Bakker, Peter and Maarten Mous (eds.) 1994. Mixed languages. 15 case studies in language intertwining. Amsterdam: Institute for Functional Research into Language and Language Use.

Singh, Rajendra. 1995. Linguistic theory, language contact, and modern Hindustani. The three sides of a linguistic story. New York: Peter Lang.

Treffers-Daller, Jeanine. 1994. Mixing two languages. French-Dutch contact in a comparative perspective. Berlin: Mouton de Gruyter.

Carol Myers-Scotton

No doubt the rationale of commissioning this review article to cover these three volumes is that they all refer to language contact phenomena. After a nod in that direction, however, the authors all go their own ways in both subject matter and approach. This is especially the case in reference to Singh's study: it *mentions* the term "language contact" often, but little detail on the subject is found therein. The same can be said for the degree to which Singh considers modern Hindustani, also an element in the volume's title. In this review, I will first consider Singh's contribution and then go on to deal in more detail with the language contact phenomena and their analyses as presented in the Treffers-Daller and Bakker & Mous volumes.

Singh, in a very repetitive and too-often magisterial fashion, devotes most of his space to arguing for the centrality of Universal Grammar (UG) in accounting for the structures found in any language. He argues just as strenuously against the value of functionalist or variationist accounts of language use patterns.

Singh seems to equate taking account of UG with espousing formalistic theories of language, i.e., generative theories. Also, he makes it clear he has in mind Labov (1980) and his associates (e.g. Poplack 1990) when he refers to variationist approaches. While most readers would agree with these characterizations of formalism and variationism, most would find Singh's view of functionalism to be rather idiosyncratic. He ignores entirely the Hallidayian school of functionalism and also that of functionalism in discourse analysis (e.g., "The Santa Barbara school"), and while he refers to Dressler (1985) with appro-

Rivista di Linguistica, 11.2 (1999), p. 367-386 *(ricevuto nel febbraio 1997)*

val, it is clear he does not include him when he uses the term *functionalism*. Instead, he seems to have in mind only those researchers who argue that social details of the contact situation play a role in determining patterns of language use. For example, he often mentions Kachru, as in the following: "most studies of these contacts assume some version of what might be called the non-autonomy thesis, defended most vigorously by B. Kachru (1980), and do not give due importance to those aspects of contact that follow from factors inherent in language design". (p.11) While Singh finds faults with all three approaches (formalism, functionalism, and variationism) he is especially scornful of functionalism and variationism. Yet, he rarely offers specific reasons. The following is a typical reproach: "It is worth emphasizing that sociolinguistic or socio-psychological investigations of loan-phonology could not even ask some of the questions that I have tried to answer, and that if they were to ask them, no known theory of language in contact could possibly provide an explanation for them" (p. 28). His only specific criticism of his view of functionalism is in reference to Pandharipande (1981), who suggests that the extent to which loan words are nativized in Marathi depends on speaker attitudes. Even here, however, his argumentation lacks detail.

He rejects out of hand quantified approaches to explaining linguistic structure ("the structure of language is not amenable to statistics" p. 117) and the view that variability is a feature of language ('...phonologies have only categorical rules. They appear variable only when one does not understand the nature of higher level prosodic principles that regulate this apparent variability" (p. 118). Pace Singh, I would argue that quantified performance data offer the best test of hypotheses about language-design when one considers the acceptability of innovating structures in language contact. In such cases, the grammaticality judgments so favored by formalist linguists often give conflicting results.

It is often difficult to follow Singh's discussion, but what follows are, as far as I understand his goals, the main ideas presented in this volume, or at least the ones I see as most worth considering. They all relate to what is perhaps Singh's major point: "...whereas generative grammar insists on too rich a hard-wiring, functionalism suggests too little...". However, this idea is little developed, although often repeated.

1) Generative linguists would do well to consider external evidence for testing models of grammar. Along this line, Singh states, "...an adequate theory of phonology automatically accounts not only

for system-internal facts but also for certain facts generally considered to be external" (p.76-77). Unfortunately, he does not go into details as to those "certain facts".

2) Functionalists' accounts of contact dwell on *language* when the relevant subject matter is *grammar*. That is, functionalists "err in ignoring the power of grammar (as a mental entity), which can provide significant explanations of crucial aspects of contact..." (p. 131).

3) Phonology is independent of morphology in crucial ways, although these are hardly spelled out. In regard to interference from L1 in a learner's variety of L2, Singh writes, "Interference, in other words, is caused only by across-the-board phonological rules of L1. So called morphophonemic rules of L1 do not cause it, and morphological interference from L1 seems not to exist, as word-formation errors in intermorphology are the results of illegal extensions of L2 word-formation rules" (p. 81). This is the most tantalizing insight in the volume. Related to this, Singh earlier states, "The learner, in other words, can only make a mistake if he has 'acquired' a pattern in the L2 to license it" (p. 77). That is, the source of an error is analogy rather than a rule-system, and Singh suggests that morphology may be different from phonology and syntax in its reliance on analogy (p. 78).

As far as language contact is concerned, Singh is being properly modest when he writes in the Prologue (p. 8) "...this book is an attempt to deal with a small part of the linguistic competence of the urban, educated MH [Modern Hindustani] speaker". Yet, the volume does live up to the main kernel in the above quotation: It is language *competence* which interests Singh. Thus, the volume includes nothing on the Indian contact situation or even the range of structural phenomena in Hindustani which result from that situation. In Chapter 1, the few Hindustani examples he does cite come from contact with English resulting in lexical borrowing and what he calls code-mixing (=intrasentential codeswitching). In other chapters, he cites more examples from other languages, not all arising in language contact; again, his goal is to discuss theory, not data, and not necessarily theory especially relevant to language contact phenomena.

