LANGUAGE CONTACT: WHY OUTSIDER SYSTEM MORPHEMES RESIST TRANSFER

Carol Myers-Scotton
Michigan State University

1. Introduction

My over-arching goal in this paper is to show how bilingual speech data are a resource in the quest for new insights about the architecture of language in its general sense. Specifically, I consider the role of structure-building morphemes in language contact data. Their distribution in these data can give us new perspectives on how language is organized.

Morpheme types differ in related ways in their participation in various types of language contact (e.g., codeswitching, creole development, mixed languages). They also differ in the extent to which they are transferred or replicated in various ways across languages. The basic premise behind this paper is that the explanation for these distributional patterns is not at the surface level, but at abstract levels of several types. That is, hypotheses about the differing roles of morpheme types in structuring the abstract organization of clauses, and differences in how they are accessed in language production offer explanations for surface distributions.

The 4-Morpheme (4-M) model of morpheme classification, in effect, is a hypothesis about morpheme types and how they differ, both at abstract levels and in surface distributions (Myers-Scotton & Jake 2000; 2001; 2008; Myers-Scotton 2002). The basic division that the 4-M model offers at one abstract level is between morphemes that are conceptually-activated by speakers’ intentions and morphemes that are structurally-assigned by a specific language’s grammar. Two types of morpheme are conceptually-activated. They are content morphemes and the “early” system morphemes that content morphemes indirectly elect to flesh out meaning. Together, they carry out the speaker’s intentions regarding the semantics of a clause, and, to some extent, its pragmatic import. In the production model proposed here, content and “early” system morphemes are activated earlier than other morpheme types. The remaining morpheme types (“bridge” and “outsider” late system morphemes) largely build structure in the clause, rather than encode

---

1 I have tried to avoid acronyms in this paper, limiting them to what I hope are fairly-well known terms. Here are those that I use, listed alphabetically: AGR= Agreement marker; CS= codeswitching; EL= Embedded Language; IP= Infl Phrase; ML= Matrix Language; MLF= Matrix Language Frame Model; NP= Noun Phrase; SM= system morpheme; 4-M= 4-M model (4-Morpheme model).
meanings. Specifically, this paper is most concerned with the nature of those late system morphemes (hereafter SMs) that are called “outsiders”; however, all morpheme types are illustrated in section 3 when the 4-M model is discussed in more detail.

The paper has four major parts. In section 2, the nature of outsider system morphemes and their role in providing abstract structure for the clausal frame is outlined. Examples here and in later sections come from codeswitching (hereafter CS). Section 3 gives an overview of the principles and models employed in the paper. In section 4, how morpheme types are accessed in language production suggests an explanation for how they participate—or not—in all contact phenomena, not just CS. In sections 5 and 6, the argument about access and its relation to surface roles is further developed.

2. Outsider System Morphemes

To make clear from the outset what counts as outsiders SMs, this section briefly describes such morphemes and exemplifies them. They are called outsiders because they look for information about their form outside the constituent in which they appear. This information may be elsewhere in the clause or in the larger discourse. For example, in German, verbs and prepositions can assign case designations in a clause, but the actual case morpheme (the outsider) appears in elements forming a constituent separate from (i.e., outside) the larger constituent containing the verb or preposition that assigns case.

One group of typical outsider system morphemes are those morphemes that carry the case designations that appear on elements in the core thematic grid (not those on adjuncts). As such, they make the argument structure of the clause clear. Examples (1a) and (1b) show accusative suffixes from one language on nouns from another language. These suffixes are outsider SMs; as such, they distinguish the position of the nouns they accompany in the argument structure of the clause. Example (1a) from Turkish-Norwegian CS illustrates a Turkish suffix marking accusative case that appears on a Norwegian noun (skap). The Matrix Language Frame model, discussed in section 3.2, identifies Turkish as the Matrix Language (ML) of this clause based on its role in supplying such outsider SMs as this suffix. As the ML, Turkish supplies the abstract grammatical frame of the bilingual clause. In (1b) from Finnish-English CS the English compound noun receives the Finnish suffix marking accusative case. Finnish is the ML.

