

Socio-psychological factors or grammatical structure?

English verbs in Swahili-English CS plus EL infinitives in CS elsewhere

Carol Myers-Scotton

myersssc3@msu.edu

Michigan State University

A Nairobi business woman commenting on her language use:

*...U-na-feel free ku-ongea English,
u-na-feel free ku-ongea Kiswahili.*

‘... You feel free to speak English, you feel free to speak Swahili.’

Why are there so many English verbs in Swahili- English codeswitching in Nairobi?

(I begin by briefly discussing the socio-psychological motivations for many English verbs appearing in codeswitching in data recorded in Nairobi in 2013. But I move on to concentrate on offering an explanation for the grammatical structure of English Embedded Language (EL) infinitives in Nairobi, but also other infinitives elsewhere.)

The socio-psychological argument: Swahili-English Codeswitching (CS) is “the unmarked choice”

CS is expected in many informal interactions, no matter what the socio-economic level or levels of English proficiency. Some of the less proficient speakers use more!

Unmarked choices are what is expected, given the specifics of an interaction.

The notion of unmarked and marked choices comes from the Markedness Model (Myers-Scotton, 1993b, *inter alia*).

Brief examples are on the next few slides.

Example of a unmarked choice. A school principal is in Nairobi on a visit. He speaks Swahili, English, and his first language. He wants to visit a friend at his workplace, an international company.

•Guard (Swa): *Ninapenda nakusaidia namna gani?*

‘In what way do you want me to help you?’

•Principal (Swa): *Ningependa kumwona Peter Murengo.*

‘I want to see Peter Murengo.’

•Guard (Swa): *Bwana Peter hayuko. Ingia na uende kwa **office of inquiries** na umngoje.*

‘Mr. Peter isn’t here. Go in to the office of inquiries and wait for him.’

•Receptionist (Eng): *Good morning. Can I help you?*

•Principal (Eng): *Good morning. I came to see Mr. Murengo.*

•Receptionist (Eng): *He’s out but he will come back soon. Please have a seat and wait for him.*

Example of a marked choice (1993a: 40). Setting: Nairobi post office.

- Clerk (Swa): *ee-sema*. ‘What do you want?’ (lit. ‘speak’)
- Customer (Swa): *Nipe fomu ya kuchukua pesa*.
‘Give me the form for withdrawing money.’
- Clerk (Swa): *Nipe kitabu kwanza*. ‘Give me your passbook first.’
- Customer (Swa): *Hebu, chukua fomu yangu*.
‘Say, how about taking my form?’
- Clerk (Swa): *Bwana, huwezi kutoa pesa leo kwa sababu hujamaliza siku saba*.
‘Mister, you can’t take out money today because you haven’t yet finished seven days [since the last withdrawal].’
- Customer (switching to Luo): ***Konya an marach***
‘Help, I’m in trouble.’
- Clerk (also switching to Luo): ***Anyalo konyi, kik inuo kendo***.
‘I can help, but don’t repeat it.’

Examples of “Codeswitching as the Unmarked Choice”

A Nairobi business woman explains that she lost her phone:

. . . *i-li-snatch-i-w-a*, *so a good friend of mine*

CL9-PAST-**snatch**-APPL-PASS-FV

a-na-ni-hurum-i-a, *a-ka-ni-buy-i-a* *simu*

3S-CONSEC-OBJ-**buy**-APPL-FV

‘. . .it was snatched, so a good friend of mine took pity on me, and he/she bought a phone for me.’

CS as the unmarked choice with many English EL bare infinitives

Note how in the following examples both Swahili and English are freely used, making CS the unmarked choice.

Also note how CS pervades across social groups from business women to market women to both young men and women.

Two Nairobi market women are just chatting, commenting on something they noticed:

Ohh, hata sikuwa ni-me-notice hiyo.

1S-PERF-notice

*Kwangu sijamaliza **rent** ya **last month**.*

Siku hizi zi-na-rush haraka.

CL10-PROGR-rush

‘Oh, even I noticed that. At my place, I haven’t yet paid the last month’s yet. These days rush [by] fast.’

A mechanic and company driver are talking about getting an ID card

*. . . kuna **program prez** [president] a-li-launch*

3S-PAST-launch

i-na-it-w-a “Huduma Kenya.”

*U-ki-**apply** kama leo, **by next week***

2S-CONDIT-apply

*u-ta-kuwa usha-**get your ID.***

ALREADY-get

‘There is a plan the president launched—it’s called “Aid Kenya.” If you apply for example today, by next week you already will get your ID.’

