J. Linguistics **22** (1986), 1–24. Printed in Great Britain

Government and code-mixing[1]

ANNE-MARIE DI SCIULLO, PIETER MUYSKEN AND RAJENDRA SINGH

Université du Québec à Montréal, Universiteit van Amsterdam and Université de Montréal

(Received 26 January 1985)

Introduction

The aim of this paper is to argue that the process of code-mixing is constrained by the government relation that holds between the constituents of a sentence. The government constraint replaces a number of specific constraints that have been proposed in the literature to account for apparently 'impossible', 'ungrammatical' or 'non-occurring' types of intra-sentential switches. Code-mixing is a form of linguistic behaviour which produces utterances consisting of elements taken from the lexicons of different languages. Some examples are given in (1).

(1) (a) *English–Spanish* (taken from Sankoff & Poplack, 1981)
Uno no podia comer carne *every day*
'We couldn't eat meat...'
(b) *French–Italian*
Perché è *mauvais*
'Because it is bad'
(c) *Hindi–English*
Bread ne nas mar diya
'The bread (erg.) ruined it'

We will not attempt to give a complete characterization of code-mixing, a phenomenon to which a vast literature has been devoted from the points of view of grammar, sociolinguistics, psycholinguistics and discourse analysis. As far as the speakers of a mixed code are concerned, the alternation between the elements from different lexicons is quite automatic and goes much beyond inter-sentential code-switching, typically observed in diglossia-type situations (cf. Ferguson, 1959).[2] Here only one code is employed at any one time or the

[1] A first version of this paper was presented at the 1980 annual meeting of the Canadian Linguistics Association, Halifax, Nova Scotia. We want to thank the participants in that meeting and later audiences, as well as the reviewers for this journal for their comments, and D. F. Drukker for his help with the Hindi transliterations. The paper was inspired by observations presented in an earlier paper by Singh (1981). Work on this paper was in part supported by an S.S.H.R.C. grant (410-82-0918 R1) to Singh.

[2] Cf. Joshi (1981, 2): 'Mixed utterances are spoken without hesitation, pauses, repetitions, corrections, etc., suggesting that intrasentential code-mixing is not some random interference on one system with the other. Rather the switches seem to be due to systematic interactions between the two systems.'

code alternation corresponds to structurally identifiable stages or episodes of a speech-event (as opposed to a single sentence). Only questions of *who* speaks *what* to *whom* and *when* enter here (cf. Fishman, 1965), not grammatical constraints.

Not only do we have to distinguish code-mixing from inter-sentential code-switching, but also from different types of other language mixture processes which typically affect structural characteristics of the languages involved. In code-mixing 'the structural integrity of the component languages' is preserved (Sankoff & Poplack, 1981), and the mixed codes remain phonologically and morphologically separate. In order to study code-mixing, therefore, we have to abstract away from possible cases of borrowing, fixed mixed expressions, relexifications, and newly formed mixed compounds.[3] These occur, unfortunately, particularly in communities where code-mixing is frequent, and hence the languages involved exert a considerable lexical influence on each other. It is not very frequent in situations where such abstracting away would not be required. The problem, however, is not an insurmountable one, since true borrowing generally involves phonological nativization (cf. Gumperz, 1976; Singh, 1981; Poplack, 1980) and speakers often have intuitions about the status of borrowed items. Additional criteria are whether a base language equivalent is in common use in the community and known to the speaker.

This paper has the following structure. In Section 1 we present a brief summary of some of the previous work on grammatical constraints on code-mixing. Section 2 contains a theoretical presentation of the core elements of our proposal, and lists a number of its grammatical consequences. Section 3 is devoted to the application of our proposal to two empirical domains: French–Italian and English–Italian code-mixing in Montreal (3.1), and Hindi–English code-mixing in urban North India (3.2). In Section 4 finally we attempt to provide a more principled explanation for our proposals.

1. Background

Some linguists have despaired of finding any structural constraints on code-mixing. Lance (1975), for example, claims that 'there are perhaps no syntactic restrictions on where the switching can occur'. But as Gumperz (1976), Pfaff (1976, 1979), Shaffer (1978), Kachru (1978, 1980), Singh (1981), and Timm (1975), among others, have shown, there clearly are some. The question, in other words, is not whether there are any structural constraints but what is the best way to characterize them, and whether they can be made to follow from an independently motivated, more general principle.

---

[3] For a discussion of the problems that arise from a refusal to do so, see Shaffer's (1978) critique of Lance, and for an early attempt to distinguish code-mixing from borrowing see Haugen (1973). Like Haugen, we also use 'lack of phonological adaptation' as a crucial criterion.

At this point we shall not enter into a discussion of all the specific constraints proposed in the literature, preferring to refer to them at specific points in the exposition in Sections 2 and 3. In any case, the majority of them are limited to a specific structure or constituent: the coordinating conjunction constraint (Gumperz, 1976); the complementizer constraint (Kachru, 1980, Singh, 1981); the specifier constraint (Kachru, 1980, Singh, 1981; Timm, 1975); the clitic constraint and the inflectional constraint (Pfaff, 1979); and the adjective order constraint (Pfaff, 1976).

The first general principle formulated to constrain code-mixing appears in Sankoff & Poplack (1981: 4): 'The equivalence constraint: the order of sentence constituents immediately adjacent to and on both sides of the switch point must be grammatical with respect to both languages involved simultaneously'. Thus, if in language 1 the order of two types of constituents or elements is A/B, and in language 2 it is also A/B, we find the possible outputs A_1/B_2 and A_2/B_1 in mixed code. If on the other hand language 1 has A/B and language 2 had B/A, no code-mixing will be possible.

The equivalence constraint, we would like to argue, is undesirable from a theoretical point of view, as well as empirically inadequate. Note that, for it to be applicable to code-mixing in natural languages, there needs to be categorial equivalence. If language 1 had the categories determiner and conjunction, for example, language 2 must have them also, otherwise it will be impossible to determine whether switching is possible at the point between conjunction and determiner. While there probably are major categories shared by all languages, there are a considerable number of categories which only occur in specific languages. And even then it is not evident that the categories in different languages will precisely correspond. In the model that we will propose below this problem is avoided because it predicts only where switching could occur, and this with respect to one linguistic system.

A second general problem with the equivalence constraint is that it is formulated exclusively in terms of linear sequence, rather than in terms of structural relations. Since we hold that most principles of grammar are formulated in terms of hierarchical relations rather than of linear order, and since code-mixing appears to involve central aspects of grammatical competence, it would be necessary from the point of view of the theory of grammar that constraints on code-mixing are structural rather than linear. We will try to formulate such constraints in the next section.

While Sankoff and Poplack's equivalence constraint goes a reasonably long way towards excluding a number of non-occurring switches in English–Spanish mixed codes as used by Puerto Ricans in New York, we will argue in Section 3 that for other types of mixed codes it makes the wrong predictions. In the case of French–Italian code-mixing in Montreal, switching should be possible at every juncture, given the equivalence constraint, since the word orders of the languages involved are rather similar. In fact, we find numerous restrictions in the case of French–Italian code-mixing, which would have to be blocked

by constraints functioning alongside the equivalence constraint. While 'over-predicting' in the Montreal case, the equivalence constraint underpredicts in the Hindi–English code-mixing situation. Given that Hindi is in many respects typical of an SOV language, Hindi–English code-mixing is predicted to be virtually non-existent. Quite a few switches of different types are possible in the Hindi–English mixed codes. A similar problem exists with the recent proposal by Woolford (1982) to constrain code-mixing in terms of the congruence of the phrase structure rules of the two languages.

Another empirical inadequacy of the equivalence constraints is that it leaves unexplained for the Spanish–English case why certain allowable switch points show hardly any or no cases of switching, why the strength of a syntactic boundary is directly proportional to the possibilities of switching. 'Those exceptional boundaries', Sankoff and Poplack observe, 'which show a relatively low rate of switching involve two closely bound syntactic elements whose relationship approaches, but does not quite enter' the domain of morphological boundedness (46). In other words, what is involved is a general principle of syntagmatic coherence, not a principle in terms of linearity. We shall argue in Section 2 that the principle of government, the syntagmatic coherence principle of traditional grammar and of recent generative grammar *par excellence*, provides a coherent and quite general account of allowable switching sites. In fact, the focus will not be on switching sites, but on relations between elements: when a government relation holds between elements, there can be no mixing; when that relation is absent, mixing is possible. No specific constraint needs to be stated to account for code-mixing restrictions. These fall out from general considerations of lexical integrity, constrained by the government condition, which hold for all uses of natural languages, not just for code-mixing.

The government principle subsumes most of the cases predicted by the equivalence constraint of Sankoff & Poplack and by the particularistic constraints in the earlier literature, and provides a principled explanation for the boundary-strength mystery referred to above. We do not want to claim at the present moment that the principle of government is the only condition, capable in itself of handling all possible restrictions existing in different code-mixing situations, but rather that it is the only universally applicable one. In specific cases, there most certainly will be additional language-particular constraints. An example may be that NP-internal agreement rules may block cases of switching within the Noun Phrase.

