PDF hosted at the Radboud Repository of the Radboud University
Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/14680

Please be advised that this information was generated on 2019-01-08 and may be subject to change.
1. Introduction

In this paper I want to argue that there are three separate patterns of code-switching within sentences (intra-sentential code-switching, sometimes referred to as code-mixing), and explore some of the ways to study these patterns through the systematic exploration of bilingual corpora, in addition to the detailed structural analysis of individual examples (van Hout and Muysken 1995). The three patterns are alternation, insertion, and congruent lexicalization. Possible examples are given in (1)-(3), taken from Spanish-English bilingual usage.

(1) ALTERNATION
Andale pues *and do come again.* (Peñalosa 1980)
'That's all right then, and do come again.'

In the case of alternation, there is a true switch from one language to the other, involving both grammar and lexicon. Thus in (1) there is no reason to assume that the Spanish first segment is embedded in the English second segment or vice versa. Alternation is of course just a special case of code-switching as it takes place between utterances in a turn or between turns.

(2) INSERTION
Yo anduve *in a state of shock* pa dos días. (Pfaff 1979)
'I walked in a state of shock for two days.'

With insertion, there is embedding. The English prepositional phrase is inserted into an overall Spanish structure. Insertion is akin to (spontaneous) lexical borrowing, which also involves one lexical unit.
CONGRUENT LEXICALIZATION
Bueno, in other words, el flight [que sale de Chicago around three o’clock].

'Good, in other words, the flight that leaves Chicago around three o’clock.'

(Pfaff 1976)

This third case is less obviously interpretable. The term congruent lexicalization refers to a situation where the two languages share a grammatical structure which can be filled lexically with elements from either language. The mixing of English and Spanish could be interpreted as a combination of alternations and insertions, but the going back and forth suggests that there may be more going on (see more extensive discussion below), and that the elements from the two languages are inserted, as constituents or as words, into a shared structure. In this perspective, congruent lexicalization is akin to style or register shifting and monolingual linguistic variation. The latter would be the limiting case of congruent lexicalization.

The interpretation of these three patterns can be:

- **structural**, in terms of labels in tree configurations;
- **psycholinguistic**, in terms of different degrees of activation of components of both languages (in the case of alternation, activation would switch from one language to another, and in the case of insertion, activation in one language would be temporarily diminished);
- **sociolinguistic**, in terms of bilingual strategies (an example would be Poplack & Sankoff’s equivalence versus insertion).

The structural interpretation is what concerns us most here. Consider the trees (1)’–(3)’, where A, B are language labels for non-terminal nodes (i.e. fictitious markers identifying entire constituents are belonging to one language), and a, b are labels for terminal, i.e. lexical, nodes, indicating that the words chosen are from a particular language.
In this situation, a constituent from language A (with words from the same language) is followed by a constituent from language B (with words from that language). Unspecified is the language of the constituent dominating A and B in (1)'.

\[(2)'
\begin{align*}
\text{A} \\
\text{B} \quad \text{insertion}
\end{align*}
\]

In the situation defined by (2)', a single constituent B (with words b from the same language) is inserted into a structure defined by language A, with words a from that language.

\[(3)'
\begin{align*}
\text{A/B} \quad \text{congruent lexicalization}
\end{align*}
\]

Finally, in (3)' the grammatical structure is shared by languages A and B, and words from both languages a and b are inserted more or less randomly.

The structural features of these three patterns can also be seen in terms of constraints. For alternation the best candidate at present is basic linear equivalence at the switch site, essentially Poplack's (1980) equivalence constraint. The issue unresolved at present is the nature of the constituent dominating both nodes A and B in (1)'. We could assume that it is shared by A and B (as in the congruent lexicalization model), or that it is from either A or B (in which case alternation would be like insertion at the sentence-margin.

For insertion the best candidate at present is Myers-Scotton's MLF-model (1993; 1995), in which the matrix language constituent order dominates and matrix language function words dominate. Researchers such as Sebba
(1995) have stressed that there needs to be some kind of categorial equivalence or congruence (Myers-Scotton's term) between the constituent inserted and the matrix language node into which it is inserted. If we term linear equivalence in Poplack's sense bilateral compatibility, here we have a case of unilateral compatibility.

For congruent lexicalization it is quite possible that there are no structural constraints, since all that is involved is insertion of words into one single syntactic structure (Clyne 1987, Giesbers 1989).

