994: 204)
- b. *meeə* *kiə – nəvəə* to identify yourself
 this one tell.PRS to identify yourself
 [This is used to identify yourself.] (32:12)
- b. *maməeəə* *kəranan* after the operation
 Isg that do.VL⁴ after the operation
 [I will do that] (39:25)
 (Senaratne 2009: 177)

⁴See Gair (1970) for the 'volitive and involitive optative' in Sinhala. Both nan and vi are attached to the stem of a Sinhala verb as finite affixes.

In (5), the juxtaposition is between a Sinhala phrase and an prepositional English phrase. The example is indicative of alternation. Both languages play an equal role.

- (6)
- a. Adverbial modification
En automatiquement klapte gij ook schoon Vlaams
'And/ automatically/ you would switch to standard Flemish.'
(Treffers-Daller 1994: 178)
 - b. Interjections
Aller a l'hopital toch niet?
'Going to the hospital/you don't mean'
(Treffers-Daller 1994: 213)
 - c. Flagging
Daar zetten ze **eah** des barriers
'There they put up eh/ barriers'
(Treffers-Daller 1994: 204)
 - d. Co-ordination
Nous on parle francais le flamand en de hele boel
'We speak French, Flemish/and all the rest'
(Treffers-Daller 1994: 207)

The framework proposed by Muysken suggests that considerable contact between typologically not-so-distant language pairs may lead to 'congruent lexicalization' (CL). The theory of CL characterizes the convergence of two grammatical systems into one that can take place between either related dialects or a dialect and a standard language. It is based on the notion of style shifting and variation. In CL, the grammar of the sentence is shared either fully or in part by the two languages. Muysken's theory of CL suggests that constraints do not apply for multi-constituent and non-constituent CM, as constraints presupposes that there is order in the sentence. In CL, all function words are switched, as there is no matrix language. Hence, CL is characteristic of a convergence of two systems into one, which is lexically accessible by both systems. Observe the following for CL from Sinhala-English CM taken from Senaratne (2009: 175).

- (7)
- a. *isteeshen ekə - Tə e - nəvaa*
station.sg NM.DF-DA come-PRS
[He/she is coming to the station.]
 - b. *ma - Tə istuɔyleave tiyə - nəvaa*
1sg-DA study leave be-PRS
[I have study leave.]

Reflecting on the social settings of the strategies, insertional CM is observed as occurring in colonial settings and in recent migrant communities whereas alternational CM is assigned to more balanced bilingual communities. In alternational CM, the speaker's equal proficiency of the languages is significant. CL is observed in closely related languages with equal prestige and with 'no tradition of separation' Muysken (2000).

3.2 Myers-Scotton (2002)- The MLF theory

Myers-Scotton (2002) defines insertion laying emphasis on a Matrix Language Frame Model, where the Matrix Language provides the morphosyntactic frame (all abstract grammatical requirements such as word order, morpheme order etc) or the base for the clause which contains code switching or mixing. According to Myers-Scotton, the Matrix Language acts as a foundation where the insertion from the Embedded Language (or guest language) is placed in.

The MLF theory explains intra-sentential CS patterns found in Swahili-English. The MLF theory is used to explain insertions, which exclude established loan words. The theory proposes that a set of abstract principles govern intra-sentential CS and forwards an abstract frame governed by a dominant language to interpret bilingual data. The MLF model makes distinctions between content morphemes and system morphemes, ML hierarchy and EL hierarchy and between the Morpheme Order Principle and the System Morpheme Principle. In a recent study, developing on the Morpheme Order Principle and the System Morpheme Principle, Myers-Scotton also introduces the 4-M model. The theory deviates from other morpho-syntactic frameworks, developed to explain rules of switching languages.

3.3 The Content vs System Morpheme distinction

The content morpheme and system morpheme distinction suggest that all the content words (basically noun forms, verbs and adjectives which are often switched), belong to the Embedded Language (EL), while all system morphemes for the frame (which included all affixes that are bound morphemes and some function words), are provided by the ML. CS conforming to this is referred to as Classic CS. Note that discourse markers are considered content morphemes in this analysis.

