Forms of Matrix Language + Embedded Language Formula in Code Switching Instances

Kanyi Charles Mugo, Nancy Anashia Ongo'nda

Full Terms & Conditions of access and use can be found at
http://hrmars.com/index.php/pages/detail/publication-ethics

Forms of Matrix Language + Embedded Language Formula in Code Switching Instances

Kanyi Charles Mugo

Department of Languages and Humanities, Mount Kenya University, Kenya
Email**:** ckanyi46@yahoo.com

Dr. Nancy Anashia Ongo'nda

School of Social Sciences, Department of Languages and Humanities Mount Kenya University, Kenya
Email: anashian@yahoo.com

Abstract

This paper presents on the forms of Matrix Language and Embedded Language in Code Switching within a multilingual society, that is, Mwea East Sub-County, Kirinyaga County Kenya where eight languages are spoken. Information for the survey was collected using questionnaire sets distributed to 300 multilingual participants, who completed them. Audio tapes were also used to capture the code switching conversation in a natural setting. The data were examined within the sketch of Myers-Scotton's Matrix Language Framework that is applicable in the identification of Matrix and Embedded language. The result of the investigation established that Code switching structures are systematic and not chaotic. This study hopes to provide insights on the code-switching phenomenon in a multilingual society. Additionally, the findings of this study will be of value for the development of code-switching studies in the structural area.

Keywords: Code Switching, Matrix Language, Embedded Language, Multilingualism

Introduction

Code Switching (CS) is viewed as a bilingual/multilingual practice that is used not only as a conversational tool, but also as a way to establish, maintain and delineate ethnic boundaries and identities. Code-switching is defined as the alternation of two languages within a single discourse, sentence, or constituent (Grosjean, 1982; Poplack, 2000; Clyne, 2000). Grosjean (1995) defines CS as shifting completely (emphasis added) to the other language for a word, a phrase, a sentence, etc. The study of CS has been carried out from linguistic (structural), sociolinguistic, pragmatic, psycholinguistic and applied linguistic perspective. The current study is on the structural features of CS and the syntactic constraints governing its operation. The structural approach in the current study tries to identify the structural features of morphosyntactic patterns underlying the grammar of CS in a multilingual society. That is, Mwea East Sub County which where eight languages are spoken

(Gikuyu, Kikamba, Kimberee, Kiembu, Kimeru, Kiswahili, English and a local pidgin called Githungu kia nguku). The eight codes in Mwea makes ML a complex matter that warrants an analysis.

In code switching studies, the dominant language is often called the matrix language, into which elements from the embedded language are inserted.

"...the participating languages are labelled in the following way. The 'base' language is called the matrix language (ML) and the 'contributing' language (or languages) is called the embedded language (EL) Myers-Scotton, 1993:20).

Chun (2001) observes that the ML is the language that plays predominant role in language production and its grammar sets the morphosyntactic frame for CS instances. ML might be identified as the first language of the speaker or the language in which the morphemes or words are more frequently used in speech. Matrix language is the language that determines the syntax of a code switching instance and its presence is obligatory in the instance. The embedded language in the CS instance can be one or more and is the code of a lesser degree of contribution in the CS instance. EL helps in completing the codeswitching instance in the formula, "matrix language plus embedded language" makes CS instance (hereafter referred to as ML + EL). According to the Blocking Hypothesis, in ML + EL constituents, a blocking filter blocks any EL content morpheme which is not congruent with the ML with respect to three levels of abstraction regarding subcategorization. "Congruence" is used in the sense that two entities, linguistic categories in this case, are congruent if they correspond in respect of relevant qualities. The three levels of abstraction are: First, even if the EL realizes a given grammatical category as a content morpheme, if it is realized as a system morpheme in the ML, the ML blocks the occurrence of the EL content morpheme. A content morpheme is often called an "open-class" morpheme, because they belong to categories that are open to the invention of arbitrary new items. They can be made-up words like "smurf", "nuke", "byte", etc. and can be nouns, verbs, adjectives, and some prepositions. A system morpheme, e.g. function words and inflections, expresses the relation between content morphemes and does not assign or receive thematic roles.