(1a) Turkish-Norwegian (Türker 2000:68)

\[
\begin{align*}
\text{skap-} & \quad \text{doldur-du-k} \\
\text{cupboard-ACC} & \quad \text{fill-PAST-1PL}
\end{align*}
\]

‘We filled the cupboard.’

(1b) Finnish-English (Lane 2006:139)

\[
\begin{align*}
\text{minä ve-in} & \quad \text{se-n} \quad \text{semmoseem paikkaa} \quad \text{joka isont-i} \\
\text{I take-PST-1S IT-ACC this-kind-of place-ILL} & \quad \text{that enlarge-PST.3S} \\
\text{tämä-n \textbf{newspaper clip}-i-n} & \\\n\text{this-ACC} & \quad \text{i-ACC}
\end{align*}
\]

‘I took it to this place that enlarged this newspaper clip.’

In (2) the English predicate adjective (natural) is case marked with a multi-morphemic suffix (-no) from Croatian; the suffix includes an outsider SM indicating nominative case. The suffix also includes what are called early SMs (for singular number and neutral gender). Croatian is the ML. In
section 5.3 the effect of outsider SMs on the accessing of multi-morphemic elements is discussed.

(2) Croatian-English (Hlavac 2003:81)
   ...sve...[am] ...ima...pundo auto, to je natural-no [naturalno] ka ima puno...
   everything...um...there’s... a lot of cars, that’s natural-SING.NEUT.NOM when there’s alot ...
   ‘...everything, um, there’s a lot of cars, that’s natural when there’s a lot...

Another typical group of outsider SMs are AGR markers for subjects or objects of verb phrases. They co-index other elements present in the clause or in the recent discourse. For example, in (3) the object clitic (te) is an AGR marker, referring to a participant in the discourse. Spanish is the source of this clitic and is the ML of the clause.

   Ø te ponemos el black suit
   1PL 2S.ACCUS put.1PL DET.M black suit
   ‘We put the black suit on you/we dress you in the black suit.’

Example (4) comes from a Shaba Swahili-French corpus in which Shaba Swahili is the ML. The subject prefix (i-) on the French verb (i-na-tu-appartenir) agrees with richesse and the object prefix in the verbal assembly (-tu-) agrees with the speakers in the discourse (batoto yake ‘his children’).

(4) Shaba Swahili-French (De Rooij 1996:186)
   Donc, (h)ii richesse y-ote (h)ii i-na-tu-appartenir shi ba-toto y-ake
   ‘So, all these riches, it belongs to us, his children.’

2.1. Explaining Outsiders

Although one can argue that all three types of SMs build structure, outsiders are distinctive in building the basic grammatical relationships within the clause. They do this in two ways. First, they build structure through co-indexing that operates across phrases. Co-indexing provides a precise indication of relations that hold beyond word order. As will become clearer when other types of SMs are exemplified, these SMs (earlies and bridges) rely on juxtaposition as their main means of building constituent structure. In contrast, outsiders clarify argument structure and allow for flexibility in syntax (e.g., movement and relation-changing rules, such as passives and raising). Second, outsider SMs knit together elements at another level. They contribute to the semantic coherence in the clause, or even in the larger discourse (cf. Myers-Scotton 2005b and Myers-Scotton & Jake 2008 forthcoming for further discussion).

2.2. Outsider SMs and Argument Structure

Emphasizing the role of outsiders in expressing argument structure makes clearer their membership in the critical architecture of language. In many languages, and to some extent in all languages, the identifying and the positioning of NP arguments of predicates rely on the surface-level morphological features of those languages. True, in some languages, word order points toward argument structure. Even so, without outsider SMs to realize argument structure through their co-
indexing role, there is no interface of the semantic-syntactic component with the morphological realization patterns and no possible surface realization.

I will argue that these outsider morphemes are the primary way that argument structure is indicated and agreement relations are maintained in a clause. To put it another way, these morphemes are the way that what I call uniform structure is signaled and maintained in bilingual clauses. These elements are what Sapir long ago (1921, reprinted 1949) referred to as those grammatical elements that express “concrete and abstract relational concepts” that “serve to relate the concrete elements of the proposition to each other, thus giving it definite syntactic form” (p.101).