3.4 The ML hierarchy vs the EL hierarchy distinction

The ML hierarchy is defined as when one language is the base and takes control of the grammar of the utterance, and the EL hierarchy is defined as when the switched elements are at the periphery of the utterance (which includes adjuncts and idiomatic expressions).

3.5 The Morpheme Order principal vs the System Morpheme Principle

The Morpheme Order Principle proposes that the surface order of the mixed constituent (which may consist of at least one word from the EL and a number of words from the ML) will be that of the ML. The System Morpheme Principle suggests that all the system morphemes, which have grammatical relations external to their head constituents will be from the ML.

3.6 The Matrix Language Frame Model (MLF)

The MLF model suggests that there is a dominant language that provides the morpho-syntactic frame for the bilingual utterance and this dominant language is named the 'matrix' language (ML). The other language that acts as a guest in the utterance is the 'embedded' language (EL). The model presumes that the CS utterance contains ML islands and EL islands.

In the MLF model, the word order of the utterance is governed by the matrix language. Several structural features of the matrix language were outlined for it to be identified. These features include the structure of the discourse (morpho-syntactic frame which is provided by the matrix language), in certain instances the branching of the sentence (which should indicate the matrix language) and the number of morphemes in an utterance where the assumption is that the matrix language provides more morphemes (Myers-Scotton 2001a). The theory presupposes that in bilingual speech production, one language is always more activated (the matrix language) than the other (the embedded language). Note that the theory is based on the asymmetry between the languages where the asymmetry acts as an indicator to the matrix language of the utterance.

Myers-Scotton describes intra-sentential CS also under the heading 'CS as an unmarked choice' or as 'classic CS'. In the MLF theory, the ML will not change in conversation in the 'unmarked CS' type. The ML will always be one of the local/native or indigenous languages which will be mixed with an international language. Singly occurring EL forms are a distinct feature of unmarked CS.

The type referred to as 'unmarked CS' is 'very common' according to Myers-Scotton (1993b) in Hispanic, African and South East Asian communities. Note that in most of these communities, one of the languages in the bilingual utterance is an internationally powerful prestigious language such as English or French that formerly belonged to the colonial ruler. Note also that in Myers-Scotton's theory, momentary borrowings are code-switches⁵. Myers-Scotton (2002: 41) explains that borrowing is the most obvious outcome of being bilingual.

⁵ This view is not shared by Muysken (2000).

An obvious sociolinguistic characteristic of borrowing is that, for the most part, this borrowing is one-way. L1 speakers of the less prestigious group (however prestige is conceived of) take into their language words from the L1 of the more prestigious language. A psycholinguistic characteristic that differentiates borrowing from code-switching is that not all the speakers who use borrowed forms need to be bilingual in the donor language.

Observe the example in (8) taken from Swahili-English mixing from Myers-Scotton (2006: 254) her example (9).

- (8) ulituma barua ya application?
 [Did you send the letter of application?]

In this sentence, the matrix language is analyzed as Swahili into which an EL form from English is inserted and the EL form is analyzed as a code-switch not a borrowing. Note however that there is also the argument that all EL forms that are found in data may not be code-switches and that they may be established borrowings.

The theory of the MLF forwarded by Myers-Scotton presumes that there is either a ML or an EL hierarchy in CS. It observes that all syntactically active system morphemes of the utterance belong to the ML. The theory is used to describe CS between language pairs that are distant to each other e.g. Swahili-English.

3.7 The Markedness Model

Myers-Scotton's (1993a) markedness model proposes that speakers have a 'sense of markedness' in the use of the linguistic codes available to him/her and that all code choices can be explained in terms of speaker 'motivations' (Myers-Scotton 1993a: 109). Accordingly, speakers make the 'unmarked' choice for safer and simpler reasons. Referring to an audio-recorded conversation at a bus stop in Nairobi, where most of the conversation takes place in Swahili, Myers-Scotton proposes Swahili as the 'unmarked choice' (1993a: 78). Describing further, Myers-Scotton suggests that the 'unmarked' code generally acquires 'fewer distinctive features' and represents 'greater frequency' (Myers-Scotton 1993a: 80). In essence, the language that is less expected to be used at a given interaction with an interlocutor represents the 'marked code choice'.

