The Matrix Language Frame Model's Grammatical Constraints on Intra-Sentential Code-Switching and Hokkien-Mandarin-English Data from Singapore[1]

Liang Chua

Abstract

The Matrix Language Frame (MLF) model puts forward five grammatical constraints on intra-sentential code-switching. This article challenges the universality of these constraints. This article is based on a study in which the constraints and the model's hypotheses relating language shift to language attrition are tested using Hokkien-Mandarin-English data from a Singaporean Chinese family (Chua, 2001a; 2001d). Hokkien is a Chinese dialect belonging to the Southern Min group. The five members of the family which provides the data are: grandfather (GF), grandmother (GM), daughter (D), granddaughter (GD) and grandson (GS). All the mixed constituents in the data conform to the Morpheme-Order Principle. Four counter-examples to the System Morpheme Principle were found and eighteen (60%) of the Embedded Language (EL) islands contravene the EL Island Trigger Hypothesis. The predictions made by the EL Hierarchy Hypothesis are also not attested in the data.

1. Introduction

This article tests Myers-Scotton's (1993) grammatical constraints on intra-sentential code-switching on my Hokkien-Mandarin-English data. Under the MLF model, these constraints are the ML (Matrix Language) Hypothesis (section 2), the Blocking Hypothesis (section 3), the Embedded Language (EL) Island Trigger Hypothesis (section 4) and the EL Hierarchy Hypothesis (section 5). The ML Hypothesis consists of the Morpheme-Order Principle (section 2.1) and the System Morpheme Principle (section 2.2). Apart from Park's (2000) study of Korean-Swedish code-switching, I know of no other systematic test of Myers-Scotton's proposals. Even Park's study omits the Blocking and EL Island Trigger Hypotheses, which he claims are to a great extent abstract in character (p. 115). My study covers all the hypotheses under the MLF model.

Although the MLF model predicts convincingly the code-switching patterns in Myers-Scotton's Swahili-English data, Bentahila (1995) has

doubts about the universal applicability she claims for it. From the figures Myers-Scotton gives, her data appears to be consistently dominated by Swahili; the vast majority of switches are for single English lexemes, with far fewer EL islands and very few whole sentences in English. Also, the variety identified as the ML is Swahili almost all of the time. Swahili is acquired informally and is 'a language of solidarity among different ethnic groups' (Myers-Scotton, 1993:11). Bentahila (1995:139) suspects that the MLF model is unlikely to cope so well with less prototypical cases, such as those in which the ML is not an informally acquired variety of community identity, but an 'international' variety learned at school. Moreover, the postulation of an ML seems less plausible in discourse where the switches are mainly between clauses rather than for small constituents and within clauses.

Backus and Boeschoten (1996:130) have similar doubts about the universality of the model. Initial application of the model to their Turkish-Dutch data indicates that the model describes insertional code-switching (i.e. the type of code-switching found in Myers-Scotton's Swahili-English data) well, but may not cope equally well with alternational code-switching (i.e. the type of code-switching that results when speakers mix the languages in a more balanced way, yielding a high frequency of clause-level switching).

My data is in many ways different from Myers-Scotton's. There is no single variety which dominates the discourse consistently. Hokkien, Mandarin and English play almost equal roles, and thus the ML changes accordingly (the ML is identified using the ML Criterion[2]). With Myers-Scotton's data, English is the EL and Swahili is the ML all of the time. There is also a large number of inter-sentential switches, something which the MLF model does not cater for. In addition, the number of mixed constituents and EL islands is much lower: there are 168 mixed constituents and thirty EL islands in eighteen hours of tape recorded material, compared to Myers-Scotton's 374 mixed constituents and 121 EL islands in about twenty hours of material. Moreover, while Hokkien and Mandarin are acquired informally in the home environment, English is learned mainly at school.

As we move through this article, we will see that Bentahila is right in suspecting that the MLF model is orientated towards one particular type of discourse, and does not deal equally adequately with the patterns attested for some other communities.

2. The ML Hypothesis

The ML Hypothesis states that as an early step in constructing ML+EL constituents, the ML provides the morphosyntactic frame of ML+EL

constituents (Myers-Scotton, 1993:82). Two testable hypotheses follow from the ML Hypothesis, and they are stated as principles:

> 'The Morpheme-Order Principle: In ML+EL constituents consisting of singly-occurring EL lexemes and any number of ML morphemes, surface morpheme order (reflecting surface syntactic relations) will be that of the ML.
>
> The System Morpheme Principle: In ML+EL constituents, all system morphemes which have grammatical relations external to their head constituent (i.e. which participate in the sentence's thematic role grid) will come from the ML.'
>
> (Myers-Scotton, 1993:83)

The Morpheme-Order Principle is dealt with in section 2.1, and the System Morpheme Principle in section 2.2.

2.1 The Morpheme-Order Principle

Myers-Scotton (1993:83) asserts that evidence indicates that the Morpheme-Order Principle holds categorically. She cites examples from her own data, the Nairobi corpus (which involves Swahili and English), as well as from data sets of other researchers involving Ewe and English; Hindi and English; Alsatian and French; and so on. She asserts that the Nairobi corpus provides strong empirical support for the Morpheme-Order Principle; out of 374 mixed constituents, only one case (p. 236) is a counter-example to this principle. She claims that the case in question is not entirely a counter-example. Rather, it is an instance of a structurally marked choice for socio-pragmatic effect. This way of explaining away counter-examples to the predictions of the MLF model has been criticised by Meechan (1995:109). She considers it 'the greatest flaw in the theory' since even those utterances otherwise prohibited by the model are permitted as long as they serve some socio-pragmatic function. It is not clear why psychologically based structural limitations can be overridden by socio-pragmatic factors. Meechan feels that this gives the model so much power that it is difficult to imagine any single case that would constitute a clear counter-example.

Here are some of the examples from the Nairobi corpus:

2.1 Mungu anaweza yote muamini ataweza kubadilisha na utakuwa na
with

ma-mbo	m-engi	*new* - ma-pya	katika		ma-isha
CL 6-things	CL 6-many	new CL 6-new	in		CL 6-life

y-ako.
CL 6-your

'God is able to do all [if] you believe he will change you and you will have many new things – new in your life.'

(Myers-Scotton, 1993:85)

In all the examples in this article, the EL is italicised and the ML is unmarked. The mixed constituent *mambo mengi new*, which is literally 'things many new', shows the morpheme order of the ML, Swahili. With most modifiers, Swahili morpheme order is head-first and in clear contrast with English's head-last order for NPs.

2.2 Akikosa mlo siku moja anakuja kudai siku ya pili- a-na-ku-l-a
3S-PRES-INFIN-eat-

	plate m-bili	z-a	murram	a-ki-kos-a	
INDIC	plate CL 10-two	CL 10-of	maize	3S-CONDIT-miss-INDIC	

moja.
one

'When a student misses a meal one day he/she comes to claim [it] on the second day. He eats two plates of maize if he missed one.'