Culicover & Jackendoff (2005) are among linguists proposing alternative models to mainstream generative grammar; they argue for “a simpler syntax” (e.g., no movement rules or derivations). Although they do not discuss the role of outsiders in argument structure, they do recognize the need for special attention to argument structure. Therefore, they posit an add-on to a semantic-syntactic interface; they call the add-on “the grammatical function (GF) tier”. This tier is added to provide the means of identifying which argument is in which grammatical role, but they do this by syntactic, not morphological, means. Still, their discussion includes this summary statement that is reminiscent of Sapir’s discussion. They write, “The assignment of arguments to positions on the GF-tier is in effect an indexing of the arguments that allows us to state the relatedness between constructions” (p. 194).

Similarly, I argue for the importance of outsider SMs in the realization of argument structure in various ways. That they accomplish their tasks may well depend on a universally present “hidden” tier, which is the way Culicover and Jackendoff characterize their Grammatical Function tier. But, of course, outsider SMs and their accomplishments are “in the open” (i.e., they are present in surface structure). Their accomplishments are the co-indexing of semantic arguments with grammatical functions.

Recognizing the particular accomplishments of these morphemes in the abstract levels of grammar is important in discussions of contact-motivated language change. The reason is that their roles help explain empirically-evident unequal distributions of morpheme types in bilingual data: outsiders rarely transfer across languages.

3. Two Models and One Principle

What follows in sections 3.1-3.3 are sketches of the theoretical frameworks used in discussing data. Discussion begins with the Uniform Structure Principle. It declares that, in languages, uniform structure is maintained across any constituent type. In bilingual clauses showing codeswitching, this principle preferences the structure of the Matrix Language. The implication of this preferencing is that if the abstract structure of one language prevails, maintaining uniform structure across constituent types is straightforward. Next, I discuss briefly the Matrix Language Frame (hereafter MLF) model of CS, providing some examples to show how CS data follow the asymmetries that the model highlights and the Uniform Structure Principle as well. Specifically, these examples show how outsider SMs regularly come from only one of the participating languages. Finally, the 4-M model is discussed and more examples show how the classification under the 4-M model is played out in codeswitching data.

3.1. The Uniform Structure Principle

Some years after devising the MLF model (1993; 1997), and after collaborating with Jan Jake on the 4-M model (2000), it became obvious to me that both models imply a particular view of constituent structure in language in general. This recognition resulted in the Uniform Structure Principle. Even though the principle was formulated last, I discuss it first because it helps make
sense of arguments in the other models about structuring potentials in bilingual corpora. In effect, this principle captures the key generalizations from both the MLF model (the notion of asymmetry) and the 4-M model (the notion that morphemes that build structure across phrases are the most critical grammatical elements in building clausal structure).

The Uniform Structure Principle strengthens the MLF model because it privileges the type of morpho-syntactic frame for CS that the model already predicted for what I have called classic CS. But this principle does more than that. It expresses a generalization about the nature of language that may seem obvious to linguists working on monolingual data, but it is not obvious at all for bilingual data. The source of grammatical structure in bilingual speech could be shared in any number of ways. Instead, a uniformity prevails that relies on an asymmetry between the participating languages. The version of the principle that follows applies specifically to CS, but empirical evidence indicates that the basic notion of maintaining uniformity in the dominant language applies in other bilingual data.

The importance of the Uniform Structure Principle is that it not only offers a principled explanation for structure in existing bilingual corpora, but—coupled with the other models—it predicts the cross-linguistic distribution of morpheme types in future corpora. All of the predictions depend on the deceptively simple notion that uniformity in structure is preferred. It is in this sense that the principle is a powerful tool. One prediction that follows naturally from the principle, of course, is that changes in basic structure are resisted; this prediction is relevant even in situations of extensive contact. Here is the Uniform Structure Principle:

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 (ML) are always preferred. Embedded Language (EL) islands (phrases from other varieties participating in the clause) are allowed if they meet EL well-formedness conditions, as well as also meeting those ML conditions applying to the clause as a whole (e.g., phrase placement). (cf. Myers-Scotton 2002, Myers-Scotton & Jake forthcoming 2008).