Note how integrated into the ML verb phrase structure these English verbs can be.

Go back to slide 8. First, note that the English verb, **snatch**, receives a prefix to mark class nine for the meaning “it” that refers to the phone and then *-li-* that marks past tense.

The English verb is also embedded in derivational suffixes. First, *-i-* the applied form, marks “movement from” here and the *-w-* is a suffix marking passive.

The second English verb **buy** receives a subject prefix for 3rd person singular and then a suffix *-i-* that marks dative or benefactive here.

Lots of English by Relatively Uneducated Youths

Two young men conversing on a matatu:

A. *Na-fanya a lot of tizi (Sheng=exercise), in fact ni-li-join
1S-PAST-join
basketball, where tu-na-practice a lot—kama ku-beba chuma iko
1S-PROG-practice*

very heavy en it really umiza me normally

hurts (missing verbal inflections on –umiza)

‘I do a lot of practice, in fact I joined basketball, where we practice a lot—such as to lift/carry metal is very heavy and it really hurts me normally.’

B. *Mazee hiyo zoezi ni noma (Sheng/slang: difficult).*

Mimi I fear coz u-ta-doze when u-me-choka sana.

2S-FUT-doze

‘My friend, get used to it--it’s difficult. As for me, I fear because you will doze when you are very tired.’

More Examples Including an English infinitive with no infinitive marker

The first example shows an example of a causative suffix *-iz-* as well as a passive marker.

The second shows a bare English infinitive without either the *ku-* that marks infinitives in Swahili. It would be used here in standard Swahili. But it is also without “to” that can mark English infinitives in English, but not in this particular English construction.

A young man and a woman talking about not having an identity card:

Unajua u-na-wez-a peneliz-w-a

penalize-PASSIVE-FV

‘You know you can be penalized.’

Business woman asked how much she uses Facebook:

Haki si-wezi estimate

1S.NEG-can-NEG estimate

‘The right amount I can’t estimate.’

Why is it so easy to inflect English infinitives with Swahili (ML) verbal affixes?

Myers-Scotton & Jake (2014) make an argument regarding why CS seems so effortless and fast.

We argued that one reason is the presence of EL nonfinite forms. They are less “costly” in various ways including response time, but that claim has yet to be tested by psycholinguists. We suggested it is because they do not carry the same “grammatical baggage” that EL finite forms would have.

However, we did not go into details on all the characteristics of EL infinitives that makes their employment in CS fortuitous.

The goal here is to offer a more general, but also more precise, explanation of the nature of infinitives in CS.

This explanation relies on the SM Principle and also the Differential Access Hypothesis, as well as the classification of morphemes under the 4-M model.

This explanation emphasizes how the features of morphemes at an abstract level affect their level of participation in CS.

The Differential Access Hypothesis

(Myers-Scotton, 2002, 2005)

The different types of morpheme under the 4-M model are differentially accessed in the abstract levels of the production process. Specifically, content morphemes and early content morphemes are accessed at the level of the conceptualizer, but late system morphemes do not come salient until the level of the formulator when large constituents are assembled.

At the abstract level, the major division in the 4-M model is between conceptually-activated morphemes and structurally-assigned morphemes.

Conceptually-activated: Content morphemes and early system morphemes (Earlies)

Structurally-assigned morphemes: Late SMs, both bridges and outsiders).

Outsiders are those that co-index critical relations in any clause; examples are agreement markers. I refer only to outsiders here and call them “Lates” in this presentation.

The use of the terms “Early” and “Late” is metaphorical. We have no evidence about timing in our model of language production.

The Matrix Language Frame (MLF) model and Its System Morpheme Principle

The System Morpheme Principle (SM Principle) states that only one of the participating languages in CS supplies the system morphemes that—under the 4-M model—are called “late outsider system morphemes” and will be called “Lates” in this presentation. This language is called the Matrix Language (ML).

Swahili is identified as the Matrix Language (ML) and English is the Embedded Language (EL).

Levels of Activation in a Model of Production

Earlies are indirectly elected at the level of the conceptualizer in a language production model and Lates are not.

This distinction makes a difference in what occurs in CS regarding EL non-finite verbs.

My colleague, Jan Jake, also a co-author of the 4-M model (cf. Myers-Scotton & Jake, 2000) contributed to the arguments that follow.

**EL Infinitives differ in their makeup regarding
System Morphemes as they are defined
in the 4-M model**

All contain a content morpheme, but not all
contain the same types of SM — are they
Earlies or are they Lates?