The sociolinguistic embedding of these three patterns, i.e. their use as bilingual strategies, can be described as follows, in a preliminary way. I'll take the three processes in the same order. The process of alternation is particularly frequent in stable communities with a tradition of language separation, but occurs as well in many other communities. It may be the most frequent and least structurally intrusive type of code-switching. Insertion is frequent in neo- or ex-colonial settings and recent migrant communities, where there is a considerable asymmetry in the speakers' dominance of both languages. A language dominance shift, e.g. between the first and third generation in an immigrant setting, may be reflected in a shift in directionality of insertion: from insertion into the language of the country of origin to insertion into the language of the host country. Finally, congruent lexicalization may be particularly associated with second generation migrant groups, dialect/standard and post-creole continua, and bilingual speakers of closely related languages with roughly equal prestige.

Psycholinguistic factors determining the choice between these different processes would include bilingual competence, level of monitoring in both languages, the triggering of a particular language by specific items and the degree of separateness of storage and access systems.

2. Diagnostic criteria

Having presented the three processes, insertion, alternation, and congruent lexicalization, I now try to present a number of diagnostic features which may be used to distinguish these three patterns. There are a number of criteria I would like to consider, and I will illustrate their application with concrete examples. Sources for language pairs are listed in Table 1:
Table 1. Sources for the language pairs in the discussion of the criteria

<table>
<thead>
<tr>
<th>Source</th>
<th>Language Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gumperz &amp; Hernandez Chavez (1975)</td>
<td>Spanish-English</td>
</tr>
<tr>
<td>Lance (1975)</td>
<td></td>
</tr>
<tr>
<td>Poplack (1980)</td>
<td>Sarnami Hindustani-Dutch</td>
</tr>
<tr>
<td>Krishna (1979)</td>
<td>Spanish-Hebrew</td>
</tr>
<tr>
<td>Berk-Seligson (1986)</td>
<td>Turkish-Dutch</td>
</tr>
<tr>
<td>Boeschoten (1990)</td>
<td>Moroccan Arabic-Dutch</td>
</tr>
<tr>
<td>Norrier (1990)</td>
<td>Moluccan Malay-Dutch</td>
</tr>
<tr>
<td>Huwaë (1992)</td>
<td>French-Dutch in Brussels</td>
</tr>
<tr>
<td>Treffers-Daller (1994)</td>
<td>Sranan-Dutch</td>
</tr>
<tr>
<td>Bolle (1994)</td>
<td>Chinese-English</td>
</tr>
<tr>
<td>Milroy &amp; Li Wei (1995)</td>
<td>Papiamentu-Dutch</td>
</tr>
<tr>
<td>Muysken, Kook &amp; Vedder (1995)</td>
<td></td>
</tr>
</tbody>
</table>

The criteria are rarely knock-down criteria by themselves, but should be used conjointly to characterize an utterance or a bilingual speech sample as a case of alternation, insertion, and congruent lexicalization. A first cluster of criteria involve *constituency*, a key notion in grammatical analysis, of course.

When several constituents in a row are switched, which together do not form one unique constituent, insertion is not plausible and congruent lexicalization is a serious possibility. A number of elements form a unique constituent if that constituent contains no other elements. With several constituents, we would have to assume multiple contiguous insertions.

- When the switched element is a single, well-defined constituent, e.g. a noun phrase or a prepositional phrase, insertion is a plausible option; this holds *a fortiori* for single words.
- When several words are switched which do not form one or more constituents together, congruent lexicalization is plausible.

Two examples where we have insertions of a single constituent are (4), where two Dutch constituents are inserted into a Malay structure, and (5):

(4) *Nou* ini *laatste avond*
    now DEM last evening
    'Now this is the last evening.' (Huwaë 1992)
(5) zib li-ya *een glas water of zo*  
get for-me a glass water or so  
'Please get me a glass of water or something.'  

Typical examples of *several constituents* switched contiguously, making insertion less likely, are (6) and (7):

(6) He was sitting down *en la cama, mirandonos peleando*, y really, I don’t remember *si el nos separò* (... on the bed, watching us fight, and ... if he took us apart.)  
    (Poplack 1980)

(7) *[Want] [continu omongj maar [omong leuk]*  
for continuously talk but talk nice  
'For she talks continuously but she talks nice.'  
    (Huwaë 1992)

In (7) *want* and *continu* do not form a constituent together. The constituency criterion needs to be applied with care in the case of idiomatic expressions and collocations such as (8):

(8) *politiek essahtan [reelj [interesseren yap1 - yor]*  
politics really does[not] interest [me] a bit  
    (Boeschoten 1990: 94)

Here *reel* and *interesseren* are inserted contiguously in grammatically separate constituents (and hence would not qualify as insertions according to the constituency criterion), but at the same time clearly belong together as part of an idiom. Incidentally, the idiom appears in atrophied form in the bilingual expression; see below. Clearly idioms and collocations warrant a separate treatment, which goes much beyond the scope of this paper, however (cf. Muysken, in prep.).

