The typology of code-switching

Pieter Muysken

*Universiteit van Amsterdam*

In this note I will briefly give a global indication of the work in progress in the field of code-switching, and I refer the reader to the series of proceedings of the ESF Network on Code-Switching and Language Contact for a more detaild perspective. I will focus on the limited area of my expertise in this large and interdisciplinary domain: the grammatical analysis of intra-sentential code-switching. The paper becomes progressively narrower in scope. Section 1 briefly outlines some of the factors that help shape the patterns of code-switching in individual communities. Section 2 focusses on some of the parameters that can be identified in the grammatical typology of code-switching, and section 3 evaluates the merits of my own approach to the phenomenon (with some bias, I am sure), in terms of the notion of goverment.

Code-switching in general is a key process in present-day linguistic behaviour, speaking generally, for the simple reason that bilingual communicative settings are becoming more and more wide-spread. There is a general tendency, in language and culture, towards globalization—in terms of movement of goods, persons, and above all, information— at the same time as towards fragmentation—affirmation of identities, at the ethnics, regional and generational level. The work on the grammatical dimensions of code-switching has gone through several stages. In a first stage there were a number of separate observations; then from the early eighties onwards, people tried to come up with universalist postulates. As the data became more varied, after 1986, there were a
number of attempts at parametrized or relativized global models, with attention paid to typological differentiation. At present, we find only tentative multi-factorial models on the base of comparative research. The reason for this retreat or evolution is that the factors that intervene are highly complex. That we must be less optimistic than before has positive aspects as well, since the need for caution is due to the much more detailed knowledge that we have of code-switching patterns.

I. Factors that intervene

A first set of factor has to do with the social definition of the bilingual situation. We have to distinguish a great number of different contexts. To name just a few:

- linguistic fronteers (Brussels, Strasburg)
- elite bilingualism (pre-soviet Russia)
- colonial language (French in Morocco)
- migrant communities (Puertoricans and Mexicans in the U.S., minority communities in Europe)
- clusters of multilingual tribal societies (Amazon basin, Australia, New Guinea)
- dialect-standard language (Germany, Italy)
- minority language islands (Basque)

These communities differ in many ways, including:

- the degree of acceptation in the community of code-switching
- attitudes towards bilingualism in general
- structures of linguistic domination
- transplanted or endogenous bilingualism
- the generational axis

Second, the interactional setting should be taken into account, it turns out in a number of studies. Context investigated include:
The typology of code-switching

- adolescent peer group informal interactions
- family conversations, e.g. at mealtimes
- classroom interactions
- functionary/citizen interchanges
- exploratory conversations between relative strangers

A third set of factors have to do with the linguistic typology of the languages concerned. Dimensions include:

- order of constituents
- morphological typology (agglutination, etc.)
- marking of the relation predicate/arguments
- phonology

In a more general sense the degree of linguistic kinship and the lexical and morphosyntactic similarity between the languages involved are important, of course.

A fourth set of factors which help determine the form of code-switching in specific situations is psycholinguistic. One relevant dimension is bilingual competence, involving possibilities such as:

- language loss
- incipient bilingualism/second language learning
- perfect, balanced bilingualism
- dementia and aphasia

Another dimension is age: are we dealing with child / adolescent / adult / old age bilingualism?

A final set of factors that need to be mentioned in this rather hectic catalogue is duration of the contact. Longterm contacts may facilitate code-switching in at least two ways:

- linguistic convergence, leading to increased equivalence
- the emergence of specific bilingual adaptation strategies

In addition, of course, longterm contacts may lead to increased tolerance for the mixed forms.
2. Parameters

Let us now turn more concretely to the phenomenon of code-switching as such. I will discuss a number of basic analytic distinctions that could be made with respect to the code-switching data that have come to the fore in recent years:

(a) A first one is the distinction between *alternation* and *insertion*. In some cases, such as (1), it seems halfway through the sentence, one language is replaced by the other. In others, such as (2), a single constituent is inserted into a frame provided for by the matrix language:

1. ... maar 't hoeft niet # li-'anna ida ëeft ana ...  
   ... but it need not (be) for when I-see I ...
   (Moroccan Arabic-Dutch; Nortier, 1990: 126)

2. kalau dong tukan bikin dong tukan bikin  
   when they always make always make
   # voor acht personen # dek orang cuma nganga dong makan  
   for eight people and then they only look they eat
   (Moluccan Malay-Dutch; Huwae, in progress)

(b) Second, we must distinguish cases where the switch involves otherwise *grammatical* constituents, as in (3), from those where the constituents from at least one language are *incomplete* or *telegraphic*, as in (4):