2. Code-mixing and government

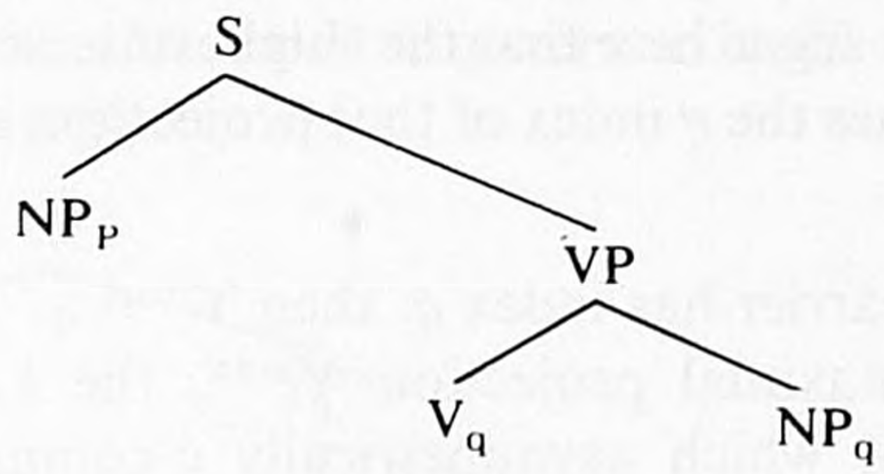
Any constraint on code-mixing should capture the fact that within a sentence elements bearing a certain type of relation to each other must be drawn from the same lexicon or, stated differently, must have the same language index q. We take the notion of language index to be a basic one; it simply marks

the words that are drawn from a particular lexicon. Base rules do not have language indexes as such associated with them, but structures may be indexed through percolation, as we shall argue in Section 4 below. Formally, then, the government requirement would be:

(2) $\ldots X_q \ldots Y_q \ldots$, where X and Y are related elements.

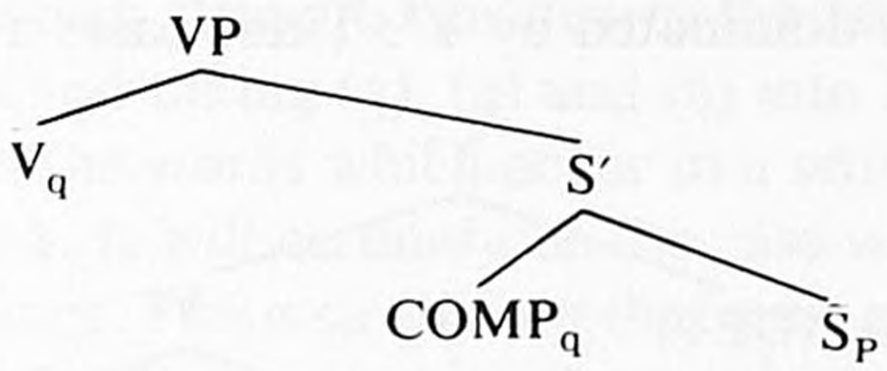
A number of questions arise immediately. (a) What kind of relationship must hold between X and Y? (b) Are X and Y constituents or terminal phrase nodes? (c) If they are not terminal nodes, how then can they have a language index assigned to them? We shall discuss these questions in turn. A first, very general, observation to be made with respect to the relation in (2) is that switching may occur between subjects and verbs, but not in the same way between verbs and objects. Schematically:

(2)′

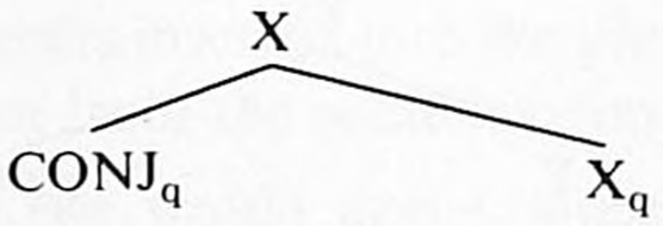


A second observation made in several separate studies is that complementizers can be in a different language from their sister S. At the same time, conjunctions are in the same language as the constituent that they conjoin to something else.

(2)″



(2)‴



The asymmetry between subjects and objects, on the one hand, and between complementizers and conjunctions, on the other, suggests that the relation between X and Y may be government: if X has language index *q* and if it governs Y, Y, must have language index *q* also:

(3) if X governs Y, $\ldots X_q \ldots Y_q \ldots$

Note that while in (3) the restriction is formulated sequentially, no sequentiality is implied. Furthermore, we do not find an *only if* relationship in (3): most discourses are characterized by elements having the same L_q index, while there is NO necessary government relation between them. For the purposes of this paper, we adopt the following definition of government:

(4) X governs Y if the first node dominating X also dominates Y, where X is a major category N, V, A, P[4] and no maximal boundary intervenes between X and Y.

Thus (3) and (4) provide an answer to question (a) and part of question (b) above. A lexical element X must c-command Y for both to have the same language index. How about Y? We assume that the general format of the X-bar expansion rule is as in (5):

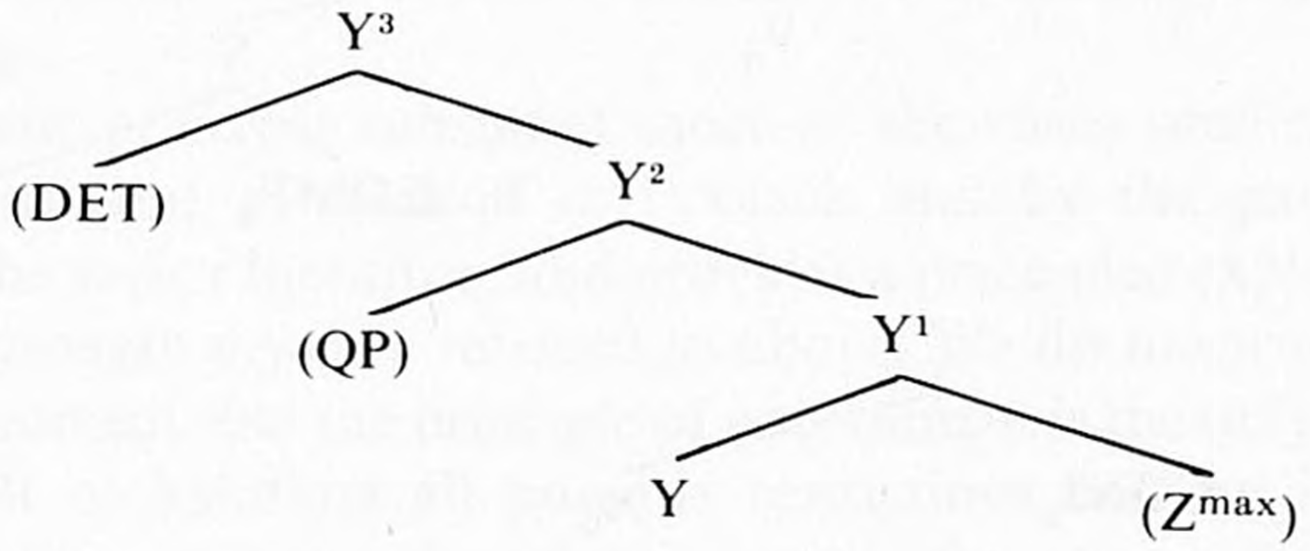
(5) $X^i \rightarrow \ldots Y^{max} \ldots X^{i-1} \ldots Z^{max} \ldots$

Of this general rule, the instantiation where *i* has the value 1 will produce strings in which a lexical item X will c-command maximal categories Y^{max} and Z^{max}. Thus the governed item in (3) will not be a terminal phrase node, but rather a maximal projection. How, then, can it have a language index *q* assigned to it: We argue here that the 'highest' lexical element in a maximal projection determines the *q* index of that projection, and call this element the L_q carrier.

(6) (a) If L_q carrier has index *q*, then $Y^{max}{}_q$.
(b) In a maximal projection Y^{max}, the L_q carrier is the lexical element which asymmetrically c-commands the other lexical elements or terminal phrase nodes dominated by Y^{max}.

The L_q carrier may be the head, when there are no lexical elements dominated by the Y^2 or Y^3 levels; a quantifier phrase (QP), when there are no lexical elements dominated by the Y^3 level; or the determiner (DET), which we assume to be dominated by Y^3. These cases are illustrated in (7).