Second, the ML also blocks an EL content morpheme in these constituents if it is not congruent with MLA content morpheme counterpart in terms of theta role assignment and thirdly Congruence between EL content morphemes and ML content morphemes is realized in terms of their discourse or pragmatic functions. Neil (2002) observes that the EL in a code switching instance can be one or more. Matrix language (ML) is a concept characteristic of bilingual and multilingual societies. As the speaker code switches, he or she has a certain code as ML in the code switching instance and since there are eight codes in the area under study, ML is a complex matter. Most researchers whose interests are on the structure and syntax of CS agree that it is a rule-governed linguistic behaviour and so has a grammar (Gumperz 1976; Hoffman, 1991; Pfaff 1979; Poplack 1980, Poplack and Meechan (1998); Myers-Scotton 1993; 1995; Silva (1994). A number of rules and constraints have been proposed. Some of the constraints are specific to certain language pairs, while some are said to be universal. Most of the claimed universal constraints have been criticized, rejected or modified.

The aspect of constraints was critical in the investigation of ML+EL forms in the current study. A Matrix Language island is a constituent composed entirely of Matrix Language morphemes. Myers-Scotton and Jake (2002) points out that when a speaker is proficient enough to make a sufficient grammatical structure in the ML, it is called classical code-switching. It is also possible that some 'speakers do not have full access to the grammatical frame of the intended ML, part of the abstract structure comes from one variety and part from another'. This type of CS is known as composite CS. The assumption of the model is that the ML provides most of the words, especially functional

words/morphemes, while the EL contributes few lexical items that must fit in properly into the appropriate slots in the sentence structure. Chan (1998:2) notes that in a more up-to-date version of the MLF model, it is proposed that an EL word (or more abstractly the lemma of this EL word) has to be congruent with its ML equivalent for it to be inserted into the code-switched sentence. Lemma is defined as the morphological and syntactic properties which a word inherently possesses. These properties determine its co-occurrence and selection restrictions (Levelt's, 1989; Diane and Michael, 1991). Neil (2002) on the other hand observes that simplification is important CS. The speaker does away with structures such as preposition and is left with the content words. In semantics, parts of meaning that are not central to communicating the required message are left out. Other characteristics are overgeneralization, transfer and convergence. A speaker transfers structures of one language to another in his or her communication.

Currently, there is much research on the subject of CS as a linguistic peculiarity typical of bilinguals and multilinguals; however, the true syntactic nature of CS is still only vaguely defined. Mwea East Sub-County includes a large percentage of native speakers of a variety of other languages that is 8 languages. Thus, the aspect of CS is bound to occur to enhance interaction. However, as much as CS is a quite normal form of bilingual interaction it requires a great deal of bilingual competence. Moreover, the complex linguistic situation in Mwea East Sub-County, calls for a linguistic research in a multiligual society. Therefore, the current study is set to investigate the forms of matrix language + embedded language formula in Mwea East Sub- County Kirinyaga County Kenya. Understanding the rules governing multilinguals in the modes of CS they use is likely to open new possibilities in the field of conversational linguistics. The current research on the forms of ML+EL will add knowledge on language in a multilingual society and provide a better understanding of CS patterns. This research, therefore, has important implication on the structural aspect of CS in a multilingual society.

Theoretical framework
Matrix Language Framework Model
The concept of the MLF model is influenced by psycholinguistic theories Grosjean, (1988), that is, differential activation of base language and guest language, the different retrieval process of closed class items and open items in Garret's (1975) speech error study and Levett (1989) in the mental lexicon linking conceptual information and grammatical function. Myers-Scotton's (1993) criteria are more structurally based. MLs provide abstract grammatical frames where ELs (Myers-Scotton, 2002:55; Myers-Scotton & Jake, 2000:2). The MLF is a production- based theory used to explain the morphological, grammatical, and syntactic coordination of various language units in code switching speech. The premise of this theoretical frame states that the matrix language (base language) exists as a dominant language frame into which the code switches are inserted as the embedded language (guest language) items. Understanding of the MLF proposed by Myers- Scotton is impossible without identifying the constraints of the code switching behaviors, and the 4m model. Myers--Scotton and Lake (2000) characterized the 4--m model as the one offering an explanation about how language production works, and how linguistic competence is linked to speakers' performance. The 4M model explains how the content morphemes (occurring more freely, participating in the thematic grid) differ in their access from functional elements, system morphemes (similar to functional, or closed--lass elements), and how system morphemes differ from each other.