Hence, Myers-Scotton's MLF model recognizes that there is an unequal participation of languages (asymmetry) in Classic CS. Joshi (1985) was one of the first to refer to the frame-building language as the ML and to the other participating language as the Embedded

Language (EL). Observe Myers-Scotton's (2006) example for CS intra-sentential CS where elements from two languages appear in the same clause. It is an example of insertional CM.

- (9) . *El le cambio los fans*
He changed the /fans
 (Myers- Scotton 2006)

According to Myers-Scotton (2006) empirical evidence shows that the grammatical structure of one language prevails in what has been coined as 'Classic CS'. In this case, Classic CS is akin to insertional CM. The MLF Model was specifically designed to explain 'Classic CS' where elements from two languages are included in the same clause, but only one of the languages is the source of the morphosyntactic frame for the clause. This means that both singly occurring words and full phrases from one language are included within a frame set by another language.

Another important consideration in Classic CS is that the speaker only needs to be proficient in the language structuring the clause so that they can follow the well-formedness of that language. A high degree of proficiency in the other language, that is termed as 'Guest language' or Embedded language (EL)' is not so critical in Classic CS. Apart from this, the MLF model recognizes the unequal participation of languages in CS.

The notion of a presence of a Matrix Language has posed many problematic areas for researchers in Contact Linguistics. Myers-Scotton (2002) also holds the view that the Matrix Language could change during a sentence from one language to another language. But this limits, as Muysken (2000) suggests, to the empirical scope of the notion of 'matrix'. He argues that in many cases there is much evidence of a 'base', 'matrix' or a 'dominant' language and in many cases there may not be (as in the case of CL where mixing takes place haphazardly). Hence, identifying the ML in an utterance has proven to be problematic for researchers in Contact Linguistics. This will bring us to the question of how to identify the presence of a Matrix Language in a mixed corpus.

What are the determining linguistic factors that reveal a Matrix Language at work in bilingual utterances? Identifying the Matrix Language is most important in the MLF model though it is an extremely difficult task. Myers- Scotton (2006) proposes that the ML can be identified as the language which provides the source for the morphosyntactic structure for the clause as in Classic CS. But how can one assume that a particular language has provided the morphosyntactic structure for the clause? Here, Muysken's (2000) typology may assist as it suggests that in certain cases, the main verb can be taken to determine the base of the ML.

A second approach is in terms of left to right parsing, where the first word or set of words determines the ML. But according to Muysken (2000), this too proves to be problematic, as most bilingual utterances start off with interjections or exclamations as the first element. These elements do not in any way determine the structure of the rest of the sentence. A third possibility is morpheme-counting. In Myers-Scotton (1993) the ML is defined as also the language that provides the most number of morphemes in the interaction. This definition too is vague as in some utterances there may be an equal number of morphemes from both languages.

Psycholinguistically, a base or an ML can be defined as the language that is most activated for the bilingual speaker. But how can one establish this as a fact? Given these difficulties it is most difficult to identify the base or the ML. But bilingual speakers have no difficulty in identifying the language that is most activated for them in their utterances especially with regard to insertion. Whenever we hear someone speak in two languages, we instinctively know if the utterance is guided by a dominant language or not. No doubt that in insertion the ML plays a significant role.

Contrary to the above hypotheses, Muysken (2000) observes that determining the matrix may be in part empirical and in part theoretical. According to him, a general valid criterion to determine the ML for a sentence or conversation is hard to find. Consider the following example where the speaker considers himself to be more dominant in Sinhala and proficient in English. Does the utterance reveal what is indicated in the self-assessment?

- (10) *ee caar ekə - Tə lisen gannə*
 that car.sg NM.DF-DA license.sg take.INF
 amaaruvunaa.
 difficult be.PST
 [It was difficult to get the license for that car.]

The language that has provided the grammar for the utterance is Sinhala. The elements preceding the insertions 'car' and license' are structurally related and therefore, insertion is plausible. However, can the insertions be really called insertions? They are wholly integrated into the base language phonologically, morphologically and syntactically.