One of Singh's most detailed discussions is on code-mixing, a subject of interest to students of language contact, of course. Put simply, the argument is that the notions of language index and government are basic: "If X has language index q and if it governs Y, y must have language index q also" (p. 41). Unfortunately, the discussion is old (Singh 1985 and DiSciullo, Muysken and Singh 1986) and, more important, it has been shown that its claims are not supported by

naturally-occurring data in other data sets. Myers-Scotton ([1993]1997: 43–44) cites several counter-examples from diverse sources, such as this one from Swahili/English in which the verb, *ulikuwa ukiongea,* governs an NP complement which is switched: *Ulikuwa ukiongea A LOT OF NONSENSE* "You were talking a lot of nonsense". Muysken (1995: 187), one of Singh's co-authors in DiSciullo et al. (1996), acknowledges that the government constraint is too strong, pointing out that "counterexamples abound". Specifically, he mentions that in the Moroccan Dutch/Arabic data set of Nortier (1990) verbal and prepositional object NPs are often in a different language from their governing verb or preposition.

Further, even though this discussion takes up a good part of chapter one (pp. 29–53), later, Singh states that "[one] is forced to conclude that although it is true that statistical principles of the sort proposed by Sankoff and Poplack (1981) do not provide a satisfactory account of the facts of the matter, the government account defended rather vigorously in chapter 1 does not either" (p. 103). To illustrate the problem with the government constraint, he cites data on Spanish/English codeswitching but leaves to the reader to figure out what there is about the data which invalidates his earlier analysis.

The entire volume could have used more careful argumentation and editing. For example, Singh's comment that interlanguage is a good source of evidence for general theory-building appears to be a useful insight. But his development of this statement is very opaque: "The study of interlanguage provides almost laboratory conditions for separating (i) a more (distributional) generalization from a linguistically significant one, (ii) a dead rule from a live one, and (iii) an unrealized potential from an unrealizable one" (p. 125). The codemixing analysis is an example of poor judgment on the part of the author and on the part of the editor, too: why ask the reader to work through an analysis which is then so cavalierly discarded? Editing, in general, is a problem. Very frequently, Singh repeats himself, word for word, a few pages later (e.g., compare p. 104 and 125). Finally, the index is woefully inadequate as only a "list of names", not content; even this list has its problems – Kachru, for example, is entered under "Kachru" and then separately under "Kachru, B." However, there is no index at all in the Bakker and Mous volume; and that in Treffers-Daller for "subjects" is rather brief.

In his lack of actual contact data and his few efforts to explicate them or detail his arguments, Singh contrasts sharply with the authors of the other two volumes whose primary focus is on data (Bakker and Mous) or both data and analysis with some preliminary

attempts at theory-building (Treffers-Daller). The volume by Treffers-Daller (T-D) treats codeswitching (CS) and borrowing in Brussels, where at least part of the population speaks both French and Brussels Dutch (BD) and at times switches between them in the same conversation. A revised version of T-D's doctoral dissertation at the University of Amsterdam, the volume retains all the markings of an excellent dissertation, but, unfortunately, also some of the shortcomings. It is obviously a product of much time and conscientious effort, dealing with both the social and the structural sides of her topic as well as delving into each topic at surprising depth. For example, she is not content to report simply on gender assignment of French nouns in BD, but discusses such assignment in Standard Dutch and Southern Dutch dialects. In addition, she briefly surveys the literature on gender assignment, especially of French nouns, in other languages. Finally, she uses evidence from both spontaneous data and a translation experiment which tested gender assignment. Both sources of data show that French nouns tend to keep the gender they have in French.

The volume's major shortcoming is that T-D does not approach her data with any unified theoretical framework to test and therefore the concluding section includes only a few rather tentative assertions, not all by any means connected to her analyses in the earlier chapters. True, she does test a few of the main claims about constraints on CS which were current in 1990 (about the time when she must have written her dissertation), but how the test results inform her analyses is not evident.

While T-D's argumentation is generally clear, it is confusing regarding the distinction between CS forms and borrowings. T-D clearly states at several points that the distinction between borrowing and intrasentential codeswitching is problematic and makes a good argument to support her claim ("[codemixing and borrowing] become very similar from a theoretical point of view. Both codemixing and borrowing may be considered, in the first place, as cases of the interaction of two lexicons" (p. 90)). As if to demonstrate similarities between codemixes and borrowings in other ways, she applies earlier ideas about a hierarchy of borrowability of lexical items (from Haugen 1953 and Muysken 1981) to come up with a hierarchy of switchability in reference to full constituents. This will be discussed further below. However, whatever similarities she finds between single words and full constituents from Language A in a grammatical frame from Language B, she makes clear that she intends to treat single words and full constituents separately. As a rationale, she first

states the obvious fact that the morphological strategies of integration which apply to single words do not apply to entire constituents, but does not consider the logical impossibility of applying such strategies (from Language B) to full constituents (from Language A inserted in Language B). Second, she states that syntactic constituents can occupy argument slots, "...whereas single words do not generally occupy argument slots, except for single nouns" (p. 93). It is the "except for single nouns" which is the kicker, since "single nouns" are the prototypical receivers of thematic roles and therefore T-D's distinction between full constituents and single words on the grounds of thematic role assignment becomes very hollow. (For an argument that single "guest" words, whether they are established borrowings or spontaneously used forms, should and can be treated similarly under a model of constraints on CS, see chapter 6 in Myers-Scotton [1993]; 1997). My claim is that both B forms and single CS forms show the same type of morphosyntactic integration into the "host" language and therefore – at least in a synchronic consideration – can be covered by the same morphosyntactically-based constraints).