(Myers-Scotton, 1993:86)

Plate mbili za murram, literally 'plates two of maize', again shows Swahili morpheme order.

2.3 Unamuangalia *movements* y-ake z-ote...
movements CL 9-her CL 10-all

'You were watching all her movements...'

(Myers-Scotton, 1993:86-87)

In *movements yake zote* 'movements her all', the morpheme order is again that of the ML.

However, Myers-Scotton's claim that the Morpheme-Order Principle holds categorically has not gone unchallenged. Park (2000) presented counter-examples to this principle in his study of Korean-Swedish code-switching. One such counter-example is example 2.4:

2.4	haksaeng-tûl-hanthe	kasô	kû	/	sônsaeng-i		*Norrland* /
	student-PL-DAT	go	DEM		teacher-NOM		Norrland

Norrland	chulsin-ilakuyo	kûlaesô /	cangnan-ûloyo /	*Ta-Taiwan*-i	
Norrland	origin-COP	so	fun-INST	Ta-Taiwan-NOM	

cô : ki	cô: i	/	*Abisko*-e	ka-sô	*skog-en*-e
there	DEM DEM		Abisko-DAT	go-and	forest-ART-DAT

ka-sô	salacy-ôss-nûnte	ikô an	nao-n-taku
go-and	disappear-PAST-but	DEM NEG	come-PRES-DECL

'The teacher is from Norrland. He would [probably] say to [my] students for fun that I (lit. the Taiwanese) disappeared in the forest in Abisko and that I have not come [back yet].'

(Park, 2000:198)

In *skog-en-e* 'in the forest', the morpheme order is not that of the ML, Korean. In fact, the morpheme order comes neither from Korean nor from Swedish, the EL. The morpheme order between the Swedish head noun *skog* 'forest' and the Swedish determiner *-en* 'the' conforms to Swedish syntax. The morpheme order of the rest of the constituent is Korean.

Having introduced the Morpheme-Order Principle, I proceed to discuss it with reference to my data.

All but one of the 168 mixed constituents in my Hokkien-Mandarin-English data from Singapore conform to the principle. The exception in question contains an internal EL island (see Myers-Scotton, 1993:151-156), and because the violation occurs within the internal EL island, it does not really count as a counter-example to the principle.

2.5	D:	*Hái shì*	school	*de*	*nà*	*gè*	*exercise*?
		or	school	NOMS	that	CL	exercise

'Or (is it the) school exercise?'

(MC 148)

According to the guidelines laid out for defining a mixed constituent in Chua, 2001a; 2001b, the entire sentence is a mixed constituent. *De nà gè exercise* (*exercise* is treated as a borrowing in Mandarin) is an EL island in its own right, but because the smallest maximal projection which contains *hái shì* 'or' is *hái shì school de nà gè exercise*, *de nà gè exercise* is also regarded as an internal EL island within the larger mixed constituent. The morpheme order in the internal EL island is clearly that of the EL, since English does not permit the demonstrative *that* to follow the possessive marker *'s*. On the other hand, the demonstrative *nà* can occur after the nominalising particle *de* in Mandarin. However, internal EL islands are exempt from the predictions which apply to mixed constituents, according to Myers-Scotton. Bentahila (1995:138) points out that the recognition of the possibility of EL islands being situated within mixed constituents, while it may allow for some otherwise unexplainable examples, seems to weaken considerably the predictions made by the Morpheme-Order Principle, and in the next section, the System Morpheme Principle.

We now turn our attention to the other 167 mixed constituents which conform to the Morpheme-Order Principle. Example 2.6 is one such case (Hokkien material appear in bold throughout this article):

2.6 GM:	**Eŋ kai to**	**tso**	**tsi tau**	*pocket*	**hɔ?**
	should	make	here	pocket	QP

'Shouldn't (they) have made (a) pocket here?'

(MC 39)

The mixed constituent of interest is *tso tsi tau pocket* 'make (a) pocket here'. The ML is Hokkien and the mixed constituent does show Hokkien morpheme order. If the morpheme order comes from the EL, English, *pocket* would precede *tsi tau* 'here'. So it is clear in example 2.6 that the ML supplies the morpheme order.

One might be inclined to point out that unlike Swahili and English, which have pronounced word order differences, Hokkien, Mandarin and English are relatively similar in word order. Hokkien and Mandarin are both varieties of Chinese; and Chinese and English are both SVO and all Chinese modifiers precede the elements they modify (Norman, 1988:160), as in English. Swahili, on the other hand, has head-first order with most modifiers. The point is that when the Morpheme-Order Principle is tested on data which involves languages with similar word orders, one might not be able to tell

whether the morpheme order comes from the ML or the EL, since they are similar anyway.

However, Chinese differs from English syntactically in many ways and these syntactic differences provide cues as to which language supplies the morpheme order. The use of classifiers is a case in point. A classifier is a morpheme co-occurring with a noun which is individuated or specified in the discourse; that is, a noun which occurs with a numeral, a quantifier or a demonstrative (Li and Thompson, 1987:95). When a Chinese noun is individuated, quantified or specified, the classifier occurs as a suffix of the numeral, the quantifier or the demonstrative.

2.7 nèi běn shū
that CL book

2.8 sì tiáo shé
four CL snake

English, on the other hand, does not make such extensive use of classifiers. Although the English demonstrative *that* precedes the noun *book*, and the English numeral *four* precedes the noun *snake*, as in Chinese, there are still differences in morpheme order between the two languages due to the use of classifiers in Chinese. Examples 2.9 and 2.10 illustrate this:

2.9 D: Tsɛ Tsɛ, nǐ gēn wǒ kàn Ma Mi nà
sister 2PS PRO for 1PS PRO look mummy that

gè *story* kàn dào nǎ yī gè *page*.
CL story look until which one CL page

'Tsɛ Tsɛ, (go) see until which page I read in my story.'

(MC 150)

There are two mixed constituents in the above example: *nà gè story* 'that story' and *nǎ yī gè page* 'which page'. It is clearly Mandarin which supplies the morpheme order in both mixed constituents. If the morpheme order did come from English, one would expect to see something like *nà story* and *nǎ page*, with the classifiers omitted, since English does not require them in these cases. The presence of classifiers in the two mixed constituents confirms that the morpheme order is that of the ML, Mandarin. The mixed constituents *yī gè triangle* 'one triangle' and *yī gè dot* 'one dot' in example 2.10 illustrate the same point (*triangle shape* is treated as an EL island):

2.10	GD:	Shì	bù	shì	yǒu,	yǒu,	yǒu,	yǒu	yī	gè	shì
		COP	NEG	COP	EXI	EXI	EXI	EXI	one	CL	COP
ə	yī	gè	*triangle shape*		hɔ,	zhè	yàng	hái		yǒu	duō
er	one	CL	triangle shape		PP	this	way	in addition		EXI	more
yī	gè	*triangle*	yǒu	yī	gè	*dot.*					
one	CL	triangle	EXI	one	CL	dot					

'Is there a triangle? And in addition to that there is a triangle with a dot.'
(MC 167)

In summary, the Morpheme-Order Principle holds in practically all the mixed constituents in my data.