Some details from our model of language production (following Levelt, 1989 and later)

Content morphemes encode the speaker's intentions regarding semantic and some pragmatic information. They are directly elected at the level of the Conceptualizer. Any early SMs that flesh out content morphemes are indirectly elected at the same level as their content morpheme heads.

Late SMs are not available at the conceptual level. Structurally-assigned morphemes (Lates) only become salient at the level of the Formulator when they participate in the makeup of the larger constituents in a clause.

Certain multi-morphemic lexemes, those containing late SMs, such as case, may be "delayed" (i.e. they are not available at the conceptual level). This may affect the lexemes' degree of participation in CS and in other ways. For example, German determiners are multi-morphemic (cf. Myers-Scotton & Jake, 2015 forthcoming).

English infinitives may include a content morpheme and a late outsider SM-- that is, “to” plus a content morpheme.

But in English, “to” is considered a Late SM. The reason for this status as a Late is that its presence or not depends on the relevant syntactic structure considered.

For example, “to” is present in some monolingual English constructions, but not every time the infinitive content morpheme appears (e.g. compare *I want to swim often* with *Let him go or He made him finish by noon.*)

When English infinitives appear in CS and English is the EL, the “to” never appears—This also suggests it is a late SM.

Recall the SM Principle. This means that Lates can come only from the ML in CS. That is, in Swahili-English CS, only a bare English infinitive (content morpheme only) can be accommodated by the Swahili frame if the construction calls for such an element from the verbal complex.

The motivation behind the SM Principle

Because CS cannot both accommodate the critical grammatical information that Lates carry from more than one language and maintain a uniform structure across a CS clause, the SM Principle disallows EL Lates.

Note that *ku-*, the Swahili infinitive marker, is a Late; it participates in the Bantu noun class system that marks locatives with renditions of *ku* or other locative markers. But in Swahili-English CS Swahili is the ML, not the EL, and therefore its Lates are allowed.

In short, if the EL infinitive marker is a late – only the infinitive marker from the Bantu language can appear.

Here is a Swahili-English example from the Nairobi corpus:

Ni-li-try ku-m-call

1S-Past-try Infin-Obj-call

‘I tried to call her.’

Thus, the explanation for so many English verbs in Swahili-English CS

Because the English infinitive in CS can contain only a content morpheme, it can appear in an ML frame in the verbal assembly with no EL grammatical information that is disallowed (i.e. the material that Lates carry or point to).

Of course this implies that the ML frame is flexible enough to accommodate and integrate such forms.

One question remains: while it is clear that the bare infinitive carries EL conceptual meaning, can it assign theta-roles as finite verbs do?

Either the answer is “yes ”

Or it is possible the ML frame is able to pull in Generalized Lexical Knowledge about the theta-roles and their following argument structure that would be consistent with the “conceptualization” contained in the infinitive form.

In either case, I argue the infinitive is a verb, not a noun.

How English EL infinitives appear in other corpora supports the argument here

Most Bantu languages in the CS literature pattern in line with the Nairobi Swahili-English corpus. Examples from Tanzanian Swahili-English and Xhosa-English follow.

Tanzanian Swahili-English (Bwenge, 2010: 32)

....wafugaji hawa saa nyingine tulikuwa tukiwalaumu kwamba wa-na-move **from one place** kwenda kwingine.

‘... sometimes we were blaming livestock keepers that they are moving from one place to another.’

Xhosa-English (Myers-Scotton, 2004, Mpulelelo file)

Abantu ba-se-**South Africa** ba-feel-a
ba-threatened ...

‘People of South Africa they feel threatened ...’

Xhosa-English Myers-Scotton (2004, Mplulelo file)

... i-**condom** ... zi-ya-distribute-w-a na-se-**clinic** ...

‘... condoms are distributed by the clinics...’

The same explanation for bare English EL infinitives in ML frames also applies for Ewe-English in Ghana.

Ewe-English (Amuzu, 2010: 162)

Nye me **believe**-na e-fe nya d̩desiad̩de o

‘I don’t believe everything he says.’

(na=habitual)

What if French, not English, is the EL?

French infinitive markers are Earlies, not Lates, so they can appear with their content morpheme heads, as in Bukavu Swahili-French

Bukavu Swahili-French (Goyvaerts & Zembele, 1992:74)

Kweli? Ndjo u-li-esquiver ile **question?**

Siku-rever ata!

‘Really? I that why you avoided that question?
I didn’t figure it out at all.’