While T-D herself does not propose an unified treatment of single "guest" words in mixed utterances, she does at least recognize that the issue cannot be ignored, even though earlier formulations of constraints, such as Poplack (1980) and DiSciullo, Muysken, and Singh (1986), seem to have done this. In this regard, T-D says, "...an easy way to invalidate potential counterexamples... is to put [sic] that these constitute cases of borrowing and not of codemixing. But this does not solve the problem, because borrowing is a rule-governed process as well. The question why a particular element from language A (considered to be a borrowing) can be placed so easily in a particular slot remains unanswered. The problem is shifted towards the formulation of constraints on borrowing" (p. 90).

Altogether, T-D does do a fine job of presenting and analyzing her data. Many of the results, such as that nouns represent 58% of French single words in BD frames or that French single forms are almost ten times as frequent in the BD parts of the transcript as are BD loans in the French parts, are not surprising to any student of language contact, but still are important additions to our knowledge base. Other data sets have consistently shown that nouns are the most frequent singly-occurring forms, whether as borrowings or CS forms; and, given what is known about the sociolinguistic profile of Brussels, with French as the higher status language, more loans *from* French are expected.

Other results, such as how French verbs are integrated into BD

frames, have more potential theoretical value. French verbs which appear as finite verbs in BD frames (and therefore which receive BD inflections) take the form: base form + er (e.g., *bless+er*). T-D argues that the infinitive is not the base form, because French verbs from other conjugations, such as *offrir*, become *offr+er*. Further, the suffix even appears combined with nouns to form verbs that do not exist in French (e.g., *fantas-er-en*) 'to fantasize' (with *-en* as the BD infinitival suffix). She notes, however, that this suffix is not attached to English verbs in Dutch/English CS in The Netherlands. T-D's explanation is that somehow the suffix itself has become productive for French verbs, but not for Germanic ones. Yet, another explanation strikes me as more likely, just because it is in line with findings from other data sets and results in a theoretical claim about the nature of lemma entries in the mental lexicon. On this view, the infinitive *is* the base form and the French infinitive includes *-er*, the morpheme for 'to' as part of its lemma entry in the mental lexicon. Evidence for this claim comes from other CS data sets: English infinitives *never* appear in mixed constituents with *to* across various data sets. In my Swahili-English data set, English verb stems (with English as the Embedded Language) appear instead inflected with the Swahili equivalent of *to*—e.g., the infinitival prefix *ku-* 'to' as in *huwezi ku-**burn your skin*** "you can't [are not able to] burn your skin"). In contrast, French verbs in Lingala/French CS (when French is the Embedded Language) appear with *-er*, followed by Lingala suffixes, (cf. Kamwangalamu 1989). The difference between how English infinitives versus French infinitives appear in CS (when English or French is the Embedded Language) is evidence that French lemma entries for infinitives are different from English ones. Those for English do not include the morpheme which is equivalent to *to*, but French lemmas for infinitives do include this morpheme. As for why the infinitival suffix is consistently *-er*, T-D herself states that "many borrowed verbs in (Brussels) Dutch stem from the class with the *-er* suffix" (p. 111). I suggest that the analogical pressure of the dominant class causes verbs from other classes to undergo reanalysis to this class, *not* that the *-er* suffix has become productive *just* for French verbs in BD frames.

In her chapter on borrowing, T-D does not comment at all on several of her findings which are of great potential interest. For example, T-D does not attempt to explain the results of her own analysis, such as the fact that French singly-occurring elements appear apparently consistently with BD inflections. In my Matrix Language Frame Model (MLF model), which had only been introduced in an early version (Myers-Scotton 1990) by the time T-D presu-

mably was writing this manuscript, I propose a System Morpheme Principle. This principle states that in mixed constituents (i.e., with morphemes from both varieties involved), all system morphemes (=inflections and some functional elements) which are syntactically active must come from the Matrix Language (BD in this case). See Myers-Scotton [1993] 1997: 83 ff. Or, to take another example, T-D reports that "prepositions are generally switched in combination with their complement" (p. 114). This means, in terms of the MLF model, that if French prepositions appear at all in a BD frame it is only in Embedded Language Islands, constituents consisting entirely of morphemes of one language. T-D suggests, following Muysken 1981, that prepositions may be difficult to switch because they are governors. Yet, this argument falls through when one considers that French verbs, which are also governors, are switched fairly frequently as singly-occurring forms (353 or 2% of the total corpus of tensed verbs are CS forms). I would argue that singly-occurring prepositions are not switched because either they are system morphemes (Myers-Scotton [1993] 1997: 83) or, if they are content morphemes (i.e., those which assign thematic roles), because of incongruence between the French and Dutch counterparts in regard to their lexical-conceptual structure (cf. Myers-Scotton and Jake 1995).

T-D devotes an entire chapter to discussing the morphosyntactic integration of borrowed adjectives. French adjectives are largely postnominal while BD adjectives are prenominal; BD adjectives are inflected only in attributive position. Given the constraints, it is understandable that most French adjectives in BD frames occur in predicative position "because no inflection from the matrix language is imposed on them in these contexts. If borrowed adjectives are placed into the attributive position, they have to be inflected according to the rules of the matrix language" (p. 172). This matter-of-fact reference to the matrix language and its effects precedes T-D's definition of the matrix language by a number of pages. She refers to it again several times and finally defines it in chapter 8, and must consider it again when she discusses an early version of the MLF model on pp. 238-39. Yet, she never discusses what it "means" to claim that there is a matrix language.