2.2 The System Morpheme Principle

The second principle that follows from the ML Hypothesis predicts that all system morphemes which have external grammatical relations in mixed constituents will come from the ML. Myers-Scotton draws support for the System Morpheme Principle from various researchers' work, including Forson, 1979; Sridhar and Sridhar, 1980; and Joshi, 1985. Forson (p. 116) observes in his Akan-English data that English (supposedly the EL) bound morphemes are not used with Akan (supposedly the ML) stems in what would be called mixed constituents in the MLF model. Sridhar and Sridhar (p. 409) state that grammatical items such as articles, quantifiers, auxiliaries, prepositions and clitics are least likely to be mixed by themselves; and Joshi argues that closed-class items in what Myers-Scotton would call mixed constituents are from only one of the languages.

However, the view that lone EL items do not take EL system morphemes is not uncontroversial. Meechan (1995:107) stresses that it has been amply demonstrated by other researchers that lone EL items do occur with EL system morphemes, and that this is precisely the situation of what is known in common parlance, in traditional linguistics and in recent empirical research as *loanwords*. Evidence from Bentahila (1995), Park (2000) and this study (all to be discussed shortly) seems to indicate that Meechan is right. I have more to say on this towards the end of this section.

Myers-Scotton (1993:102) reports that there is only one counter-example to the System Morpheme Principle out of the 374 mixed constituents in the Nairobi corpus. She also claims that the literature on code-switching offers additional empirical support for the principle, with no clear counter-examples to her knowledge. She offers examples from her own data as well

as from other researchers' data involving Moroccan Arabic and French; French and English; Korean and English; and so on.

Here are some of the examples from the Nairobi corpus:

2.11 U-na-anza ku-*behave* kama watu
2S-NON-PAST-begin-INDIC INFIN-behave as people

wa huko wa-na-vyo-*behave*.
of there 3PL-NON-PAST-MANNER-behave

'You will begin to behave as people from there behave.'

(Myers-Scotton, 1993:103)

The verb *behave* is inflected with Swahili system morphemes to give *ku-behave* 'to behave' and *wa-na-vyo-behave* 'as they behave'. None of the system morphemes is from the EL.

2.12 Mmathe wa hiyo hao alikuwa akilia joo vile vitu zi-me-*spoil*-i-they-PERF-

w-a.
spoil-θ-PASS-INDIC

'The mother of that house was crying oh how things were spoiled [for her].'

(Myers-Scotton, 1993:103)

The verb *spoil* is again inflected with ML system morphemes to give *zi-me-spoil-i-w-a* 'they were spoiled [for her]'.

2.13 Eh, unalipwa, lakini ile *scale* y-a chini kabisa...
scale CL 9-of below

nilitoka Eldoret nikaja Nakuru na *hope* y-a ku-fanya *interview*.
hope CL 9-of to do interview

'Yes, you are paid, but the lowest scale completely... I went from Eldoret to Nakuru with [the] hope of doing [the] interview.'

(Myers-Scotton, 1993:104)

The mixed constituents involving the nouns *scale* (*scale ya chini* 'scale-of-low') and *hope* (*hope ya ku-fanya* 'hope-of-to-do') include system morphemes from the ML. None of the system morphemes is from the EL.

Despite Myers-Scotton's claim that there are no known counter-examples to the System Morpheme Principle, Bentahila (1995:138) found many problematic examples in her Arabic-French data. Isolated Arabic system morphemes such as complementizers, possessive markers, determiners and the locative preposition *ƒ* occur in discourse clearly dominated by French. One such example is example 2.14:

2.14 Je sens *bi'anna* je suis vieux pour encore faire des études.

'I feel that I am old to do more studies.'

(Bentahila, 1995:138)

Bi'anna is the Arabic complementizer.

2.15 Je me réveille le matin: la premièr des choses à faire, comme un militaire, le lit *djali*.

'I wake up in the morning: the first thing to do, like a soldier, my bed.'

(Bentahila, 1995:139)

Djali is the Arabic possessive marker.

Also, recent evidence from Korean-Swedish code-switching data (Park, 2000) as well as from my own data (to be discussed later in this section) has brought to light counter-examples to the System Morpheme Principle. There are about a hundred EL morphemes occurring with EL system morphemes in Park's data. These include finite verb forms, inflected adjectives and definite nouns. Examples 2.16 and 2.17 are two such cases. In both examples, Korean is the ML and Swedish, the EL.

2.16	*ettusentvåhundra*	*bank-en*-e	iss-ô
	one thousand two hundred	bank-ART-DAT	exist-SE

'Do you have one thousand two hundred [crowns] in the bank?'

(Park, 2000:159)

In *bank-en-e* 'in the bank', the Swedish noun *bank* is first inflected with the Swedish definite article *-en* and then with the Korean nominal ending *-e*.

2.17	[ômma]-to	iss-ko	/	kûlaeyaci	*enkl-are*-ya	mwô
	mom-also	exist-and		only so	easy-COM-COP	I mean

hana- / amu	ttaena	hana-myôn	toe-canha	
one- any	time	one-if	will do-TAG	

'Same for Mom. If it is so, it is easier. One [birthday] will do, right?'
(Park, 2000:179)

In *enkl-are-ya* 'is easier', the Swedish adjective *enkl* is first inflected with the Swedish ending *-are* and then with the Korean *-ya*. Also in example 2.4 given earlier, the Swedish definite noun *skog* in *skog-en-e* 'in the forest' is inflected with the Swedish definite article *-en* and the Korean nominal ending *-e*.

We now turn our attention to the results of testing the System Morpheme Principle on my data. Six EL system morphemes were found to occur in mixed constituents, two of which occurred in internal EL islands. Internal EL islands are supposed to be exempt from predictions applying to mixed constituents. The first of the two such cases is given below:

2.18	D:	Auntie	**hɔ**,	auntie also teach *tā*	*de*
		auntie	PP	auntie also teach 3PS PRO	NOMS

children *dú*	*shū*,	you know?
children study	book	you know

'Auntie also teaches her children (how to) study, you know?'
(MC 68)

The mixed constituent of interest is *teach tā de children dú shū* 'teach her children (how to) study', which contains the internal EL island *tā de* 'her'. *De* is the nominalising particle (or genitive morpheme) in Mandarin and also a possessive and a system morpheme according to Myers-Scotton's definition of system morphemes. The other case where an internal EL island with an EL system morpheme occurs within a mixed constituent is example 2.5. *De* is again the offending EL system morpheme.