The speaker in (11) considers himself a native speaker of Sinhala, but more proficient in English. According to the MLF theory, the morphosyntactic structure of the utterance can be analyzed as Sinhala. The collocation 'company vehicle' is an insertion from English. However, the speaker's self-assessment of language proficiency cannot be determined from the utterance.

- (11) a. *mama ethanata giyee company vehicle eken*
 [I went there in the company vehicle.]
 b. *eka initial stage eka*
that /initial stage / one
 [That is the initial stage]

Both (11a) and (11b) are examples for insertional CM. However, observe the difficulties in determining the dominant language using the MLF Model. In (9b) the speaker too chose Sinhala as the L1 but English is used 80% in informal discourse (self-assessment). What is the ML in (11b)? Is it a sentence where the collocation 'initial stage' is nested in a frame, set by Sinhala? Morpheme counting too will not assist to determine the ML as there is an equal number of morphemes provided by both languages. The conversation that was recorded was completely in English and this was the only instance where the speaker chose to mix.

Theories put forward by Myers-Scotton (2002) and Muysken (2000) recognize that in insertion there is a dominant language that influences the patterns of mixing. Muysken (2000) makes a distinction between insertion and alternation that corresponds to Auer's view (1995). Where Myers-Scotton (2002) discusses that in 'Classic CS' the MLF model has largely explained the role of a dominant language at play in the speech of the bilingual, focusing that insertion takes place within a clause and alternation takes place between sentences. Muysken (2000) argues that alternation too can take place within a clause. The same structures are analyzed by Myers-Scotton as Classic CS.

Much of the controversy in CM stems from the presence of 'other language' single words in bilingual data. The two structural approaches by Muysken (2000) and Myers-Scotton (2002) toward single lexical items in code-mixed data stem from the symmetrical and asymmetrical relationships between the languages concerned.

From the structural analyses proposed, the theory of CL (Muysken 2000) provides a satisfactory observation of random mixing patterns present in bilingual discourse as revealed in the data in this study. There are some indications that the bilingual makes use of both grammars when CM. In many instances, there are indications of a dominant base language effect especially regarding single word mixes (mainly in insertional CM). This stems from the argument that the needs of the bilinguals' languages are different depending on the interlocutor, topic and domain. The fourth problem is undoubtedly the status of the mixed varieties that have originated as a result of language contact and language change.

4. METHODOLOGY

This study uses Muysken's (2000) typology of CM and Myers-Scotton's (2002) theory of the MLF Model to analyze data in the Sinhala-English bilingual corpus. Both techniques are employed to determine the mixed strategies and the dominant language in mixed data and whether the patterns indicate the influence of a dominant language.

A group of 20 speakers were chosen for the study. Self-assessment questionnaires were used to determine the use of two languages in speech by the informants. The influence or non-influence of L1 of speakers was also determined by self-assessment questionnaires. The speakers were chosen based on their language use in specific domains (a minimum of 6 questions out of ten ticked as using both Sinhala and English in conversation). The intention of the questionnaire was to determine the L1, gender, language of instruction, profession, ethnicity, language used in different domains and attitudes towards languages and language use in society. The self-assessment was used to show that the dominant language of utterances. Self-assessments were also used to choose speakers for the study. Only those speakers who chose Sinhala as the L1 were chosen as the analysis was between Sinhala and English language mixing. As Sinhala was used excessively in informal conversations, the investigator decided to use speakers whose L1 was Sinhala⁶ Spontaneous speech productions of informal discussions on general topics were taped and transcribed.

Muysken's typology of CM was applied to identify the variety of mixing strategies employed by the bilinguals. This study will prove that Muysken's theory is the best that explains the number of mixing strategies employed by the Sinhala-English bilinguals

While the MLF Model is used to explain the presence of a dominant language or grammar at work in the bilingual mental lexicon when they are engaged in intra-sentential CM, this study will reveal how the MLF proves problematic in defining certain bilingual utterances as the dominant language cannot be determined. However, the theory of MLF explains mixing concerning a dominant language, especially in insertional CM.

