Bilingualism: Language and Cognition 8 (1), 2005, 1–22 © 2005 Cambridge University Press doi:10.1017/S1366728904002068

Codeswitching and generative grammar: A critique of the MLF model and some remarks on "modified minimalism"*

JEFF MACSWAN
Arizona State University

This article presents an empirical and theoretical critique of the Matrix Language Frame (MLF) model (Myers-Scotton, 1993; Myers-Scotton and Jake, 2001), and includes a response to Jake, Myers-Scotton and Gross's (2002) (JMSG) critique of MacSwan (1999, 2000) and reactions to their revision of the MLF model as a "modified minimalist approach". The author argues that although a new structural definition of the Matrix Language (ML) makes the MLF model falsifiable, its empirical predictions are inconsistent with the facts of codeswitching (CS). The author also identifies significant theoretical problems associated with the MLF model, and suggests that it be rejected on empirical and theoretical grounds, and on grounds of scientific parsimony. In addition, the author contends that JMSG's critique of the Minimalist approach to CS rests on conceptually significant misreadings of published research in CS and in the theory of syntax, and that JMSG fail to motivate their central claim, namely, that a Minimalist approach to CS cannot succeed without incorporating the ML construct. It is further shown that an analysis of the CS facts which JMSG consider as their test case may be straightforwardly pursued within the Minimalist Program without appealing to the ML construct or any other aspect of the MLF model. The author concludes with the recommendation that research on CS depart from the formulation of general CS-specific constraints like the MLF model and engage in the task of analyzing language contact phenomena in terms of independently motivated constructs of linguistic theory.

The Matrix Language Frame (MLF) model of Myers-Scotton (1993) and colleagues (Myers-Scotton and Jake, 2001; Jake, Myers-Scotton and Gross, 2002) has enjoyed considerable popularity among researchers concerned with the grammatical properties of codeswitching (CS) since it first appeared. Simply put, the MLF model theorizes that the "morphosyntactic frame" of a code-switched utterance is set by a "matrix language" (ML), with "embedded language" (EL) morphemes from another language inserted into this frame. Recently, Myers-Scotton and colleagues (Jake, Myers-Scotton and Gross, 2002; henceforth, JMSG) have attempted to translate this basic approach into the language of modern syntactic theory, introducing what they call "a modified minimalist approach". JMSG argue that linguistic theory (as instantiated, in particular, in the Minimalist Program) must be augmented by Myers-Scotton's (1993) ML/EL distinction if it is to succeed in capturing the full range of CS data. In the course of arguing their case, JMSG produce a critique of "earlier minimalist approaches to codeswitching", focusing on the work of MacSwan (1999, 2000).

In the present article, I provide a theoretical and empirical critique of the MLF model, taking the opportunity to respond to JMSG's critique of MacSwan (1999, 2000). I argue that JMSG's criticisms rest on misreadings of my work and of other published research in CS and in syntax. I show, further, that JMSG fail to motivate their basic claim (that the ML construct is required for the successful analysis of CS data), and I provide evidence that it is false. I argue that the MLF model should be rejected for both theoretical and empirical reasons.

The article is organized in three major sections. To provide a context for subsequent discussion, the first section presents an overview of the basic architecture of the Minimalist Program and of the application of this body of research to CS, opening with some remarks on the nature of a linguistic theory. A second major section presents an evaluation of the classic MLF model and some revisions in JMSG. The third section provides a response to JMSG's criticisms of MacSwan (1999, 2000), noting some additional problems with their proposed "modified minimalist" version of the MLF model. A final conclusions section summarizes some major points.

* For useful comments on an earlier version of this manuscript, I am indebted to three anonymous *BLC* reviewers, Katja Cantone, Sonia Colina, Kellie Rolstad, and Kara McAlister. I am also grateful to Marcel den Dikken, Janet Fuller, Elly van Gelderen, Helena Halmari, Jim Hlavac, Janet Kaaya, Hilda Koopman, Carson Schütze and Rose Thomas for responding to e-mail queries, and to unnamed consultants who assisted with data queries. Kara McAlister provided additional assistance with data collection. Of course, I alone am responsible for the content and views expressed. Aspects of this research were supported by a grant from the National Academy of Education, with funding from the Spencer Foundation.

Address for correspondence
Arizona State University, PO Box 872011, Tempe, AZ 85287-2011, USA
E-mail: *macswan@asu.edu*

1. Linguistic theory and the grammar of codeswitching

1.1. The nature of a linguistic theory and appropriate data in CS

A formal grammar is a DEFINITION of a language; it specifies the (infinite) class of well-formed expressions of the language under analysis. As Chomsky (1957, p. 2) classically remarked,

> The fundamental aim in the linguistic analysis of a language L is to separate the *grammatical* sequences which are the sentences of L from the *ungrammatical* sequences which are not sentences of L and to study the structure of the grammatical sequences. The grammar of L will thus be a device which generates all of the grammatical sequences of L and none of the ungrammatical ones. [italics in original]

Similarly, the grammar of CS will generate all of the well-formed expressions which invoke elements contributed by more than one language, and none of the ungrammatical ones.

A linguistic theory so constructed will PROJECT a class of grammatical expressions, infinite in range, based on a finite number of observations of well-formed and ill-formed expressions. Note that the set of observations which makes up this body of evidence has a specific character – namely, it must provide samples of what CAN occur (the grammatical expressions) as well as what CANNOT occur (the ungrammatical expressions). While positive evidence will allow the linguist to theorize about "the grammatical sequences of L", only negative evidence can identify the ungrammatical ones. Thus, without both kinds of evidence, it will not be possible for the linguist to verify whether a proposed theory is empirically adequate or not.

The topic of "negative evidence" in the context of child language acquisition is related but distinct from the issue noted here. The absence of negative evidence available to children acquiring a first language has been used as a component of the argument from the poverty of the stimulus to show that acquisition is guided by a biological endowment (Lightfoot, 1982; Chomsky, 1986). Unlike children acquiring a first language, however, linguistic theorists require negative evidence to test whether their model correctly bans ill-formed expressions.

In CS, a persistent controversy concerns whether naturally-occurring data or experimental data (sentence judgments, in particular) are more appropriate sources of evidence for constructing theories. While some researchers have strongly rejected all experimental evidence (e.g. Mahootian, 1993; Mahootian and Santorini, 1996), others have voiced a strong preference for it (e.g. Toribio, 2001). Naturalistic data collection has the important advantage of placing CS in a more realistic context, but it has a great disadvantage as well: It does not provide instances of ill-formed sentences – or if it does, they are not so labeled. This is a significant problem if one is interested in constructing an explicit theory of a bilingual's linguistic competence, because such a theory must generate all of the well-formed utterances in a bilingual's repertoire and none of the ill-formed ones. Without disconfirming evidence, we would have no way of knowing whether a theory sanctions unacceptable utterances. We cannot confidently assume that the absence of a form in naturalistic data means that the structure is not permitted; it may be absent because it CANNOT occur, or it may be absent because it HAS NOT occurred. On the other hand, experimental data may be corrupted by prescriptivist attitudes, a lack of real-world context, or other factors (see Schütze, 1996 for extended discussion). Thus, a balanced perspective is preferred – one which advocates careful consideration of all linguistic data, along with the inherent messiness and unique limitations associated with both approaches – but negative evidence is indispensable to a generative theory of linguistic competence.

Another dimension of controversy regarding CS data concerns age of onset of exposure to the languages known to bilingual participants. Because there is reason to believe that second language may be represented differently in the mind/brain of non-native speakers than it is for natives, CS research should be based as much as possible on the language of simultaneous bilinguals for whom CS is not stigmatized (MacSwan, 1999, pp. 29–36). Johnson and Newport (1989) found that the earlier children acquired their second language, the more native-like their ultimate attainment, with second language learners ages 3–7 indistinguishable from natives, and young adolescents much better than older L2 learners. We should therefore prefer data from simultaneous bilinguals, but may on occasion resort to data from early sequential learners or others when the former is not practically available.

Let us now review some aspects of the Minimalist Program and some consequences for CS research.

1.2. The Minimalist Program

The Minimalist Program (MP) was developed in the context of an emerging conception of syntactic theory in which PARAMETERS were restricted to the lexicon rather than operating on syntactic rules (Borer, 1984; Chomsky, 1991). Restricting parameters to the lexicon means that the I-language is fixed and invariant, and linguistic variation results from feature checking of the morphological properties (abstract and concrete) of the lexicon. In the MP, there are two central components of the syntax: C_{HL}, a computational system for human language, presumed to be invariant across languages, and a lexicon, to which the idiosyncratic differences observed across languages are attributed.

Thus, phrase structure is derived from the lexicon in the MP. An operation called SELECT picks lexical items from the lexicon and introduces them into a LEXICAL ARRAY (LA) (formerly, a NUMERATION), an assembled subset of the lexicon used to construct a derivation. Another operation, MERGE, takes items from the LA and forms new, hierarchically arranged syntactic objects. The operation MOVE applies to syntactic objects formed by Merge to build new structures. Hence, phrase structure trees are built by the application of Merge to items selected for the derivation, and elements are rearranged within the trees by the application of Move, triggered by feature checking. Chomsky (2000, 2001) subsequently introduced the operation AGREE, which establishes a relation (such as agreement or case checking) between a lexical item and a formal feature. In this approach, Move is taken to be the combination of Merge and Agree.

Movements within the structure are driven by FEATURE CHECKING, and may be of two types: A head may undergo HEAD MOVEMENT and adjoin to another head (forming a complex head), or a maximal projection may move to the Specifier (or Spec) of a head – XP MOVEMENT; in either case, the element moves for the purpose of checking lexically encoded features. In addition, its movement may be OVERT or COVERT. Overt movements are driven by STRONG FEATURES and are visible at PF (phonetic form) and LF (logical form). Covert movements, driven by WEAK FEATURES, are visible only at LF.

Principles of Economy form an important aspect of the MP. One such principle, Full Interpretation (FI), requires that no element lacking a sensorimotor interpretation be admitted at PF; applied at LF, FI entails that "every element of the representation have a (language-independent) interpretation" (Chomsky, 1995, p. 27). Thus, uninterpretable features, which are CASE and the set of φ-features (that is, PERSON, NUMBER and GENDER), must be checked and deleted by LF. Interpretable features are the categorial features and the φ-features of nouns.

A derivation is said to CONVERGE AT AN INTERFACE LEVEL (PF or LF) if it satisfies FI at that level; it CONVERGES if FI is satisfied at both levels. A derivation that does not converge is also referred to as one that CRASHES. If features are not checked, the derivation crashes. Derivations that crash are ill-formed constructions.

Figure 1 illustrates the relationships between components and operations in MP-style syntax (adapted from MacSwan, 2000). It is important to emphasize that Figure 1 illustrates a theory of linguistic competence, not a theory of production. As such, the arrows in the diagram are intended to convey steps in the computation. While an MP approach might well be embedded in a parsing model (as in Stabler, 1997, 2001), it is not itself such a model.

It may be helpful to consider some illustrations of how feature checking drives movement in the MP, a matter of concern in JMSG's paper. As mentioned, movement may be of two types; in head movement, a head (or X^0) moves

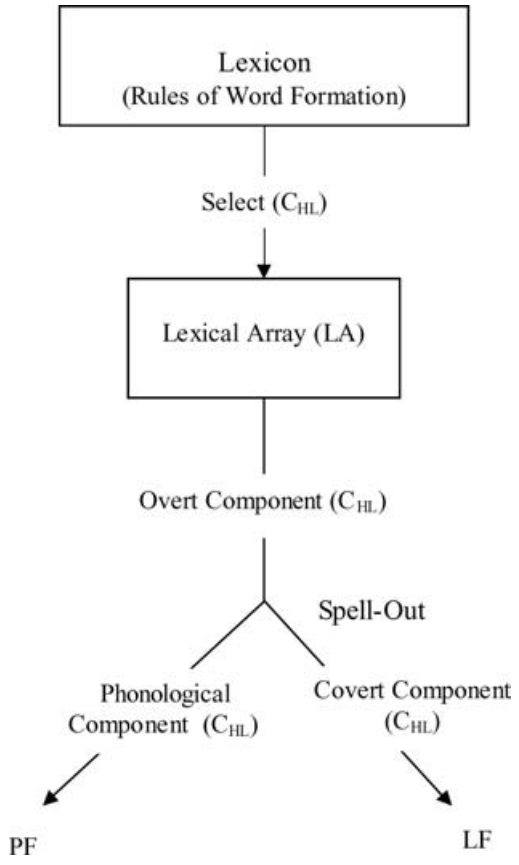


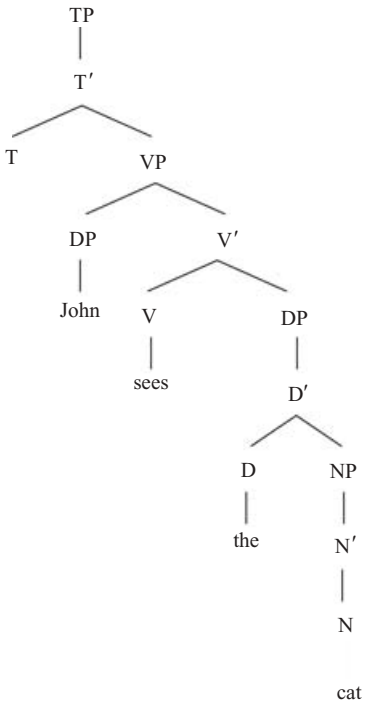
Figure 1. The Minimalist framework.

by adjunction to another head, forming a complex X^0, and in XP movement, an XP moves to the Specifier position of another XP. Movement is driven by a need to check features, and the configuration into which the element moves for feature checking constitutes its CHECKING DOMAIN.