What if Spanish is the ML? Spanish-English
(D.Smith, 2013 ms.)

Voy a send-er te un **email**

‘I’m going to send you an email.’

What if Arabic is the ML? (Chakrani, 2014 P.C.)

Marrakesh Arabic

Alash **block-er-ti-ni?**

block-er-2S-1S

‘Why did you block me? (soccer or on the road)’

An example when the infinitive marker of both the ML and EL do appear

The Chiluba infinitive marker is a Late and the French one is an Early. Chiluba-French (Kawamwangalu, 1987:169)

Ndi m-ku-téléphoner bwa ku-ku-**informer** ne je nai pas oublié problème u-mu-vuw-a m-n-soumettre m=1S; n=CL9 (3S)

‘I’m calling you to tell you that I haven’t forgotten the problem [that] you submitted to me.’

Because Chiluba is the ML, the SM Principle does not affect it. The French marker is an Early and therefore also is not affected by the SM Principle.

An English EL infinitive when French is the ML

In Camfranglais, a mixed language that is a combination of French with English and elements from various indigenous Cameroonian languages, English as the EL also provides a bare infinitive.

Camfrangalis (French-English) A.L. Nchare (2010: 27, 42)

- Plus on **djoum**-ait dans la relation, mois elle do-ait.

‘The deeper we got into the relationship, the lesser (sic) she did it.’

- ...Connaissante le kô que nous know-ons.

‘... Knowing the country the way we know it.’

Consider mixed languages briefly in terms of the 4-M model and the SM Principle

In theory, in mixed languages, Lates from more than one language can occur as long as the second language has Lates only in one sub-system of the mixed language. However, only few of the languages that some classify as mixed languages actually do this. Consider Copper Island Aleut as one of the few.

Most of the so-called mixed languages that do accommodate system morphemes from more than one language only have Early SMs from a second language (e.g. the noun phrases of Michif include Earlies).

Of course, in a few cases, a few or single Lates from the EL or second language do slip through, but the generalization of the SM Principle still holds for the majority of cases in both CS and mixed languages.

There remain two other possible ways to integrate an EL infinitive into an ML frame.

First, the ML language can accept bare infinitives that are preceded by a “nativizer” element (from the ML). That is, the bare infinitive form appears, without the infinitive marker—whether it would be an Early or a Late SM.

Second, “do verb constructions” are employed.

The use of a nativizer

Here is an example:

In Brussels Dutch-French CS, a nativizer *-eer* (based on French *-er*) appears following the bare French verb and preceded by any Brussels Dutch inflections (Treffers-Daller, P.C., cited in Myers-Scotton & Jake, 2014: 519):

Hij ... pakt zijn vest en hij caress-eer-t zijnen hond terwijl

‘He takes his cardigan and strokes his dog ... ‘

What does the nativizer accomplish?

The nativizer, usually an ML Early SM, integrates the EL infinitive into the ML to the point that ML inflection can then be suffixed to the infinitive. The nativizer indicates the default class or conjugation if the ML has more than one conjugation or other verb classes. In effect, the nativizer adds an Early SM to the infinitive.

Consider –yi- a nativizer on Spanish infinitives when Mosestén is the ML (Sakel 2004:6, cited in Myers-Scotton & Jake, 2014: 520)

Digamos nuestros padres jike **suffrir-yi-in**
'We would say that our fathers suffered.'

Another Nativizer

Spanish-English CS (Pfaff, 1979 cited in Backus, 1996:220 as a nativizer))

- Los hombres me **trust**-ear-on
- ‘The men trusted me.’

The omnipresent “do verb construction”

This strategy makes it possible to include an EL infinitive and the conceptual material is carried in a CS clause with no adjustments to the infinitive.

The structure of such constructions is well-known: all TMA inflections and any derivational suffixes on a “do” verb from the ML, with the EL infinitive as part of the predicate. This strategy is very freely used across different CS corpora.

Two examples of “do constructions”

- Chichewa-English (Simango, 1995 ms 1.4, p.5)

Mu-na-ka-chit-its-a **vacuum** galimoto.

2S-pst-consec-do-caus-FV vacuum car (vacuum is a verb)

‘And you had the car vacuumed.’

English nonfinite infinitive with a causative suffix following *chita*, the “do” verb from Chichewa.

- Turkish-Norwegian (Türker, 2000:117)

... Niye søk-e yap-ma-di-n?

why apply-INF do-Neg-Past-2S yap=“do”

‘... Why didn’t you apply?’