3. ma ta # voor de zekerheid # ham even nihur ke... 
   but to be sure I bent over...
   (Sarnami Hindustani-Dutch; Kishna, 1979: 69)

4. politiek essahtan # reet interessen # yapI – yor  
   politics really does (not ) interest (me) a bit
   (Turkish/Dutch; Boeschoten, 1990: 94)

In (4) the Dutch idiom is only partly present: *me* `me' and *geen* `not one' are left out.
The typology of code-switching

(c) A further important distinction is that between smooth switching, in which the transition between the two languages is seamless, as in (5), from flagged, specially marked switches, as in (6).

\[(5) \quad \text{Zib li-ya} \quad \text{een glas water of zo} \\
\quad \text{get for me} \quad \text{a glas of water or something} \\
\quad \text{(Moroccan Arabic-Dutch; Nortier 1990: 131)}\]

\[(6) \quad \text{tu sais, l'affutleur de scies hein,} \quad \text{# daan gink bij die} \\
\quad \text{you know, the knife-grinder eh,} \quad \text{he went to the} \\
\quad \text{beenhouwers, de zager} \\
\quad \text{butchers, the sawyer} \\
\quad \text{(French-Dutch in Brussels; Treffers 1990: 266)}\]

In (6), the discourse marker *hein* separates the two languages, French and Dutch.

(d) A crucial distinction is that between *clause-central* and *clause peripheral* code-switching. Is the switch-point at the heart of the clause between verb and infinitival complement, as in (7), or rather marginal to the core proposition, as in (8), where we have a left-dislocated French subject noun phrase, followed by a Dutch subject pronoun:

\[(7) \quad \text{ka-ne9ref} \quad \text{# Surinaams zingen} \\
\quad \text{I-know} \quad \text{Surinamese sing} \\
\quad \text{(Moroccan Arabic-Dutch; Nortier 1990: 131)}\]

\[(8) \quad \text{les étrangers} \quad \text{# ze hebben geen geld hè} \\
\quad \text{the foreigners} \quad \text{they don't have any money do they?} \\
\quad \text{(French-Dutch in Brussels; Treffers 1990: 271)}\]

(e) A final set of distinctions I will mention here has to do with the degree of adaptation of the switched elements. Some elements are completely adapted, as in (9), where Turkish affixes appear on Dutch elements. In (10) there is a phonetically ambiguous pronun-
ciation at the switch-point, and in (11) the Dutch verb could either be an infinitive or a non-adapted plural form of the verb:

(9) ben kamer-im-i opruimen yap-ar-ken
    I room-my-ac tidy while-doing

(10) weet jij # waar/where # Jenny is
    do you Know where Jenny is
    (English-Dutch; Cram and Van Gelderen 1984)

(11) tta huma # reserveren # kull-ši
    also they reserve everything
    (Nortier 1990: 144)

In (12), finally, it is clear we are dealing with an infinitive, which is the way Dutch verbs function in in variant form in Malay:

(12) Co ko6 # vertolken # akan dolo
    Hey you interpret it just
    (Moluccan Malay-Dutch; Huwae, in progress)

This concludes my brief survey of some of the distinctions one might want to make to analyze code-switching phenomena.

3. Government: the fate of a constraint

In 1981 I worked in Montreal together with two Canadian colleagues, Rajendra Singh and Anne-Marie DiSciullo, and we noticed that many cases of switching in French/Italian and Hindi/English code-switching obeyed a general constraint of the type in (13):

(13) DiSciullo, Muysken, Singh (1986):
    X and Y cannot have different language indices iff X governs Y

A constituent «is in » a certain language if it has the corresponding language index. Government is the general relation of dependency between a head and a complement.
The typology of code-switching

The constraint in (13) predicts that switches of the types in (14) are ungrammatical:

(14) a. Verb / Object
    b. Preposition / NP-complement
    c. Verb / Clausal complement

Indeed we find many switches which do not involve these configurations. However, (13) fails badly for a large number of cases such as those in (15):

(15) a. sempre vicino a quella / machine
    always near to that / machine

    b. ha ricevuto il / diplôme
    has received the / diploma

unless we take the noun phrases *quella machine* and *il diplôme* to be Italian, despite *machine* and *diplôme* being French, on the basis of the Italian demonstrative and article. Therefore, the notion of Lq-carrier or language index carrier was introduced in our original article, as in (16):

(16) The highest (non-lexical) node in a projection determines its language index

On the basis of this notion, the first element in the constituents in (17) determines the language index of the whole constituent:

(17) a. [ Determiner / Noun ]
    b. [ Quantifier / Noun ]
    c. [ Subordinator / Clause ]

Notice now that the example in (15) are a case of type (17a).