(7)



[4] The distinction between N, A, V, P and other categories is defended in work on categories such as Jackendoff (1977) and Van Riemsdijk (1978). The distinction is made in terms of being an open or a closed class, having a full projection, etc. In later work (e.g. Chomsky, 1981) INFL is added to the list of governors, particularly with respect to the subject. This extension is explored for code-mixing in a preliminary way in the work of Klavans (1983). She notes that there are restrictions on switching in pro-drop languages, as in (i):

(i) * pro_{sp} $works_{eng}$

To incorporate this observation, we would have to extend the definition in (4) to include INFL. The trouble is that other subject/verb switches are possible, as in (ii):

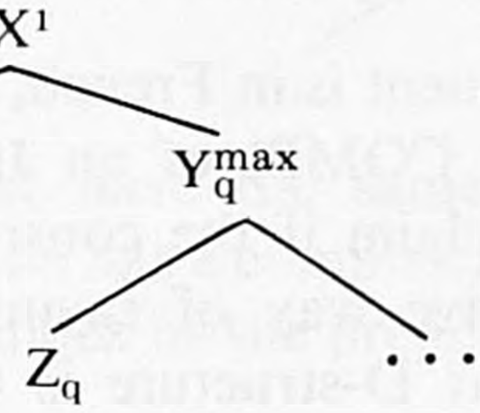
(ii) $Mary_{eng}$ $trabaja_{sp}$

We shall leave this problem, noted by Pfaff (1979) and Woolford (1982), for further research.

We assume, in (7), that the lexical head is selected as L_q carrier in case there is both a head present and its complement(s), given (6b).

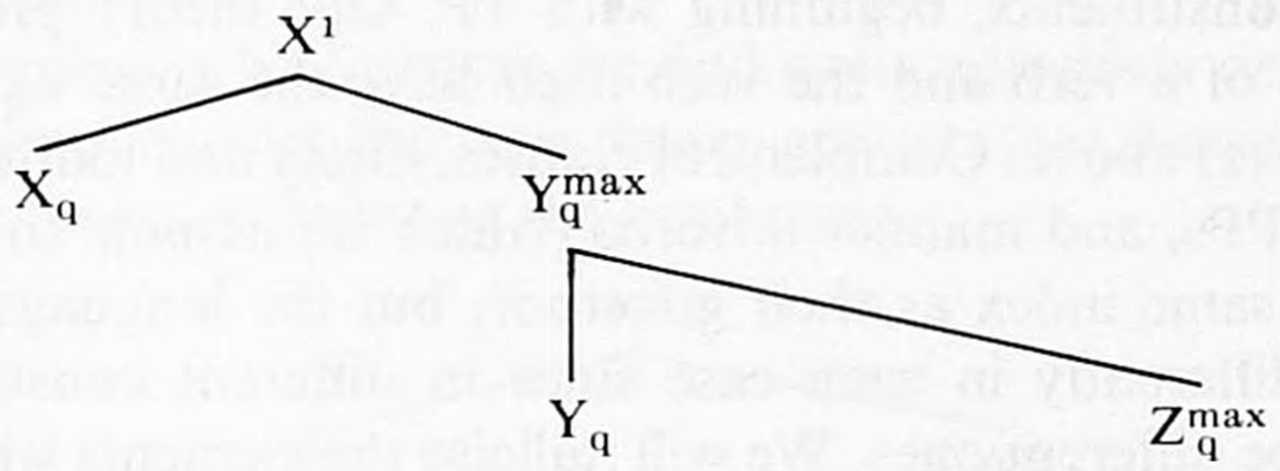
Given (3), (4) and (6), we can now give a more precise characterization of the way code-mixing is constrained by the principle of government. At least the L_q carrier of a governed category must have the same L_q index as its governor:

(8)



In those cases in which the L_q carrier of the governed category (Z in (8)) is the head of that category and hence a governor itself, we get a chain of co-indexed elements, as in (9).

(9)



Then again, if Z^{max} in (9) has its head as L_q carrier, the chain continues.

Trees (8) and (9) exemplify which elements must carry the same language index, in a given configuration and taking (3), (4) and (6) into account.[5] In ordinary cases of language use, the words which occur in a sentence will be all drawn from the same lexicon. It will certainly be the case when speaker and hearer speak the same language. This is so obvious that most grammarians haven't bothered to formulate constraints such as the one in (10).

(10) All elements inserted into the phrase structure tree of a sentence must be drawn from the same lexicon.

Fortunately so, we would argue, since (10) is too strong, and should be replaced by (3). Thus code-mixing can be seen as a rather ordinary case of language use, requiring no specific stipulation. Whenever the syntagmatic coherence principle of government does not hold, the lexical elements may be drawn from different lexicons, if social setting, participants in the conversation, topic of conversation, and cultural intentions would make that

---

[5] What we are suggesting is that code-mixing follows the constraints it does because, *ceteris paribus*, it pretty much has too. What is interesting about particular cases of code-mixing is the interplay of language-particular parameters that allow leaks in the constraint proposed in this paper (or, conversely, the development of additional constraints that must be obeyed).

desirable. Before going on to discuss specific instances of the government relation, we should specify the level of grammar at which the constraint formulated in (3) holds: we will assume it to be S-structure. Arguments can be given on the basis of switches such as:

(11) *L'échantillon*$_i$ [che fanno e$_i$ . . .
'The sample that they create . . .'

Here the italicized constituent is in French, moved from the object position (or linked to it through COMP) of an Italian clause. This would be a counter-example to our claim if the constraint were to be formulated at D-structure. An alternative way of saying that the constraint holds at S-structure rather than at D-structure is to state that language indexes, contrary to referential indexes, are not assigned to chains, but to individual lexical items, i.e. to positions in a chain dominating lexical material.

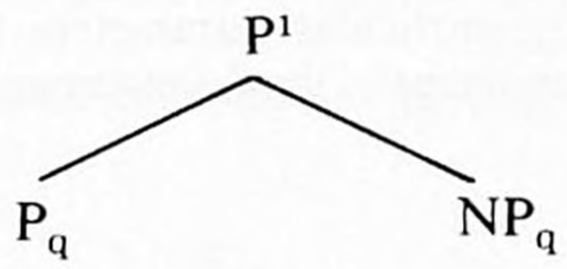
We will now show what implications the principles outlined above have for specific constituents, beginning with VP. Our theory predicts that the complements of a verb and the verb itself have the same L_q index, as was mentioned in (2) above. Complement clauses, direct and indirect object NPs, complement PPs, and manner adverbs (which we assume to be APs) must all carry the same index as their governor, but the language index will be determined differently in each case since in different constituents the L_q carriers will be different ones. We will italicise the elements which must have the same L_q index in the following English examples:

(12)	(a)	I *saw that* he left	V COMP
	(b)	I *saw the* man	V DET
	(c)	I *went to* Rome	V P
	(d)	I *went very* quickly	V Q

In (12a) *that* is the L_q carrier of the complement clause, in (12b) determiner *the* is the L_q carrier of the direct object NP, in (12c) the preposition *to* is the L_q carrier of the complement PP, and in (12d), finally, *very* is the L_q carrier of the manner adverb phrase. Again, our theory predicts that in actual code-mixing situations, these underlined elements will always be drawn from the same lexicon. In fact, (12a) corresponds to the case of the complementizer condition, stated independently by Kachru (1980) and Singh (1981), ensuring that the complementizer of a complement clause is in the same language as the matrix verb, not as the complement clause itself necessarily. In Section 3 we shall see to what extent our predictions are borne out.

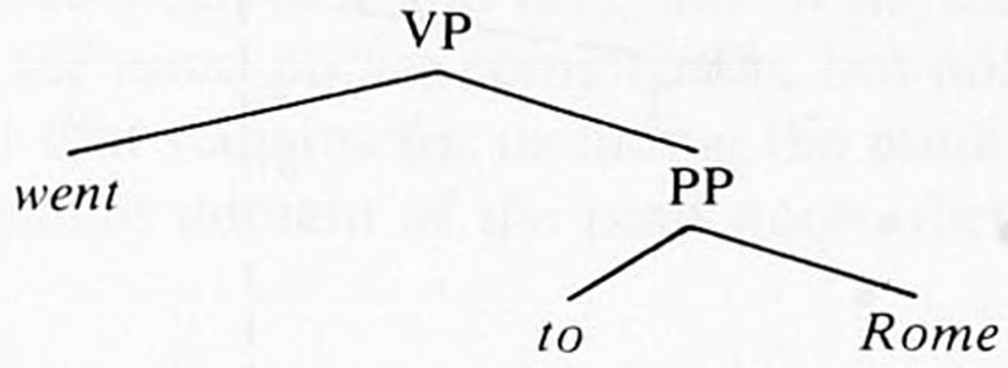
Similarly, the complements of a preposition must have the same index as the preposition itself. This involves, most often, an NP complement, of course:

(13)



Since the prepositions often will be the L_q carrier of their maximal projection, PP, prepositions will tend to be involved in government chains as in (14).