Also, this study will reveal that the English-Sinhala corpus contains mixing not only within clauses but also between sentences and sometimes within words which cannot be explained by the theory of MLF. The mixed types are best explained using Muysken's (2000) typology of CM.

This reveals that Sri Lankan bilinguals engage in all types of mixing namely insertion, alternation and CL in their mixed conversations. The most noteworthy mixing type is CL

⁶Bilinguals whose L1 was Sinhala was also chosen to determine whether the base language in the utterance was the L1 of the speaker.

patterns where word-internal mixing takes place. It is CL that is used by Sinhala-English bilinguals to create borrowings from English in Sinhala utterances. Data from CL result in Sinhalizations, borrowings and Hybrids (Senaratne 2009) which are not discussed in this study.

5. ANALYSIS

An analysis of the data reveals that all types of mixing patterns are prevalent in the Sinhala-English corpus. The following is an example for alternational CM.

- (12) She has always tried to put me down whenever she can
haem ðisseemð she will say something bad about me.
 [She has always tried to put me down whenever she can **always** she will say something bad about me.]

In (13) there is a collocation and the verb in English. The sentence follows Sinhala word order and not English. While the sentence begins in one language and ends in another, this cannot be analyzed as alternation for many reasons the most important being that the morphosyntactic frame for the utterance has been provided by Sinhala. According to the MLF theory, the dominant language is Sinhala. The verb of the utterance is a bilingual verb (use kara) according to Muysken's (2000) typology. Based on Muysken's (2000) typology, 'polythene banners' is a collocation and an insertion in the utterance that belongs to Sinhala. In (13), the dominant language is identifiable.

- (13) I never thought it was going to be so easy I mean *poddak hithanna*
 [I never thought it was going to be so easy I mean **just think**]

In (14), elements preceding the switch and elements following the switch are not structurally related therefore, alternation is plausible. English appears to be the dominant language based on the MLF, however, as both grammars are in tact, this is not justifiable.

- (14) a. He was not interviewed, interview *eka tibee nee*
 [**He was not interviewed** the interview was not held.]
 b. I don't do anything here *nikan idagana innava*
 [**I don't do anything here** I just sit and wait.]

In (14a) and (14b), mixing has taken place at sentence boundaries. The switches are not structurally related. Therefore, alternation is plausible in both instances. The dominant language is not identifiable in both examples. The utterances reveal that the speakers are proficient in both languages, which is a case for alternation.

The examples in (15) reveal CL patterns of mixing. In (15a) the speaker uses both CL and insertion. Observe the structural differences of the mixed elements. In the nativized element, which is a borrowing, the English noun is accompanied by a Sinhala suffix. It is phonologically, morphologically and syntactically integrated into the borrowing language. However, the same speaker, when mixing the English noun 'vehicle' adopts a different strategy. The noun is accompanied by a nominalizer which structurally differentiates code-mixes from borrowings in Sinhala (Senaratne 2009: 248). In (15b), the speaker uses the English noun 'duty' and nativizes it with a Sinhala suffix. This too is an example for borrowing.

- (15) a. *apee compaeniyen vehicleekak gannə puluvan*
 [We can get a vehicle from our company.]
 b. *eeə apee diuṭi – yə nee*
 [That is our duty no.]

The analysis reveals insertion as the dominantly used mixing strategy of the speakers. It is evident that the mixed data represents all phenomena: insertion, alternation and CL patterns. It also reveals that bilingual Sri Lankans use all types of mixing in the course of conversation. Many examples were analyzed as cases for insertional CM justifying Muysken's observation that insertion occurs in colonial settings. However, a number of examples were also analyzed as alternational CM. Data also revealed a dominant or a base language at work in some instances, most often in insertional CM.

In the English-Sinhala corpus, data revealed that the ML was always one of the participating languages. Interestingly, speakers who were more proficient in English chose Sinhala as their L1. Data also revealed that insertion occurs with single, content, nested forms. In alternational CM, utterances follow a pattern of non-nested forms. There is no syntactic relationship between the mixed elements.