Examples (5) and (6) are dramatic examples illustrating the extent to which uniform structure from the ML is maintained in CS. In example (5) from Acholi-English CS, all the content morphemes come from English. Still, Acholi is the ML, the source of the clause’s morpho-syntactic frame. The evidence is that the grammatical elements that knit together the clause come from Acholi, not English, which is the EL. The EL’s role is described in 3.2. Note especially the Acholi associative me in the Noun Phrase + Noun Phrase construction. These examples make it very clear that explanations for how CS is organized must lie beyond surface structure.

(5) Acholi-English (Myers-Scotton 2002:132)

costs me education i private institutions tye high...

‘Costs of education in private institutions are high...’

Again, in Xhosa-English CS, shown in (6), the associative element (a bridge SM under the 4-M model) comes from the ML (Xhosa), but the nouns in this construction come from the EL (English). Also, uniform structure is maintained through the subject prefix i- on the verb; it co-indexes the verb with its subject, crime.
(6) Xhosa-English (Myers-Scotton 2005b, Xhosa-English corpus)

_... i-rate y-e crime i-nyuk-ile_

CL9-rate CL9-ASSOC crime CL9-go up-PERF

‘...[the] rate of crime is high.’

Jake et al. (2002) provides quantitative evidence of how the ML is maintained in a CS corpora. In a Spanish-English corpus for which Spanish is the ML, there are 161 well-formed mixed constituents (i.e., they include both Spanish and English elements). Most of these (151 or 93.8%) are mixed Noun Phrases (NPs) with Spanish determiners and English nouns (e.g., _el career fair_). There are no mixed NPs with English determiners; English determiners occur only in 21 EL islands (e.g., _the first one_). These constituents are well-formed phrases entirely in the EL—often adjuncts—whose placement meets the requirements of the ML (cf. Myers-Scotton 2002:139-153 on EL islands). Determiners are early SMs under the 4-M model.

### 3.2. The Matrix Language Frame (MLF) Model

What does the MLF model accomplish? In a word—and the word is “asymmetry”—the MLF model offers a synthetic characterization of the constraints that apply to bilingual distributions in CS. A few disclaimers clarify limits on the model. First, the model applies only to classic CS; this is CS in which the source of the morpho-syntactic frame of the bilingual clause clearly and consistently is only one of the participating languages. Thus, the model was not designed to explicate all language contact phenomena. Still, portions of the model do apply to what I call composite CS, as well as to other contact phenomena. In composite CS, generally one language is the identifiable source of most of the abstract grammatical structure of the clause—but just not all of it. That is, composite CS is defined by the fact that _both_ participating languages are sources of the abstract structure of the clause. Certainly, I argue that the model’s notion of asymmetry does apply across bilingual data in general. Finally, the model was intended to apply only to participating varieties that are not mutually intelligible, although it may apply otherwise as well.

The key feature of the MLF model is that it differentiates both the participating languages and morpheme types at a number of abstract levels. First, it differentiates two morpheme types: content and system morphemes. Thus, the model does not use more familiar terminologies, such as the lexical vs. functional distinction, or the open vs. closed class distinction. The reason is that a classification based on participation in the thematic grid better characterizes morpheme distribution in the data considered. Content morphemes are defined as assigning or receiving thematic roles; system morphemes are defined as performing neither of these functions. Nouns and verbs are prototypical content morphemes, and affixes and some function words that do not occur alone, such as determiners and clitics, are examples of system morphemes.

Second, the model emphasizes asymmetry by claiming a dominant role in the bilingual clause for only one of the participating languages. To this end, the two main principles of the model are, in effect, hypotheses about the differing roles of the participating languages in the bilingual clause. They are the Morpheme Order Principle and the System Morpheme Principle. They propose that in what is now called classic CS one language only is the source of morpheme order and of one type of system morpheme. The original formulation of the System Morpheme Principle did not name this morpheme type, but only described it. The name, outsider SMs, came later from the 4-M model. As already noted, what is distinctive about this morpheme type is that it makes the argument structure in a clause transparent.