The last example is also a case where the two inserted Dutch elements do not form *a unique constituent*, and hence could not be an insertion (apart from their status as an idiom). There are a number of cases like that. In (9) *pa' 'for'* and *el 'the'* belong to a prepositional phrase together, but this constituent also includes English *hamburger meat*.

(9) Yeah, but I buy 'em mostly *pa'el hamburger meat*.  
    ... for the  
    (Lance 1975)

We could still analyze this as an insertion if we think of *[pa'el hamburger meat]* as an inserted Spanish constituent, which itself contains *hamburger meat* as an inserted English constituent. A similar analysis is possible for (3), repeated here as (10):

(10) ...
We could think of the expression [flight [que sale de Chicago [around three o'clock]]] as an English insertion, which contains [que sale de Chicago [around three o'clock]] as a Spanish insertion, which again contains [around three o'clock] as an English insertion. This would make it possible to save an insertion analysis for this example, but this going back and forth is limited to specific types of code-switching in a limited set of language pairs, and may more profitably be thought of as something qualitatively different from insertion, in this case congruent lexicalization.

The same holds in a number of examples presented in (11)-(15). In (11)-(13) the language switch is in the middle of the clause:

(11) verdeeldheid ben de [duidelijk merkbaar] [tussen] division PST COP clearly noticeable between

hindustani nanga a blakaman
Hindustani and the black people.
'Divisions were clearly noticeable between Hindustani and the black people.' (Bolle 1994: 82)

(12) den sten e kon [tapu a voorgrond] [in plaats] fu den dron
PL voice PR come to the foreground instead of PL drums
'Voices come to the foreground instead of drums.' (Bolle 1994: 82)

(13) Ik was ik had mea .. leao heb ik drie jaar gedaan en meao
I was I had Mea .. Leao have I three years done and Meao
Vier jaar [want aku zakken jang laatste jaar] four years [for I fail-INF FOC last year]
'I was I had done three years of Leao and four years of Meao for I failed the last year.' (Huwaë 1992)

In the last example, there is a Dutch clause followed by a Malay clause (containing zakken 'fail' and laatste jaar 'last year' as inserted elements from Dutch but having an unmistakable Malay structure, visible in the function words aku 'I', jang 'focus' and the infinitive ending on the verb). However, the switch comes after the conjunction want 'for' rather than before it.
Quite common are cases like this, where the sentence starts in one language, and continues in another language, but the switch site does not correspond with a natural constituent boundary:

(14) xess-na [m9a bestuur] [praten] (Nortier 1990)
    we-must with board speak
    'We must speak with the board.'

(15) E puntanan aki no [tin niks] [mee te maken] (Kook 1993)
    'The points here not [have nothing] [to do with it].'

Having dealt with constituency in some detail, I now turn to the criterion of *peripherality*. When the switched element is at the periphery of an utterance, alternation is a clear possibility.

(16) les étrangers, ze hebben geen geld hè (Treffers-Daller 1994: 207)
    the foreigners they have no money, huh?
    'The foreigners, they don't have any money, do they?'

Often, a switch involves a left- or right-dislocated element, or two conjoined clauses. These cases clearly qualify as alternations.

A third criterion is *nestedness*:

- When the switched string is preceded and followed by elements from the other language, which are structurally related, it is a case of insertion.

- However, when the switched string is preceded and followed by elements from the other language, which are not structurally related, it is a case of alternation or congruent lexicalization.

Consider the contrast between (17) and (18):

(17) Es una little box así y ya viene … (Lance 1975: 145)
    'It is a little box like this and it comes already'…

(18) (A) Se me hace que (B) I have to respect her (C) porque 'ta …
    older
    'It appears to me that … because [she] is'…

In (17) *little box* is an English noun phrase inserted into a Spanish indefinite expression (determiner phrase), while in (18) *I have to respect her* is
not similarly nested, since (A) and (C) are not at all related.