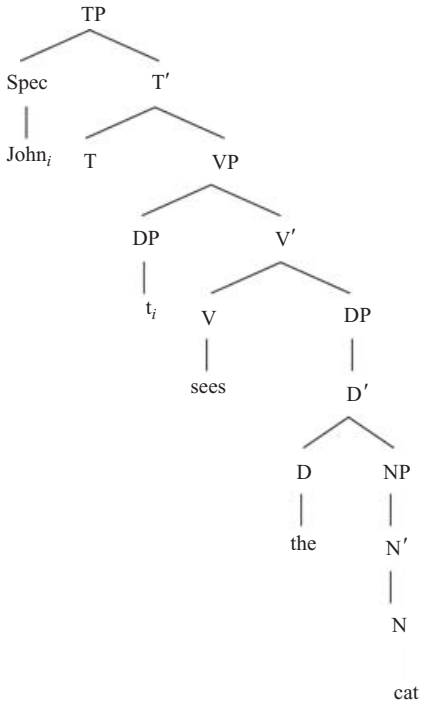
XP movement is illustrated in (1). The tree in (1a) is formed by successive application of Merge, which uses lexically-encoded categorial features to build a classical phrase structure representation. The subject bears a case feature. T is a nominative case assigner which attracts the case feature of the subject to its Specifier position, bringing the full DP along. Because case is an uninterpretable feature, it must be checked and deleted before LF, or the derivation will crash. Hence, the DP is attracted to T, and moves to its Checking Domain, the Specifier of TP, as shown in (1b).[1]

[1] There are some terminological differences and departures from earlier generative models. TP, or Tense Phrase, replaces IP, Infinitival Phrase; DP, or Determiner Phrase, is headed by a determiner and dominates NP, as illustrated.The subject originates in a VP-internal position, following Koopman and Sportiche (1991), raising to the specifier position of T to check its case feature. *t* is the trace of movement, co-indexed by *i*, as in classical approaches. Chomsky (2000, 2001) adopts a bare phrase structure approach, dispensing with intermediate bar levels when they are not relevant to output conditions. As is

(1) a.



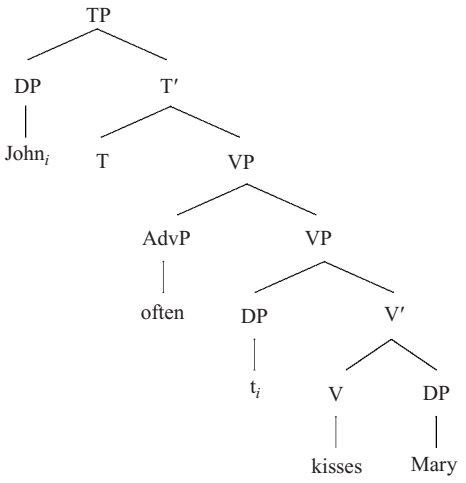
b.



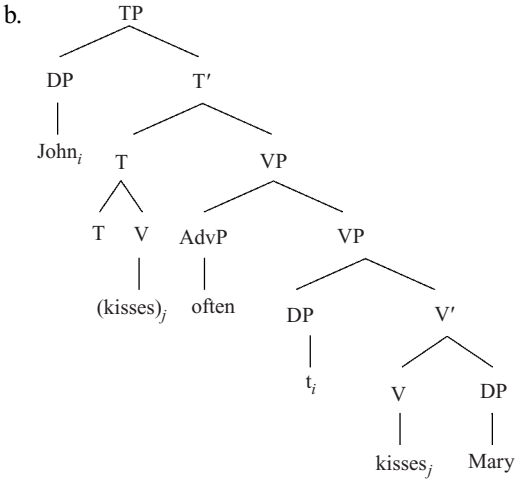
Now consider an example of head movement, illustrated in (2). The successive application of Merge results in the formation of a base structure; as in (1), the subject DP moves out of the VP shell to check its case feature. The resulting structure is shown in (2a). V, a head, moves to T by head adjunction in order to check and delete its φ-features, as shown in (2b).

traditionally done, we have simplified aspects of the structure in (1) (and elsewhere) which are not relevant to the present context.

(2) a.



b.



Feature strength (WEAK, STRONG) is used in the MP to account for crosslinguistic variations in word order. Notice, for instance, the contrast in (3).

(3) a. John often kisses Mary
b. John completely lost his mind
c. Jean embrasse souvent Marie
d. Jean perdit complètement la tête

In English, VP-adverbs precede verbs, but in French they follow them. We might assume, then, that in English V moves to T COVERTLY, attracted by T's WEAK φ-features; this is represented in (2b) with the use of parentheses around the verb, illustrating that the phonetic features of the V have been left behind. By contrast, in French, T has STRONG φ-features, resulting in OVERT movement. In this case, all of V's features raise, with the result that it occurs before its adverbial modifier in (3) (see Emonds, 1978; Pollock, 1989; Chomsky, 1995). (For more details on the MP generally, see Uriagereka (2000).)

1.3. The grammar of codeswitching

As has been frequently noted, the question of where one language ends and another begins is more a matter of

politics than of linguistics. As Chomsky (1995, p. 11) has put it,

> what we call "English", "French", "Spanish", and so on, even under idealizations to idiolects in homogeneous speech communities, reflect the Norman Conquest, proximity to Germanic areas, a Basque substratum, and other factors that cannot seriously be regarded as properties of the language faculty.

This is a significant issue for linguists interested in the study of bilingualism. It suggests that the notion of "a language" should play no role in the formal system employed to account for the data under analysis. Traditionally, CS is taken to be the study of "language switching" and the mechanisms which constrain it, where "language membership" is identified by the syntactic, morphological, and phonological properties of a word or morpheme. In a more basic sense, we might reasonably think of research on CS as the study of how these linguistic systems interact in the mind/brain of a bilingual such that morphological and phonological coding of a lexical item is possible in some contexts and not in others.

The leading aim of the MP is the elimination of all mechanisms that are not necessary and essential on conceptual grounds alone; thus, only the MINIMAL theoretical assumptions may be made to account for linguistic data, privileging simpler and more elegant accounts over complex and cumbersome ones. These assumptions would naturally favor accounts of CS which make use of independently motivated principles of grammar over those which posit rules, principles or other constructs specific to it. Applied to CS, this research program might be stated as in (4), where the minimal CS-specific apparatus is assumed (MacSwan, 1999, p. 146; 2000, p. 43):

(4) Nothing constrains codeswitching apart from the requirements of the mixed grammars.

Notice that (4), an articulation of a particular research program, does not imply that there are no unacceptable code-switched sentences. In (4), *constrain* is used in a technical or theoretical sense, and as such implies that there are no statements, rules or principles of grammar which refer to CS. Put differently, (4) posits that all of the facts of CS may be explained just in terms of the principles and requirements of the specific grammars – and embedded principles of Universal Grammar – involved in each specific utterance under consideration.

The idea that a sound account of CS will appeal to no CS-specific constraints or mechanisms is a natural one in linguistic research. It may be traced back at least as far as Woolford (1983), who noted that a theory of CS consistent with leading assumptions in linguistic theory would include no CS-specific rules. Mahootian (1993) similarly emphasizes that the proper account of CS will make no appeal to CS-specific mechanisms.

Adhering to (4), and drawing upon linguistic research in the MP, MacSwan (1999, 2000) posits that lexical items may be drawn from the lexicon of either language to introduce features into the numeration (or lexical array), which must then be checked for convergence in the same way as monolingual features must be checked, with no special mechanisms permitted. In this lexicalist approach, grammatical requirements are carried along with the lexical items of the respective systems as they enter the derivation. Thus, it makes sense to formalize the grammar used for CS as the UNION of the two lexicons, with no mediating mechanisms.

As indicated in Figure 1 above, at Spell-Out the derivation is split, with features relevant only to PF sent to the phonological component where the derivation is then mapped to π (or PF), and interpretable material treated by further application of the syntactic component is mapped to λ (or LF). The specific differences between the syntactic and phonological components of the grammar become particularly relevant to the analysis of CS data. As mentioned, in an important respect, research on CS is concerned with how subsystems of our linguistic knowledge interact to constrain language switching.

A salient difference between syntax and phonology is that phonological rules are ordered with respect to one another and vary cross-linguistically, as noted by Bromberger and Halle (1989). (In Optimality Theoretic terms, "constraints" are "ranked", or ordered in relative importance, and these rankings vary cross-linguistically; see Kager, 1999.) We have been assuming that CS is formally the UNION of two (lexically-encoded) grammars, where the numeration may draw elements from the union of two (or more) lexicons. Each lexical item imposes certain requirements on the derivation in terms of the encoded features, and syntactic operations need take no notice of what particular language a lexical item is associated with. However, suppose that a PF component PF_1 contains rules ordered such that R1 precedes R2 and R3 precedes R4, and suppose that in PF_2 rules are ordered such that R1 follows R2 and R3 follows R4. Then the union of PF_1 and PF_2 will have no ordering relations for R*n*. In other words, under union (resulting from CS), the PF components would not be able to meet their requirement that they have (partially) ordered rules or constraints. To prevent this breakdown, language mixing within the phonological component must be ruled out.

These observations may be stated more formally in terms of the PF Disjunction Theorem, shown in (5) (MacSwan, 1999, 2000).

(5) *PF Disjunction Theorem*

(i) The PF component consists of rules/constraints which must be (partially) ordered/ranked with respect to each other, and these orders vary cross-linguistically.

(ii) Codeswitching entails the union of at least two (lexically-encoded) grammars.
(iii) Ordering relations are not preserved under union.
(iv) Therefore, codeswitching within a PF component is not possible.

We might think of (5) as an instantiation of the Principle of Full Interpretation (FI), the requirement that every object have a sensorimotor interpretation to qualify as a legitimate representation, a kind of "interface condition". If phonological systems cannot be mixed, then CS at PF generates "unpronounceable" elements which violate FI. Because (5) is a deduction based on the nature of the linguistic system and not itself a principle of grammar, it is termed a "theorem" rather than a "principle". It is an epiphenomenon, not a mechanism.

The question of what constitutes an input to the phonological system is significant here. The intuitive notion is that a change in phonological systems may not occur in the course of a PHONOLOGICAL DERIVATION, understood as the application of a sequence of (partially) ordered rules which are associated with one another by input-output relations. In addition, the relevance of inflectional morphology to the phonological component further suggests that changing phonological systems in the context of such morphophonological material would disrupt the mapping to PF, and hence is also disallowed. Furthermore, following Chomsky (1995), MacSwan (1999, 2000) posited that heads or X^0s are inputs to phonology, suggesting that mixing within an X^0 (say, between bound morphemes) will also be ill-formed. Hence, the PF Disjunction Theorem predicts that CS will not occur in contexts involving phonological derivations, that inflectional material from one language will not be coded by the phonology of another language, and that CS will not occur internally within an X^0.

Furthermore, note that head movement has the effect of creating "complex heads", as may be in seen in (2b) where the head T immediately dominates heads T and V. Thus, if X^0s are inputs to phonology, and if head movement creates complex X^0s, then it is further predicted that CS in the context of head movement is prohibited. Indeed, Chomsky has recently proposed that head movement actually falls within the phonological component rather than the narrow syntax (Chomsky, 2000, p. 146; 2001, p. 37).