Still in this chapter, T-D brings up Poplack's Equivalence Constraint (Poplack 1980), correctly pointing out that it is not clear if the constraint refers to switch points (i.e., both languages must have equivalent word orders at switch points) or to the switchability of individual items (i.e., must the order of N and Adj be equivalent within a constituent?). T-D comes back to Poplack's Equivalence

Constraint in chapter 8 and concludes that this constraint "does make correct predictions in many cases... if one assumes that equivalence needs to hold at the constituent level only" (p. 226). That is, if differences in order within a constituent (e.g., order of N and Adj) are not considered as cases of non-equivalence, the constraint holds.

However, T-D concludes that the other constraint which Poplack proposes, the Free Morpheme Constraint, does not make correct predictions for the Brussels data. According to this constraint, a switch may not occur between a bound morpheme and a lexical form unless the lexical form is phonologically integrated into the language of the bound morpheme. Yet, French adjectives occur morphosyntactically integrated into BD frames (i.e., with a BD suffix and in BD word order). Also, T-D says that "syntactic constraints" (never specified) do not make the correct predictions, either, since they would predict against postnominal French adjectives in positions before Dutch nouns, and such examples do occur. The only cases which seem to bother T-D are N+A combinations, such as *créme glacée* "ice cream" since they follow French, not BD word order. For the MLF model, these combinations are no problem since they are considered as Embedded Language Islands, not singly-occurring words. Some of these combinations are what I call internal Embedded Language Islands; that is, they are in mixed constituents headed by a Dutch determiner (e.g., *de* as in *de créme glacée*). See Myers-Scotton [1993] 1997: 152–54.

Chapter 7 describes a convergence phenomenon of potential theoretical interest well beyond French/BD contact: BD's Verb Second Rule is being replaced by a Verb Third Rule when a French adverb appears in the sentence in initial position. The most common position for French adverbs in mixed sentences is in the leftmost position before the subject, with the verb following (103 out of 426 cases). Only 3 appear in the normal position for BD adverbs (directly preceding the finite verb). These "adverbs" may well be considered discourse markers, and their discourse roles may have much to do with their apparent influence. In fact, T-D does suggest that they may be what she calls "sentence adverbs", rather similar to interjections, and she points out that interjections do not trigger Verb Second in Dutch (p.178).

T-D also notes that word order in subordinate clauses is not influenced by French; however, under the current version of the MLF model (Myers-Scotton 1997) these subordinate clauses are analyzed as monolingual CPs. As such, they are headed by complementizers from the same language as the clause. Once one recognizes that the

constituent in question is a monolingual Dutch CP, it is no surprise that the word order is not influenced by French (or vice versa). It is only when a given CP is bilingual (i.e., with morphemes from both languages) that one language would be expected to influence the other. At the time T-D was writing this book, the sentence was considered the unit of analysis. Yet, reflection indicates that the proper unit of analysis for considering syntactic constraints on CS is the CP which has morphemes from two (or more) languages, not the sentence. First, the sentence is a poorly-defined unit. Second, any sentence may have morphemes from two languages, but still be made up of monolingual CPs (i.e., one CP from language X and one CP from language Y). In this case, morphemes from the two languages are not in contact at a level which provides a non-trivial test for syntactic constraints on CS.

Chapter 8 represents T-D's clearest attempt at making a theoretical contribution; this is where T-D introduces her hierarchy of switched constituents and explains the hierarchy in relation to argument structure. Stating that her approach is different from that of earlier CS researchers who concentrated on switch points between constituents, she considers the constituents themselves. As noted above, it is not until this chapter that T-D defines the Matrix Language (ML). Following Klavans (1983), who operates under the influence of GB theory, T-D defines the ML language as the language of the "inflection bearing element of the verb", relating the construct of matrix language to INFL. While the centrality of INFL may be well motivated in certain generative syntactic theories that have been applied to monolingual data, it has no obvious motivation as a critical unit in explaining bilingual data. Once more, a definition of the ML based on INFL does not give the correct predictions in all CS data sets. For example, in the Arabic/English corpus discussed in Jake and Myers-Scotton 1997, English is the Embedded Language (EL). Many IP EL Islands occur. This means that the "inflection bearing elements" in these islands (i.e., the inflections on the verb) are all from English, as well as the verb stems. Yet, Arabic is clearly the ML in these data, as a full analysis of the entire mixed CP indicates. Most crucially, in mixed constituents within the same CP as the IP EL Island, all the function words and inflections come from Arabic, even when they occur with an English content word).

How T-D uses her designation of the ML is not clear to me since she discusses the ML most fully when referring to monolingual constituents. In the MLF model, the ML versus EL distinction is only relevant in mixed constituents. Still, T-D's findings regarding the

nature of these monolingual constituents are of interest. Her implied hypothesis is that "constituents that are arguments of a verb or of a preposition are switched less easily than constituents that are not arguments" (p. 226). And T-D does demonstrate, for example, that although more NPs than any other type of constituent are switched (49 or 41% of the full constituents switched), many of these particular NPs are of types which do not participate in the thematic grid of the clause. For example, of switched NPs, 37% are "dislocated" (e.g., *les étrangers* in [195] *Les étrangers, ze hebben geen geld* "The strangers, they don't have money"); only 8% are objects of verbs and only 6% are subjects of verbs. PPs are in second place behind NPs as a frequently switched type of constituent. This finding also supports T-D's hypothesis that constituents which are easily switched are peripheral to argument structure, since many of the PPs are adverbial (adjuncts, not arguments of the verb).