A criterion of a different nature is selectivity. If a switched fragment forms a constituent selected by an element in the fragment in the other language, insertion or congruent lexicalization are a good possibility. If not, alternation is a plausible option. In (19) there is a direct dependency relationship between the inserted Dutch elements *iets* 'something', *reageren* 'react', and *heel dom* 'very dumb' on the one hand and the Malay structure on the other:

(19) *Dus* kal *elke keer* kal matje bilang *iets* a sing *reageren*
thus when each time when she say something I NEG react
*of aku bikin eh* *heel dom*
or I act eh very dumb
'Thus each time she says something I don't react or I act very dumb.' (Huwaê 1992)

Such a relationship is absent in (20), where the Spanish relative clause simply modifies the English noun phrase these kids:

(20) *We've got all these kids here right now, *[los que están ya criados aquí]*
'..., those that already grew up here'...
(Gumperz and Hernandez Chavez 1975: 157)

In (20) alternation is a plausible option, but not in (19).

Yet a further question that may be posed is whether there is linear equivalence between the two languages. If so, alternation and congruent lexicalization are a clear possibility; otherwise, insertion is most plausible. Notice, of course, that this observation is due to David Sankoff and Shana Poplack (Nait M'Barek and Sankoff 1988; Poplack and Meechan 1995).

I will illustrate the effects of linear equivalence with an interesting case where Dutch has been in contact with two languages with different word orders: Sarnami Hindustani (SOV) and Sranan (SVO), and the results have been different. Dutch has SOV (and particle-verb) order in subordinate clauses, but in main clauses without a modal or auxiliary the finite verb moves to second position, thus creating a superficial SVO (and verb-particle) order. In Sarnami Hindustani mixed verbal compounds, generally the Dutch particle-verb order occurs, as in (21), although there are a few cases like (22) with verb-particle order.
We also have cases where a Dutch particle is combined with a Sarnami Hindustani verb (but with Dutch lexical semantics), and these are always particle-verb:

(23) voor aːve voor-komen 'occur, exist' (Kishna 1979)
    over hove ja: over-gaan 'finish the year'

Dutch verbs have been integrated into Sranan, but always with the verb-particle order, and the particle appears in an adverbial (or serial verb; cf. Seuren's analysis cited in Myers-Scotton 1993) position following the direct object:

(24) a. mi e leg i uit (*leg uit i) (Bolle 1994: 94)
    1sg ASP lay 2sg out
    'I explain to you.'

    b. druk en door (*druk door en)
    press 3sg through
    'push it through'

Again, there are cases where a Dutch particle is combined with a Sranan verb (but with Dutch lexical semantics), and these are always verb-particle in order:

(25) a. fa Draver o syi uit in '92? (Bolle 1994: 93)
    how Draver FU look out in '92
    'What will Draver look like in '92?'

    b. a kranti sa ben gi uit na a kop Coroni
    the paper FU PST give out with the headline Coroni
    'The paper would be published with the headline Coroni.'

Yet a further criterion is length. The more words a switched fragment contains, the more likely that it is alternation. From a psycholinguistic perspective it is plausible that activation of a matrix language decreases as number of words in the intrusive language is larger. Related to length is
complexity. The more complex structure a switched fragment contains, the more likely that it is a case of alternation rather than insertion. In (26) the fragment *donde tenía una casa* is a full clause, and hence it would be difficult to treat it like an insertion or a case of congruent lexicalization.

(26) Right to 104th Street *donde tenía una casa* which were furnished rooms (... where I had a house ...).

(Poplack and Sankoff 1979)

Quite a complex, but potentially useful, criterion is bidirectionality. A pattern of apparent *bidirectionality* of insertion, the absence of a clearly defined matrix language, suggests that we have a shared structure, and hence congruent lexicalization. This was illustrated above in (3) (=10).

So far, we have only considered sentence-internal criteria. However, the embedding in discourse can be indicative of type of switching as well. Consider a mixed clause starting in language A and ending in language B. If the preceding utterance is in A, and the following clause is in B, alternation is a plausible analysis. An example that illustrates this line of argumentation is the following interchange:

(27) A: Yeo hou do yeo contact. (Milroy and Li Wei 1995: 147)
    We have many contacts
G: We alway have opportunities heu xig kei ta dei fong gao
    We always have opportunities to get to know people from wui di yen.
    other churches.