Apart from the two cases just mentioned, four other mixed constituents clearly contravene the System Morpheme Principle. They involve two occurrences of the English aspect marker *-ing* and two occurrences of the

tense marker *-ed*. Aspect and tense morphemes are treated as system morphemes in the MLF model. These are the sentences in which they occur:

2.19 GM: **Ko** **m** **mien** *parking.*
in addition NEG need parking

'In addition, (there is) no need for parking (the car).'

(MC 4)

-Ing is an EL system morpheme in *m mien parking* 'no need for parking (the car)'.

2.20 D: **Li** **kã** *fried* **bo?**
2PS PRO dare fried QP

'Do you dare fry (food)?'

(MC 58)

In the mixed constituent *kã fried* 'dare fry', *-ed* is an EL system morpheme.

2.21 D: Shàng cì lǎo shī *allowed* nǐ meń zài
last time teacher allowed 2PP PRO at

nà biān děng mā?
there side wait QP

'Did the teacher allow you to wait there the last time?'

(MC 91)

In *allowed nǐ mén zài nà biān děng* 'allowed you to wait there', *-ed* is the English tense marker and therefore an EL system morpheme.

2.22 D: **Hi** **khi** **e** *shining* **e** **hã?**
that type can shining NOMS QP

'(Is it) that type which shines?'

(MC 131)

-Ing is the EL system morpheme in *e shining.*

I agree with Meechan (1995) that lone EL items do take EL system morphemes. Disregarding the MLF model and the concept of mixed constituents for a moment, *parking, fried, allowed* and *shining* seem to be perfectly ordinary loanwords on their own. Inflected English loanwords abound in everyday Singaporean speech and many of them would be said to be from the EL under the MLF model. There seems to be no reason why a loanword should not carry with it an inflectional element.

To conclude, it is beyond doubt that the System Morpheme Principle does not hold categorically.

3. The Blocking Hypothesis

The third prediction to be tested is the Blocking Hypothesis, which states that 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 subcategorisation (Myers-Scotton, 1993:120). The three levels 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. Second, the ML also blocks an EL content morpheme in these constituents if it is not congruent with an ML content morpheme counterpart in terms of thematic role assignment. At a third level, congruence between EL content morphemes and ML content morphemes is in terms of their discourse or pragmatic functions. Discussion of this type is beyond the scope of the current volume.'
>
> (Myers-Scotton, 1993:121)

I agree with Myers-Scotton that third-level type incongruencies are beyond the scope of grammatical constraints on code-switching, and I will not attempt to test them on my data. She claims that she is not alone in suggesting that subcategorisation restrictions may explain constraints on code-switching utterance formation. Researchers such as Bentahila and Davies (1983), Muysken (1990, 1991) and Azuma (1991) have done the same.

Pronouns could be an example of first-level type incongruency. One language may have pronominal forms which are agreement clitics (system morphemes) while the other language may have free-form pronouns (content morphemes). If the ML has pronominal clitics, then no EL free pronoun may be substituted for a clitic in mixed constituents (Myers-Scotton, 1993: 121).

There are no grammatical categories to my knowledge which show first-level type incongruency between Hokkien, Mandarin, Cantonese and English. Pronouns, for example, are free forms in Chinese and English. There are therefore no instances of first-level type incongruency in the data, nor do I expect to find any.

The English preposition *for* is another case in point of subcategorisation incongruency (first- and second-level type incongruencies). *For* is a content morpheme because it is [-Quantification][+Thematic Role-Assigner]. It assigns the thematic role of beneficiary or goal to John in *I bought the book for John. For*, however, is incongruent with Swahili morphemes. In Swahili, the verb assigns the thematic roles of beneficiary or goal. A counterpart for *for* could be the suffix *-i-* or *-e-* (depending on vowel harmony) known as the applied form. This Swahili suffix is a system morpheme, and therefore is incongruent with the English *for*. The presence of the Swahili applied form requires a switch from English to Swahili:

3.1 Labda yeye hana vitabu vyake *father*
father

a-li-m-*buy*-i-a akapoteza vyote.
3S-PAST-3S/OBJ-buy-APPL-INDIC

'Maybe he doesn't have his books [which his] father bought for him, and he lost all of them.'

(Myers-Scotton, 1993:123)

The lack of congruence between *for* and a Swahili counterpart is implied in the non-occurrence in the Nairobi corpus of a sentence such as example 3.2, but the occurrence of example 3.3:

3.2 *Nikamwambia anipe ruhusa niende ni-ka-*check* *for*
1S-CONSEC-check for

wewe.
you

'And I told him he should give me permission so that I go and check for you.'

(Myers-Scotton, 1993:124)

3.3 Nikamwambia anipe ruhusa niende ni-ka-*check for you*.

(Myers-Scotton, 1993:124)

Meechan (1995:107) criticises Myers-Scotton for examining only counter-examples for congruence violations, with considerable danger that the resulting explanations are ad hoc. The reader gets no indication as to how many of the unproblematic mixed constituents also show some degree of non-congruence. She goes on to assert that virtually all language pairs may exhibit some mismatch across lexical items (Muysken, 1991), and some non-congruence in any such comparison is conceivably more likely than not. Furthermore, the vagueness surrounding the definition of congruence also allows a wide range of explanations for counter-examples.

In a later work (Myers-Scotton and Jake, 1995), the notion of congruence is referred to in terms of levels or subsystems of complex lexical structure, where the three levels are: lexical-conceptual structure (semantics and input for pragmatic readings); predicate-argument structure (relations between verbs and prepositions and their arguments); and morphological realisation patterns (surface requirements for well-formedness, including word order). The three levels are present simultaneously although not necessarily activated at the same time. The authors claim that their proposals about lexical structure are motivated by Talmy (1985) and Jackendoff (1990), amongst others. I find that the notion of congruence is given a more detailed exposition and illustrated with more examples in Myers-Scotton, 1993 than in Myers-Scotton and Jake, 1995, and therefore I test the Blocking Hypothesis on my data based on the 1993 version of congruence.