The existence of constraints stems from the fact that code switching has proven to be rule--governed (Poplack, 1980). MLF makes a distinction between content and system morphemes is crucial in identifying the ML. Content morphemes, e.g. nouns, verbs, adjectives and some prepositions, express semantic and pragmatic aspects and assign or receive thematic roles. These are essential to convey messages in communication. System morphemes, e.g. function words and inflections, express the relation between content morphemes and do not assign or receive thematic roles. They are essential in building grammatical frames. The early system morphemes are usually activated at the lemma level, which means that th ey are basic components of the mental lexicon, the fundamental knowledge of an individual. However, unlike content morphemes, early system morphemes do not receive or assign the matic roles; they are activated at the lemma level because lemmas underlying the content morphemes point to them (Jacobson, 2001). Early system morphemes contribute to the for mation of the conceptual structure connected to the lemma, which means that they are con ceptually activated, similarly to content morphemes (Jacobson, 2001). Some of the prepositi ons that are used in the English language are classified as early system
morphemes

Research Methodology
Research Design
A descriptive design was used to guide the current study. This research has a purpose to examine the forms of ML+EL concept in code switching.
The qualitative analysis identifies the code switching patterns according to syntactic categories.
The question of the study is: What are the forms of ML +EL formula in Mwea East Sub- County multilingual society?
The data analysis process involved a thorough investigation of the syntax of coded phrases, s entences, or single words to determine the type of morphemes used in each of the respondents' utterances.

Data Analysis
Forms of ML+EL
The purpose of the study was to investigate the forms of ML+EL Concept in CS instances. Through the application of the MLF model the study found out that when an intra-sentential CS occurs, the distribution of two languages is asymmetrical. Myers-Scotton's (1993) criteria are more structurally based. MLs provide abstract grammatical frames where ELs are inserted.

Syntactic patterns of CS
The findings of the present study on the syntactic patterns of ML+EL forms support the theory of congruence between ML and EL. The MLF is the dominant language that sets the morphosyntactic frame for ML and EL constituents. The EL in Mwea East Sub County was content words with system morphemes. This was in line with the MLF Content-system morpheme distinction as postulated by (Myers-Scotton, 1993). The findings of the current study show that content morphemes, e.g. nouns, verbs, adjectives and some prepositions, expressed semantic and pragmatic aspects and assigned or received thematic roles. System morphemes on the other hand, e.g. function words and inflections, expressed the relation between content morphemes. The system morphemes were essential in

building grammatical frames. The examples in Table 1 below show code switches with content and system morphemes and how they determine the forms of ML+EL.

Data	Languages involved	Form of Ml+EL formula
(i) Nindilamu_LAIKI_te... [I … him/her…] ..*UkinaciBRINGI* […and then you … them]	Gikuyu / English Kimbeere / English	Intrabound morpheme Intrasentential
(iii)…**gu_TAIM_anga** [… to wait for something to happen]	English/ Gikuyu	Simplificational

Table 1 above shows that the content system morphemes determine the structure of ML+EL forms. The survey data also showed that the content morphemes bear the fundamental lexical meaning ascribed to particular speech item. The prevalence of certain morphemes in CS indicates the striving for congruence in the structure of the ML; in cases when the subsequent congruence of structure is impossible, the speakers choose to continue CS to avoid ungrammatical utterances.

Manifestation of the 4M Model in the Use of ML and EL
The content-system morpheme opposition of the MLF model is refined and an extended version "4-M model" as proposed by Myers-Scotton & Jake (2000; 2001). The first category of 4M model is Content morphemes.Content/system morphemes are distinguished according to whether they assign/receive a thematic role or not. Content-morphemes assign/receive a thematic role and are activated at the lemma level. They are directly elected according to the speaker's intention. The content system morphemes were informed of intraword and intrasentential CS as shown in Table 2 below. The content words switched in EL had a thematic role while the system did not.