We have been speaking freely of the LEXICONS (plural) of the distinct languages involved in CS, but this assumption, too, requires justification. After all, a simpler assumption might hold that there is only one lexicon, and bilingualism is, in some sense, an illusion. Because phonological rules are sensitive to inflectional content, and because the rules of word formation are presumed to be internal to the lexicon in Minimalist syntax, we are faced with two possible assumptions: Either (a) there is a single lexicon, and each lexical item is marked for a specific set of phonological and morphological rules which yield the appearance of one language or another; or (b) the lexical items in a bilingual's repertoire are mentally compartmentalized in some sense, with a specific set of phonological and morphological rules associated with each "lexical compartment". The second alternative requires fewer mechanisms, since the morphophonology is associated with sets of elements to which it applies rather than to individual member. We therefore will assume the latter to be correct. The model sketched here is represented graphically in Figure 2 (adapted from MacSwan, 2000).

We earlier noted that "languages", as such, are theoretically spurious entitles in the sense that "French" or "German" are political rather than grammatical constructs. We now introduce the notion of separate lexicons, one for each language known to a bilingual, raising questions about the previous point. In doing so, however, we have used the term "language" to refer to a collection of formally substantiated morphological and phonological rules which applies to a specific set of lexical items, contained in a LEXICON. Thus, we seek to identify the formal locus of linguistic differences associated with what we commonly call "French", "Spanish", or "Chinese", and to hold these responsible for the facts observed in bilingual CS.

It is important to note that the PF Disjunction Theorem is intended to subsume Poplack's well-known Free Morpheme Constraint given in (6), with some important differences.

(6) *The Free Morpheme Constraint* (Sankoff and Poplack, 1981, p. 5)
A switch may not occur between a bound morpheme and a lexical item unless the latter has been phonologically integrated into the language of the bound morpheme.

The Free Morpheme Constraint is concerned with accounting for the sharp ungrammaticality of expressions such as those given in (7), in which a switch in phonological systems occurs between a stem and a bound grammatical morpheme (MacSwan, 1999, p. 222).[2]

(7) a. *Juan está *eat*-iendo.
Juan be/3Ss eat-DUR
"Juan is eating."

b. *Juan *eat*-ió.
Juan eat-PAST/3Ss
"Juan ate."

c. *Juan *com*-ed.
Juan eat-PAST
"Juan ate."

d. *Juan eat-*ará*.
Juan eat-FUT/3Ss
"Juan will eat."

[2] Please refer to the appendix for the definition of abbreviations used in glosses here and elsewhere.

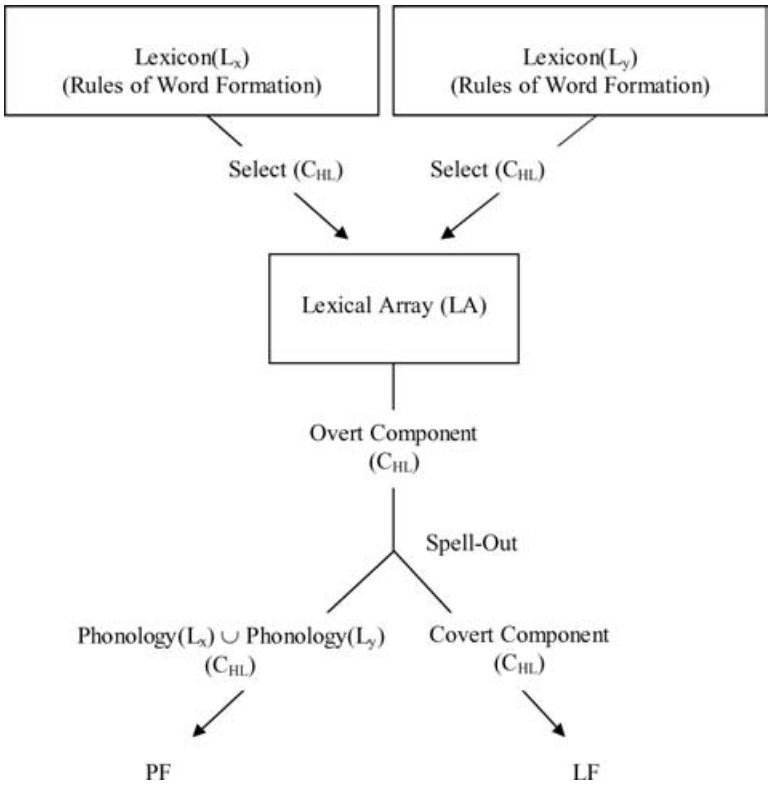


Figure 2. A minimalist approach to codeswitching.

Note, however, that when the stem is phonologically integrated into the language of the inflectional morpheme, as in (8), where the English stem *park* is borrowed into Spanish, no ill-formedness results.

(8) a. Juan está parqueando su coche.
Juan be/3Ss park-DUR his car
"Juan is parking his car."
b. Juan parqueó su coche.
Juan park-PAST/3Ss his car
"Juan parked his car."
c. Juan parqueará su coche.
Juan park-FUT/3Ss his car
"Juan will park his car."

The same facts hold for novel borrowings, which linguists and lexicographers have long called NONCE BORROWINGS. Suppose that the stem *eat* in (7) is phonologically integrated into Spanish, and represented orthographically as *it*. Borrowed words in Spanish generally take *a* as a thematic vowel with an intervening *-e-* (Harris, 1991, 1996), as illustrated in (8). The result is (9), judged to be well-formed by Spanish–English bilinguals (MacSwan, 1999, pp. 222f.; see also Muysken, 2000, p. 54).

(9) a. Juan está iteando su pozole.
Juan be/3Ss it-DUR su pozole
"Juan is eating his pozole."
b. Juan iteó su pozole.
Juan eat-PAST/3Ss su pozole
"Juan ate his pozole."
c. Juan iteará su pozole.
Juan e/3Ss eat-FUT/3Ss su pozole
"Juan will eat his pozole."

The PF Disjunction Theorem thus significantly overlaps with the Free Morpheme Constraint in terms of range of empirical prediction, but there are important differences. The PF Disjunction Theorem predicts switching of phonological systems between a stem and an affix to be ill-formed, but only if the affix is attached presyntactically (by rules of word formation internal to the lexicon). Hence, phrasal affixes are not affected. In addition, the PF Disjunction Theorem predicts CS in head-movement contexts to be ill-formed, moving beyond Poplack's constraints. Further, unlike the Free Morpheme Constraint, the PF Disjunction Theorem is not a principle of grammar, but is a deduction from elementary facts regarding the nature of rule ordering in phonology and syntax. In this sense, it adheres to (4).

Poplack's Free Morpheme Constraint has been controversial. While it is attested in numerous corpora (Bentahila and Davies, 1983; Berk-Seligson, 1986; Clyne, 1987; MacSwan, 1999), others claim to have identified some counter-examples (Nartey, 1982; Bokamba, 1989;

Myers-Scotton 1993; Halmari, 1997; Chan, 1999; Hlavac, 2003). However, in presenting counter-examples, researchers have often given too little attention to the specific phonological, morphological and syntactic characteristics of the examples cited, making it difficult to determine whether they are in fact violations. Examples in which an other-language stem has been phonologically integrated into the language of an inflectional affix, as in (8) and (9), do not constitute counter-examples to either the Free Morpheme Constraint or the PF Disjunction Theorem. Indeed, they provide confirming evidence. Below, we define the formal characteristics of genuine counter-examples to the PF Disjunction Theorem in the context of some proposed counter-examples advanced in JMSG.

Analyses of CS data supporting the approach advanced here may be reviewed in MacSwan (1999, 2000), where it is concluded that the facts of CS may be explained just in terms of the principles and requirements of the specific grammars used in each specific utterance under consideration, as proposed in (4).

Let us now examine the MLF model and consider its empirical and theoretical efficacy.

2. A critique of the MLF model

The basic mechanisms of the MLF model are given in (10) and (11).

(10) *The System Morpheme Principle*
All syntactically relevant system morphemes must come from the ML.
(Myers-Scotton, 1993, p. 7; cf. JMSG, p. 73)

(11) *The Morpheme Order Principle*
Morpheme order must not violate ML morpheme order.
(Myers-Scotton, 1993, p. 7; cf. JMSG, p. 73)

"Syntactically relevant system morphemes" are defined as "late outsider system morphemes" – that is, morphemes which "look outside of their maximal constituents for information about their form and distribution" (JMSG, p. 74; cf. Myers-Scotton and Jake, 2001, p. 100).

Naturally, the so-called "matrix language" (ML) must be clearly identifiable if (10) and (11) are to have real empirical content. In earlier work, Myers-Scotton (1993) proposed a "frequency-based criterion", claiming that the ML was the language which contributed the greater number of morphemes to the discourse (p. 68), and that "the ML may change across time, and even within a conversation" (p. 69). This way of casting the ML made it difficult to know, for any given utterance, which language functioned as the ML and which as the EL. Concern over the vagueness of the definition of the ML has been voiced in Bentahila (1995), Muysken and de Rooij (1995), Backus and Boeschoten (1996), MacSwan (1999, 2000), and Muysken (2000).

In more recent work (Myers-Scotton and Jake, 2001; JMSG), this weakness has been addressed with the introduction of a structural definition of the ML: "The ML may change within successive CPs, even within a multi-clausal sentence, but we stress that the ML does not change within a single bilingual CP" (JMSG, p. 73). This structural definition of the ML permits the MLF model to make empirical predictions that can be verified by inspecting the structure of code-switched utterances. In short, below the CP, the MLF model predicts that all "late outsider system morphemes" will be from one language only (System Morpheme Principle), and all constituents will be in the order required by the ML (Morpheme Order Principle).

As an additional nuance, Myers-Scotton and colleagues stipulate that an "EL island" may occur below the CP: "[A]s well-formed maximal constituents in the EL, [EL islands] are not inflected with ML system morphemes, although they occur in positions projected by the ML, following the Morpheme Order Principle" (JMSG, p. 77). Hence, EL islands are essentially lawful violations of the System Morpheme Principle because they contain grammatical morphemes that are not in the ML, but an EL island must be a maximal projection and must remain true to the Morpheme Order Principle (that is, its position within the utterance must be dictated by the ML) (also see Myers-Scotton, 1993, pp. 137–147).

2.1. The System Morpheme Principle

The System Morpheme Principle in (10) predicts that all late outsider system morphemes within a CP will be contributed by one language only, labeled the ML. Consider (12), in which both examples involve identical DPs.

(12) a. ¿Tus *coworkers* no tuvieron vacaciones
your coworkers no have/3P/PAST vacation-PL
todavía, verdad?
yet right
"Your coworkers haven't had a vacation yet, right?"
b. Tus *coworkers haven't had a vacation yet, right*?
your coworkers haven't had a vacation yet right
"Your coworkers haven't had a vacation yet, right?"

An agreement relation holds between the D in (12a) and its nominal complement, making *tus* a late outsider system morpheme. The structure violates the System Morpheme Principle because system morphemes are contributed by both languages. Note that [*tus coworkers*] is not an EL island, as it involves mixed constituents below the DP, its maximal projection. Note that both structures in (12) are well-formed regardless of how the ML is identified, suggesting that the ML construct does not play a role in determining grammaticality.

Further, consider the contrast between (13) (from Di Sciullo, Muysken and Singh, 1986) and (14) (from Belazi, Rubin and Toribio, 1994).

(13) Oui, alors j'ai dit que *si potev* aller comme ça.
yes so I have said that REF could walk like that
"Yes, so I said that we could go like that."

(14) *The students had *visto la película italiana*.
"The students had seen the Italian movie."

In (13), an Italian aspectual and its reflexive clitic are used in an otherwise French construction. Note that the Italian aspectual, being the highest V, selects a French VP complement. The V is (or contains) a late outsider system morpheme because it is related to C by Merge (a subcategorization relation), which is outside its maximal projection. Therefore, (13) contains late outsider system morphemes contributed by both French and Italian, a violation of the System Morpheme Principle; [*si potev*] cannot be analyzed as an EL island because [*aller comme ça*] is contained within its VP. Thus, the System Morpheme Principle predicts (13) to be ill-formed, contrary to the facts. Note that (13) is structurally similar to (14), and yet its acceptability contrasts with it. Yet, the MLF model predicts (14) to be well-formed, since its VP [*visto la película italiana*] is an EL island (Spanish) embedded in an otherwise ML CP (English).

These facts suggest that the System Morpheme Principle is not descriptively adequate. It should be noted, however, that Myers-Scotton (1993) and JMSG allow "internal EL islands", defined as "a constituent in the EL made up of EL morphemes following EL morpheme order, but smaller than a maximal projection" (JMSG, p. 76). This stipulation would sanction (12), for example, as well-formed constructions in the MLF model, because in (12a) *coworkers* may be analyzed as an "internal EL island" and in (12b) *tus* may be so analyzed. However, notice that essentially any configuration of morphemes may be analyzed in these terms – (14) notwithstanding – predicting all CS constructions to be well-formed, contrary to the facts once again.