The argument that in insertion there is a dominant language at work does not necessarily mean that the speaker is less proficient in the other participating language. The complexity arises as speakers switch from insertion to alternation during the course of conversation. Hence, it is possible that one phenomenon leads to another. All the informants selected for this study were code mixers who employed both insertional and alternational CM. They were bilingual speakers of both Sinhala and English.

Based on the patterns of mixing, and since the dominant language cannot always be determined, it is most difficult to determine the language proficiency of speakers as speakers incorporated many mixing patterns in their utterances. Determining the dominant

language in certain utterances also proved problematic as a result of the mixing patterns employed by the speakers and their complexities.

Data also revealed that word internal mixing takes place in the Sinhala-English mixed corpus. These are analyzed as depicting CL mixing patterns. Borrowings are created using what Muysken (2000) analyses as CL patterns of mixing. CL patterns can also be observed in English utterances where Sinhala modifiers carry English suffixes. These are completely new words in the bilingual's repertoire and are not discussed in this study.

6. CONCLUSION

In the Sinhala-English corpus, there are English nouns that are inserted with an article particle. The data reveals that code-mixers use a nominalizer that accommodates and actually contributes to the mixing of an extensive number of English nouns, NPs and verb stems. This study demonstrates how the presence of this nominalizer indicates the mixing strategy used by the bilingual. Similarly, the presence of case markers, which are separate pronominal elements (as in Tamil-English data) is also indicative of the mixing strategy. The asymmetry in the mixing patterns is revealed in insertional CM.

Furthermore, the 'nonce borrowing' theory is incompatible with complex mixes such as 'educational systems' and 'Hindi songs' (Muysken 2000: 79). Contrary to analyses that propose these as borrowings, Muysken suggests that these examples are compatible with NP insertions and are indicative of insertional CM rather than borrowing. In the Sinhala-English corpus, a number of such examples are prevalent in the repertoire of the bilingual speaker.

In alternation, mixing is straightforward. Alternation is predominantly revealed in conjoined sentences from Sinhala and English, as illustrated in the data. Often the mixed string reveals an A B pattern where the utterance begins in language A and ends in language B. However, complication arises as single words, not just phrases indicate alternation as in the case of interjections, tags, quotatives and particles. Nonetheless, distinguishing alternation from insertion is less difficult a task than distinguishing borrowing from insertion. In alternation, the two languages are syntactically not related and the strings from the two languages are juxtaposed according to Muysken's (2000) CM theory. One of the distinct features of alternational CM is that it reveals switches that are at the periphery of utterances. Hence, an important indication of alternation is the position occupied by a word or phrase in the mixed utterance. When words are mixed at the periphery of utterances for more than referential purposes, alternation is plausible than insertion and borrowing.

Word internal mixing, prevalent in the Sinhala-English bilingual corpus, successfully captures Muysken's observation that borrowing patterns are indicative of CL. In CL, an important contribution has been made to the study of bilingual data in all settings, most importantly with regard to borrowings. Muysken (2000: 123) recognizes the significance when he observes:

Though congruent lexicalization is mostly a phenomenon occurring at the sentence level, we also find several cases of it in borrowing. These may puzzle language contact researchers if borrowing is only regarded from the perspective of insertional mixing.

Another important observation in Muysken's typology is the analysis of mixed compounds. Muysken includes hybrid compounds from Germanic languages as indicative of CL where a bidirectional pattern is revealed. The hybrid formations categorized under Sri Lankanisms (Senaratne 2009:239), prevalent in the Sinhala-English corpus hence are indicative of borrowing patterns through CL, based on Muysken's (2000) theory. However, there is no bidirectional pattern involved in the Sinhala-English data. Accordingly, CL not only acknowledges that mixing can be both constrained and unconstrained, but also admits the possibility that it can be a productive and creative process given the bilingual's ability for innovation.

In conclusion, the structural differences that are prevalent in insertion and alternation do not provide a clear indication of the language proficiency of speakers in English-Sinhala CM as speakers use both phenomena in conversation. As a result of the complexities of the mixing types, determining the dominant language of utterances is not always easy.