At first glance, there appear to be counter-examples to the Blocking Hypothesis (second-level type incongruencies) in my data. Further investigation reveals that they do not contravene the principle at all. They consist of mainly English verbs which have different subcategorisation requirements from their Chinese counterparts, and are therefore incongruent with them. English verbs typically subcategorise for a subject argument and transitive verbs require an object argument in addition. In Chinese, the subject argument of verbs can often be omitted. This is partly because Chinese is a *pro*-drop language; that is, it allows the subject of a finite clause to remain unexpressed. In addition to that, Chinese allows the object to be omitted in certain contexts. In the data, there are several striking cases of singly-occurring English verbs with their required subject or object arguments nowhere to be seen. Table 3.1 lists them:

Table 3.1 Suspect incongruent EL content morphemes in mixed constituents

	Missing subject	Missing object
Verbs	*Improve*	*Choose* (two occurrences), *improve, remind, try*
Adjective	*hungry*	-

However, one needs to take into account the fact that all the speakers in the study speak a different variety of English from Standard English, i.e. Singapore English, and Singapore English subcategorisation requirements are different from Standard English subcategorisation requirements. Like Chinese, Singapore English permits *pro*-drop and allows the object to be omitted in certain contexts. This is because of the status of Singapore English as a contact variety:

> 'One of the characteristics of SgE is that it shows the transference of features from the various speech varieties spoken by the different ethnic groups. ...the dominant substratum influence on Singaporean English syntax as spoken by those with lower levels of education, and by others in informal situations, is Chinese. This is also true of semantics, such as the semantics of verbs of movement and, to a great extent, phonology.'
>
> (Ho and Platt, 1993:8)

Subcategorisation requirements are certainly included in the grammatical aspects of Chinese evident in Singapore English, amongst many others. Once this is recognised, what appear to be subcategorisation incongruencies between the EL content morphemes in Table 3.1 and their Chinese counterparts no longer exist. Example 3.4 is one such example:

3.4	GD:	Nǐ	*remind*	de	shí hòu	tā
		2PS PRO	remind	NOMS	time	3PS PRO

jiù	zuò	zài	nà	biān	yī zhí	tīng	tīng	tīng	lɔ.
already	sit	at	there	side	all the time	listen	listen	listen	PP

'When you were reminding (me), he was sitting there listening all the time.'

(MC 179)

In *nǐ remind* 'you remind (me)', *remind* seems to be incongruent with its Mandarin counterpart *tí xǐng*. *Remind* subcategorises for a subject and an object argument, but with *tí xǐng* in the same context, both can be omitted. In example 3.4, the object argument of *remind* is nowhere to be seen, and this seems to violate its subcategorisation requirements. However, it should be the Singapore English subcategorisation requirements which are taken into account, and indeed in Singapore English, *remind* in the above context can occur without an object argument. This means that Singapore English *remind* is congruent with *tí xǐng*. Example 3.5 illustrates the same point:

3.5	D:	**To ai**	*improve*	**bo**	**bue**	**sai.**
		have to	improve	otherwise	NEG	can

'(I) must improve (my English).'

(MC 110)

Although Standard English subcategorisation rules require *improve* in *ai improve* 'have to improve (my English)' to have a subject and an object argument, Singapore English does not, in the above context. This makes Singapore English *improve* congruent with its Hokkien counterpart *tsin pɔ*.

We have seen that there are Standard English verbs which show subcategorisation incongruency with their Chinese counterparts. As for Singapore English verbs, or in fact Singapore English content morphemes, I do not know of any which would show similar incongruency. Unlike the Nairobi corpus, in which prepositions, pronouns and verbs show incongruencies between Swahili and English, similar incongruencies involving the same categories are not exhibited in my data. As for Hokkien and Mandarin, they are both varieties of Chinese and have practically identical subcategorisation requirements.

There are no instances of either first-level type or second-level type incongruencies in the data, but that is because such incongruencies do not exist between the languages concerned anyway, rather than because the predictions of the Blocking Hypothesis are accurate.

4. The EL Island Trigger Hypothesis

Before introducing the EL Island Trigger Hypothesis, I want to take stock of the results of testing the predictions of the MLF model on my data so far. The data conforms to the Morpheme-Order Principle; counter-examples to the System Morpheme Principle were found; and as for the Blocking Hypothesis, the data is unsuitable for testing.

Myers-Scotton (1993:138) reveals that under the MLF model, EL islands are less studied than mixed constituents are. She adds that the hypothesis is merely a preliminary suggestion regarding EL islands. Her analysis is based on the 121 EL islands in the Nairobi corpus. She cites examples from her own data, but there is limited discussion of other data sets.

According to the EL Island Trigger Hypothesis, activating any EL lemma or accessing by error any EL morpheme not licensed under the ML or Blocking Hypotheses triggers the processor to inhibit all the ML accessing procedures and complete the current constituent as an EL island (Myers-Scotton, 1993:139). EL islands may be produced in two ways:

> '1. If an EL morpheme implicating non-ML morpheme order in a constituent is accessed as the initial element in a constituent, this triggers processing of the entire constituent in the EL, thereby forming an EL island. This prediction is a corollary to the Morpheme-Order Principle applying to ML+EL constituents and permitting only ML morpheme order in these constituents. ...
>
> 2. If any EL system morpheme, or an EL content morpheme not showing correspondences to an ML content morpheme, is accessed, ML procedures are inhibited, and the entire constituent of which the EL morpheme is a part must be produced as an EL island. This restriction complements the System Morpheme Principle and the Blocking Hypothesis. ...'
>
> (Myers-Scotton, 1993:139-140)

This hypothesis predicts that in Swahili-English code-switching, if an English adjective is accessed before its head noun, then the adjective+head noun must form an EL island. This is because Swahili requires head-first order in NPs with adjectives. This pattern is attested in example 4.1:

4.1	Hujasikia kutoka		*next week*	wafanya kazi	wa
	you not yet hear from		next week	workers	of
serikali	hawatakuwa	wakienda		kazini	*on*
government	they will not be	they be going		to work	on

Saturdays.
Saturdays

'Haven't you heard that from next week government workers will not be going to work on Saturdays?'

(Myers-Scotton, 1993:137)

4.2 Wache mimi nielekeee tauni, tukutane *this evening at*
let us meet this evening at

the usual place.
the usual place

'Let me go so that I may reach town, let's meet this evening at the usual place.'

(Myers-Scotton, 1993:140)

In *at the usual place*, the English preposition *at* assigns thematic role to the locative NP. In Swahili, it is the verb stem which does so. *At* is thus incongruent with Swahili morphemes, according to the Blocking Hypothesis. Furthermore, *at* is a content morpheme, but its Swahili counterpart is a system morpheme. Having accessed *at*, the processor must go on and complete the PP as an EL island.

The hypothesis has been challenged by Meechan and Poplack (1994), who found that at least one type of non-congruence, i.e. the lack of categorical equivalence between languages, cannot be linked to an increased incidence of EL islands.