As part of this discussion, T-D introduces Muysken's (1991) ideas about "switching under equivalence", with switching possible only when the element is considered to be equivalent in both languages by speakers of those languages. From T-D's brief discussion of the subject, equivalence seems to refer to realizations at the surface level; e.g., morphological realization patterns. For example, she states that if Language A has prepositions and Language B has postpositions, it may be impossible to switch prepositions "because there is no equivalent category in the other language" (p. 246).

In her closing statements, T-D focuses on "categorical equivalence" and her hierarchical approach (which constituents are more likely to be switched) as among the most promising ideas to pursue in future CS research. Certainly, Muysken's general idea that equivalence in some sense affects possible CS structures has merit, but at least in his 1991 publication and in a more recent version (Muysken 1995), he does not give many details. In one of his more explicit statements (1995: 192-3), he says, "There can be equivalence of categories (lexical elements, phonemes, phrase structure nodes, morpho-syntactic features) or of relations between categories, in structuralist terms. The latter are either syntagmatic (e.g., word-order or agreement rules) or paradigmatic (equivalence oppositions)".

The notion of equivalence has always been reflected in my own work (cf. the Blocking Hypothesis in Myers-Scotton [1993] 1997: 120ff.) where I acknowledge the role of Muysken and others in recognizing the role of subcategorization restrictions. However, I now emphasize *congruence* at an abstract level, rather than primarily in terms of surface equivalences (cf. recent work by Janice Jake and

myself, starting with Myers-Scotton and Jake, 1995). We refer to *congruence* in terms of abstract lexical structure at three levels. While the term *lexical structure* is used, the three levels go well beyond the semantics of words. The levels are (1) lexical-conceptual structure, which does refer to the semantic and pragmatic features, as well as thematic roles of morphemes or words; (2) predicate-argument structure, which refers to relations between verbs and prepositions and their arguments; and (3) morphological realization patterns, which refers to surface structures, including word order. Note that while Muysken speaks of "switching under categorical equivalence" as only *one* of three major strategies for intrasentential CS, we argue that congruence *at a more abstract level than equivalence at the category level* figures in *all* CS. The only exception would be for Muysken's third strategy, paratactic switching, because this refers to switching of such elements as tags that do not participate in the thematic grid. That is, congruence *also* figures in whether or not singly-occurring EL forms may occur in a constituent morphosyntactically framed by the ML; this is what Muysken refers to as "morphologically shielded switching" and what I discuss as the switching resulting in some ML + EL constituents or mixed constituents, first so characterized in Myers-Scotton 1990.

Jake and I suggest that incongruence at any of these levels may determine whether CS between two languages consists largely of mixed constituents or of EL Islands. For example, we argue that, in a Spanish/English data set, incongruence at the lexical-conceptual structure between Spanish lemma entries and English entries for nouns blocks English EL Islands (full constituents in English only). Instead, mixed constituents headed by Spanish elements in the DET slot followed by English NPs are common. Incongruence arises because *phi*-features become salient in Spanish at the level of lexical-conceptual structure, but not in English until the level of morphological realization patterns (Myers-Scotton and Jake 1997). Of course if English were the ML, not Spanish, the results are predicted to be quite different because the *phi*-feature constraints of English do not produce the same blocking effect as the Spanish ones. Elsewhere (Jake and Myers-Scotton 1997) we contend that English IP EL Islands are frequent in Arabic/English CS because of incongruence between Arabic and English verbs at the lemma level; that is, the only way to access an English verb in a mixed CP which is framed by Arabic as the ML is to produce an island (i.e., with English inflections). Also at issue is the internal structure of mixed constituents; e.g., in some language pairs, switching to an EL verb is only possible

if there is a *do*-support verb from the ML present (e.g., Japanese/English CS or Turkish/Dutch CS).

In a final comment on T-D's volume, let me note that a major virtue is that even though T-D herself draws few theoretically-oriented conclusions, she provides data (e.g., on changes in word order) and identifies directions for research (e.g., equivalence) which are now being explored by other researchers.

As the sub-title indicates, the Bakker and Mous volume is a collection of 15 case studies. To me, the appellation "case study" itself implies a descriptive, not a theoretical, bias. In this particular instance, the lack of explanation, which an emphasis on description often entails, is not necessarily a bad thing. This is so because the chapters provide descriptions of a language contact phenomenon which many linguists have claimed does not happen: the emergence of a mixed language. In such a language, the bulk of its lexicon almost always comes from one source, but its grammar from another source. Thus, given the fact that a "new" phenomenon is being recognized, the emphasis on description is justified. In fact, as the editors write in their "introduction", it is likely that this is the first book devoted to mixed languages ever published. More recently, another volume edited by Sarah Thomason (1997), which covers some of the same subject matter, has appeared, indicating that it is a subject attracting researchers today. The Bakker and Mous (B&M) volume resulted from a workshop on mixed languages held at the University of Leiden in 1993. That B&M could gather together linguists with descriptions of 15 different cases which qualify in some way as "mixed languages" is itself of great *theoretical* interest, given earlier claims about the impossibility of such varieties.