JMSG attempt to recast the System Morpheme Principle as the UNIFORM STRUCTURE PRINCIPLE, defined as follows:

> A given constituent type in any language has a uniform abstract structure and the requirements of well-formedness for this constituent type must be observed whenever the constituent appears. In bilingual speech, the structures of the matrix language are always preferred. (JMSG, p. 72)

The Uniform Structure Principle opens with a clear truism: Well-formedness requirements apply to any constituent type (DP, VP, D, T, or whatever) to determine its structure. The structure might be "uniform" in the sense that it does not randomly vary, and it might be "abstract" in the sense that the structural requirements are not directly observable. But how does it follow from these facts that "the structures of the matrix language are always preferred" – in other words, that the System Morpheme Principle generally holds? That claim is simply asserted without argument.

The authors additionally attempt to glean support for the System Morpheme Principle by citing Chomsky's (2001, p. 2) reference to the Uniformity Principle: "In the absence of compelling evidence to the contrary, assume languages to be uniform, with variety restricted to easily detectable properties of utterances". According to JMSG, in order for the values of uninterpretable features in bilingual speech to "satisfy the Uniformity Principle by being easily detectable, system morphemes must come from the ML" (p. 85). However, the Uniformity Principle, which has a long history in linguistic research, is not a principle of grammar, and could not be; rather, it is a principle of linguistic analysis. It guides researchers to regard the underlying structure of diverse languages to be uniform unless there is learnable (detectable) evidence to the contrary. Hence, the theoretical support which the System Morpheme Principle sought from Chomsky (2001) is similarly unavailable.

2.2. The Morpheme Order Principle

Now consider (11), the Morpheme Order Principle, which predicts that morpheme order within a single CP will not violate the morpheme order of the ML, or of the language contributing the late outsider system morphemes. Mahootian (1993) noticed that in Farsi–English CS, word order will reflect Farsi (OV) if a Farsi verb is present and will reflect English (VO) if an English verb is present, as illustrated in (15).[3]

(15) Tell them you'll buy *xune-ye jaedid* when
tell them you'll buy house-POSS new when
you sell your own house.
you sell your own house
"Tell them you'll buy a new house when you sell your own house."

The example in (15) appears to be consistent with the MLF model; its word order is English, the ML, and [*xune-ye jaedid*] may be considered an EL island. However, we may enrich the structure slightly by adding an English PP within the object DP; because the DP then contains elements from both English and Farsi, it no longer qualifies as an EL island. The resulting structure, judged to be well-formed,[4] thus contains system morphemes

[3] For discussion of word order facts in CS, see MacSwan (2003, 2004).

[4] The Farsi–English example in (16), a modification of Mahootian's (1993) datum shown in (15), was judged to be well-formed by a

contributed by both languages; (16) is therefore a counter-example to the Morpheme Order Principle.

(16) Tell them you'll buy *xune jaedid* with blue shutters when you sell your own house.
"Tell them you'll buy a new house with blue shutters when you sell your own house."

Similarly, consider the Spanish–English data in (17), judged to be well-formed by Spanish–English bilinguals (reported in MacSwan, 2004).

(17) a. ¿Funciona *the computer* de tu hermano
function-3Ss the computer of your brother
en la oficina?
in the office
"Does your brother's computer in the office work?"
b. ¿Quiere *the boy* del pueblo
want-3Ss the boy from-the town
algo de comer?
something to eat
"Does the boy from the town want something to eat?"

Note that [*the computer*] in (17a) and [*the boy*] in (17b) cannot count as EL islands because each contains an other-language prepositional phrase within its maximal projection. Here again we see violations of the System Morpheme Principle, involving a system morpheme of the EL mixed in with system morphemes of the ML.

Also consider the examples in (18) (MacSwan, 1999). In (18a), the ML would appear to be Spanish, and the VP [*wants to go*] an EL island. Similarly, English is the ML in (18b), and [*quiere ir*] an EL island. But both are ill-formed, despite their adherence to the Morpheme Order Principle.

(18) a. *El no *wants to go*.
he not want to go
"He doesn't want to go."
b. *He doesn't *quiere ir*.
he doesn't want/3Ss go/INF
"He doesn't want to go."

In Spanish–Nahuatl CS, however, this configuration is ungrammatical in the case of Nahuatl negation, but not in the case of Spanish negation, as shown in (19) (MacSwan, 1999, 2000). Again, the VP in both instances could be analyzed as an EL island, and the structures should be well-formed according to the MLF model, contrary to the facts.

(19) a. *No *nitekititoc*.
no ni-tekiti-toc
not 1S-work-DUR
"I'm not working."
b. Amo *estoy trabajando*.
not be/3Ss work-DUR
"I'm not working."

Apart from making incorrect empirical predictions in the case of (19a), the MLF model is silent about the asymmetrical acceptability of cases such as (19), just as it was regarding the contrast noted in (13) and (14).

Also consider (20) (MacSwan, 1999). Here the ML is presumably Nahuatl, with EL islands as subject DPs, predicting both examples to be well-formed – again, contrary to the facts.

(20) a. *Tú *tikoas* *tlakemetl*.
tú ti-k-koa-s tlake-me-tl
you/SING 2S-3Os-buy-FUT garment-PL-NSF
"You will buy clothes."
b. Él *kikoas* *tlakemetl*.
él 0-ki-koa-s tlak-eme-tl
he 3S-3Os-buy-FUT garment-PL-NSF
"He will buy clothes".

Another significant problem for the MLF model is (7), discussed earlier, in which Spanish uniformly serves as the ML, yet the presence of an EL stem renders the examples ill-formed, even though these constructions comply with the Morpheme Order Principle and the System Morpheme Principle.

The examples just considered recall Chomsky's definition of a grammar, discussed at the outset: A descriptively adequate grammar must generate all the well-formed utterances and none of the ill-formed ones, and must be constructed in light of both positive and negative linguistic evidence. The MLF model licenses numerous ill-formed CS constructions, but its proponents fail to notice this fact because there is an exclusive reliance on positive evidence (Myers-Scotton, 1993, 2002; JMSG).

As important as these empirical limitations, however, are the theoretical problems associated with the MLF model. As we will emphasize below, in the context of responding to some of JMSG's criticisms of MacSwan (1999, 2000), the MLF model introduces mechanisms and concepts into the analysis of CS which are foreign to the theory of grammar. We will argue that these limitations suggest a defect in principle, moving well beyond the need for greater empirical adequacy, and that further tinkering is unlikely to solve the deep-rooted problems of the MLF model.

33-year-old Farsi–English bilingual who learned Farsi as an infant and English as an adolescent, and who resides in Central Arizona.

3. A response to JMSG's critique of the minimalist approach to CS

JMSG attempt to reconcile the MLF model with recent work in generative grammar in the context of a critique of "earlier minimalist approaches to codeswitching", focusing on the work of MacSwan (1999, 2000). Some aspects of this effort, which they call "a modified minimalist approach", have already been mentioned; here we discuss in detail the authors' proposed minimalist theories – including their central idea, that the ML/EL distinction is a necessary component of any successful minimalist account of CS data. We do so in the context of responding to some of JMSG's criticisms of MacSwan (1999, 2000).

3.1. The alleged uniform ban on singly occurring code-switches

JMSG attribute to me the view that "singly occurring code-switches"[5] are uniformly banned. They devote considerable energy to this point, a major focus of their paper. However, I DO NOT BELIEVE THIS TO BE TRUE, and it is nowhere stated or suggested in my work. Indeed, I do not ever use the phrase "singly occurring code-switches", and I report more than 62% of the utterances in my published corpus of Spanish–Nahuatl CS data to be of this type, and provide analyses of such forms from Spanish–Nahuatl, Spanish–English, French–Italian, and other language pairs (MacSwan, 1999).

Nonetheless, JMSG charge,

> Most applications of MP to CS to date (e.g. MacSwan 1999, 2000) attempt to account for switching from one monolingual constituent to another; they tend to dismiss the more theoretically challenging problem of accounting for singly occurring forms from one language within a constituent structured by another. (JMSG, p. 70)

And further,

> Some applications of MP to CS claim that in bilingual constituents most singly occurring forms from another language are simply part of monolingual constituents, regardless of inflection (MacSwan, 1999); any features such singly occurring EL forms have that are different from those of the ML are not relevant to the derivation. The upshot of a position such as MacSwan's is that how these singly occurring forms are integrated into a bilingual structure does not need to be accounted for under a CS model. (JMSG, p. 71)

Although I cannot be certain of the origin of this error, the confusion may have arisen from a misunderstanding of the PF Disjunction Theorem and its consequences for linguistic theory. The PF Disjunction Theorem makes two major predictions, as discussed earlier: (a) CS may not occur word-internally, where WORD is understood as a lexical head (X^0) whose morphological composition has been determined internally within the lexicon; and (b) CS may not occur below a complex X^0, where BELOW is understood in terms of hierarchical syntactic structure, typically represented as a branching tree, as shown in (1) and (2) above. The broader generalization is that CS is prohibited within anything which an X^0 dominates (a word or other X^0s). As for adjacent items, the syntactic reflexes of the PF Disjunction Theorem are limited to the prediction that CS cannot occur if such adjacent items can be described in these structural terms.

Nevertheless, it should be reasonably clear from both documents (MacSwan, 1999, 2000) that I do not believe, as JMSG assert, that singly occurring code-switches "are simply part of monolingual constituents, regardless of inflection" (p. 71). Quite the contrary: I adopt the traditional view that a borrowed word is an item which has been "lexically, syntactically, morphologically and phonologically incorporated into a host language" (MacSwan, 1999, p. 59),[6] as illustrated in (8) above. I use the term NONCE BORROWING to refer to a non-established borrowing, illustrated in (9). For the purposes of a synchronic theory of language contact, the distinction between BORROWING and NONCE BORROWING is unimportant: The difference in meaning depends on a word's history – inaccessible to a linguistic system represented in the mind/brain of an individual.

JMSG contend that the MLF model accounts for both borrowing and CS without the need for a separate theory of borrowing. In the MLF model, it is maintained that the phonological properties of morphemes are unconstrained, and that the principles developed to account for the distribution of content and system morphemes under the MLF model account for both kinds of language contact phenomena (Myers-Scotton, 1993; JMSG), avoiding the "weaker and less explanatory" accounts which resort to borrowing (JMSG, p. 72).

Note, however, that the term BORROWING (as well as NONCE BORROWING) is used to define a particular array of empirical properties (primarily phonological and morphological); it is therefore a definition of a set of facts, and not a theory about them. With respect to these properties, the theoretical challenge in contact linguistics is to

[5] Although JMSG do not explicitly define the term, I use the examples cited in their paper to infer the obvious meaning of an utterance in which an individual word appears with at least one other-language word immediately before it and immediately after it, in linear order.

[6] The definition needs to be slightly more nuanced, but that is not important for present purposes. The definition quoted here is actually for words described as "fully borrowed", while "partial borrowings" are more curious entities which may have the phonological matrices of one language, but the semantic or syntactic features of another. See MacSwan (1999, sections 2.3.2 and 5.3.1.7), for discussion.

account for the grammaticality contrast illustrated in (21).[7]

(21) a. Mañana voy a lonchar ([lončaɾ]) con Mario.
b. *Mañana voy a *lunch*-ar ([lʌnčaɾ]) con Mario.
c. *Mañana voy a lunchar ([lʌnčaɾ̃] or [lʌnčəɾ])
tomorrow I.be.going.to lunch-INF
con Mario.
with Mario
"Tomorrow I'm going to have lunch with Mario."

Here we have an English-origin verb stem, *lunch*, with three different phonological representations: (21a) is phonologically pure Spanish, including flapped /r/ and /o/ substituted for caret (ʌ) in English *lunch* ([lʌnč]); (21b) occurs with mixed English–Spanish segments involving a verb stem pronounced in English and associated inflectional suffix in Spanish; and (21c) is phonologically pure English, on both the English-origin verb stem and the Spanish suffix. While (21a) is a common expression in Southwest U.S. Spanish, (21b) and (21c) are strongly deviant.

The PF Disjunction Theorem captures this contrast by presenting a deduction from the nature of the rule system in the phonological component which bans word-internal switching of phonology, strongly associating morphological and phonological operations and distinguishing them from those of the syntactic system – a move that is consistent with Minimalist theory. The contrast in (21) is a fact about language contact which any reasonable proposal – particularly one which claims to account for both CS and borrowing – must take into account. However, contrasts such as (21) remain mysterious in the MLF model: The structures are morphologically and syntactically identical, and hence are uniformly licensed or banned by the System Morpheme Principle (10) and the Morpheme Order Principle (11), both of which are silent about the phonological properties which appear to have a significant impact on the acceptability of (21).