Myers-Scotton changed her mind about the EL Island Trigger Hypothesis and the EL Hierarchy Hypothesis (see the next section) in the *Afterword* to the paperback edition of Myers-Scotton. In 1993, she stated that EL islands could be either obligatory or optional, but in 1997 she wrote that all EL islands are obligatory. She proposed a new EL Island Hypothesis to replace the EL Island Trigger Hypothesis and the EL Hierarchy Hypothesis in Myers-Scotton, 1993:

> 'EL Island Hypothesis: When there is insufficient congruence between the lemma underlying an EL content morpheme and its ML counterpart at one or more of the three levels of lexical structure, the only way to access the EL element is in an EL island.'
>
> (Myers-Scotton, 1997:250)

The three main sources of incongruence resulting in EL islands are identified as: incongruence at the lexical-concept level, either in terms of semantic or pragmatic implications; incongruence at the level of morphological realisation patterns; and incongruence regarding what information about morphological realisation patterns is entered in the lemma which supports the content head of a phrase. I decided to stick to the two

1993 hypotheses rather than to test the 1997 EL Island Hypothesis. Testing other researchers' hypotheses and theories requires one to fully grasp and understand the hypotheses and theories concerned. The book-length explanation (Myers-Scotton, 1993) gives a much stronger exposition of the EL Island Trigger Hypothesis and the EL Hierarchy Hypothesis than the paper-length descriptions (Myers-Scotton, 1997; Myers-Scotton and Jake, 1995; and Jake and Myers-Scotton, 1996). I feel much more confident about testing the 1993 hypotheses and it would not be fair to the later EL Island Hypothesis if I attempted to test it without thoroughly understanding it.

There are thirty EL islands in total in my data. Twelve (40%) of them conform to the EL Island Trigger Hypothesis in that the initial element in the constituent implicates non-ML morpheme order. Examples 4.3 and 4.4 are two such cases:

4.3	GF:	**Li**	**kin**	**ni**	*last*	*year*	**hã?**
		2PS PRO	this	year	last	year	QP

'Is this year your last year?'

(ELI 2)

The equivalent of the EL island *last year* in Hokkien would be *lɔ be tsi ni*, with the classifier *tsi* between *lɔ be* 'last' and *ni* 'year'. English does not require the use of a classifier here and nothing is expected to occur between *last* and *year*. Thus, *last* can be said to implicate non-Hokkien morpheme order.

4.4	D:	Nĭ	kàn dào	tā	*how to fried the, the*
		2PS PRO	see	3PS PRO	how to fried the the

yóu	*tiáo,*	*you know?*
oil	strip	you know

'You (will) see how he fries the oil fritter, you know?'

(ELI 52)

The first element *how* in *how to fried the, the yóu tiáo, you know* (*yóu tiáo* is treated as a borrowing) implicates non-Mandarin morpheme order in that it requires the infinitive *to* to follow it. In Mandarin, the verb *zhà* 'fry' can follow the adverb *zěn me* 'how' immediately without an intervening infinitive.

All of the EL islands which conform to the hypothesis are of the type illustrated by the two examples above, i.e. they are triggered by an initial element in the constituent which implicates non-ML morpheme order. There is none of the type which is triggered by lack of congruence. This follows from the results of testing the Blocking Hypothesis in the previous section, where it was found that there are no incongruencies, whether first-level or second-level types, between the languages concerned.

Eighteen (60%) of the EL islands contravene the EL Island Trigger Hypothesis. Here are some of them:

4.5	D:	**I naŋ**	**ka ki**	*keep record.*
		3PP PRO	REF PRO	keep record

'They keep record themselves.'

(ELI 62)

The object follows the verb in SVO languages like Chinese and English. The EL island *keep record* shows ML morpheme order, and could not have been formed because it implicated non-ML morpheme order. *Keep* also corresponds to its Hokkien counterpart *siu* in terms of subcategorisation requirements; both subcategorise for a subject and an object argument. In addition to that, both *keep* and *siu* are [-Quantification][+Thematic Role-Assigner], and therefore content morphemes. In fact, there is a sentence in the data where *siu* and *record* occur next to each other:

4.6	D:	**I naŋ**	**ka ki**	**siu**	*record*	**e**	**la.**
		3PP PRO	REF PRO	keep	record	NOMS	PP

'It's for them to keep record themselves.'

(MC 61)

Example 4.6 shows that *keep* and *siu* are substitutable for each other, and must therefore be congruent with each other.

4.7	GD:	Lǎo shī	gēn	wǒ mén	jiǎng	tā	méi	yǒu
		teacher	with	1PP PRO	say	3PS PRO	NEG	EXI

provide food	gěi	nǐ	chī.
provide food	give	2PS PRO	eat

'Teacher told us that he does not provide food for you to eat.'

(ELI 177)

Provide food shows the same morpheme order as its Mandarin counterpart *tí gōng shí wù*. *Provide* also has the same subcategorisation requirements as *tí gōng*; and both are [-Quantification][+Thematic Role-Assigner].

The EL Island Trigger Hypothesis accounts poorly for my data. Only 40% of the EL islands in the data are formed as predicted by the hypothesis; the remaining 60% are unaccounted for. It is beyond the scope of this article to attempt to account for them.

5. The EL Hierarchy Hypothesis

The comment Myers-Scotton made about the tentative nature of the EL Island Trigger Hypothesis also applies to the EL Hierarchy Hypothesis. After all, EL islands are not as well-studied as mixed constituents are under the MLF model. The analysis is based on the 121 EL islands in the Nairobi corpus. Examples from the Nairobi corpus were cited, but there were few from other data sets. We will see later in this section that the hypothesis runs into problems in Park's (2000) data as well as in my data.

According to Myers-Scotton, the hierarchy is suggested by data from the Nairobi corpus, and by theoretical claims in other areas of linguistics and data from other studies. For example, the idea that the most central constituents of a sentence, the subject NP and the VP, are the least accessible to EL islandhood comes from theories of functionalism in grammar (for example, Hopper and Thompson, 1984; Givón, 1979, 1989), as well as from psycholinguistic studies. Treffers-Daller (1991a, 1991b) also developed a similar hierarchy based on her Flemish-French data.

The EL Hierarchy Hypothesis consists of two sub-hypotheses:

> '1. The more peripheral a constituent is to the theta-grid of the sentence (to its main arguments), the freer it is to appear as an EL island.
> 2. The more formulaic in structure a constituent is, the more likely it is to appear as an EL island. Stated more strongly, choice of (any) part of an idiomatic expression will result in an EL island.'
>
> (Myers-Scotton, 1993:144)

The hypothesis places the most accessible constituents for islandhood at the top of the hierarchy, and proposes that an implicational relationship from bottom to top is in force (i.e. if there are subject NPs as EL islands,

there are also object NPs as islands and so on). What follows is the Implicational Hierarchy of EL Islands:

> '1. Formulaic expressions and idioms (especially as time and manner PPs but also as VP complements);
> 2. Other time and manner expressions (NP/PP adjuncts used adverbially);
> 3. Quantifier expressions (APs and NPs especially as VP complements);
> 4. Non-quantifier, non-time NPs as VP complements (NPs, APs, CPs);
> 5. Agent NPs;
> 6. Thematic role- and case-assigners, i.e. main finite verbs (with full inflection).'
>
> (Myers-Scotton, 1993:144)

The argument behind the hypothesis is that the central constituents carry the main semantic weight of the sentence. Since the ML has more psycho-sociolinguistic dominance in the discourse, the central constituents should be either in ML islands or possibly in mixed constituents. Thus, to allow elements which are peripheral to the core of the communicative intention to appear in the EL as islands seems a likely corollary. However, the claim that peripheral constituents are also peripheral to the core of the communicative intention is disputable because they often carry essential information (for example, time and manner PPs). A structural explanation seems more convincing: peripheral constituents are peripheral to argument structure and hence can be switched more easily.