What is a mixed language? B&M state, "We propose the term 'language intertwining' for the process forming mixed languages showing a combination of the grammatical system (phonology, morphology, syntax) of one language with the lexicon of another language" (p. 4). However, it is only to varying degrees that some of the case studies meet this characterization, and I will return later to the explanatory power of "intertwining".

The "best" mixed language may well be Ma'a or Mbugu, spoken in a small area in Northeastern Tanzania and discussed by Maarten Mous (pp. 175-200). Ma'a is well-known to students of language contact since it is at least mentioned in many overviews and is discussed extensively in Thomason and Kaufman (1988), the widely-read volume detailing examples of externally-motivated language changes. Ma'a has a lexicon for which a Cushitic origin seems likely in many

cases and a grammar which has even clearer parallels in the neighboring Bantu languages.

Mous's chapter has many virtues: it is based on recent field work (1991-93), it gives a more complete inventory of the grammatical system than heretofore available, and it can consider the evidence against the backdrop of several recent hypotheses about how Ma'a came about (i.e., not only Thomason and Kaufman, but also Brenzinger 1987 and Myers-Scotton 1993, 1997).

Mous prefers to view Ma'a as one of two *registers* of the same language; he refers to what others call Ma'a as "Inner Mbugu" and to the more Bantu-like variety as "Normal Mbugu". Except for their lexicon, they are nearly identical. For the most part, their grammars have Bantu sources. Where there is a possible Cushitic origin for functional elements, this is only the case for those forms which would be called content morphemes (as opposed to system morphemes) in the Myers-Scotton ([1993], 1997) opposition between content and system morphemes. Under this opposition, content morphemes may come from the EL (Embedded Language), but not system morphemes; in this case, Cushitic is the EL. The content morphemes in question are possessive pronouns and they are classified as content morphemes because they can receive thematic roles vs. system morphemes which cannot receive these roles.

The major difference between the two varieties is in the lexicon. First, there are two forms for many entries, one belonging to Inner Mbugu and the other to Outer Mbugu. Second, Mous reports that individuals vary in their knowledge of Inner Mbugu forms, with some younger speakers knowing/using only an Outer Mbugu form. Also, certain semantic domains may have only one entry; for example, Mous cites one such domain as non-domestic animals and states that forms in this case are of Bantu origin, predominately from Shambaa (a local Bantu language) or Swahili (also a Bantu language, but the national lingua franca in Tanzania as well).

When discussing the possible origin of Ma'a/Mbugu, Mous states that because speakers of Inner Mbugu today also speak Normal Mbugu "I assume that it has always been like that for Mbugu" (p. 199). Why he assumes that the state of affairs could not have changed is not clear. But then he goes on to claim that, in fact, change has happened, but a rather curious change: he argues that Inner Mbugu was consciously created by speakers of Normal Mbugu to set themselves apart from their Bantu neighbors. The probability of conscious creation in general will be discussed below. Here, let me say that a more plausible explanation is that Inner Mbugu as a state *preceded*

Normal Mbugu, a conclusion supported by the fact that many younger speakers do not know Inner Mbugu. That is, I would argue that a "turnover" from a Cushitic variety to a Bantu variety as the Matrix Language (i.e., the language structuring the grammatical frame) created Inner Mbugu; also, a movement toward more lexical items from Bantu is the natural progression, given the Bantu milieu in which Mbugu people find themselves. I would agree that they maintain some Cushitic lexical items to set themselves apart from Bantu speakers, but to assume that they *added* these words rather than *retained* them begs the question of where they "found" them in order to add them. Maintenance rather than addition seems like a more natural state of affairs. (For an extended version of my argument, see Myers-Scotton 1998).

The other mixed language of most interest in the volume is Michif, a language spoken by descendants of French speaking fur traders from Quebec and Acadia and their Amerindian wives. The other editor, Bakker, is the author of the chapter on this language. Michif is spoken by some members of the Métis ethnic group; those speaking it live mostly in Manitoba and Saskatchewan in Canada or in the neighboring U.S. states of North Dakota and Montana. Bakker estimates that there are only 1,000 speakers of Michif, but other estimates range up to 5,000. Almost all speakers are older than 60; while they all learned Michif as their first language, all also know English as well.

Simply described, Michif has been characterized as a language with French noun phrases and Cree verbs. But the situation is not quite so neat since some French and English verbs are in use with Cree inflectional morphology, according to Bakker. Also, the French copula and the French verb *avoir* "to have" are used. While Michif sentence structure is much closer to Cree than to French, Bakker notes, noun phrase structure is as in French.

Michif differs from a creole in that neither French nor Cree structures in Michif have lost much of their grammatical complexities. Also, while there is some influence from Cree on the phonology of French forms, for the most part both Cree and French forms have retained their own phonology.

Bakker argues that social circumstances would lead to a language with a Cree grammatical system (the mother's language) and a French lexicon (the father's language). He uses Thomason and Kaufman's (1988) discussion of outcomes of language contact as his point of reference in explaining Michif. Thomason and Kaufman refer to four possible outcomes: language maintenance, language shift, and

pidginization/creolization. In addition, in cases of maintenance, there can be massive grammatical replacement (as the ultimate stage of lexical borrowing). Bakker rejects any of these outcomes as either a characterization or explanation for Michif. He writes, "I would argue that there is a process of 'language intertwining' which may look like 'massive grammatical replacement', but it is not. It is a fifth process, differing from all four mentioned in T&K" (p. 24). Bakker goes on to define intertwining as "the combination of the grammatical system (phonology, morphology, and syntax) of one language with the lexicon of another language" (p. 24). And, of course, many of the mixed languages discussed in this volume (e.g. Mbugu or Ma's just mentioned above) have such a dichotomy.