(14)



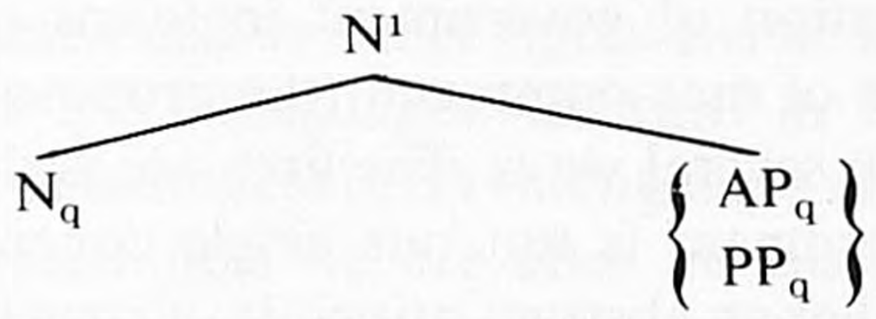
Here *Rome*, *to* and *went* must have the same index. If, however, the complement NP of the preposition has a complex internal structure, only the L_q carrier must have the same index as the preposition:

(15) a sonata *for two* violins

Our theory predicts that the complement of a preposition, through its L_q carrier, will be drawn from the same lexicon as the preposition itself in code-mixing situations.

A third governor is N. Suppose we find some adjectives on the N^1 level, as well as PP complements, then our theory predicts that these elements must have the same language index as the head noun.

(16)



Both Pfaff (1976) and Sankoff and Poplack, researching Spanish–English code-mixing, note that adjectives are most often switched outside the immediate domain of the noun they modify, for example in substantive use, as predicates, and when an adverb intervenes between the noun and the post-nominal adjective. Given that in English adjectives are pre-nominal, while in Spanish they tend to occur post-nominally, both blame the effect on word-order clash, in accordance with the equivalence constraint. Note, however, that the same result follows from our theory, adopting (16), which has the additional advantage of explaining why an intervening modifying adverb will make the switch possible: It functions as L_q carrier, even though this offends the word order of English as much:

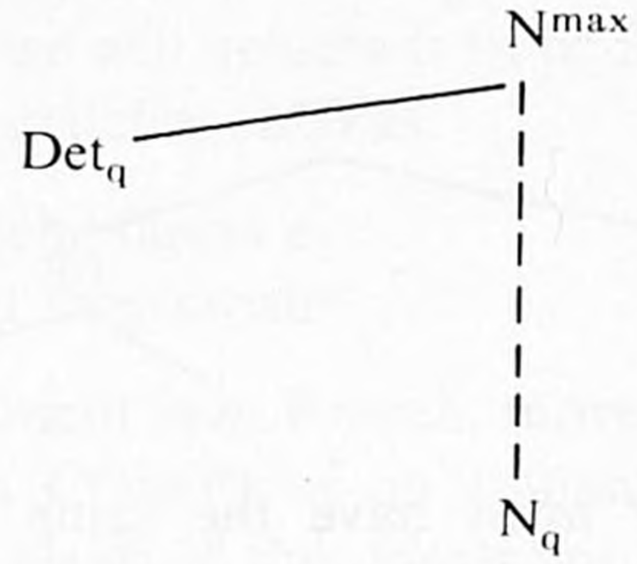
(17) es eso color como [[muy *dark*] *maroon*] (Pfaff, 1976: 256)
'It's that colour like very dark maroon'

Here *muy* modifies *dark*, making the inner AP Spanish, and the AP *muy dark* modifies *maroon*, making the matrix AP Spanish, for the purpose of code-mixing.

For the purpose of our analysis the definition proposed by Aoun & Sportiche (1983) for government, in terms of the maximal projection of the governor, will not do. In the uncontroversial case of the noun phrase as the maximal projection of the noun, the definition of Aoun & Sportiche predicts

that all elements in its domain, including the determiner, will be governed by the head:

(18)



In (18), Det would have the same L_q index as the head noun. In fact, switches between the determiner and the head are exceedingly frequent. Pfaff (1979) documents 759 DET N internal switches, for instance.

Note that adopting Aoun & Sportiche's definition for government will make the notion of L_q carrier vacuous, if we continue to consider that code-mixing is constrained by the relation of government. It is easy to understand why under this definition, as applied to code-mixing, the majority of attested switches would be ungrammatical: all heads in a given configuration (e.g. VP) would have the same L_q index.

Defending the notion of government in terms of minimal c-command rather than in terms of max-command (the proposal of Aoun & Sportiche, 1983) can be done in several ways. The first one is the weakest and involves the claim that government is not one single concretely defined structural relationship, but rather an abstract principle of structural dependence, which through the interaction with other principles of the grammar yields a family of dependences, such as Theta-government, Case-government, Proper government, anaphoric c-command (cf. Koopman, 1984, where this option is explored). In fact, this option is already implicit in Chomsky's adoption of the Aoun & Sportiche proposal (1981) with respect to max-government to explain the distribution of PRO, and of minimal c-command for proper government, to explain the ungrammaticality of (i):

(19) * whose$_i$ did you see [e_i attempt to leave]

If proper government were defined in terms of max-command, *attempt* would properly govern the trace; it it does, however, (19) cannot be blocked with the Empty Category Principle, which states that all traces must be properly governed.

The second type of defence of minimal c-command would be that all cases of government are defined in this way, and that Aoun & Sportiche (1983) are simply mistaken. Note, first of all, that max-command yields the wrong predictions for the binding theory within noun phrases. Consider (20) and (21):

(20) * each others'$_i$ pictures of the artists$_i$
(21) the artists'$_i$ pictures of each other$_i$

These examples should be equally ungrammatical under a max-command definition since *the artists* and *each other* max-command each other, and referential noun phrases cannot have a co-referential binding element. In fact, there is an asymmetry between (20) and (21), due to the fact that the noun phrase specifier binds the noun phrase complement, but not vice versa. The only serious argument that remains for including the noun phrase specifier position in the government domain of the head noun, then, is the absence of PRO here:

(22) * PRO books (meaning: somebody's books)

By deducing that PRO must be ungoverned (it cannot have a governing category since it is both anaphoric and pronominal), Aoun & Sportiche (1983) explain the ungrammaticality of (22) under the max-government definition. Many other possible explanations are available for blocking PRO in (iv): the specifier position must be case marked, etc. Admittedly, these explanations may be more ad hoc than the original ones, but the assumption that PRO is both anaphoric and pronominal is far from uncontroversial (cf. Bouchard, 1984; Sportiche, 1983).

A third line of argument for the analysis we are presenting is the adoption of the notion of directionality of government (Stowell, 1981). If we assume that government is parametrized as either rightward or leftward in different languages (rightward in SVO languages, leftward in SOV languages), it follows that prenominal determiners in SVO languages are ungoverned. This is the main empirical result that we are after at this point. This line of argument would leave the intuitively attractive notion of max-command intact, of course. While several options remain open, we feel that the definition of government that we have adopted in (4) above is far from *ad hoc*, and has a solid base in correct linguistic theory.

We shall not discuss the case of A as governor here, but refer briefly to some problematic cases. A first one involves conjunctions. It has been noted by Gumperz (1976) as well as by others that coordinating conjunctions appear in the same language as the clause they link to a preceding whole. Assume that the structure of conjunctions is as in (23).

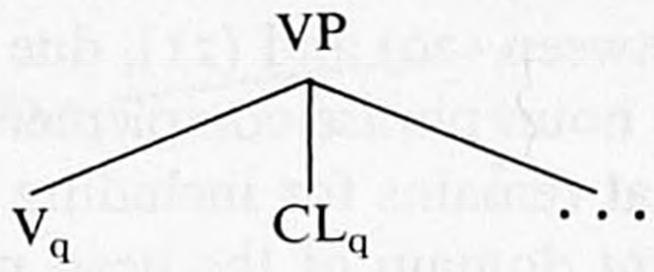
(23)

From our definition of government, it does not follow that CONJ and S′ in tree (23) have the same index. Note, however, that conjoined elements do not govern each other. Thus in (23) the conjunction does not function as an L_q carrier with respect to a language index imposed by an external governor.

A second case so far not discussed is clitics. Pfaff (1979) notes that clitic pronouns are always in the same language as the element to which they are attached. It may be possible, under a syntactic view of cliticization, to say

simply that clitics are dominated by VP, and hence governed by the verb, as in (24):

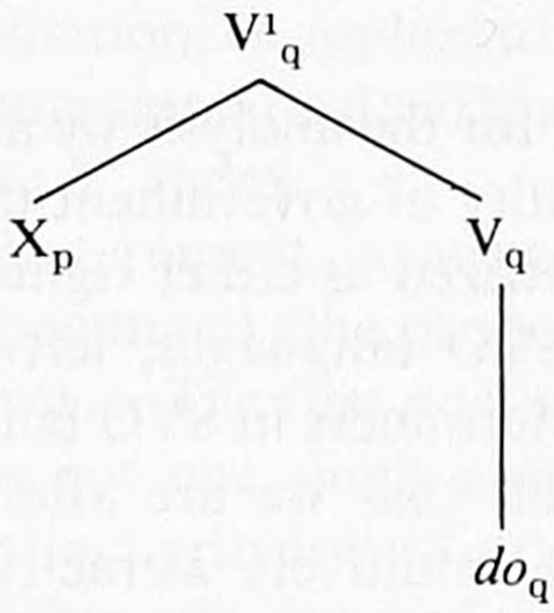
(24)



Precisely the same analysis could be made for nominal clitics, of course. In those cases the head noun would be the governor. Alternatively, we could claim that cliticization is phonological and that clitics form part of the same lexical entry as the verb in (24). In this case the co-indexing would be an automatic consequence of lexical insertion.