It turned out, however, that (14) still had a number of undesirable consequences, even for the original data considered by DiSciullo, Muysken, and Singh (1986). First, a definition of government in which the noun governs the whole noun phrase, its maximal projection, would rule out cases like (17a), which are extremely common. The same would occur with respect to adverbs,
which often are within the verb phrase, albeit not direct complements of the verb. At the same time, switched adverbs, as in (18), are extremely frequent:

(18) la lascia / toujours / sulla tavola  
She leaves it / always / on the table

Therefore a limited definition of government, involving only the immediate domain of the lexical head: its complements, must be adopted.

Second, in the configurations in (17) one might want to say that the determiner governs the noun, etc., particularly in the view that became popular in the mid-eighties holding that determiners and subordinating complementizers are functional heads. Similarly, the auxiliary or the finite tense marking on the main verb is often assumed to govern the subject. Still, there are many subject / verb phrase switches, as in (19):

(19) la plupart des canadiens / scrivono «c»  
the majority of Canadians / write «c»

If all these cases are taken into account, a more limited definition of government, restricting it to the relation between a lexical head (i.e. N, V, A, P, and excluding functional heads such as determiners) and its immediate complements (excluding other elements within the constituent), should be adopted for code-switching. A convenient term for this relation is given in Chomsky (1986), L-marking. Thus (14) could be formulated as in (20):

(20) Muysken (1988)  
X and Y cannot have different language indices iff X L-marks Y

As it turns out, however, even this restricted version runs into grave difficulties. Particularly damaging for the constraints in (14) and (20) are the data presented in Nortier (1990), who studied Dutch-Moroccan Arabic code-switching in detail. Her findings are completely in contradiction with the government constraint. First of all, we get large number of cases of switching between a verb and its complements:
The typology of code-switching

(21) a. čib li-ya / een glas water of zo  
    get for-me / a glass of water or so (7)

b. anaka-ndir / intercultureel werk  
    I I-am doing / intercultural work (14)

c. wellit / huisman  
    I-became / «houseman» (10)

d. 9end-i / vijf Surinaamse kinderen  
    I-haye / five Surinamese kinderen (8)

We get seven cases of switching between indirect and direct object (21a), no less than 14 cases of switching between verb and direct object (21b), and ten cases involving a predicative after a copula-type verb, (21c). The status of cases such as (21d) is less clear; we may have a post-poned subject here rather than an object, since literally it says: «there are for me...». I should also mention the occurrence of 97 switches of object noun phrases involving a single noun.

The data in (21) are particulary damaging since switching between subject and verb is, if anything, less frequent in Nortier's corpus, than switching between object and verb. Compare the cases in (22):

(22) a. humaya / vergelijken de mentaliteit met de Islam  
    they compare the mentality with the Islam (3)

b. l'islam kûll-u / is echt liefde  
    the-islam all-of-it / is truly love (3)

c. le-mgarba / strak, hè, stroef  
    the-Moroccans / tight, huh, rough (17)

d. hna ka-yseku / zowel Marokkanen als Nederlanders  
    here they-are-living / both Moroccans and Dutch (1)
In (22a) a Moroccan Arabic subject is combined with a Dutch verb phrase (three cases), and in (22b) with a Dutch copular predicate (three cases). More frequent are examples where a Moroccan Arabic subject appears with a Dutch non-verbal predicate without a copula, as in (22c) (17 cases). (22d), with a Dutch postponed subject, may well be similar to (21d).

We also find 15 cases where a Dutch noun phrase is the complement of a Moroccan Arabic preposition, as in (23).

(23) a. u dewwezna f / zelfde tijd
    and we-spent in

    b. ka-yxxess bezzaf dyal / generaties voorbijgaan
        it-must much of

    c. muwaddafa kama / maatschappelijk werkster
       appointed as

In addition, there are 101 cases of switch between a preposition and single noun.
These data clearly show that the government constraint, even in the revised form of Muysken (1988), cannot be maintained. The distribution of noun phrases is much wider than predicted.

4. Conclusion

The above material was meant to give an illustration of the type of work that can be done and is done in the grammatically oriented study of code-switching processes. The story of the demise of the government constraint is illustrative of this research tradition; in fact, a number of models have been proposed in recent years, and a yet larger number of plausible models still needs to explicitly formulated and tested.
References