In psycholinguistic terms, we find that languages are activated in different degrees for insertion and alternation. In CL, both are activated in similar degrees. In insertion, activation of one language would be temporarily diminished and in the case of alternation, activation would switch from one language to another. Psycholinguistic factors determining the choice between these different processes include bilingual proficiency, the triggering of items, the degree of separateness of languages and the access systems according to Muysken (2000).

Sociolinguistically, CM is a popular code. However, there is a tendency to regard CM as a linguistic phenomenon that destroys the purity of languages though it is an integral part of language change. Interestingly, it is used by proficient bilinguals who wish to signal their membership in the communities of speakers of both languages, not just one. Thus CM is indicative of dual membership. Research in Contact Linguistics has evidence that much CM, as a medium of communication, creates positive attitudes rather than negative

ones (especially with regard to the speakers) and little CM reveals a polarization within society. Myers-Scotton (2002) observes that CS is identified with in-group membership and reflects solidarity in that community. Hence, it is typically an in-group mode of communication rather than one used with strangers or acquaintances. This mode of communication often reflects the identity of that speech community and is indicative of how speakers view themselves. The mixed code is evidence that speakers use both codes as salient indices of the values they incorporate in their identities. This is in accordance with claims on discourse patterns in the third world (Pakir 1989) where an indigenous variety is used in CM with the language of the former colonial power. It is also observed that the dominant language of the bilingual utterances is often, not the language of prestige and power offering upward mobility to its speakers in the larger community.

REFERENCES

- Auer, P. (1995) The pragmatics of code-switching: a sequential approach. In Milroy and Muysken, 115-35.
- Belazi, H. M., J. Rubin Edward, and Almeida, J. T. (1994) Code-switching and X-Bar theory: The functional head constraint. *Linguistic Inquiry*, 25 (2): 221-37.
- Gumperz, J. J. (1982) *Discourse strategies*. Cambridge: Cambridge University Press.
- Grosjean, F. (1982) *Life with two languages. an introduction to bilingualism*. Cambridge, Mass.: Harvard University Press.
- Joshi, A. K. (1985) Processing of sentences with intrasentential code-switching. In D. R. Dowty, L. Karttunen, A. M. Zwicky eds., *Natural language parsing*. Cambridge: Cambridge University Press, 190-205.
- Kachru, B. B. (1978) Code-mixing as a communicative strategy. In J. Alatis ed., *International dimensions of bilingual education*. Washington, D.C: Georgetown University Press, 107-24.
- Muysken, P. (2000) *Bilingual speech: a typology of code-mixing*. Cambridge: Cambridge University Press.
- Myers-Scotton, C. (1993) *Duelling languages: grammatical structure in code-switching*. Oxford: Clarendon Press.
- Myers-Scotton, C. (2001) Structural uniformities vs. community differences in code-switching. In R. Jacobson ed., *Modes of switching selected codes: a global perspective*. Berlin: Mouton de Gruyter.
- Myers-Scotton, C. (2002) *Contact linguistics: bilingual encounters & grammatical outcomes*. Oxford: Oxford University Press.
- Myers-Scotton, Carol, 2004, 'Precision Tuning of the Matrix Language Frame Model of Code Switching', *International Yearbook of European Linguistics, Sociolinguistica*, Max Niemeyer Verlag publishers.
- Myers-Scotton, C. (2006) *Multiple voices: an introduction to bilingualism*. Malden MA: Blackwell Publishers
- Pakir, A. (1989) 'Linguistic Alternates and Code Selection in Baba Malay', *World Englishes*, 8:379-88.
- Poplack, S. (1980) Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: toward a typology of code-switching. *Linguistics*, 18: 581-618. Also in Li Wei (2000) *The Bilingualism Reader*. Routledge, 9: 222-57.
- Poplack, S. and Meechan, M. (1995) Patterns of language mixture: nominal structure in Wolof-French and Fongbe-French bilingual discourse. In Milroy and Muysken, 199-233.
- Senaratne, C. D. (2009) *Sinhala-English code-mixing in Sri Lanka*. A sociolinguistic study. Netherlands Graduate School of Linguistics. LOT publications.
- Treffers-Daller, Jeanine, 1994, *Mixing two languages, French- Dutch contact in a comparative perspective*, Berlin: Mouton de Gruyter