Table 2
Content system morpheme in Intraword and intrasentential switching

Data	Languages involved	Form of Ml+EL formula
i) *KANGWENDERUA KERITU NIATUmishi?* [is my daughter to be married off by leaders?]	Kiembu, Kiswahili	Simplification
(ii)…*nDAVAINDI nthaka ikivinyuria* [I … young men running.]	English/Kimbeere	Intrasentential`
iii)…**niuka_ANDJAST_aga.**_AM COMING_ [you will be…]	Gikuyu, English	Intrasentential

Table 2 shows that the content morphemes are perceived as morphemes functioning for conveying the semantic content; they are mostly found among nouns, adjectives, verbs, and some prepositions. Table 3 and 4 further shows the 4M model.

Table 3
Code switches in Content words 1

Data	Languages involved	Form of Ml+EL formula
i) *wamaanisha nini*? ’*NIKIITI*’ *ni nini*? [What do you mean? 'Nikiiti' is what? –eating food no nini]	Kiswahili/English	Intrasentential`
(ii)…*wakani***twara** […they took me]	Gikuyu/ Kiswahili	Intrasentential`
iii)*BIRITHON*i**ini** [in prison]	Gikuyu/ English	Simplificational

Table 4
Code Switches in content words 2

Data	Languages involved	Form of Ml+EL formula
i) *NINASAVIRE* [I suffered]	Kiembu / English	Intraword
(ii)KUTHUKUMA NA VITII ni wega muno […working very hard is good]	Kikamba/Gikuyu	Intrasentential`
iii)**NYINDO KIRO MICIRO** *na SONDA* [I like eating rice and …]	Githungu kia nguku/English /Kiswahili	Intraword

The findings of Table 2, 3 and 4 of the study shows that most code switches occurred with content morphemes: nouns, verbs, and code switched adjectives. The fact signals the comfort with which participants resort to CS in content morphemes, which can occur due to a number of reasons. One reason may between of the cultural origin, i.e., the participants wanted to convey a specific meaning about a specific object inherent in their culture (Agikuyu, Kimbeere or Kiembu). Second, the code switches in content morphemes also signal the convenience in using certain terms or words in the mother tongue that may be unknown in the second language.

The types of CS that determined the ML+EL can further be summarized as:

(i) Intrabound morpheme form of ML+EL formula
The intrabound morpheme demonstrated the aspect of late outsider system morphemes. Outsider morphemes differ from bridge morphemes in that they "depend on grammatical information outside of their own maximal projection" (Myers-Scotton & Jake, 2000:100). They are structurally assigned at the positional/surface level. For example the 3rd person singular –s is a late outsider morpheme. This is the ML+EL formual where a bound morpheme intergrates two or so structures borrowed from two different languages.For example, **mu*LAIK*ire** (Gikuyu for 'was liked') is a bound morpheme that includes Gikuyu and English.

'**mu**' is a Gikuyu morpheme for third person singular. '_LAIK_' is an English word 'like' that has been forced to take the Gikuyu grammatical structure while '**ire**' is the Gikuyu morpheme for past tenseness replacing the English regular verb past tense marker 'ed'

The whole structure '**mu**_LAIK_**ire**' is itself a bound morpheme comprising of English and Gikuyu structures. It is worth noting that _ire_ is an early system morpheme**.** If a system morpheme is activated at the lemma level, it is an early system morpheme. Early morphemes do not play a thematic role but they contribute to the mapping of the conceptual structure to the lemma like content morphemes. Myers-Scotton & Jake (2000) defines early morphemes as follows: Early system " are always realized without going outside of the maximal projection of the content morpheme that elects them" and "their form depends on the content morpheme with which they occur". Examples of early system morphemes in English are determiners, plural-s, some prepositions, etc. For example the Gikuyu morpheme _ire_ contributes to the tense of the sentence_**.**_

(ii) Simplificational form of ML+EL formula
This is the form of ML+EL formula where the languages involved in the codeswitched instance lose certain structures, a behavior that makes the whole instance be realized in a manner acceptable to the code that is determining syntax. An exemplification of this is found in the word '_BIRITHON_**ini'.**The word in English is 'in prison'. It includes Gikuyu and English. The English sound /p/as in /pri/does not existin Gikuyu. In Gikuyu it has been realized as /f/which is written as/b/ .Gikuyu realises 'insideness' at the end of the noun and so its insideness morpheme '**ini**' comes to at the end of the word prison. /s/in Gikuyu is realized as /sh/ and therefore the English structures are simplified and then the codeswitching instance is realized as 'birithoniini'.