To maintain his definition of intertwining and to justify placing Michif in the same grouping with Mbugu, Bakker must argue that while Michif looks rather different from Mbugu, which basically has *all* its grammar from a Bantu source, it is not different. He asserts that Michif "has a Cree grammatical system and a French lexicon" and only looks different "because of the typological nature of Cree". (p. 25). As if to avoid saying that the French determiners in noun phrases are evidence of French grammatical elements, he states that nominal marking in Cree is rare. He also argues against considering the Cree verb stems as lexical elements by arguing that "the Cree verb should *as a whole* [italics added] be taken as part of the grammatical system of the language, and therefore the whole verb phrase, and not only the morphological markers, are from Cree in Michif" (p. 25).

In addition to the chapters Mous and Bakker, there are 13 other chapters which describe varieties which show varying degrees of mixed origins. These include Romani, Town Frisian, Maltese, Mednyj Aleut, Shelta, Javindo, Island Carib, Amarna-Akkadian, Media Lengua, Callahuaya, Ilwana, Petjo, and KiMwani. Many of the discussions are quite short, presumably because no more data exists about the variety in question. In some cases, such as Muysken's discussion of Media Lengua, the appellation "mixed language" definitely seems to fit. Spoken in Central Ecuador, Media Lengua has a Quechua grammatical base, but largely a Spanish lexicon. However in such cases as Drewes' treatment of Maltese, it seems clear that the variety in question is not a mixed language – according to B&M's definition (given above). Rather, Maltese remains an Arabic dialect, with heavy borrowing from Italian and English, but only of content morphemes. That is, its grammatical system fits the Arabic mold. I see this as evidence that there has not been any turnover of the ML,

and I would predict that if CS occurs in this community, Maltese is always the ML. That is, there is no evidence of a turnover in the ML and this is what separates Maltese from "true" mixed languages. In other cases, one questions the extent to which the variety discussed really exists and is in use as a fully-constituted linguistic system rather than as a few set sentences (e.g., Romani varieties).

Finally, the most interesting mixed language is surely Mednyj Aleut (discussed by Golovko). This variety emerged at the end of the 19th century on Mednyj (Copper) Island in the Aleutians and does not neatly fit the characterization of "lexicon from one language, grammar from another"; unfortunately, however, not a great deal of data exist. While the majority of the lexicon does come from Aleut, verb stems have Russian inflections. Other elements of the grammatical system also come from Russian, but not all, it seems. I suggest that Mednyj Aleut can be explained as the result of an arrested ML turnover, as discussed below.

I applaud the editors for having assembled this volume; it brings together data which need to be considered in any discussion of language contact which claims to be comprehensive. However, I have several quarrels with the main theoretical assumptions regarding how these languages came about. First, while it is true that social factors are the stimulus for the formation of mixed languages, I remain unconvinced that social factors actually determine the design of linguistic systems. Similarly, *intertwining* is not an explanation for the makeup for a mixed variety; it is only a description of the state of affairs. Thus, as I argue elsewhere (e.g., chapter 8 in Myers-Scotton, [1993] 1997 and Myers-Scotton, 1998), I suggest that an ML turnover is both the explanation and the mechanism which results in a mixed language. To explain my position briefly: The antecedents to an ML turnover must be (i) bilingualism and codeswitching (CS) in a community and then (ii) a shakeup in the social situation in favor of the "infringing language" in the community such that this language begins to take over as the ML (the language setting the morphosyntactic frame in mixed constituents) in CS. With time, the new variety (which *is* CS between the L1 and the L2, but with the L2 now setting the grammatical frame) becomes the community's main language, generally replacing the L1. I do not suggest that an ML turnover happens suddenly, but rather various versions of a "composite ML" (composed of abstract lexical structure from both participating languages) structures utterances over the course of the change. At the same time, the speakers must feel sufficient loyalty to their L1 to retain much of the lexicon of the L1 in this new community language;

this is what keeps the situation from being a case of total language shift. Thus, in summary, I accept intertwining *as a description* of mixed languages, but hypothesize that CS, followed by an ML turnover, is both *the explanation and the mechanism* which brings about the structural state of affairs found in mixed languages. I also would argue that there is no "separate process of language intertwining" (p. 25), but rather that all forms of bilingual speech (e.g., language shift, attrition, pidginization/creolization, and interlanguage in second language acquisition) are the products of the same processes, with outcomes differing because of social and psychological factors. That is, ML turnover (and the formation of a composite ML along the way) figures in all of these, but goes to various degrees of completion because of external factors. For example, in the case of Michif, I suggest that Cree was the original ML (in a CS variety including Cree and French) and that a turnover to French as a new ML began, but was arrested, leaving behind French NPs.

A second quarrel I have with the volume is the idea expressed by some of the authors that the resulting varieties were somehow consciously put together. For example, Golovko makes this statement about Mednyj Aleut: "I think CIA was 'constructed' by adults" (p. 117) and Mous also refers to conscious creation in reference to Inner Mbugu. Also, some of the authors equate mixed languages with the making up of a secret language. Mixed constituents in CS data sets around the world constitute indirect counter-evidence because they are a mixed product and generally are produced with no conscious awareness on the part of its speakers. In such constituents, all of the syntactically-active grammatical features (i.e., system morphemes in my terminology) come from one language, the ML, and content morphemes may come from either the ML or another language, the EL. If speakers were consciously producing these mixed constituents, how can we explain that CS speakers all over the world produce mixed constituents of the same composition? Surely the uniformity in the structure of mixed constituents in CS is evidence of universally-present strategies under unconscious cognitive control, not the result of "on the spot invention". Further, the fact that most of the 15 cases cited in the B&M show similar structures also argues for cognitively-based strategies. Finally, the differences *across* the mixed languages can be explained by the process of ML turnover. That is, not all communities where mixed languages have developed/are developing need be at *the same stage of ML turnover*.