Furthermore, JMSG suggest that I call singly occurring code-switches "borrowings" as a way of insulating them from linguistic analysis, so that I may happily proceed without the need to work on the "more theoretically challenging" task of constructing a theory which accounts for these forms as well as others in the CS literature. But this is false. The PF Disjunction Theorem IS an account of such items, with a clear prediction of their distribution: A word's "language membership" is epiphenomenal of lexically-encoded morphological and phonological features; a word has membership in a language L and not L′ if it has the morphological and phonological characteristics of L. Hence, rather than claiming that such items are borrowed "regardless of their inflections" (JMSG, p. 71), my analysis advances the notion that their inflectional material constrains the selection of phonological systems used in the mapping to PF, as illustrated in (21).

The bulk of JMSG's proposed counter-examples to the PF Disjunction Theorem are of the singly occurring variety. These do not constitute counter-examples to the PF Disjunction Theorem, and therefore specific discussion of these examples is not warranted. However, JMSG do present a set of five candidate counter-examples which I will discuss directly.

3.2. JMSG's proposed counter-examples to the PF Disjunction Theorem

A counter-example to the PF Disjunction Theorem should exhibit switching of phonological systems below X^0 – that is, either within a lexical item – possibly morphologically complex, formed by rules of affixation internal to the lexicon – or between adjoined syntactic heads. JMSG present five examples which have the appearance of legitimate counter-examples, and which therefore merit individual discussion.

First, let us consider a Croatian–English example which JMSG gleaned from Hlavac (2000, p. 392), published as Hlavac (2003, p. 207).

(22) . . . i tako pak-ujem one . . .
and so pack-1S/PRES one
kontejner-e i tako dalje . . .
container-M/PL/ACC and so on . . .
"And so [I] pack those — containers and so on . . . "

According to JMSG, the example in (22) "represents a phonologically unintegrated noun in Croatian-English CS" (p. 76). However, Hlavac (2000, 2003) explicitly represents *kontejner-* and *pak-* in (22) as involving English-origin stems which have been PHONOLOGICALLY INTEGRATED into Croatian,[8] so that the phonology of the stem matches the phonology of the inflectional affix, as the PF Disjunction Theorem predicts. We may therefore disregard JMSG's proposed counter-example in (22) as originating in a misinterpretation of Hlavac's work; to the contrary, (22) is an instance of positive evidence which confirms the PF Disjunction Theorem.

[7] Grammaticality judgments for (21) and (31) were obtained from two early sequential Spanish–English bilinguals, now young adults, who learned English at age 5 and reside in Central Arizona.

[8] See Hlavac (2000, p. 79; 2003, p. 36) where the transcription system employed is explicated. Words which are phonologically unintegrated into Croatian are represented in Hlavac's work with English orthography; in (22), *kontejner* is spelled in Croatian, indicating that it is pronounced using Croatian phonology, not English. Hlavac (2000, p. 116; 2003, p. 66) even uses *kontejnere* in this example in his analysis of specific phonemic substitutions of English vowels with Croatian vowels under conditions of phonological integration. Note that JMSG (p. 74) introduce square brackets around these stems ([*pak*] and [*kontejner*]), traditionally used to suggest phonetic transcription, but brackets are not used in Hlavac's text.

Next, citing co-author Gross's (2000) dissertation, JMSG (p. 76) report,

> Bilingual speakers of German and English living in the United States may inflect an English past participle with a German past participle prefix *ge-*; however, English phonology is retained. For instance, in his study of six adult immigrants, Gross (2000) finds examples such as (*ge-cured*). The German rule for word-final obstruent devoicing does not apply to the English past participle. That is, *cured* has a voiced final stop. Note also that the verb stem retains its English phonology [k^h*uwrd*].

This is an informative example, since past participles are generally assumed to be formed internally within the lexicon. The example would suggest that CS may occur internally within an X^0, violating the PF Disjunction Theorem.

Regrettably, however, the *gecured* datum cannot be found in the work cited (Gross, 2000). Gross (2000) gives only one example of a mixed past participle, shown in (23). Although phonetic details are not presented, the orthography suggests that the English verb stem *fix* has been phonologically integrated into the language (Middle Bavarian German) of the inflectional morphology (with final devoicing of the obstruent represented as *t*),[9] as predicted by the PF Disjunction Theorem.

(23) Do hon sie schon die Telefonleitung *kfixt* khobt. (Gross, 2000, p. 56)
"They had already fixed the telephone line."

The *gecured* datum may therefore be dismissed as spurious.[10] However, note that this item, were it available, would also constitute a counter-example to the MLF model. Past participial morphology (German *ge-* . . . *t*, English *-ed/-en*) counts as a "late outsider system morpheme" because it has grammatical relations external to its maximal projection which determine its form (the Merge or subcategorization relation of the main V, which selects it). The mixture of English and German participial morphology in *gecured* would therefore violate the System Morpheme Principle (10).

Nonetheless, in an effort to probe the plausibility of JMSG's assertion regarding this form, we presented the *gecured* example to two German–English native bilinguals, interviewed separately,[11] who expressed a strongly negative reaction to this and other items produced with final voicing of /d/ as [d], as in English; these consultants found the *gecured* example acceptable only if its phonology conformed entirely to German. These results are also consistent with Fuller's (1999) findings, exemplified in (24) and (25); here we see devoicing of final obstruents occurring with English verbs if and only if *ge-* is present; the orthography suggests that the stems are phonologically mapped to German in (24b) and (25b), consistent with the PF Disjunction Theorem. Data in (24a) and (25a) are traditional code-switches.

(24) a. Mer hen *farmed* mit Geil.
we have farmed with horses
"We farmed with horses."
b. Mer hen *'bout three years* ge-farm-t.
we have [a]bout three years farm-ed
"We farmed about three years."

(25) a. Ein jung *Bu*, *Greg*, *moved* vun Missouri.
a young boy Greg moved from Missouri
"A young boy, Greg, moved [here] from Missouri."
b. Er is nach Florida ge-move-t.
he be/3Ss to Florida move-d
"He moved to Florida."

Finally, Cantone (2003) reports the examples in (26) from her Italian–German corpus, patterning with (24a). These contrast in acceptability with forms involving any mixture of German and Italian part participial morphology, such as **gecomprato* "bought" and **gestata*, judged to be ill-formed by Cantone (personal communication), a simultaneous Italian–German bilingual.

(26) a. a-adesso è *gelandet*.
now is/3Ss ge-land-ed
"Now it is landed."
b. noi abbiamo *gewonnen*.
we have won
"We won."

Hence, examples of German–Italian CS further corroborate the claim that the German part participial form *ge-* . . . *t* cannot co-occur with a phonologically unintegrated stem from another language, contrary to JMSG's assertion.

In conclusion, the *gecured* example, while potentially interesting, does not appear in the reference cited, and therefore cannot be inspected. However, the available evidence indicates that the conclusions reached by JMSG regarding CS in German past participial contexts are false, and that the presence of German verbal morphology triggers German phonological coding on an incorporated stem, consistent with the PF Disjunction Theorem.

Let us next consider a Swahili–English example which JMSG (p. 74) gleaned from Myers-Scotton's unpublished Nairobi corpus:

(27) na sisi tujaribu kutenda yale tunayofikiri kwamba tumeyasomea, halafu m-tu-*evaluate*.

[9] English past tense /d/ would be devoiced in this context as well.

[10] The item may be part of Gross's larger unpublished corpus, or may be an informal recollection, approximately cited as his dissertation. Given the detailed claims about the unusual phonetic shape of this word, and that it is cited as an "instance" of a more general pattern, we reasonably expect to see a non-anecdotal report involving at least one documented occurrence.

[11] The consultants are young adult German–English simultaneous bilinguals, ages 18 and 20, who have actively used both languages since birth and who reside in Central Arizona.

> "And we would try to do those things which we thought we were trained for, then you should evaluate us."

JMSG note that *evaluate* is pronounced with English phonology, not Swahili. Hence, it appears that Swahili agreement prefixes *m-* and *-tu-* have affixed to an English-origin verb stem pronounced with English phonology, an apparent violation of the PF Disjunction Theorem and at odds with other evidence so far considered.

To illuminate the structure of (27), consider the facts in (28) and (29), with grammaticality judgments rendered by a Swahili–English bilingual.[12] Here, as in (27), the English stems (*evaluate*, *call*) are pronounced with English phonology.

(28) a. Halafu m-tu-*evaluate*.
then you-us-evaluate
"Then you evaluate us."
b. *Halafu a-tu-*evaluates*.
then 3PS-us-evaluates
"Then he/she evaluates us."
c. *Halafu m-(li)-tu-*evaluated*.
then you-(PAST)-us-evaluated
"Then you evaluated us."

(29) a. Halafu m-tu-*call*.
then you-us-call
"Then you call us."
b. *Halafu a-tu-*calls.*
then 3PS-us-calls
"Then he/she calls us."
c. *Halafu m-(li)-tu-*called*.
then you-(PAST)-us-called
"Then you called us."

Notice that the constructions are acceptable if and only if the English stem is barren of any inflectional material.

The facts in (27)–(29) may be reconciled by appealing to independently motivated analyses of grammatical structure. First, let us assume with Cocchi (2000) that Swahili agreement morphemes are clitics rather than agreement affixes, as one would find in the Romance languages. Clitic constructions represent an active area of linguistic research, but a common and natural idea is that clitics are Ds which adjoin to verbs as a byproduct of V-to-T raising, resulting in the formation of a complex X^0 which includes them. In the context of analyzing a variety of English verb phenomena, Pollock (1994) proposed that English bare forms do not undergo V-raising: Having no overt morphological features to check, bare V remains in situ and never adjoins to T. We posit that V raising does not occur in the case of the English bare stems in (27)–(29), following Pollock's (1994) proposal. As a result, the clitic does not incorporate with T, so that no mixed-language complex X^0 is formed in cases involving bare stems. However, inflected verbs do require checking, and hence raise to T; clitics are incorporated as inflected V climbs to T, forming a mixed-language complex X^0. The latter are predicted to be ill-formed by the PF Disjunction Theorem, as attested.[13] We conclude that (27) is not a counter-example to the PF Disjunction Theorem, as the presence of the bare stem does not correspond to CS within a complex X^0.

Now consider a Brussels Dutch–French example which JMSG report from Treffers-Daller's (1994, p. 152) data:

(30) Da's ne *sympathiqu*-e.
that's a likeable (one)
"That's a likeable one."

Here the French adjective *sympathique*, pronounced in French, is used with a Dutch agreement suffix *-e*, pronounced in Dutch as schwa.

Zwart (1996) argues that the Dutch adjectival agreement suffix *-e* is a phrasal affix akin to English *-'s* (genitive case). The phrasal nature of English *-'s* is evident in expressions like [*the man from Nebraska*]*'s hat* and [*Tom and Mary*]*'s house*. Thus, on Zwart's analysis, Dutch *-e* in (30) is a phrasal affix, and does not involve presyntactic word formation or head movement.

Consider further the examples of Spanish–English CS in (31)[14] and of Croatian–English CS in (32) (Hlavac, 2003, p. 165); here an English genitive *-'s* occurs after an other-language N.

(31) a. Su novia*'s* coche está nuevo.
his girlfriend-GEN car is new
"His girlfriend's car is new."
b. Mi cuñado*'s* *motorcycle is in the driveway*.
my brother-in-law-GEN motorcycle is in the driveway
"My brother-in-law's motorcycle is in the driveway."

(32) . . . imam, moja mamin*'s* sestra je tu i . . .
I have my mother's sister is here and . . .
sve moj tata*'s* *family* je sve u Zagreb . . .
all my dad's family is all in Zagreb

These examples are well-formed, contrasting with (7), (21b) and (21c) above. If we assume that CS is banned within a lexical item formed by rules of word formation internal to the lexicon, or instances of head adjunction,

[12] Grammaticality judgments for (27)–(29) were obtained from a Swahili–English adult bilingual who resides in Los Angeles and teaches Swahili at a research university. She is a native speaker of Meru, and learned Swahili in primary school and English in secondary school.

[13] Of course, T must discharge its φ-features even when the verb stem is bare. See MacSwan (1999, section 5.2.2.4), where Schütze's (1997) Accord Maximization Principle is invoked in this connection regarding other data.

[14] See note 7.

then (31) and (32) are predicted to be well-formed, as attested, since these items involve phrasal (XP-level) affixation. These observations reconcile (30) with other data so far considered, and suggest that (30) is not a counter-example to the PF Disjunction Theorem.

Finally, JMSG offer (33), taken from Halmari's (1997, pp. 80, 156) American Finnish–English corpus:

(33) a. Mä oon sii-nä *green costume*-i-ssa.
I am it-INES green costume-SF-INES
"I am in that green costume."
b. Onks sulla vähän *napkin*-eitä.
have-Q you-ADE some napkin-PL/PART
"Do you have some napkins?"