The notion of L_q carrier gives an interesting result for cases involving alien verbs. Here we often find a verbal complex consisting of a native L_q head, the equivalent of the dummy verb *do*, while the alien verb is either in a nominal or an infinitive form. Together they form a small V^1, as in (25).[6]

(25)



It is not always clear, however, whether configurations such as (25) are a product of borrowing or of code-mixing.[7]

The discussion in this section so far has been focused on configurations in which code-mixing would be excluded. Where, then, do we predict mixing to be possible? In the following list, possible mixing will be indicated through the use of the subscripts *p* and *q*. Mixing contexts include:[8]

(26) (a) $NP_q \quad VP_p$
(b) $AUX_q \quad VP_p$
(c) $V_q \quad DET_q \quad N_p$
(d) $P_q \quad DET_q \quad N_p$
(e) $NP_q \quad Copula_q \quad AP_p$
(f) $V_q \quad QP_q \quad A_p$
(g) $V_q \quad COMP_q \quad S_p$
(h) $S'_q \quad CONJ_p \quad S'_p$

---

[6] Cf. Wehrli (1981), where it is argued that *faire V* forms a small V^1 in the unmarked case in French.

[7] The evidence we have from Hindi–English code-mixing (cf. 3.2 below) suggests, however, that what is involved is borrowing.

[8] Of course, this list could easily be extended. Scheme (22) includes just some of the more common switching sites.

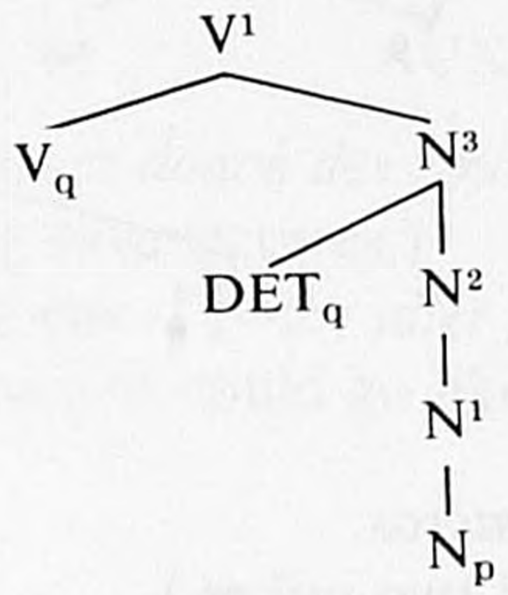
3. Case studies

In the following, we shall see that this theory generally makes the right predictions for the French–Italian–English and Hindi–English code mixes.

3.1 *French–Italian–English*

The data in (27)–(37) show that the government constraint makes the right predictions for the French–Italian–English case.[9] In (27) for instance, mixing occurs between a DET and a N. According to our theory the DET is the L_q carrier and should agree in language index with the governor of the NP, and in fact the DET carries the same language index as its governing V. The examples (28) instantiate structure (27).

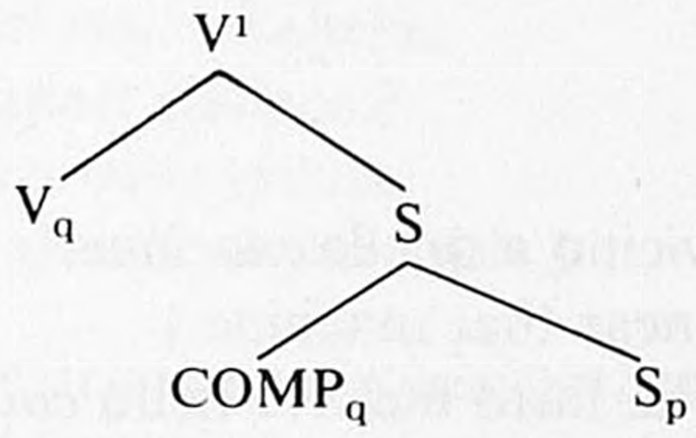
(27)



(28) (a) Ha portato un *cadeau*.
((He) brought a present.)
(b) Ha ricevuto il *diplôme*.
((She) received the diploma.)
(c) Io posso fare i *chèques*.
(I can do cheques.)

In (29) mixing occurs between COMP and S. This case follows from our theory since the COMP is the L_q carrier and shares the language index of the governing V, as in the following structure:

(29)

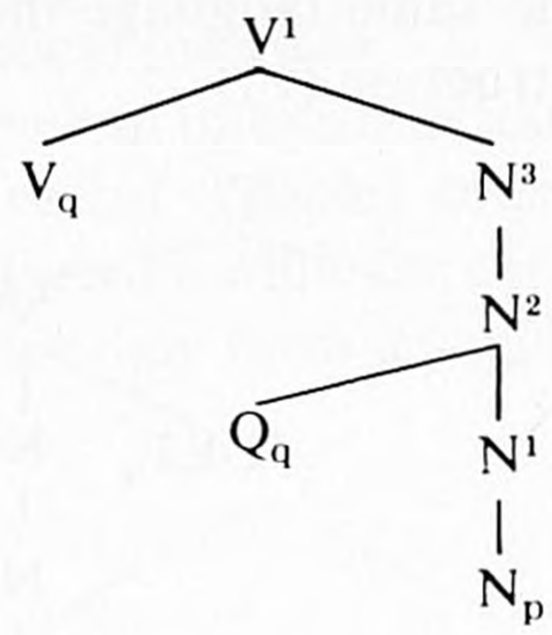


[9] The phrases are taken from the corpus of the research programme on multilingual (Italian–French–English) interaction (cf. di Sciullo *et al.* 1975), collected with the aid of the Social Sciences and Humanities Research Council of Canada between 1973 and 1974 in Montreal, by the second author of this paper. The following interviews were used for this study: AMR 1.1.1, AMR 3.1.2, AMR 3.6.2, AMR 6.3.1, AMS 1.1.1, AMS 3.6.1, AMS 4.7.1, AMS 4.7.2, AMS 7.8.2. We thank Henrietta Cedergren and Paul Pupier for their permission to refer to the corpus.

(30) (a) Basta che *marche*.
((It) suffices that (it) works.)
(b) E l'altro dice come *s'appelle?*
(And the other one says how is it called?)
(c) Dice quando *paye*, *all right*!
((He) says when (it) pays, all right!)

In (31), mixing occurs between Q and N. Again these cases are allowed in our analysis: the Q is the L_q carrier and agrees in language index with the governing V, as in the following structure:

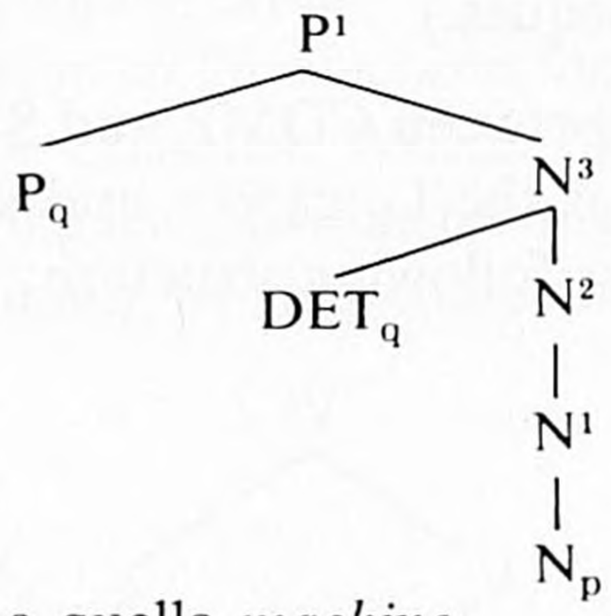
(31)



(32) (a) Portava due *micros*.
((She) brought two mikes.)
(b) Metteva tanto *maquillage* sulla faccia.
((She) put a lot of make-up on her face.)

When the governor is a P, the theory still holds. In (33), the DET is the L_q carrier and has the same language index as the governing P.

(33)



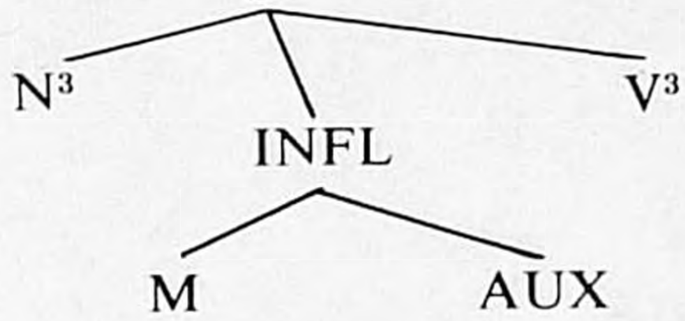
(34) (a) Sempre vicino a quella *machine*.
(Always near that machine.)
(b) Allora, l'ha fatto mettere nello *coin*.
(So, (he) forced him to go in the corner.)
(c) A finire dentro a un *bureau*...
(To end in an office...)