Concluding Remarks

The main theoretical argument presented here is that the claim/finding that Early SMs and Late SMs differ at abstract levels is behind the different forms that EL infinitives take in Swahili-English CS and in CS elsewhere.

Two hypotheses, one testable and the other not, underlie the argument. These are the SM Principle of the MLF model (testable) and the Differential Access Hypothesis.

Still, even without hard evidence about the abstract levels in language production, rather precise predictions are possible because Earlies vs. Lates are easy to distinguish, given the 4-M model's definitions of its four morpheme types based on their surface level representations. These augment the abstract definitions that have been highlighted here. Other arguments based on differences in typological features or such constructs as "morphological boundedness" offer more limited explanations or even predictions because of their inherent vagueness.

More Concluding Remarks

Whether specific predictions can be supported in other CS corpora depends on the nature of the specific ML's verbal system. But, to illustrate, here are two predictions that are supported: (1) In Swahili-English CS, the only infinitive marker that may occur comes from the ML (Swahili) because English "to" is a Late. (2) In CS pairs with French as the EL, the French infinitive can appear with its infinitive marker intact because it is an Early.

In Swahili-English CS, certainly, the cachet of an English verb is one important factor that is behind the high frequency of English EL infinitives in ML slots. Verbs have a certain "dynamic" quality; this may even add more connotations to the overall message. But because nonfinite verbs, as content morphemes, are conceptually-activated and carry largely just the speaker's intentions regarding semantic and pragmatic information, they slip into the ML frame without impinging on the ML frame.

Thanks to everyone whose data are cited here. Special thanks to my research colleague, Jan Jake, for putting on her thinking cap once again.

References

- Amuzu, A. (2010). *Composite codeswitching in West Africa: The case of Ewe-English codeswitching*. Saarbraten: Lambert.
- Backus, A.M. (1996). *Two in one. Studies in Multilingualism*, Vol. 1.
- Bwenge, C.M. (2010). *The tongue between, Swahili & English in Tanzanian parliamentary discourse*. Munich: LINCOM.
- Chakrani, B. (2014). Moroccan Arabic (Marrakesh) examples. P.C.
- Goyvaerts, D & T. Zembele. (1992). Codeswitching in Bukavu. *Journal of Multilingual and Multicultural Development* 13:71-81.
- Kamwangamalu, N.K. (1987). French/Vernacular code-mixing in Zaire: Implications for syntactic constraints. Chicago Linguistic Society. Cited in Myers-Scotton & Jake (2014).
- Levelt, W.J. M. (1989). *Speaking, from intention to articulation*. Cambridge, MA: MIT Press.
- Myers-Scotton, C.M. (1993a). *Duelling languages, grammatical structure in codeswitching*. OUP.
- Myers-Scotton, C.M. (1993b). *Social motivations for codeswitching: Evidence from Africa*. OUP.
- Myers-Scotton, C.M. (2002). *Contact Linguistics, bilingual encounters and grammatical outcomes*. OUP.

- Myers-Scotton, C.M. (2005a). Supporting the Differential Access Hypothesis: Code switching & other contact data. Kroll, J. & A. de Groot (Eds.), *Handbook of psycholinguistics*, pp. 326-348. OUP.
- Myers-Scotton, C.M. (2005b). Xhosa-English CS corpus. Unpubl.
- Myers-Scotton, C.M. & J.L.Jake. (2000). Four types of morpheme: Evidence from aphasia, codeswitching and second language acquisition. *Linguistics* 38, 6: 1053-1100.
- Myers-Scotton, C.M. (2013) New Nairobi Swahili-English corpus. Unpubl.
- Myers-Scotton, C.M. & J.L. Jake (2014). Nonfinite verbs and negotiating bilingualism in codeswitching: Implications for a language production model. *Bilingualism, Language and Cognition* 17,3: 511-525.
- Myers-Scotton, C.M. & J.L. Jake (2015,to appear). Implications from codeswitching for models of language production. In J. Schweiter (Ed.) *Cambridge handbook of bilingual processing*. CUP.
- Nchare, A.L. (2010). The morphosyntax of Camfranglais and the 4-M model. Phd diss., New York University: Department of Linguistics.
- Pfaff, C. (1979). Constraints on language mixing: Intrasentential code-switching and borrowing in Spanish/English. *Language* 55: 219-318.
- Sakel, J. (2004). *A grammar of Mosestén*. Berlin: Mouton de Gruyter.
- Treffers-Daller, J. (2013). P.C.