Here CS appears to occur between the Ns (*napkin*, *costume*) and their case affixes.

Finnish has attracted considerable recent interest because of its extensive case system and atypical word order. The Finnish case system involves sixteen distinctions, more broadly categorized as STRUCTURAL (nominative, genitive, partitive, accusative) and SEMANTIC (translative, essive, inessive, elative, illative, adessive, ablative, allative, absessive, comitative, instructive, prolative) cases. In addition, Finnish uses subject–verb agreement, possessive agreement, case and number agreement in the DP, and number agreement in nominal predicate clauses; it has relatively free word order, and uses both postpositional and prepositional phrases (see Holmberg and Nikanne, 1993).

The structure of Finnish remains an active area of linguistic research, with many aspects still very controversial. However, some interesting proposals regarding the structure of the Finnish case system have been advanced. Nikanne (1993), for example, has analyzed Finnish semantic cases like the inessive as postpositional phrases, proposing that a phonetically null postposition heads an N morphologically marked as inessive. We might reasonably propose, in an effort to maintain a Universal Base consistent with recent proposals, that *-ssa* in (33a) is a lexical P, and that its DP complement overtly raises to the Specifier position of PP to check some formal feature F, semantic in nature. Vaninikka (1993) similarly proposes an analysis of the Finnish genitive, with the case-marked DP base-generated in the Specifier position of a PP. We can avoid the difficulty of accommodating a base-generated Specifier position within the MP by assuming a structural analysis similar to the one proposed for the inessive: A genitive case ending is a lexical P which attracts its DP complement to its Specifier position. We might in fact extend this analysis to other Finnish cases as well, perhaps entertaining the possibility that Finnish is a pronominal argument language in the sense of Jelinek (1984) and Baker (1996). Thus, following a modification of Nikanne's (1993) analysis, we assume that the inessive construction in (33a) is a postpositional phrase. We further provisionally assume that the analysis might be extended to other Finnish cases, including the partitive. If correct, then the examples in (33) do not violate the PF Disjunction Theorem, as the case endings appear to involve phrasal affixes.

In sum, JMSG's proposed counter-evidence to the PF Disjunction Theorem has rested on a misinterpretation of the proposal, which does not predict that singly occurring code-switches are uniformly banned, and on misinterpretations of data in Hlavac (2000) and Gross (2000). A careful analysis of the syntactic properties of the three legitimate candidate counter-examples in JMSG has shown that the examples given do not appear to evidence CS below X^0, and therefore are not counter-examples of the PF Disjunction Theorem.

Certainly, however, careful and informed critique of the proposed relationship between phonology, morphology and syntax in bilinguals is welcome, and we can be confident that further inquiry will lead to refinement of the PF Disjunction Theorem as originally formulated. As is widely admitted, there has been little focused research on the phonological properties of CS, and the relationship of morphology to other aspects of grammar remains an active area of research. Continued study, informed by our evolving understanding of relevant data and the theory of grammar, may point the way to significant future discoveries, leading to reformulation of existing notions. However, the MLF model and its recently proposed revisions do not illuminate the path, as we further show below.

3.3. "Absolute" versus "non-distinct" feature checking in JMSG

JMSG attribute to me the view that "feature mismatches block switches, unless there is almost complete congruence of features in mixed constituents" (p. 70). By contrast, their study is purported to show that "when congruence between the features of the participating languages is missing, there are two results: the derivation does not crash, and it is resolved in favor of the ML" (*ibid*.). Although others have used the term "congruence" (Stenson, 1990; Myers-Scotton, 1993; Sebba, 1998), it does not appear in my work. According to Myers-Scotton (1993, p. 120), "congruence" obtains between two linguistic "entities" under the condition that "they correspond in respect of relevant qualities". I will therefore take JMSG's characterization of my view to indicate that I believe CS may only occur if almost all features match. However, I believe, as is assumed generally in the theory of syntax, that ALL uninterpretable features must "match" – understood in the relevant technical sense, described below – not "almost all". Although I take grammaticality to be quite generally determined by this mechanism – again, following the general consensus in current approaches to syntax – I

nowhere stipulate that grammaticality in CS is determined by this property. Rather, my view is that nothing constrains CS apart from the requirements of the mixed grammars (see (4) above); hence, ill-formedness in CS must be explained in terms of independently motivated and theoretically plausible conflicts among linguistic primitives, whatever they might be. Since features are, by and large, universal, conflicts of grammaticality in CS could rarely be plausibly attributed to "feature mismatch".

However, the crux of JMSG's point is different still. They claim that in current approaches to syntax, feature checking is no longer "absolute", as presented in MacSwan (1999, 2000), but has evolved into a gentler mechanism, now satisfied with only partial fulfillment. Thus, they "[do] not support a version of absolute feature matching as applied to CS (as in MacSwan, 1999, 2000)" (JMSG, pp. 83f.), but rather they endorse the view "that mismatches involving non-distinctness of features do not block a derivation" (p. 80). JMSG argue that the contrast between ABSOLUTE and PARTIAL feature checking presented in their paper derives from an evolution in Chomsky's work, in which he has moved away from a view of feature checking as "mismatch" to feature checking as "non-distinctness" (JMSG, p. 80). The text which JMSG invoke to justify their position is taken from Chomsky (2001, p. 5): "Match is not strictly speaking Identity, but Non-distinctness: same feature, independently of value". In JMSG's exegesis, "This means that on the surface, at the level of PF, the uninterpretable features of system morphemes of the ML remain and enter into other checking relations in the larger bilingual CP" (p. 78). These ideas are used to explain findings in a corpus of Spanish–English CS data in which mixed DPs occur 70% of the time, with single-language DPs the rest of the time. Thus, on JMSG's reading, Chomsky (2001) is advocating that uninterpretable features do not necessarily require deletion to ensure well-formedness, so that feature checking becomes a more-or-less kind of thing, departing from the absolutist feature matching requirements associated with earlier proposals (JMSG, p. 70).

Whatever JMSG's views of feature checking may be, they are quite incorrectly attributed to Chomsky or any other Minimalist syntactician. For Chomsky and others working within the MP, important conceptual and empirical consequences arise around the nature of features and their values, and the question of what constitutes feature matching: Same feature, independent of its value; or same feature with matching values. Recall that in (1b) the DP moved from its VP-internal position to check its case feature in the Specifier position of T. However, the case feature might have a variety of values, nominative or accusative, among others. Finite T is a nominative case assigner. If a DP with a [+nominative] case feature moves to the Specifier position of T, its Checking Domain, it is able to discharge its case feature – an uninterpretable feature, which must be checked for convergence. However, if T is nonfinite, as in the embedded clause in (34), then a DP with a [+nominative] case feature cannot discharge its feature. This accounts for the grammaticality contrast noted in (34).

(34) a. I don't know [$_{CP}$ whether [$_{IP}$ to go to the party]].
b. *I don't know [$_{CP}$ whether [$_{IP}$ Mary to go to the party]].

Here we have an instance of NON-MATCH: The [+nominative] case feature of [$_{DP}$ Mary] in (34b) remains unmatched, as there is no available case checker; therefore, [$_{DP}$ Mary] cannot discharge its uninterpretable case feature, and the derivation crashes.

However, MISMATCH of features is a formally different matter from NON-MATCH. Suppose a DP bearing a [+accusative] case feature enters into the Checking Domain of finite T in a configuration such as (1b). Here the features are identical: T is a CASE ASSIGNER, and D a CASE BEARER. However, the case feature has different values. Although finite T assigns nominative case, D bears a case feature with the value [+accusative]. If we postulate that the case feature is simply unchecked in this scenario, then opportunities for D with [+nominative] case to discharge its case feature later in the course of the derivation would abound, an unwanted result. Therefore, in the interest of barring subsequent (less optimal) convergent derivations, Chomsky (1995, p. 309) proposed (35).

(35) Mismatch of features cancels the derivation.

Notice, then, that MISMATCH of features is to be distinguished from NON-MATCH of features in Chomsky (1995): "We distinguish mismatch from non-match: thus, the Case feature [accusative] mismatches F′ [assign nominative], but fails to match F′ = I of a raising infinitival, which assigns no Case" (p. 309). Hence, Chomsky (1995) distinguishes different matching scenarios depending on feature identity and value: MATCH obtains if a feature [accusative] is compared with [assign accusative]; MISMATCH obtains if a feature [accusative] is compared with [assign nominative]; and NON-MATCH obtains if a feature [accusative] is compared to null.

In light of these observations, we return to JMSG's interpretation of Chomsky's (2001) remark: "Match is not strictly speaking Identity, but Non-distinctness: same feature, independently of value". The terms IDENTITY and NON-DISTINCTNESS are capitalized in Chomsky's text to signify that their technical rather than everyday-use meanings are intended. The distinction between Identity and Non-distinctness relates to the question of whether matching obtains for (a) features and their values (Identity) or for (b) features regardless of their values (Non-distinctness). Identity for α and β means α is the same as β (that is, Equality, or =, holds between α and β) (Gamut, 1991); by contrast, Non-distinctness is used in the general sense of distinctive feature theory, so that features are DISTINCT if and only if they are the same

but differ in value (e.g. [+voiced], [–voiced]; [+present], [–present]), and are NON-DISTINCT if and only if the features are the same whatever their values might be. These formal distinctions may be traced back at least to Chomsky (1965): "We say that two segments are *distinct* just in case one is positively specified with respect to a feature with respect to which the other is negatively specified" (p. 81, italics in original; cf. p. 181). This reading is in fact quite clear from the text cited by JMSG; as stated, the reference is intended to define Match as obtaining in the case of the "same feature, independently of value". Thus, in Chomsky (2001), feature value assignment ("matching") may take place between an unvalued case feature and a case feature with a specified value, whatever it may be. Hence, JMSG arrive at incorrect conclusions regarding the nature of feature checking, apparently by relying on some nontechnical reading of the notions of MATCH, MISMATCH, IDENTITY, and NON-DISTINCTNESS.

Furthermore, note that JMSG's theory of feature checking as essentially optional – or a matter of being "preferred" (Uniform Structure Principle, JMSG, p. 72) or "dispreferred" (Bilingual NP Hypothesis, JMSG, p. 79) – severely undermines the empirical force of the mechanism, rendering it powerless to account for basic facts. For instance, within JMSG's theory of feature checking, the contrast in (34) remains unexplained. In (34a), the embedded subject is able to check its case feature with T, but in (34b) it is not, resulting in a non-convergent derivation. Under JMSG's construal of case checking, both should be fine, since JMSG regard feature checking to be essentially optional. Similarly, on JMSG's view, uninterpretable φ-features of T and D need not necessarily be checked, so that virtually any agreement configuration should be acceptable (**John love Mary*; **el niños bonita voy a los escuela*, "the beautiful children go to school"), contrary to the facts. These empirical and theoretical flaws in JMSG's feature checking theory cannot be easily repaired, and further motivate the rejection of both the MLF model and its attempted minimalist translation.

We next consider whether the ML construct is necessary to any successful analysis of CS data, as JMSG assert; we show that in fact the ML does no work in their own analysis, and that the facts they consider may be accounted for straightforwardly within the MP with no reference to the ML/EL distinction.

3.4. The "modified minimalist" thesis: Is the ML necessary?

For JMSG, "failing to recognize an ML [Matrix Language] limits both the depth and scope of any minimalist explanation of CS" (p. 70). This is their central claim:

> In this paper, we argue that only a modified minimalist approach can account for the substantial body of CS data. This paper recognizes some value in the minimalist approach, but argues that such an approach may even only succeed partially if it incorporates a basic asymmetry between the languages participating in CS. That is, only one language is the source of the morphosyntactic frame structuring the bilingual CP in classic CS. (JMSG, p. 69)

The authors further assert that "a parallel exists between system morphemes and morphemes realizing uninterpretable features, and between content morphemes and morphemes realizing interpretable features" (p. 72). More narrowly, JMSG appear to have in mind the class of "late outsider system morphemes" discussed earlier. These notions in hand, JMSG propose the BILINGUAL NP HYPOTHESIS to account for facts observed in a Spanish–English dataset they present:

> The system morphemes in mixed NPs come from only one language, called the ML. An asymmetry between mixed NPs and full NPs from the EL obtains: full EL NPs are dispreferred because their system morphemes (and their uninterpretable features) do not match other system morphemes and their uninterpretable features elsewhere in the bilingual CP. (JMSG, p. 79)

Here we see the application of JMSG's version of feature checking theory, in which features do not necessarily need to be checked for convergence. According to the Bilingual NP Hypothesis, EL islands are dispreferred because their features do not necessarily need to be checked for convergence. Thus, JMSG find that examples of English Ns with Spanish Ds, as shown in (36a), occur 70% of the time in their data, while bare English nouns occur without determiners about 21% of the time (see (36b)), and so-called "EL islands" occur about 9% of the time (see (36c)) (JMSG, pp. 81f.). According to the authors, this distribution is roughly what one would expect if feature checking is not "absolute" but rather "non-distinct".