I conclude as I began: there is little basis for comparing these volumes since their goals are so different. As linguists, we all should

be most interested in theory-building because our proper goal is to arrive at explanations of linguistic structure. Ironically, the most theoretically-oriented of the three volumes (that by Singh) offers the least "grist for the theory-building mill". One might wish that the other two volumes attempted more theoretical insights, but their data bases are of the sort and extent that can stimulate theory-building. In different ways, all three volumes are worth consideration.

Address of the Author

Linguistics Program, University of South Carolina, Columbia, SC 29208.

Bibliographic References

BRENZINGER, M. (1987), *Die Sprachliche und Kulturelle Stellung der Mbugu (Ma'a)*, M.A. Thesis, University of Cologne, Cologne, France.

DISCIULLO, A. M., P. MUYSKEN & R. SINGH (1986), "Government and code-mixing", *Journal of Linguistics* 22: 1-24.

DRESSLER, W.U. (1985), "Explaining natural phonology", *Phonology Yearbook* 1: 29-50.

HAUGEN, E. (1953), *The Norwegian Language in America: a Study of Bilingual behavior*, Philadelphia, University of Pennsylvania Press.

JAKE, J. L. & C. MYERS-SCOTTON (1997), "Codeswitching and compromise strategies: implications for lexical structure", *International Journal of Bilingualism* 1: 25-39.

KACHRU, B. (1980), "Socially realistic linguistics: the Firthian tradition", *Studies in the Linguistics Sciences* 10: 85-112.

KAMWANGAMALU, N. (1989), *Code-mixing across Languages: Structure, Functions and constraints*, PhD Dissertation, University of Illinois, Urbana.

KLAVANS, J. L. (1983), "The syntax of code-switching: Spanish and English", *Selected Papers from the 13th Linguistic Symposium on Romance Languages* 14, Amsterdam, Benjamins: 213-232.

LABOV, W. (1980), *Locating Language in Time and Space*, New York, Academic Press.

MUYSKEN, P. (1981), "Spaans en Quechua in Ecuador", *Interdisciplinair Tijdschrift voor Taal- en Tekstwetenschap* 1/2: 124-138.

MUYSKEN, P. (1991), "Needed: a comparative approach", *The Symposium on Code-switching in Bilingual Studies: Theory, Significance, and Perspectives*, Strasborg, European Science Foundation Network on Code-switching and Language Contact: 253-272.

MUYSKEN, P. (1995), "Code-switching and grammatical theory", *One Speaker, Two Languages: Cross-disciplinary Perspectives on Codeswitching*, in Leslie MILROY and Pieter MUYSKEN (eds.), Cambridge, Cambridge University Press: 177-198.

MYERS-SCOTTON, C.M. (1990), "Intersections between social motivations and structural processing in codeswitching", *Papers for the London Workshop on Constraints, Conditions and Models*, Strasborg, European Science Foundation Network on Code-switching and Language Contact: 57-82.

MYERS-SCOTTON, C.M. ([1993] 1997), *Dueling Languages: Grammatical Structure in Codeswitching*, Oxford, Oxford University Press (1997 paperback with new 'Afterword').

MYERS-SCOTTON, C.M. (1998), "One way to dusty death: the matrix language turnover hypothesis", in L. GRENOBLE and L. WHALEY (eds.), *Endangered Languages*, Cambridge, Cambridge University Press: 289-316.

MYERS-SCOTTON, C. & J. JAKE (1995), "Matching lemmas in a bilingual language competence and production model: evidence from intrasentential codeswitching", *Linguistics* 33: 981-1024.

MYERS-SCOTTON, C. & J. JAKE (1997), "Codeswitching and the nature of lexical entries", Paper presented at Colloque: 'Alternance des langues at Apprentissage', St. Cloud, France, February 6-8. To appear in *Plurilinguismes* 14.

NORTIER, J. (1990), *Dutch and Moroccan Arabic in contact: Codeswitching among Moroccans in the Netherlands*, Dordrecht, Foris.

PANDHARIPANDE, R. (1981), "On the nativisation of the lexicon: the case of Marathi", *Linguistics* 19: 987-1011.

POPLACK, S. (1980), "Sometimes I'll start a sentence in Spanish and termino in Español: toward a typology of codeswitching", *Linguistics* 18: 581-618.

POPLACK, S. (1990), "Variation theory and language contact: concepts, methods and data", *Papers for the Workshop on Concepts, Methodology and Data*, Strasborg, ESF Network on codeswitching and language contact: 33-66.

SANKOFF, D. & S. POPLACK (1981), "A formal grammar for code-switching", *Papers in Linguistics* 14: 3-45.

SINGH. R. (1985), "Grammatical constraints on code-mixing: evidence from Hindi-English", *Canadian Journal of Linguistics* 30: 33-46.

THOMASON, S. G, ed. (1997), *Contact Languages: a Wider Perspective*, Amsterdam, Benjamins.

THOMASON, S. G. & T. KAUFMAN (1988), *Language Contact, Creolization and Genetic Linguistics*, Berkeley, University of California Press.