If constraint (3) predicts that no mixing occurs when government is involved, it does not predict that mixing must not occur when government is absent. Mixing may occur between the subject and the VP as in (35), but it is not necessarily the case; in (28c) above and in (30b) the subject and the V have the same language index.

(35) *La plupart des canadiens* scrivono 'c'
(Most Canadians write 'c'.)

The cases of (37) are also allowed in our theory if we assume that there is an INFL node in Italian dominating modals and auxiliaries, and that it is dominated by S; modals and tense auxiliaries do not govern the V, which can have a different language index. In (38), we have to assume that no government relation holds between the copula and the predicate adjective.[10]

(36)



(37) (a) No, *parce que* hanno *donné des cours*.
(no, because they gave lectures.)
(b) *Oui, alors j'ai dit que* si potev *aller comme ça*.
(Yes, so I said that we could go like that.)

(38) Perché è *mauvais*.
(Because it is bad.)

Contrasting with the large number of switches falling within the constraint (3), there is a small number of counterexamples in the corpus. In (39) and (40) the governed category does not share the L_q index of the governor.

(39) (a) Non voglio *smokemeat*. [smok mit]
((I) don't want smokemeat.)
(b) Ma c'era una ragazza come *gerante*. [žerante]
(But there was a girl as manager.)
(c) Perché hanno fatto una via *express* sotto. [ɛkspres]
(Because they made an expressway under.)

(40) (a) Ma ci stanno dei *smart* italiani.
(But there are smart Italians.)
(b) La lascia *toujours* sulla tavola.
((She) leaves it always on the table.)

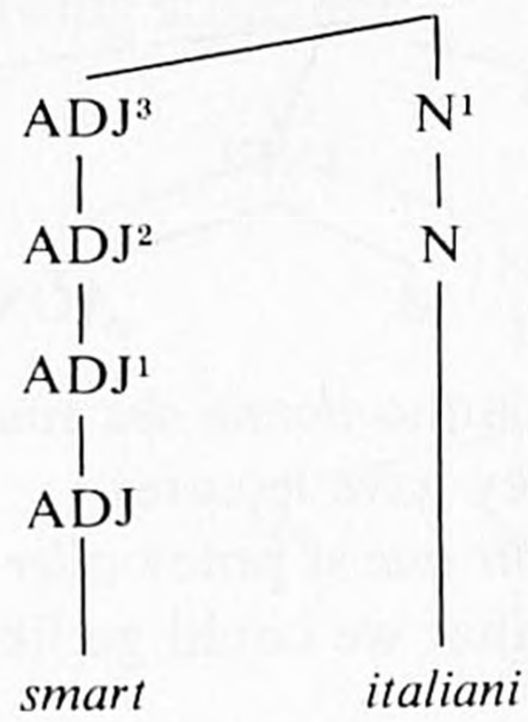
It may be the case that in (39) it is not code-mixing that is involved, but mainly borrowing. This is suggested by the phonological adaptation of these lexical items that could give them the status of nativized items. Furthermore, these items are recurrent in the speech of Italian immigrants in Montreal and frequent in the corpus. Their frequency also suggests that borrowing is involved and not mixing.

---

[10] Another possibility would be to analyse the copula as a non-governing V, in which case our analysis would hold as well. We shall not discuss this hypothesis here.

The status of examples (40) is more problematic, and depends on the analysis of ADJ and ADV. If the ADJ *smart* and all other restrictive modifiers branch from N^2, and the ADV *toujours* with other non-restrictive modifiers branches from S,[11] (40) are not counterexamples to our theory. There is no government relation between the heads and the complements in structure (41), according to our definition of c-command. If ADV and ADJ branch from X^1 types of categories, as suggested in Di Sciullo (1981), (40) are counter-examples to our theory.

(41)



It appears then that when applied to the French–Italian–English cases of code-mixing, this theory gives interesting results.

3.2 *Hindi–English*

Hindi–English code-mixing in urban North India (cf. Singh, 1981), contains far more switch sites than the word-order-sensitive linearity constraint of Sankoff and Poplack would lead us to believe, as the basic word-order of Hindi is SOVAux (cf. McGregor 1977 and Kachru, 1980, and for some further refinements, Di Sciullo, 1981). Its interest, however, is more than just negative: most of the particularistic constraints proposed for it – by Gumperz, Kachru and Singh – can be seen to follow from the L_q-government constraint we propose in this paper.[12]

Some of the constraints on Hindi–English mixing are pretty straightforward. The sentences in (42) bear out the predictions made by the analysis summarized

---

[11] As in Jackendoff (1977).

[12] The data for Hindi–English code-mixing is mostly judgmental in nature since its speakers, at least the ones used as informants for this study, are in general quite capable of judging the grammaticality of a string. Thus Joshi (1981), in his study of Marathi–English code-mixing, notes that 'participants seem to have fairly consistent judgments about the acceptability of code-mixed utterances' and that 'judgments about the "acceptability" seem to be invariant with respect to the amount of code-mixing a given participant does". The judgments reported here are quite compatible with published non-judgmental data (cf. Verma, 1976). Although we do not discuss the Marathi–English data that form the subject matter of Joshi's paper, we should like to point out that the facts discussed by him are easily accounted for by the constraint proposed here.

in (26). In (42) the conjunction must agree in its L_q index with the L_q index of the second conjunct. Since both conjuncts are in Hindi, *and* and *but* yield ungrammatical results:[13]

(42) (a) pradhān mantrī aur rājā sāhib āye the
and
prime minister king sir come (perf. part.) aux
The Prime Minister and the king had come.

(b) maĩ us ko akhbār de to detā magar diyā nahĩ
but
I him newspaper give (cond.) gave not
I could have given him the newspaper but didn't.

In (43) the strings with the complementizer not bearing the language index of the governing V are ungrammatical, as predicted by the analysis provided in (26):

(43) (a) I told him that *rām bahut bimār hai*
*ki
Ram very sick aux
I told him that Ram was very sick.

(b) mujhe lagtā hai ki *Ram will come tomorrow*
*that
me seem aux
I feel that Ram will come tomorrow.

What Singh refers to as the specifier constraint also follows from the L_q-government principle. The grammaticality of the strings in (44) depends crucially on whether the appropriate specifier bears the same language index as the verb:

(44) (a) vahā̃ pā̃c sundar laṛkiyā̃ paṛh rahī haĩ
*five
there beautiful girls read prog aux
Five beautiful girls are reading there.

(b) rām ājkal kuch *hard drinks* pine lagā hai
*some
Ram these days drink begin (perf. part.) aux
These days, Ram has begun to drink hard stuff.

(c) maĩ ne us kā *dissertation* paṛhā
*his
I read
I read his dissertation.

---

[13] When two items appear in the same position, they are to be interpreted as equivalent. Thus an asterisk effectively indicates the language from which the item cannot be drawn.

(d) ham apnī *laboratory* becẽge
our
We sell (fut.)
We will sell our laboratory.

In prepositional phrases, the sorts of contrast exemplified by (45) (a) and (b) below are also predicted by the principles discussed earlier. (45b) is grammatical because the L_q index of *kuch*, Hindi, agrees with the L_q index of *se*:

(45) (a) dam se
force
with
With force.
(b) kuch *force* se
some with
With some force.

A slightly more interesting confirmation of the theory developed here is provided by cases where the English verb is nativized (though not in the phonological sense) by the addition of an inflected form of the Hindi dummy verb *karnā* ('to do'). The nominal or infinitive form of the English verb forms a small V^1 with the native head *kar* precisely as in (25). Consider the string in (46):[14]

(46) maĩ yah *prove* kar saktā hū
*∅
I this do can aux
I can prove this.

Manner adverbs like *quickly* and *reluctantly* don't, as we would predict, and as the sentences in (47) show, mix well with Hindi verbs:

(47) (a) *milan *reluctantly* gayī
went
Milan went reluctantly.
(b) *puṣpā *quickly* bāt kartī hai
word do aux
Pushpa talks quickly.

They do, however, mix extremely well with Hindi verbs when they are accompanied by qualifying particles such as *zarā* 'a little' and *bahut* 'a lot, very'. These particles save the switch because they bear the same language index as the verbs. The sentences in (48) are therefore fully grammatical in Hindi–English mixes.

---

[14] Notice that the explanation offered here also throws some light on one of the most oft-noted facts of Hindi: the fact that the Hindi structures with *karana* and *hona* (*to be*) are typically made up of a word of Persian or Arabic origin plus the dummy verb.

(48) (a) milan zarā *reluctantly* gayi
a little went
Milan went somewhat reluctantly.
(b) puṣpā bahut *quickly* bāt kartī hai
word do aux
Pushpa talks very quickly.

Other adverbs, however, don't behave quite as straightforwardly as manner adverbs. They actually divide themselves into two classes: those that occur rather freely with Hindi verbs (*unfortunately*, *surprisingly*, *frankly*, etc.) and those that don't (*yesterday*, *tomorrow*, etc). There is, as the sentences in (49) show, never any problem with adverbs of the first type:

(49) (a) *unfortunately*, rām kal nahī̃ āyā
yesterday not came
Unfortunately, Ram did not come yesterday.
(b) *frankly*, rām bahut bevaqūf hai
very stupid copula
Frankly, Ram is very stupid.