(36) a. ¿Pero tú te refieres a *tus* *coworkers?*
but you 2S/REF refer/2S/PRES to your coworkers
"But are you referring to your co-workers?"

b. Me dieron *medium* y yo necesito *large*.
1S/REF give/3PL/PRET medium and I need/1/S/PRES large
"They gave me [a] medium and I need [a] large."

c. Después tengo que hacer *a lot of* *paperwork.*
then have/1S/PRES COMP do/INF a lot of paperwork
"Then I have to do a lot of paperwork."

However, notice that the Bilingual NP Hypothesis, which is responsible for the analysis of the Spanish-English data, uses feature checking, rather than the

ML construct, to derive its results. An asymmetry is said to result, following from the fact that the uninterpretable features of the NP do not match the uninterpretable features elsewhere in the CP. Whether or not "system morphemes" are co-extensive with the class of morphemes bearing uninterpretable features – and it is not, as we shall see directly – has no impact, since "system morphemes" are not defined in reference to the ML/EL distinction. Furthermore, the assertion that these classes of entities are the same suggests that one of them – the problematic "system morpheme" construct – is dispensable. Thus, JMSG's own analysis is unaffected if the notion of the ML is dispensed with entirely, as it is not among the actual mechanisms used to derive the results (setting aside the far more troubling problems associated with the JMSG's checking theory, embedded in their analysis). We conclude that JMSG's own analysis demonstrates the failure of their principal thesis, namely, that the ML is essential to the analysis of CS data.

Although the matter is not central to our discussion, we note that JMSG's assertion that "system morphemes" parallel morphemes with uninterpretable features does not hold. As mentioned, the uninterpretable features are case and the set of φ-features (that is, PERSON, NUMBER and GENDER), and interpretable features are the categorial features and the φ-features of nouns. While case and φ-features generally have grammatical relations with elements outside of their maximal projections which determine their form, so do categorial features and the φ-features of nouns. For instance, in *John has seen Mary*, the categorial feature of *seen* is selected by *have*, the main verb; the grammatical relation (called Merge in MP, subcategorization in earlier generative frameworks) determines the form of *seen* and holds between a head and its complement, a maximal projection. Hence, categorial features are also "late outsider system morphemes", as defined in JMSG and other work. Similarly, the φ-features of nouns, which are interpretable features, must "look outside their maximal projection" to determine their form in expressions like *las mesas* (Spanish) "the tables". Thus, uninterpretable features do not correspond with "late outsider system morphemes".

Yet another way in which we might evaluate JMSG's main thesis is to pursue an analysis of the CS facts observed in their dataset in terms of assumptions consistent with syntactic theory, adhering to (4) above. If we are able to do so without reference to the ML/EL distinction, we once again show that their main thesis – that any minimalist analysis of CS must appeal to the EL/ML distinction to "even only succeed partially" – is false.

JMSG report data in which an interesting asymmetry in Spanish–English CS appears, such that a Spanish determiner may precede an English N (inflected or not) (e.g. *los teachers* 'the teachers'), but that an English determiner does not occur before a Spanish N (e.g. **the casa* 'the house'). These facts are consistent with grammaticality judgments reported in Lipski (1978).

Based on a corpus of English–Spanish CS data from Gibraltar (Moyer, 1993) in which the same pattern is attested, Moro (2001) develops a Minimalist account of these DP-related facts. Moro notes that the composition of the set of φ-features differs for English and Spanish.[15] In Spanish, GENDER and NUMBER are morphologically marked on determiners and nouns, but in English only PERSON and NUMBER are marked while GENDER is absent. In monolingual and bilingual DP contexts, then, the feature matrices identified in (37) hold, with (37d) corresponding with the ill-formed case.

(37) a. For D, $\varphi = \{$person, number, gender$\}$;
for N, $\varphi = \{$person, number, gender$\}$
(Spanish D, Spanish N)

b. For D, $\varphi = \{$person, number$\}$;
for N, $\varphi = \{$person, number$\}$
(English D, English N)

c. For D, $\varphi = \{$person, number, gender$\}$;
for N, $\varphi = \{$person, number$\}$
(Spanish D, English N)

d. For D, $\varphi = \{$person, number$\}$;
for N, $\varphi = \{$person, number, gender$\}$
(English D, Spanish N)

As a generalization, we note that the configurations in (37) are well-formed if the φ-set of N is included in the φ-set of D.

Chomsky (2000, 2001) proposes a Minimalist architecture in which uninterpretable features (such as the φ-features of D, but not of N) enter the derivation without values specified. The operation Agree values and deletes these features from the narrow syntax. "We take deletion to be a 'one fell swoop' operation, dealing with the φ-set as a unit", notes Chomsky (2000, p. 124). "Its features cannot selectively delete: either all delete, or none."

In Moro's analysis, expressions like **the casa* crash because the N's φ-features PERSON, NUMBER and GENDER attempt to value and delete the D's features PERSON, NUMBER and GENDER in "one fell swoop"; however, the GENDER feature cannot succeed, as there is no corresponding feature in the English determiner (see (37d)). As a result, the derivation fails to converge, as the uninterpretable features of D are not deleted. In (37c), on the other hand, the φ-features of N (person, number), being included in D's φ-set, successfully value and delete D's uninterpretable features "in one fell swoop". The same analysis holds for monolingual (37a) and (37b).

[15] MacSwan (1999, section 5.2.2.6) presents arguments to this effect regarding Spanish and Nahuatl gender.

We thus show that a Minimalist analysis of CS data – specifically, of the data analyzed in JMSG – may succeed without reference to the ML construct, demonstrating again that JMSG's "modified minimalist" thesis is incorrect. Also notice that the analysis presented here comes at no cost. The mechanisms are independently available in the grammar, with no appeal to classical CS-specific constraints. By contrast, JMSG invoke a special-purpose typology of morphemes (which moves well beyond Levelt's (1989) system), a distinction between embedded and matrix languages, a System Morpheme Principle, a Morpheme Order Principle, a Uniform Structure Principle, a Bilingual NP Hypothesis, a Feature Mismatch Hypothesis, an ML Feature Hypothesis, an unorthodox construal of feature checking theory, and other mechanisms, all specific to the analysis of CS and utterly foreign to the theory of grammar. Hence, even if the results of the two approaches considered were empirically equivalent – and they are not – then the "modified minimalist" approach and associated MLF model would be dispreferred on grounds of scientific parsimony alone.

3.5. Is the MLF model part of the theory of linguistic competence?

The classic distinction between LINGUISTIC COMPETENCE, or knowledge of language, and LINGUISTIC PERFORMANCE, the use of linguistic knowledge in day-to-day situations, usefully informs both general linguistic theory and the theory of CS, permitting researchers to formalize linguistic theories in terms of a small number of primitives which recursively generate infinitely many structures, accounting for "the creative aspect" of language (Chomsky, 1965). While core linguistic theory has focused on the theory of competence, a number of intriguing facts are better illuminated in terms of performance factors. Consider, for example, (38) and (39), adapted from Stabler (1992, pp. 85–87).

(38) a. The horse raced past the barn fell.
b. [$_{CP}$ [$_{DP}$The horse [$_{RR}$ (which was) raced past the barn]] fell].
c. [$_{CP}$ The horse [$_{VP}$ raced past the barn]] fell.

(39) a. The house was built by Jack.
b. The house [the malt lay in] was built by Jack.
c. The house [the malt [the rat ate] lay in] was built by Jack.
d. The house [the malt [the rat [the cat killed] ate] lay in] was built by Jack.

The structure in (38a), known as a garden-path sentence, is interpreted as well-formed if parsed as in (38b), where *raced past the barn* is processed as a reduced relative clause and *fell* as the main verb; however, if *raced* is interpreted as the main verb, the structure is parsed as in (38c), where *fell* has no structural interpretation. In (39), the structure becomes increasingly difficult to process as more and more clauses are center-embedded. While even (39d) is "grammatical" in the sense of a theory of competence – because it adheres to the grammatical rules underlying relative clauses – it would nonetheless be flagged as unacceptable by the performance system, perhaps for reasons associated with finite memory resources. We have every reason to believe that code-switched sentences could be similarly affected, with some apparent grammaticality effects actually resulting from performance factors.

Myers-Scotton (1993) originally characterized the MLF model as a "production-based model", a fitting description in light of its extensive use of the architecture of Levelt's (1989) *Speaking* model. However, Myers-Scotton (2001) has since emphasized that the model is as much a theory of competence as it is a theory of production (p. 29); JMSG stress further that the ML/EL distinction itself "stems from general linguistic competence" (p. 72).

To demonstrate that the mechanisms of the MLF model are part of general linguistic competence, JMSG must show that they are indispensable for the analysis of bilingual data, or independently motivated by a consideration of monolingual data. We have seen that the demonstration was unsuccessful for bilingual data, and not at all promising for monolingual data, given the consequences of JMSG's feature checking theory for (34) and other facts.

Yet, JMSG maintain that "there is always an ML in monolingual language as well as bilingual language" (p. 88). As evidence, the authors turn to research on speech errors, central to Levelt (1989). Although there may indeed be evidence from speech error research that monolingual sentence production involves the construction of something like a "language frame", as the authors indicate, such evidence does not suffice to make the point that the MLF model has independent justification, contrary to JMSG's claims. Garrett (1975) and others, for instance, report speech errors such as "Make it so the *apple* has more *trees*", in which the grammatical morphemes associated with the speaker's intended expression ("Make it so the tree has more apples") remain in place even though the lexical content words have been inverted. We might readily assume that bilingual language production similarly involves such production mechanisms. This would lead us to expect code-switched speech errors in parallel to this, in which inflectional morphemes do not get inverted with lexical content words, as in, "Make it so the apple has *más árboles*" "Make it so the apple has more trees". However, note that there is no concurrent implication that a ML/EL contrast is operative in the linguistic system of either monolinguals or bilinguals, or that all system morphemes

should come from only one language in the course of sentence production, or even that only one language should be activated in generating the bilingual "language frame". These provisions are strictly stipulations of the MLF model, and do not follow from the evidence noted. Nor are they implied by Levelt's (1989) *Speaking* architecture.

Thus, while a theoretically informed theory of language production might meaningfully be used to explain interesting facts in monolingual and bilingual data alike – of the sort illustrated in (38) and (39), for instance – there appears to be nothing in either the architecture or the evidential basis of the *Speaking* model (Levelt, 1989) to motivate the System Morpheme Principle, the Morpheme Order Principle, or other tenets of the MLF model.

We now turn to a summary of our primary conclusions.

4. Conclusions

We set out to evaluate the MLF model and some revisions and extensions of it in JMSG. We found a number of counter-examples of its two primary mechanisms, the System Morpheme Principle and the Morpheme Order Principle. The MLF model was found to incorrectly identify well-formed items as ill-formed, and ill-formed items as well-formed. The proposed Uniform Structure Principle (USP), which recasts the System Morpheme Principle, was analyzed as a stipulation which in no way followed from the premises given, and it was noted that the authors incorrectly relied on the Uniformity Principle cited in Chomsky (2001) for additional theoretical support for the USP.

We have also addressed JMSG's critique of MacSwan's (1999, 2000) Minimalist approach to CS, noting that the authors incorrectly interpret the approach as uniformly disallowing singly occurring code-switches. Five candidate counter-examples of the PF Disjunction Theorem advanced in JMSG were reviewed; two of these were found to be incorrectly cited from the source documents, and the remaining three were analyzed as cases which do not involve CS in word-internal or head-movement contexts, and which are therefore not counter-examples of the PF Disjunction Theorem. The MLF model was further criticized for its silence regarding the relevance of phonological aspects to the analysis of CS data, and it was noted that JMSG based much of their discussion on a misinterpretation of checking theory in Chomsky (2000, 2001). It was noted that JMSG's version of checking theory as optional emptied the construct of its empirical content.

JMSG's main thesis, embodied in their Bilingual NP Hypothesis, asserted that a Minimalist approach to CS "may even only succeed partially if it incorporates a basic asymmetry between the languages participating in CS" (p. 69). We showed that JMSG's own analysis of Spanish-English CS in DP contexts demonstrates that the ML/EL distinction is entirely dispensable, as the construct played no role in the results presented. Further, following Moro (2001), we presented an analysis of JMSG's data using a Minimalist approach which made no reference to any aspect of the MLF model, showing that a Minimalist approach may indeed succeed without resorting to such notions.