While *contact* in A's line is probably an insertion (it is somewhat integrated in not showing English plural marking), the English stretch at the beginning of G's turn could be thought of as triggered by the earlier English, and an alternation with respect to the rest of the turn.

We can also consider the *structural position* of the switch: If the switch takes place at a major clause boundary, alternation is a plausible option. Internal to a phrase either insertion or congruent lexicalization are more plausible. An example is (27) above, and also (28):

(28) ... maar 't hoeft niet li-'anna ida seft ana ...
    ... but it need not (be) for when I-see I ... (Nortier 1990: 126)

Another indication that we are dealing with an alternation is *flagging* by a pause or a particle such as *euh* in (29). If the switch is flagged, alternation between codes is a plausible option.
(29) Daar zetten ze euh des barrières.
There put they euh the barriers
'There they put up the barriers.' (Treffers-Daller 1994: 219)

In the case of single elements, the lexical category of the switch plays a role. Content words such as nouns and adjectives are likely to be insertions, discourse particles and adverbs may be alternations, and single function words may be cases of congruent lexicalizations. There is no need here to give examples of inserted single nouns, etc. An example from Italian/Sicilian dialect switching is the following:

(30) prima di tutto u portiere non c'è, sarà minimo cinque sei anni
first of all the porter not there-is, will be minimum five six years
'First of all there has been no porter for at least five or six years.'

Here the Sicilian dialect definite article u is used in an otherwise Italian sentence, something easiest interpretable as congruent lexicalization. A case where we may think of alternation is the use of Moroccan Arabic discourse markers in Dutch, as studied by Nortier:

(31) Ik ben dokter wella ik ben ingenieur.
'I am doctor or I am engineer.' (Nortier 1990: 142)

A final criterion is adaptation: if a switched fragment is modified morphologically or syntactically, insertion or congruent lexicalization are most plausible. Many examples of syntactic modification involve deletion:

(32) politiek essahtan reet interesseren yapl - yor (= (8))
politics really does[not] interest [me] a bit
(Boeschoten 1990: 94)

In (32) the Dutch idiom is only partly present: me 'me' and geen 'not one' are left out. In (33) the omission of the article (marked with a 0) is ungrammatical in both languages in the mixing pair:

(33) I ande ez 0 kotel hamaravi, hair atika, ai nasyo mi madre
And where is Western Wall, the Old City, there was born my mother.
(Berk-Seligson 1986)

There are morphological conflicts between Spanish and Hebrew with respect to the question of whether the article should be phonologically sepa-
rate or attached to the noun. The way the conflict is resolved, Berk-Seligson suggests, is by simply omitting it.

In (34) a case is given of a morphologically adapted switch: kamer presumably is inserted into a Turkish nominal constituent.

(34) ben kamer-im-i opruimen yap-ar-ken
     I room-my-ac tidy while-doing
     'While cleaning my room.' (Boeschoten 1990: 90)

This concludes my presentation of a number of criteria that can be used to distinguish the three code-switching patterns or processes. The three patterns clearly differ in the degree to which both languages contribute to the grammatical structure of the utterance. In Table 2 the results are summarized.

Table 2. Diagnostic features of the three patterns of code-switching

<table>
<thead>
<tr>
<th></th>
<th>insertion</th>
<th>alternation</th>
<th>congruent lexicaliz.</th>
</tr>
</thead>
<tbody>
<tr>
<td>single constituent</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>several constituents</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>non-constituent</td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>peripheral</td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>nested a b a</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-nested a b a</td>
<td></td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>selected element</td>
<td>+</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>linear equivalence</td>
<td></td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>length</td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>complexity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bidirectionality</td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>embedding in discourse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>major clause boundary</td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>flagging</td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>lexical category</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>function word</td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>adverb, conjunction</td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>adaptation</td>
<td>+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
When we start applying criteria such as these, a complex picture of code-switching emerges, involving different switching patterns. However, the patterns are only partly separate, since intermediate cases may exist, as will be discussed below. Furthermore, individual bilingual speech corpora may be characterized by several patterns at once, although in different proportions. It is clear that general constraints of the type proposed by DiSciullo, Muysken and Singh (1986) or the authors cited above can hold at best for part of the material available to the research community.