Thirty-six (29.5%) of the EL islands in the Nairobi corpus are time adverbials (for example, *next Saturday, every morning* and *after four months*), many of which are almost formulaic (for example, *next weekend* and *on Saturdays*). Almost twelve (10%) are set expressions (for example, *old habits die hard, in fact* and *for personal purposes*). Most of the other islands are VP complements, many of which are complements of a copula consisting of an intensifier adverb + an adjective (for example, *very fast, very late* and *very surprised*).

Park (2000) has found that the EL Hierarchy Hypothesis does not hold in his data. Over 40% of his EL islands are main arguments of the sentence such as subject NPs, object NPs and VPs. Only a small number of EL islands are formulaic or idiomatic expressions. The number of peripheral constituents of the sentence (for example, adverbial NPs and PPs) is also relatively small.

Similarly in my data, the hypothesis encounters problems.

Using mainly the terms *subject, verb, object, complement* and *adverbial* (when inapplicable, *noun, NP, possessive pronoun* and *topic* are used) to label the thirty EL islands in my data yields the following results:

Table 5.1 EL islands as constituents (a)

Constituent(s)	Number (%)
Object	7 (23.333%)
Verb + object	5 (16.667%)
Adverbial	4 (13.333%)
Subject complement	2 (6.667%)
Verb	2 (6.667%)
Noun complement	1 (3.333%)
NP	1 (3.333%)
NP + verb	1 (3.333%)
Possessive pronoun	1 (3.333%)
Subject	1 (3.333%)
Subject + object	1 (3.333%)
Subject + subject complement	1 (3.333%)
Subject + verb	1 (3.333%)
Topic complement	1 (3.333%)
Verb complement	1 (3.333%)

Since some of the EL islands consist of more than one constituent (for instance, *tsia be* 'eat porridge' consists of verb and object), listing them individually (for example, *Verb + object* constituents are counted under *Verb* and under *Object* separately) offers a clearer picture:

Table 5.2 EL islands as constituents (b)

Constituent	Number (%)
Object	13 (33.333%)
Verb	9 (23.077%)
Adverbial	4 (10.256%)
Subject	4 (10.256%)
Subject complement	3 (7.692%)
NP	2 (5.128%)
Noun complement	1 (2.564%)
Possessive pronoun	1 (2.564%)
Topic complement	1 (2.564%)
Verb complement	1 (2.564%)
Total	39 (100%)

Core constituents such as subject (*dù zǐ* 'stomach', *nǐ* 'you' and so on), verb (*tsia* 'eat', *keep* and so on) and object (*be* 'porridge', *the, the yóu tiáo* 'the oil fritter' and so on) account for more than 66% of all EL islands. This runs counter to the prediction that core constituents do not often appear as EL islands. Adverbials (*in the end* and *no wonder*), on the other hand, though predicted to occur frequently as EL islands, only make up 10.256% of the EL islands in my data. Formulaic expressions (*in the end* and *no wonder*; part of the 10.256%) are rare with only four occurrences.

The pattern in the data seems to be the opposite of what Myers-Scotton predicted; core constituents appear more often than peripheral constituents do in EL islands. Also, there are few occurrences of formulaic expressions. This pattern is very much similar to what Park (2000) found in his data. It is premature to make claims about the universality of this observation. Future evidence from other data sets will shed more light on the issue.

6. Conclusions

Most of the early studies of code-switching did not concern themselves with formulating general constraints. Labov (1971:57), in a discussion of what constitutes a linguistic system, cites an example of Spanish/English code-switching and notes that it must be described as the irregular mixture of two distinct systems. Although Gumperz (1982:2) claims that the mixture is not random, he observes that the motivation for code-switching seems to be stylistic and metaphorical rather than grammatical. Lance (1975:43), too, suggests that there are perhaps no syntactic restrictions on where switching can occur. There have since been various attempts at formulating grammatical constraints on code-switching, Sankoff and Poplack's (1981) study of Puerto Rican Spanish/English code-switching being the first.

One of these attempts is the MLF model. Taking into account the problems with its theoretical constructs (such as the ML and the content morpheme/system morpheme distinction), as well as the counter-examples to its grammatical constraints other researchers and myself have found, the model cannot be said to have universal applicability.

The fact that none of the sets of grammatical constraints put forward thus far (including the Two-Constraint Model (Sankoff and Poplack, 1981) and the Government Model of Code-Switching (Di Sciullo, Muysken and Singh, 1986) stands without problems or counter-examples casts doubt on the possibility of formulating a set of truly universal constraints. Languages can be typologically so different that it is doubtful that a single set of constraints can apply to them universally.

However, I believe that formulating grammatical constraints on code-switching is not an end in itself. Just as important is what the constraints tell us about the psycholinguistic processes involved in code-switching. Take, for instance, the System Morpheme Principle. Given the numerous counter-examples to the principle other researchers and myself have found, the psycholinguistic processes which Myers-Scotton claims to be underlying are probably not valid. It is probably not true that when the frame of mixed constituents is being built, all the basic linguistic procedures specify selecting the ML and it is the ML which supplies the system morphemes. Formulating grammatical constraints and testing them on different data sets are both worthwhile endeavours: by formulating constraints, we make hypotheses about the psycholinguistic processes involved, and by testing them, we know which hypotheses are more likely to be valid.

This study has found counter-examples to the System Morpheme Principle and to the EL Island Trigger Hypothesis. Although the data conforms to the Morpheme-Order Principle, it contravenes the EL Hierarchy Hypothesis.

It is premature to make conclusive statements about the MLF model's grammatical constraints on intra-sentential code-switching, since few researchers have tested them systematically. The Morpheme-Order Principle needs to be further tested in order to ascertain how well it accounts for different data sets. The System Morpheme Principle appears to be problematic, having been attacked by counter-examples from a variety of data sets. It looks as if the principle will have to be abandoned eventually, unless the definition of content/system morphemes can be shown to be at fault and replaced with a more suitable set of criteria. Both the Blocking Hypothesis and the EL Island Trigger Hypothesis have received less attention than the other constraints and it would be wise to subject them to further testing before drawing any conclusions. As for the EL Hierarchy Hypothesis, both Park's and my data suggest patterns opposite to those predicted by the hypothesis. It would be interesting to find out how widespread these patterns are.

Notes

1. My gratitude goes to my supervisor, Professor Suzanne Romaine, for her comments on the doctoral thesis on which this article is based; and to the University of Oxford and St. Hugh's College, for funding the field trip to Singapore.
2. Methodological issues such as distinguishing between borrowed forms and code-switched forms, identifying the ML and determining the EL islands and mixed constituents are discussed in Chua, 2001a, 2001b, 2001c.