Finally, while noting the potential relevance of a parsing model to CS, we argued that the MLF model not only was not a theory of competence, contrary to JMSG's assertions, but that none of its principles followed from the basic architecture or evidential base of Levelt's (1989) *Speaking* model, from which it originated.

As an alternative to the MLF model, the Minimalist approach of MacSwan (1999, 2000) was reviewed. Arguing on conceptual grounds that CS research might reasonably pursue the thesis that nothing constrains CS apart from properties of the mixed grammars involved, this approach posits that lexical items may be drawn from the lexicon of either language to introduce features into the derivation, with feature checking proceeding as normal, no CS-specific mechanisms admitted. Noting the sociopolitical nature of "languages", it was further suggested that research on CS might be best conceptualized as the study of how linguistic subsystems interact in the mind/brain of a bilingual such that morphological and phonological coding of a lexical item is possible in some syntactic contexts and not in others.

Finally, noting that the MLF model introduces numerous constructs and mechanisms which play no role in the theory of grammar, and that empirically and theoretically sound alternatives exist which make no use of such mechanisms, we suggested that the MLF model should be rejected on grounds of scientific parsimony as well.

Careful and informed critique of the relationship between phonology, morphology and syntax in bilinguals is of crucial importance, and we can be confident that the analysis of CS data and proposed details within the Minimalist approach discussed here will be refined and reformulated as our understanding of data and theory advances. The linguistic study of bilingualism is in its infancy, and considerably much stands to be learned as researchers move away from CS-specific constraints and on to the challenging task of analyzing language contact phenomena in terms of the mechanisms of linguistic theory alone.

References

Backus, A. & Boeschoten, H. (1996). Review of *Social motivations for code-switching: Evidence from Africa* and *Duelling languages: Grammatical structure in code-switching*. *International Journal of Applied Linguistics*, 6 (1), 129–153.

Baker, M. C. (1996). *The polysynthesis parameter*. New York: Oxford University Press.

Belazi, H. M., Rubin, E. J. & Toribio, A. J. (1994). Code switching and X-Bar Theory: The Functional Head Constraint. *Linguistic Inquiry*, 25 (2), 221–237.

Bentahila, A. (1995). Rewiew of *Duelling languages: Grammatical structure in codeswitching*. *Language*, 71 (1), 135–140.

Bentahila, A. & Davies, E. E. (1983). The syntax of Arabic-French code-switching. *Lingua*, 59, 301–330.

Berk-Seligson, S. (1986). Linguistic constraints on intrasentential code-switching: A study of Spanish–Hebrew bilingualism. *Language in Society*, 15, 313–348.

Bokamba, E. G. (1989). Are there syntactic constraints on code-mixing? *World* English*es*, 8, 277–292.

Borer, H. (1984). *Parametric syntax: Case studies in Semitic and Romance languages*. Dordrecht: Foris

Bromberger, S. & Halle, M. (1989). Why phonology is different. *Linguistic Inquiry*, 20 (1), 51–70.

Cantone, K. (2003). Evidence against a third grammar: Code-switching in Italian–German bilingual children. Paper presented at the 4th International Symposium on Bilingualism, Arizona State University, Tempe, 3 April–3 May. To appear in J. Cohen, K. McAlister, K. Rolstad & J. MacSwan (eds.), *ISB4: Proceedings of the 4th International Symposium on Bilingualism*. Somerville, MA: Cascadilla Press.

Chan, B. H.-S. (1999). Aspects of the syntax, production and pragmatics of code-switching with special reference to Cantonese- English. Ph.D. dissertation, University College London.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin and use*. New York: Praeger.

Chomsky, N. (1991). Linguistics and cognitive science: Problems and mysteries. In A. Kasher (ed.), *The Chomskyan turn*, pp. 26–54. Cambridge, MA: Blackwell.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. (2000). Minimalist inquiries: The framework. In R. Martin, D. Michaels & J. (eds.), *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, pp. 89–155. Cambridge, MA: MIT Press.

Chomsky, N. (2001). Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A life in language*, pp. 1–51. Cambridge, MA: MIT Press.

Clyne, M. (1987). Constraints on code switching: How universal are they? *Linguistics*, 25, 739–764.

Cocchi, G. (2000). Free clitics and bound affixes: Towards a unitary analysis. In B. Gerlach & J. Grijzenhout (eds.), *Clitics in phonology, morphology and syntax*, pp. 85–119. Amsterdam: John Benjamins.

Di Sciullo, A.-M., Muysken, P. & Singh, R. (1986). Government and code-switching. *Journal of Linguistics*, 22 (1), 1–24.

Emonds, J. (1978). The verbal complex V′-V in French. *Linguistic Inquiry*, 9 (1), 49–77.

Fuller, J. M. (1999). The role of English in Pennsylvania German development: Best supporting actress? *American Speech*, 74 (1), 38–55.

Gamut, L. T. F. (1991). *Logic, language, and meaning* (vol. 2). C hicago & London: The University of Chicago Press.

Garrett, M. (1975). The analysis of sentence production. In G. Bower (ed.), *Psychology of learning and motivation* (vol. 9), pp. 133–177. New York: Academic Press.

Gross, S. (2000). The role of abstract lexical structure in first language attrition: German in America. Ph.D. dissertation, University of South Carolina.

Halmari, H. (1997). *Government and code switching: Explaining American Finnish*. New York: John Benjamins.

Harris, J. (1991). The exponence of gender in Spanish. *Linguistic Inquiry*, 22 (1), 27–62.

Harris, J. (1996). Spanish imperatives: Syntax meets morphology. Paper presented at LSRL XXVI, Mexico City.

Hlavac, J. (2000). Croatian in Melbourne: Lexicon, switching and morphosyntactic features in the speech of second-generation bilinguals. Ph.D. dissertation, Monash University.

Hlavac, J. (2003). *Second-generation speech: Lexicon, code-switching, and morpho-syntax of Croatian-English bilinguals*. Berlin & New York: Peter Lang.

Holmberg, A. & Nikanne, U. (1993). Introduction. In Holmberg & Nikanne (eds.), pp. 1–20.

Holmberg, A. & Nikanne, U. (1993) (eds.). *Case and other functional categories in Finnish syntax*. Berlin & New York: Mouton de Gruyter.

Jake, J., Myers-Scotton, C. & Gross, S. (2002). Making a minimalist approach to code-switching work: Adding the Matrix Language. *Bilingualism: Language and Cognition*, 5 (1), 69–91.

Jelinek, E. (1984). Empty categories, case, and configurationality. *Natural Language & Linguistic Theory*, 2, 369–412.

Johnson, J. S. & Newport, E. L. (1989). Critical period effects in second-language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60–99.

Kager, R. (1999). *Optimality theory*. Cambridge: Cambridge University Press.

Koopman, H. & Sportiche, D. (1991). The position of subjects. *Lingua*, 85 (2–3), 211–258.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Lightfoot, D. (1982). *The language lottery: Toward a biology of grammars*. Cambridge, MA: MIT Press.

Lipski, J. (1978). Code-switching and the problem of bilingual competence. In M. Paradis (ed.), *Aspects of bilingualism*, pp. 250–264. Columbia, SC: Hornbeam Press.

MacSwan, J. (1999). *A Minimalist approach to intrasentential code switching*. New York: Garland Press.

MacSwan, J. (2000). The architecture of the bilingual language faculty: Evidence from code-switching. *Bilingualism: Language and Cognition*, 3 (1), 37–54.

MacSwan, J. (2003). The derivation of basic word orders in code switching. Paper presented at the 4th International Symposium on Bilingualism, Arizona State University, Tempe, 3 April–3 May.

MacSwan, J. (2004). Code switching and linguistic theory. In T. K. Bhatia & W. Ritchie (eds.), *Handbook of bilingualism*, pp. 415–462. Oxford: Blackwell.

Mahootian, S. (1993). A null theory of code switching. Ph.D. dissertation, Northwestern University.

Mahootian, S. & Santorini, B. (1996). Code switching and the complement/adjunct distinction: A reply to Belazi, Rubin and Toribio. *Linguistic Inquiry*, 27 (3), 464–479.

Moro, M. (2001). The semantic interpretation and syntactic distribution of determiner phrases in Spanish/English code-switching. Paper presented at the 3rd International Symposium on Biligualism (ISB3), 17–24 April, Bristol, UK.

Moyer, M. (1993). Analysis of code-switching in Gibraltar. Ph.D. dissertation, Universidad Autonoma de Barcelona, Bellaterra.

Muysken, P. (2000). *Bilingual speech: A typology of code-mixing*. Cambridge: Cambridge University Press.

Muysken, P. & de Rooij, V. (1995). Review of *Social motivations for code-switching: Evidence from Africa* and *Duelling languages: Grammatical structure in code-switching*. *Linguistics*, 33, 1043–1066.

Myers-Scotton, C. (1993). *Dueling languages: Grammatical structure in code switching*. Oxford: Clarendon Press.

Myers-Scotton, C. (2001). The matrix language frame model: Developments and responses. In R. Jacobson (ed.), *Code-switching worldwide II*, pp. 23–58. Berlin: Mouton de Gruyter.

Myers-Scotton, C. (2002). *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.

Myers-Scotton, C. & Jake, J. L. (2001). Explaining aspects of code-switching and their implications. In J. Nicol (ed.), *One mind, two languages: Bilingual language processing*, pp. 84–116. Oxford: Blackwell.

Nartey, J. S. (1982). Code-switching, interference or faddism? Language use among educated Ghanaians. *Anthropological Linguistics*, 24, 183–192.

Nikanne, U. (1993). On assigning semantic cases in Finnish. Holmberg & Nikanne (eds.), pp. 75–78.

Pollock, J.-Y. (1989). Verb Movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20 (3), 365–424.

Pollock, J.-Y. (1994). Checking theory and bare verbs. In G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi & R. Zanuttini (eds.), *Paths towards Universal Grammar: Studies in honor of Richard S. Kayne*, pp. 293–310. Washington, DC.: Georgetown University Press.

Sankoff, D. & Poplack, S. (1981). A formal grammar for code-switching. *Papers in Linguistics*, 14, 3–45.

Schütze, C. T. (1996). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: The University of Chicago Press.

Schütze, C. T. (1997). Infl in child and adult language: Agreement, case and licensing. Ph.D. dissertation, MIT.

Sebba, M. (1998). A congruence approach to the syntax of code-switching. *International Journal of Bilingualism*, 2 (1), 1–19.

Stabler, E. P. (1992). *The logical approach to syntax: Foundations, specifications, and implementations of theories of government and binding*. Cambridge, MA: MIT Press.

Stabler, E. P. (1997). Derivational minimalism. In C. Retoré (ed.), *Logical aspects of computational linguistics*, pp. 68–95. Berlin & New York: Springer.

Stabler, E. P. (2001). Recognizing head movement. In P. De Groote, G. V. Morrill, C. Retore (eds.), *Logical aspects of computational linguistics*, pp. 245–260. Berlin & New York: Springer.

Stenson, N. (1990). Phrase structure congruence, government, and Irish-English codeswitching. *Syntax and Semantics*, 23, 167–192.

Toribio, A. J. (2001). On the emergence of bilingual code-switching competence. *Bilingualism: Language and Cognition*, 4 (3), 202–231.

Treffers-Daller, J. (1994). *Mixing two languages: French-Dutch contact in a comparative perspective*. Berlin & New York: Mouton de Gruyter.

Uriagereka, J. (2000). *Rhyme and reason: An introduction to Minimalist Syntax*. Cambridge, MA: MIT Press.

Vaninikka, A. (1993). The three structural cases in Finnish. In Holmberg & Nikanne (eds.), pp. 129–162.

Woolford, E. (1983). Bilingual code-switching and syntactic theory. *Linguistic Inquiry*, 14 (5), 520–536.

Zwart, J.-W. (1996). On the status and position of PPs inside APs in Dutch. Ms., University of Groningen.

Received January 16, 2004
Revisions received August 4, 2004; September 12, 2004
Accepted September 21, 2004

Appendix. ***Abbreviations used in glosses***

1S	first person subject agreement (unspecified for number)	FUT	future tense
1SP	first person plural subject agreement	GEN	genitive case
1SS	first person singular subject agreement	INES	inessive case
2S	second person subject agreement (unspecified for number)	INF	infinitive marker
2SP	second person plural subject agreement	NSF	noun suffix (absolutive)
2SS	second person singular subject agreement	PART	partitive case
3OP	third person plural object agreement	PAST	past tense
3OS	third person singular object agreement	PL	plural
3S	third person subject agreement (unspecified for number)	PRES	present tense
3SP	third person plural subject agreement	PRET	preterite tense
3SS	third person singular subject agreement	REF	reflexive clitic
ACC	accusative case	SF	suffix
DUR	durative aspect		