Table 3 gives a very preliminary characterization of some of the major corpora studied in terms of these patterns:

Table 3. Some corpora from recent code-switching studies characterized in terms of the three processes

<table>
<thead>
<tr>
<th></th>
<th>alternation</th>
<th>insertion</th>
<th>congruent lexicalisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swahili-English</td>
<td>some</td>
<td>much</td>
<td>rare</td>
</tr>
<tr>
<td>Tamil-English</td>
<td>few cases</td>
<td>much</td>
<td>none</td>
</tr>
<tr>
<td>Brussels</td>
<td>much</td>
<td>some</td>
<td>few</td>
</tr>
<tr>
<td>Ottawa</td>
<td>much</td>
<td>few</td>
<td>few</td>
</tr>
<tr>
<td>Strasbourg</td>
<td>some</td>
<td>some</td>
<td>some</td>
</tr>
<tr>
<td>New York</td>
<td>some</td>
<td>some</td>
<td>some</td>
</tr>
</tbody>
</table>

I will now illustrate the criteria in Table 2 in more detail using the Moroccan Arabic Dutch corpus described by Nortier (1990). Before doing so, a remark is in order about the level of analysis. Sentence grammar and Discourse grammar (governing the distribution of conversational particles) may be relatively autonomous; there is very frequent language choice disparity between these systems (cf. De Rooij 1995). On the criteria in Table 2, Nortier’s corpus can be characterized as:
Table 4. Features of the Nortier corpus of Moroccan Arabic-Dutch switching

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>single constituent</td>
<td>frequent</td>
</tr>
<tr>
<td>several constituents</td>
<td>some cases</td>
</tr>
<tr>
<td>non-constituent</td>
<td>few cases</td>
</tr>
<tr>
<td>peripheral</td>
<td>often</td>
</tr>
<tr>
<td>nested a b a</td>
<td>frequent</td>
</tr>
<tr>
<td>non-nested a b a</td>
<td>few cases</td>
</tr>
<tr>
<td>selected element</td>
<td>often</td>
</tr>
<tr>
<td>linear equivalence</td>
<td>frequent</td>
</tr>
<tr>
<td>length</td>
<td>long fragments occur</td>
</tr>
<tr>
<td>complexity</td>
<td>complex fragments occur</td>
</tr>
<tr>
<td>bidirectionality</td>
<td>infrequent</td>
</tr>
<tr>
<td>embedding in discourse</td>
<td>not known</td>
</tr>
<tr>
<td>major clause boundary</td>
<td>frequent</td>
</tr>
<tr>
<td>flagging</td>
<td>infrequent</td>
</tr>
<tr>
<td>lexical category</td>
<td>many</td>
</tr>
<tr>
<td>function word</td>
<td>few</td>
</tr>
<tr>
<td>adverb, conjunction</td>
<td>quite a few</td>
</tr>
<tr>
<td>adaptation</td>
<td>frequent through omitted articles</td>
</tr>
</tbody>
</table>

It thus reveals itself fairly similar to the materials described by Poplack (1980) for New York and by Gardner-Chloros for Strasbourg: all processes occur, although insertional patterns dominate.

Let us further analyze the directionality of switching in Nortier's material, since it yields further insights into the types of code-switching in the corpus. Consider Tables 5 and 6. In Table 5 I have interpreted the switching of NPs as insertional (using the criteria from Table 2), identified a number of alternational switch types, and finally looked at alternational switches at the discourse level (between coordinate clauses; Nortier did not quantify switches between turns in her study).
Table 5. Directionality of multi-word switches in Moroccan Arabic (MA)/Dutch (DU) code-switching (based on Nortier’s Table 7.4, p.126)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>%MA &gt; DU</th>
<th>%DU &gt; MA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(a) insertion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V/NP</td>
<td>40</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Subject NP/...</td>
<td>23</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>P/NP</td>
<td>15</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Inside NP</td>
<td>22</td>
<td>95.5</td>
<td>4.5</td>
</tr>
<tr>
<td><strong>(b) alternation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>apposition/dislocation</td>
<td>29</td>
<td>79.3</td>
<td>20.7</td>
</tr>
<tr>
<td>around adverbs</td>
<td>42</td>
<td>78.6</td>
<td>21.4</td>
</tr>
<tr>
<td>.../PP</td>
<td>25</td>
<td>72.0</td>
<td>28.0</td>
</tr>
<tr>
<td>main/subordinate S</td>
<td>37</td>
<td>70.3</td>
<td>29.7</td>
</tr>
<tr>
<td><strong>(c) discourse</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>between coordinate S</td>
<td>42</td>
<td>35.7</td>
<td>64.3</td>
</tr>
</tbody>
</table>

Table 5 shows that insertional mixing is unidirectional, with Moroccan Arabic as the matrix language and Dutch as the intruding language. In contrast, alternational mixing is bi-directional in this corpus. This holds *a fortiori* for discourse-level alternations, where the dominant direction is from Dutch to Moroccan Arabic.