(iii) Intraword ML+EL formula
In this form of ML+El formual the codeswitched ininstance is reflected by a singleword. The behavior was discovered in the course of the research. Two or more languages are involved in the same word. An example of this form of ML +EL formula is the word "_se_**thera**". It is one word that includes Gikuyu and Kiswahili. In Kiswahili the word is _"chezea "_ meaning "play for" in English while in Gikuyu the word is "**thakira**". Gikuyu realizes /z/ and /s/ usually as //. Continuity morpheme as /ae/- _chezea-_ is realized as /era/ and therefore Kiswahli is forced to accept and fit in the Gikuyu syntax. "_Se_**thera**" shows Kiswahili and Gikuyu as the two languages involved in a codeswitching instance. The code switching instance shows the intraword ML+EL formula. Intraword ML+EL formula is the form of ML+EL formula where the behavior is shown within a single word (as pointed out earlier).

(iv) Reduplicational –intersentential form of ML+EL formula
Reduplication is the repetition of the same word and the meaning of the combined repetition is different from the meaning of each one of the individual words. Combined repeat of the same word is a usual behavior in most of the languages spoken in the research site. Reduplicational-intersentential form ML+EL is a behavior whereby a word is combined with itself or repeated in a codewsitched instance that involves two or more languages. Examples of this behavior"**coka coka na thutha.** _Wavota_ **gukinya** IVETI (reverse a little bit. You might hit against the old women).

Another example of reduplicational form of Ml+El formula is in the sentence "_Kimundiu niathiica thiicagire_. _BE GO GOING_ **onawe**" (Kimundiu decided to go…) Kimbeere and Gikuyu accept reduplicaxional behavior but English does not. However since the synax is controlled by Kimbeere, English is forced to accept a behavior (reduplication) which is remote to it (English). As a result, the

speaker in the codeswitched instance says "*BE GO- GOING*" a reduplicational behavior which is a carry over from the instance which includes Gikuyu and Kimbeere. (*Kimundiu niathiicathiicagire*) which allow reduplication.

(v) Interword ML+EL formula
Interword ML+EL formula is an ML+EL structure that reflects or involves two or more words at a level below sentence level in the language hierarchy. The formula involves two or more languages. The research discovered this as a usual linguistic behavior in Mwea as representing multilingual societies. It (behavior) however had not been investigated by earlier reseachers.
An example of an instance showing ML+EL interword formula is **kingitu*FITI***(can fit us)
This structure involves Gikuyu and English and it covers three words-'**kingi** (it can) Gikuyu'**tu**'(us) and '*FITI*'(fit) English(which is realized as '*FITI*' in Gikuyu which is controlling syntax. Gikuyu does not allow a consonant at word-end position unless its followed by a vowel.

Another example is '**ningu*ARAU***'which involves two languages which are Gikuyu (**ningu**-am to) and English (allow) which is realized as '*ARAU* (in Gikuyu)
As observed in the two examples, the syntax of the instance in each case is controlled or determined by Gikuyu, a language that has more morphmes than English in the instance.

(vi) Intrasentential form of ML+EL formula
This is a behavior or a CS instance that is within a sentence. Two or more languages are involved. This form of codeswitching instance is in existing literature. Its presence is confirmed by the current research. An example of this type of ML+EL formual is '**NYINDO KIRO MICIRO** *na* *SODA*'. This instance involves a local pidgin, Kiswahili and English. In 'Githungu kia nguku' (A local pidgin), the communication is "I like eating rice.....' *na* 'is Kiswahili for 'and'.

(vii) Intersentential ML+EL formula
ML+EL can exist also among sentences. This is a form of ML+EL formula that covers two or more sentences involving at least two codes- knowledge on this behavior exists in existing literature and so its discovery in the site confirms its presence

Conclusion
The purpose of this paper was to investigate the structure CS instances used by the participants in multilingual setting of Mwea East Sub-County. CS is a phenomena which appears in multilingual community where two or more languages or language varieties are used in a single conversation, in this process even within a sentence. From this study, we conclude that CS- is a natural, creative and innovative way of communication in a multilingual society. The 4M model is reflected in the choice of ML and EL by participants. It is obvious that content morphemes and early system morphemes are mostly used in the EL, which simplifies integrating the structure of the EL into ML used, and increases the congruence of CS forms of 8 languages in the current data. The findings also reveal that the types of CS determined the Formular of ML+EL. The types of CS that were established in the current study were: Intersentential, Interword, Simplificational, Intrasentential, Intraword, Reduplicational Intersentential and Intrabound morpheme.