Adverbs like *yesterday* and *tomorrow*, on the other hand, just don't mix with Hindi verbs irrespective of their position, as the ungrammaticality of the strings in (50) shows:[15]

(50) (a) *mujhe sudeś se *tomorrow* milnā hai
me with meet aux
I have to meet Sudesh tomorrow.
(b) *mujhe *tomorrow* sudeś se milnā hai
me with meet aux

If we want to account for these facts in (47)–(50) in terms of our theory of government, we must claim that manner adverbs are governed, but can be modified by a Hindi particle that serves as L_q-index carrier, sentential adverbs are ungoverned, and time adverbs are governed, without being able to be modified by a Hindi particle. This corresponds reasonably well to standard assumptions about adverbs.

The main problem for our analysis is that of subjects in Hindi. The theory of government by lexical categories (but excluding INFL) that we adopt predicts that there cannot be a switch between verb and object (and this is borne out by the code-mixing data as well; see (44)), but that there can be a switch between the subject and the verb phrase. The examples in (51) are ungrammatical, however:

(51) (a) **the new mayor* kal dillī jāyegā
tomorrow Delhi go (fut.)

---

[15] The adverb problem discussed here is not peculiar to Hinglish. Lexical differences amongst adverbs of the 'same class' make it somewhat difficult to provide an unambiguous characterization of their geometry, even in English.

The new mayor will go to Delhi tomorrow.

(b) **his uncle* jāne-wāle hāi
go (agentive) copula
His uncle is about to leave.

(c) **some diplomats* jā rahe haĩ
go prog aux
Some diplomats are leaving.

The subject, in other words, behaves as if it were governed. The only time it does not is when it is a bare nominal, as in the sentences (52):

(52) (a) [breḍ] ne nāś kar diyā
erg. ruin do give
The bread ruined it.

(b) [kophi] ne kamāl kar diyā
erg. miracle do give
The coffee did wonders.

(c) [ṭren] ~ [ṭiren] calī gayī
move went
The train left.

The fact that our informants when asked to pronounce the sentences above used distinctly Hindi phonology, as evidenced by retroflexion, epenthetic vowel insertion, and substitution of /o/ for /ɔ/, suggests that what is involved is borrowing, and that an alternative explanation must be found for the fact that the subject in Hindi–English mixing behaves as if it were lexically governed.

At present we have no definite explanation. One possibility would be to assume that Hindi is non-configurational and that there is no syntactic VP. In that case, there would not be any difference in this respect between subjects and objects: both would be governed. Another possible avenue of research would be to assume that the restriction is due to the fact that Hindi is an ergative language, in the perfective aspect. This is less attractive because the ergative particle *ne* exempts only bare nominals and not, as (53) shows, full NPs, unless of course they contain the appropriate L_q carrier to assign them the desired L_q index:

(53) tumhāre *coffee-cake* ne beimān banā diyā
**your*
erg. dishonest make aux
Your coffee-cake made me a dishonest man.

Whatever the correct explanation for the 'governed' nature of the subject the observed L_q dependency does not violate the general principle proposed in this paper. It merely shows that specific grammars may impose additional constraints. In the case under consideration, Hindi imposes a language-specific constraint that has the effect of making the subject governed in the relevant

sense. Such language-particular constraints, however, do not override the general constraint. They serve to complement it and not to violate it, at least in the Hindi–English case. Consider, for example, the constraint, discussed in some detail in Singh (1981), that within the NP the DET and the QP must bear the same language index. It does not violate the general constraint. Language-particular constraints, in other words, merely add particularistic prohibitions. They do not seem to necessitate the suspension of the general principle.[16]

4. Conclusions

In the previous sections we have shown that the constraint of government as formulated in (3) and (6) gives a reasonable set of predictions with respect to possible and impossible mixes. So far, however, we have not posed two crucial questions with respect to our analysis:

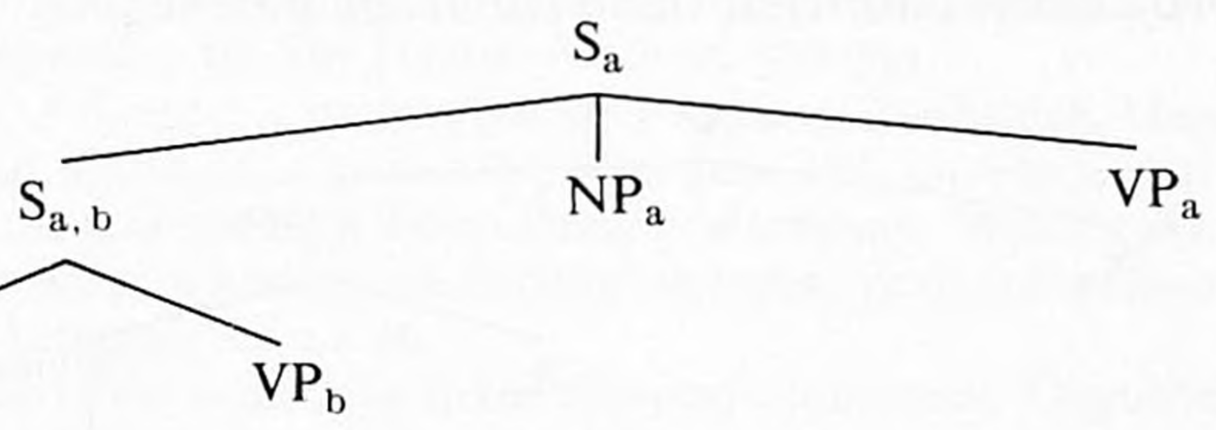
(54) (a) Why the particular definition of government in terms of immediate c-command and why the definition of L_q carrier?
(b) What kind of restrictions are there on the mixing of phrase structure rules?

We will see that an answer to (b) will lead to an answer to (a).

Our discussion in Sections 1, 2 and 3 has been formulated in terms of lexicon rather than of phrase structure. Lexicalized terminal nodes were assumed to have L_q indexes, an obvious result of the process of lexical insertion. What about non-terminal nodes? While the lexicon is most visible, it is not possible to avoid referring to the abstraction of phrase structure. We will argue that the conception of phrase structure nodes having language indexes as well will bring us closer to attaining explanatory adequacy.

Consider a case where an adverbial clause in language *b* is subordinate to a clause in language *a*. Now the rule producing the subordinate clause will correspond to the grammar of *b*, the rules of the matrix to that of *a*:

(55)

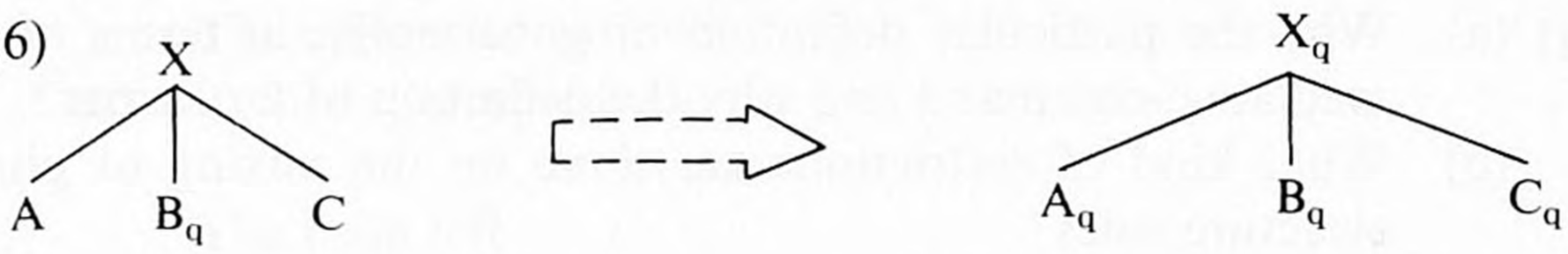


[16] There are other language-specific problems and constraints that need to be worked out. Our preliminary investigation indicates, for example, that within the VP the V and the Aux must bear the same language index. Precisely what status should be assigned to Aux in Hindi is not clear. It should be clear, however, from the examples cited in this section (3.2) that the aux and the verb must bear the same language index. There is also some problem with adjectives. We expect adjectives to be L_q-governed by the N. The problem is that almost any English adjective can be used with a Hindi head if the particle *-wala* is added to it. The problem with *-wala* is that it does not always 'nativize' the adjective phonologically. Some informants also accept adjectives that end in *-ful* in these constructions.

One node, $S_{a,b}$, could be called the L_q neutralization site: it carries two indexes. The question of constraints on code-mixing now becomes: what are the possible L_q neutralization sites? A first requirement would be, clearly, that the node is characterizable in the phrase structure of both languages involved. S in (55) clearly fulfils that requirement. We may think that the theory of grammatical categories, part of X-bar theory, provides us precisely with the list of universally defined categories. It is always possible, however, that some language-specific category is also shared by two languages.Then these could also be neutralization sites.