The data from single word switches confirm these findings independently:
Table 6. Directionality of single word switches in Moroccan Arabic (MA)/Dutch (DU) code-switching (based on Nortier’s table 7.15, p. 141)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>%MA &gt; DU</th>
<th>%DU &gt; MA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(a) major</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adjectives</td>
<td>26</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>verbs</td>
<td>11</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>numerals</td>
<td>4</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>nouns</td>
<td>286</td>
<td>99.7</td>
<td>.3</td>
</tr>
<tr>
<td><strong>(b) minor</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pronouns</td>
<td>6</td>
<td>83.3</td>
<td>16.7</td>
</tr>
<tr>
<td>adverbs</td>
<td>41</td>
<td>70.7</td>
<td>29.3</td>
</tr>
<tr>
<td>prepositions</td>
<td>8</td>
<td>62.5</td>
<td>37.5</td>
</tr>
<tr>
<td><strong>(c) discourse</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>conjunctions</td>
<td>20</td>
<td>30</td>
<td>70</td>
</tr>
</tbody>
</table>

Table 6 shows that for single word switches, major categories follow the insertional pattern (from Moroccan Arabic to Dutch), while minor categories (a much more limited category) are switched in both directions. This is not likely to be alternational, but may be due to congruent lexicalization. With conjunctions, the reverse order predominates (except for subordinating conjunctions).

I will conclude by arguing that the differences between the three types are gradual rather than absolute. Between *insertion* and *alternation* there is a transition zone since insertion of longer fragments leads to more complete activation of the second grammar. There are different modules potentially involved in the insertion: phonetic shapes, lexical meaning, morpho-syntax, which involve activation of the second grammar to different degrees.

Since only some higher-level structures may be shared, there is a gradual transition as well between *alternation* and *congruent* lexicalization: alternation where only the top node or nodes are shared (the sentence or clause node), and congruent *lexicalization* when all or most nodes are shared between the two languages.
Finally, there can be a gradual shift from one base language to shared structure and on to the other base language, possibly varying with individual bilingual competence and over time. This implies particularly that in many immigrant communities, *insertion* of new items and expressions into the home language can evolve into *congruent lexicalization* and then possibly into *alternation* (with set phrases and expressions).

**Acknowledgements**

This paper was given in various stages of development at the Summer School on Code-Switching and Language Contact at the Fryske Akademy in Leeuwarden/Ljouwert (September 1994, in a poster conjointly with Roeland van Hout), at the Romanisches Seminar, Universität Basel (January 1995), at the LAUD Symposium, Gerhard Mercator Universität Duisburg (February 1995), at a post-graduate seminar of the Holland Institute of Linguistics (March-June 1995), and as a presentation in the Theme Group on Bilingualism at the Netherlands Institute of Advanced Studies (NIAS). I am grateful for comments and critical remarks from the participants, notably Peter Bakker, Rik Boeschoten, Roeland van Hout, Georges Lüdi, Martin Pütz, and Carol Pfaff. Of course, anyone familiar with the field will recognize various intellectual contributions by others.

**References**


Giesbers, Herman  

Gumperz, John J. and Eduardo Hernández-Chavez  

Hernández-Chavez, Eduardo et al. (eds)  

Hout, Roeland van and Pieter Muysken  

Huwaë, Rosita  

Kishna, Sita  

Lance, David M.  

Milroy, Lesley and Li Wei  

Milroy, Lesley and Pieter Muysken (eds)  
1995 *One Speaker, Two Languages*. Cambridge: Cambridge University Press.

Muysken, Pieter  

Muysken, Pieter (in prep.)  

Muysken, Pieter, Kook, Hetty and Paul Vedder  
(subm.)  

Myers-Scotton, Carol  


Nait M’Barek, Mohammed and David Sankoff  

Nortier, Jacomine  

Peñalosa, Fernando  

Pfaff, Carol  
Pfaff, Carol

Poplack, Shana

Poplack, Shana and Marjory Meechan

Rooij, Vincent de

Sebba, Mark

Treffers-Daller, Jeanine