References

Azuma, S. 1991. "Two-level processing hypothesis in speech production: evidence from intra-sentential code-switching." Paper presented at *the 27th Chicago Linguistic Society Meeting*, May.

Backus, A. and Boeschoten, H. 1996. "Review of Myers-Scotton, 1993." *International Journal of Applied Linguistics* 6: 129-153.

Bentahila, A. 1995. "Review of Myers-Scotton, 1993." *Language* 71: 135-140.

Bentahila, A. and Davies, E. D. 1983. "The syntax of Arabic-French code-switching." *Lingua* 59: 301-330.

Chua, L. 2001a. "Language shift in a Singaporean Chinese family and the Matrix Language Frame model." D.Phil. thesis, University of Oxford.

——— 2001b. "Identifying the Matrix Language and determining the Embedded Language islands and mixed constituents in data from Singapore." Submitted for publication.

——— 2001c. "Distinguishing between borrowed forms and code-switched forms under the Matrix Language Frame model in data from Singapore." Submitted for publication.

——— 2001d. "Language shift and language maintenance in a Singaporean Chinese family." Submitted for publication.

Comrie, B. ed. 1987. *The World's Major Languages.* London: Croom Helm.

Di Sciullo, A. M., Muysken, P. and Singh, R. 1986. "Government and code-mixing." *Journal of Linguistics* 22: 1-24.

Dowty, D. R., Karttunen, L. and Zwicky, A. eds. 1985. *Natural Language Parsing.* Cambridge: Cambridge University Press.

Forson, B. 1979. "Code-switching in Akan-English bilingualism." Ph.D. thesis, University of California at Los Angeles.

Givón, T. 1979. *On Understanding Grammar.* New York: Academic Press.
——— 1989. *Mind, Code and Context: Essays in Pragmatics.* Hillsdale, New Jersey: Erlbaum.
Gumperz, J. J. 1982. *Discourse Strategies.* Cambridge: Cambridge University Press.
Hernández-Chavez, E., Cohen, A. and Beltramo, A. eds. 1975. *El languaje de los Chicanos.* Arlington, Va: Centre for Applied Linguistics.
Ho, M. L. and Platt, J. T. 1993. *Dynamics of a Contact Continuum. Singaporean English.* Oxford: Clarendon Press.
Hopper, P. and Thompson, S. 1984. "The discourse basis for lexical categories in Universal Grammar." *Language* 60: 703-752.
Hymes, D. ed. 1971. *Pidginization and Creolization of Languages.* Cambridge: Cambridge University Press.
Jackendoff, R. 1990. *Semantic Structures.* Cambridge, Massachusetts: The MIT Press.
Jake, J. L. and Myers-Scotton, C. 1996. "Verbs in Arabic-English code-switching and lexical structure." Paper presented at *Symposium on Code-Switching and Lexical Structure, Linguistic Society of America,* January.
Joshi, A. 1985. "Processing of sentences with intra-sentential code-switching." In *Dowty et al. (eds.),* pp. 190-205.
Labov, W. 1971. "The notion of 'system' in creole languages." In *Hymes ed.,* pp. 447-472.
Lance, D. 1975. "Spanish/English code-switching." In *Hernández-Chavez et al. eds,* pp. 138-153.
Li, C. N. and Thompson, S. A. 1987. "Chinese." In *Comrie ed.,* pp. 83-105.
Meechan, M. 1995. "Review of Myers-Scotton, 1993." *Anthropological Linguistics* 37: 106-109.
Meechan, M. and Poplack, S. 1994. "Orphan categories in bilingual discourse: a comparative study of adjectivization strategies in Wolof-French and Fongbe- French." Paper presented at *New Ways of Analyzing Variation in English,* Stanford University, October.
Muysken, P. 1990. "Concepts, methodology and data in language contact research: ten remarks from the perspective of grammatical theory." In *Papers for the Workshop on Concepts, Methodology and Data,* pp. 15-30. Strasbourg: European Science Foundation.
——— 1991. "Needed: a comparative approach." In *Papers for the Symposium on Code-Switching in Bilingual Studies: Theory, Significance and Perspectives,* pp. 253-272. Strasbourg: European Science Foundation.

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

——— 1997. "Afterword." In *paperback edition of Myers - Scotton, 1993,* pp. 240-259.

Myers-Scotton, C. and Jake, J. L. 1995. "Matching lemmas in a bilingual language competence and production model: evidence from intrasentential code switching." *Linguistics* 33: 981-1024.

Norman, J. 1988. *Chinese.* Cambridge: Cambridge University Press.

Park, H. S. 2000. "Korean-Swedish code-switching. Theoretical models and linguistic reality." Ph.D. thesis, Uppsala University.

Sankoff, D. and Poplack, S. 1981. "A formal grammar for code-switching." *Papers in Linguistics* 14: 3-46.

Shopen, T. ed. 1985. *Language Typology and Syntactic Description.* New York: Cambridge University Press.

Sridhar, S. N. and Sridhar, K. 1980. "The syntax and psycholinguistics of bilingual code-mixing." *Canadian Journal of Psychology* 34: 407-416.

Talmy, L. 1985. "Lexicalization patterns: semantic structure in lexical form." In *Shopen ed.*, pp. 51-149.

Treffers-Daller, J. 1991a. "Towards a uniform approach to code-switching and borrowing." In *Papers for the Workshop on Constraints, Conditions and Models,* pp. 259-279. Strasbourg: European Science Foundation.

——— 1991b. "French-Dutch language mixture in Brussels." Ph.D. thesis, University of Amsterdam.

Appendix 1
Abbreviations

GF	grandfather
GM	grandmother
D	daughter 2
GD	granddaughter
GS	grandson
R	researcher
ML	Matrix Language
MLF	Matrix Language Frame
EL	Embedded Language
MC	mixed constituent
ELI	EL island

CL	classifier
COP	copular
EXI	existential marker
INT	interjection
NEG	negator
NOMS	nominalising particle
PERF	perfective marker
PP	pragmatic particle
QP	question particle
1PS PRO	first person singular pronoun
2PS PRO	second person singular pronoun
3PS PRO	third person singular pronoun
1PP PRO	first person plural pronoun
2PP PRO	second person plural pronoun
3PP PRO	third person plural pronoun
REF PRO	reflexive pronoun

Appendix 2
Hokkien symbols

	Approximate IPA equivalent
p	p
ph	p^h
b	b
t	t
th	t^h
ts	c
tsh	c^h
k	k
kh	k^h
g	g
m	m

n	n
ŋ	ŋ
s	s
h	h
l	l
a	ʌ
ã	ʌ̃
ai	ʌi
au	ʌu
e	e
ə	ə
ɛ	ɛ
i	i
ĩ	ĩ
iã	iʌ̃
o	o
ɔ	ɔ
u	u
ua	uʌ
uã	uʌ̃
ue	ue
ui	ui