References

Chan, B. (1998). Functional heads, Cantonese phrase structure and Cantonese-English code-switching. UCL working papers in Linguistics 10.
http://www.phon.ucl.ac.uk/home/PUB/WPL/98papers/chan.pdf

Chun, E. (2001). The Construction of white, black and Korean. A merican Identities through African American Vernercular English. Journal of Linguistic Anthropology 11(1) 127-128.

Clyne, M. (2000). Constraints on code-wwitching: how universal are they? In Li WEI. The Bilingualism Reader. UK: Routledge.

Creswell, J. (2003). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches.* London: Sage

Diane, L., & Michael, H. (1991). *An introduction to second language acquisition research*: Longman. Britain.

Garret, M. F. (1975). The Analysis of Sentence Production. In G. Bower (ed.) *Psychology of Learning and Motivation, 9*: 133-77. New York: Academic Press.

Grosjean, F. (1982). *Life with two language. An introduction to bilingualism.* Cambridge, MA: Harvard University Press.

Grosjean, F. (1988). Exploring the Recognition of Guest Words in Bilingual Speech. *Language and Cognitive Progress. 3:* 233-74

Grosjean, F. (1995) A psycholinguistic approach to code -switching: the recognition of guest words by bilinguals. In: Milroy, Lesley and Muysken, Pieter, (eds). *One speaker, two languages. Cross-disciplinary perspectives on code-switching*, p. 259-275.

Gumperz, J. (1976). The sociolinguistic significance of conversational code-switching. In Papers on language and context, ed. by Jenny Cook-Gumperz and John Grumperz, Working Paper 46, Berkeley: University of California.

Jacobson, R. (2001). *Codeswitching Worldwide II.* Berlin: Walter de Gry

Hoffman, C. (1991). *An introduction to bilingualism*. Available on https://www.google.com/books.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation.* Cambridge, Mass: MIT Press.

Myers-Scotton, C. (1993). *Duelling Languages. Grammatical Structure in Codeswitching.* Oxford: Clarendon Press.

Myers-Scotton, C., and Jake, J. L. (1995). Matching lemmas in a bilingual language competence and production model:evidence from Intrasentential Codeswitching . *Linguistics 33* :981-1024

Myers-Scotton, C., and Jake, J. L. (2000). Testing a Model of Morpheme Classification with Language Contact Data. *International Journal of Bilingualism 4*/1, 1-8

Myers-Scotton C. (2002). Contact Linguistics: Bilingual Encounter and Grammatical Outcomes. Oxford: Oxford University Press.

Myers-Scotton, C., and Jake, J. L. (2001). Explaining Aspects of Code switching and Their Implication. In journal of

Myers-Scotton, C., and Lake, J. L. (2000). Four types of morpheme: evidence from aphasia, codeswiching and second language acquisition: an interdisciplinary of Language Sciences. Available at http://findarticles.com/p/articles/mi_hb195/is_6_38/ai_n28809900

Neil, T. (2002). *People skill*: Palgrace Macmillan. New York.

Pfaff, C. (1979). Constraints on language mixing: intrasentential code switching and borrowing in Spanish/ English. *Language 55,* 2, 291 – 318

Poplack, S. (1980). "Sometimes I'll start a Sentence in Spanish y terminó en español: Toward a typology of codeswitching." *Linguistics 18,* 581-618.

Poplack, S., & Mechan, M. (1998). How languages fit together in Code Mixing. International Journal of Bilingualism 2, 2. (127-138.

Poplack, S. (2000). Sometimes I'll start a sentence in Spanish y termino en Espanol: toward a typology of code-switching. In Li WEI. The Bilingualism Reader . UK: Routledge.

Silva, C. (1994). *Language contact and change*:Clarendon press. New York.

Stange, K. C., Miller, W. L., Crabtree, B. F., O'Connor, P. J., Zyzanski, S. J. (1994). Multi-method research: approaches for integrating qualitative and quantitative methods. J Gen Intern Med.