Suppose we make the second requirement that a neutralization site may have no lexical sister. This suggestion is plausible since it corresponds to the fundamental observation of immediate government. For the purposes of our discussion government has been something like having a lexical sister. Why then this second requirement? Because a lexical sister unambiguously indexes the node dominating it and its constituents:

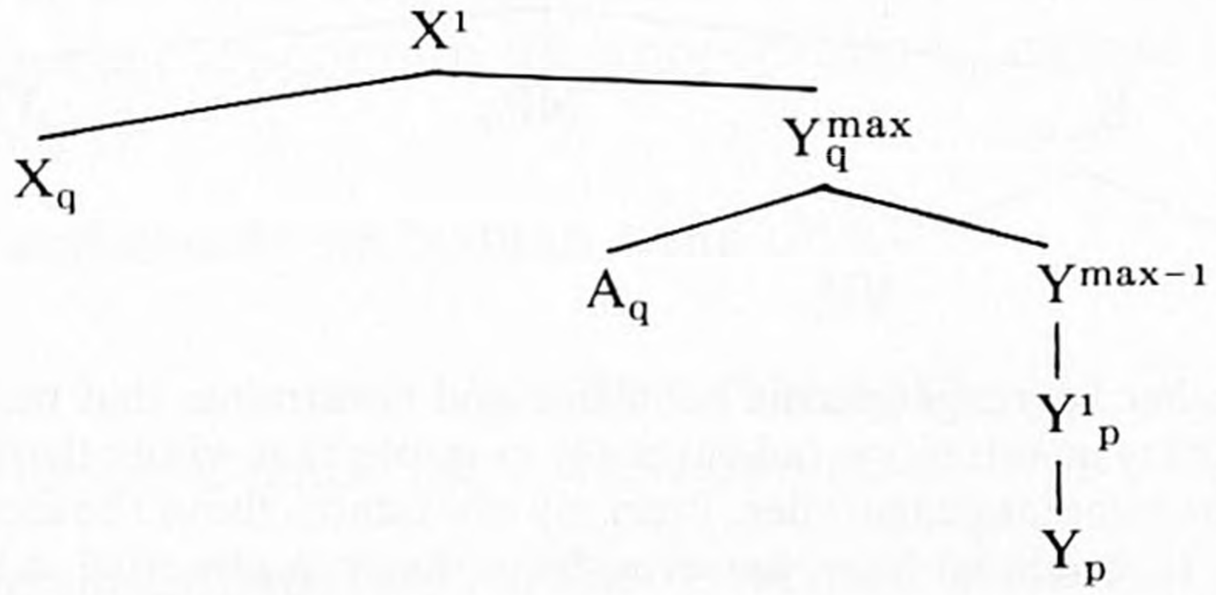
(56)



Only those nodes which are not unambiguously indexed (having no lexical sister) can be neutralization sites.

The fact that the notion of government needed to constrain code-mixing involves immediate c-command thus follows from the logic of our indexing procedure.Unambiguous L_q indexing involves nodes which have lexical items as daughters or sisters. Similarly, the notion of L_q carrier is imposed by the indexing procedure, and hence need not be defined separately. Governed maximal projections are unambiguously indexed by their lexical sister. If there is switching internal to the projection, at least one element in it, and in fact at least the 'highest' lexical element in it (which we have previously defined as the L_q carrier) must be co-indexed with maximal projection node. Lower nodes in the projection can then have different indexes.

(57)



We must ask ourselves whether A_q in (57), commonly an element like a determiner or a complementizer, counts as a lexical sister, or whether lexical sisters must be governors (N, A, V, P) as in the earlier definition. The earlier

definition is probably the more desirable one, as can be seen from a consideration of (57). Suppose that the 'highest' lexical element in (57) is in fact a sister of Y^1. Since Y^1 immediately dominates Y_p, it carries the unambiguous index *p*. Suppose then that A_q unambiguously assigns its index *q* to its sister Y^1, and we have a clash. Therefore we will adopt the following convention:

(58) (a) A major category assigns an index both to the node dominating it and to its sisters;
(b) A minor category assigns an index only to the node dominating it, not to its sisters.

A moment's reflection tells us that this convention is entirely reasonable. Minor categories are not lexical governors but they will have the same index as their dominating node since their definition is always language-specific.

In this way we have answered the questions regarding phrase structure and regarding the definitions of government and L_q heads posed at the beginning of this section. We hope to have shown that the general restrictions on code-mixing need not be stated anywhere, but arise from general conventions on language indexing.

REFERENCES

Aoun, Y. & Sportiche, D. (1983). On the formal theory of government. *The Linguistic Review* **2.3**. 211–236.

Bouchard, D. (1984). *On the content of empty categories*. Dordrecht: Foris.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Di Sciullo, A. M., Van Amerongen, A., Cedergren, H. & Pupier, P. (1975). Etude de l'interaction verbale chez les montréalais d'origine italienne. In *Cahiers de Linguistique de l'Université du Québec*, V.

Di Sciullo, A. M. (1981). L'hypothèse étoile. Mimeo, Université de Montréal.

Ferguson, C. (1965). Diglossia. *Word* **15**. 325–40.

Fishman, J. A. (1965). Who speaks what language to whom and when. *La Linguistique* **2**. 67–88.

Gumperz, J. J. (1976). The sociolinguistic significance of conversational code-switching. *University of California Working Papers*, **46**. University of California Berkeley.

Haugen, E. (1973). Bilingualism, language contact, and immigrant languages. In T. Sebeok, (ed.), *Current trends in linguistics*, **10**. The Hague: Mouton, 505–591.

Jackendoff, R. (1977). *$\bar{X}$ Syntax: a study of phrase structure*. Cambridge, Mass.: M.I.T. Press.

Joshi, A. (1981). Some problems in processing sentences with intrasentential code-switching. Paper presented at the University of Texas Parsing Workshop, March 1981.

Kachru, B. (1978). Toward structuring code-mixing: an Indian perspective. *International Journal of the Sociology of Language* **16**. 28–46.

Kachru, B. (1980). Role of the vernacular in the bilingual's repertoire. Unpublished manuscript. Urbana: University of Illinois.

Kachru, B. (1983). *The Indianization of English; the English language in India*. Delhi: Oxford University Press.

Kachru, Y. (1980). *Aspects of Hindi grammar*. Delhi: Yamona.

Klavans, J. L. (1983). The syntax of code-switching: Spanish and English. To appear in *Proceedings of the linguistic colloquium on Romance languages* **14**.

Koopman, H. (1984). *The syntax of verbs*. Dordrecht: Foris.

Lance, D. M. (1975). Spanish–English code-switching. In Hernandez-Chavez, E., Cohen, A. *et al.*, *El lenguaje de los Chicanos*. Washington: Center for Applied Linguistics. 138–153.

Lightfoot, D. (1981). The history of Noun Phrase movement. In Baker, C. L. & McCarthy, J. (eds.), *The logical problem of language acquisition*, Cambridge, Mass.: M.I.T. Press. 86–119.
McGregor, R. S. (1977). *Outline of Hindi grammar*. Delhi: Oxford University Press.
Pfaff, C. (1976). Code-switching and syntactic change. In Steever, S. B., Walker C. A. & Mufwene, S. S. (eds.), *Papers from the Parasession on Diachronic Syntax*, Chicago Linguistic Society, Chicago. 248–259.
Pfaff, C. (1979). Constraints on language mixing. *Lg* **55**. 291–318.
Poplack, S. (1980). Sometimes I'll start a sentence in English y termino en español. *Linguistics* **18: 7–8**. 581–618.
Riemsdijk, H. C. van, (1978). *A case study in syntactic markedness*. Dordrecht: Foris.
Sankoff, D. & Poplack, S. (1981). A formal grammar for code-switching. *Papers in Linguistics* **14**, 3.46.
Schaffer, D. (1975). Syntactic consideration in distinguishing code-switching from borrowing. Paper read at the Queen's College Conference on Language Learning, June 1975.
Schaffer, D. (1978). The place of code-switching in linguistic contacts. In Paradis, M. (ed.), *Aspects of bilingualism*. Columbia, South Carolina: Hornbeam Press. 265–274.
Singh, R. (1981). Grammatical constraints on code-mixing. *Recherches Linguistiques à Montréal* **17**. 155–163.
Stowell, T. A. (1981). The origins of phrase structure. PhD dissertation, MIT.
Timm, L. A. (1975). Spanish–English code-mixing: el porque y how-not-to. *Romph* **10;4**. 473–482.
Verma, S. K. (1976). Code-switching: Hindi–English. *Lingua* 38. 153–165.
Wehrli, E. (1981). Restrictions sur la cliticisation dans les constructions causatives. Mimeo: MIT.
Wentz, J. (1977). Monolingual codes; some remarks on the similarities of bilingual and monolingual code-switching. *CLS* **13**. 706–713.
Woolford, E. (1983). Bilingual code-switching and syntactic theory. *LIn* **14**. 520–536.