In satisfying these two principles, one language is identified as the language that builds the morpho-syntactic frame of the clause. This language is called the Matrix Language (ML). Thus, the heart of the MLF model is that it identifies the ML and its roles and, in so doing, limits the role of
the other participating language, the Embedded Language (EL). Generally, the EL can only provide either content morphemes in mixed constituents or EL phrase-level constituents (EL islands), or both. There are some exceptions to this statement involving early SMs (see section 5.1) and EL bridge SMs (see section 5.2). But these exceptions are few and, because they are not outsiders, they are not counter-examples to the System Morpheme Principle.

The following examples illustrate the asymmetry between the roles of the ML and the EL that the principles predict for CS.

Example (7) illustrates how the Morpheme Order Principle is satisfied. It shows how word order in the adjective phrase (m-moja stingy sana) comes from the ML (Swahili in this case).

(7) Swahili-English  (Myers-Scotton, Nairobi corpus 1988)

Tena hu-w-a    m-tu    m-moja  styngy  sana
Once more HABIT.COP.FV    CL1-thing    CL1-one stingy very
hu-bani    Ø-pesa    z-ake...
HABIT-sit on    CL10-money    CL10-3S. possessive
> Once more, he is one very stingy person [who] sits on his money...

Example (8) shows how the EL (English) verb take receives ML (Swahili) inflections for subject (a-) and also object (-i-). The object prefix agrees with beer, which is realized as a class 9 noun. These prefixes are both the type of morpheme that the System Morpheme Principle states come from only one of the participating languages (i.e., the ML). They are outsiders because they co-index the Verb Phrase with other elements in the clause or the larger discourse. Note also hours tano, which follows Swahili order.

(8) Swahili-English  (Myers-Scotton Nairobi corpus 1988)

Kweli,    Ø-beer    a-na-i-take    kwa    Ø-hours    tano
truly    CL9-beer    3S-NON-PST-CL9.OBJ    for    CL10-hour-s    five
‘Truly, he was taking [drinking] beer for five hours.’

Example (9) illustrates an EL island from French that occurs in a morpho-syntactic frame from Arabic, the ML. Such islands meet well-formedness conditions for the EL.


ya-t-haka-w    wahed    les histoires ...
3.M-P-tell-PL    INDEF    DEF    stories
‘They tell each other some [fantastic] stories....’

3.3. The 4-M Model

The 4-M model was introduced briefly in section 1. It is a model of morpheme classification of general applicability across languages, even though some researchers have wrongly assumed it is a revision of the MLF model. Its only relation to the MLF model is that its provisions refine the distinctions between morpheme types made in the MLF model.

The 4-M model distinguishes four morpheme types in three ways. First, only content morphemes
receive or assign thematic roles. But second, arguably the most important premise of the 4-M model is that morphemes are divided as either conceptually-activated (content morphemes and early SMs) or structurally-assigned (bridge and outsider late SMs). Third, structurally-assigned morphemes divide into two groups, bridges and outsiders.

3.3.1. Content morphemes are directly activated by a speaker’s intentions and early system morphemes are indirectly-elected to add specification or to modify a content morpheme’s meaning. Both become salient at the level of the mental lexicon in language production and this is a reason why early SMs have their name. Examples of early SMs include plural markers and determiners (although they are in multi-morphemic elements in some languages). They also include derivational affixes as well as satellites in phrasal verbs (e.g., for in look for the dog). In addition to semantic and pragmatic content, content morphemes and some early SM morphemes carry some information about phrase structure, so there is no strict divide between some aspects of syntax and semantics.

As is evident in the examples in this paper, EL content morphemes occur frequently in CS corpora; the reason seems to be that they express the speaker's intentions more precisely than an ML counterpart. Elsewhere other researchers and I have pointed out that EL nouns are much more frequent than EL verbs. One can argue that verbs, as thematic role assigners, have a more critical role than nouns; therefore, they come largely, or entirely, from the ML in most CS corpora. Only a few early SMs from the EL occur. This finding, and that about verbs, follows from the predictions of the Uniform Structure Principle, of course, because maintaining as much ML structure as possible makes the overall clausal structure as uniform as possible. True, EL affixes for plural are fairly frequent in some corpora, but if they build anything, it is meaning, not clausal structure.

One EL early SM that does occur in some CS patterns is the satellite preposition of phrasal verbs. This type of morpheme is illustrated in example (10). But the more dominant pattern in CS corpora is for early SMs to come the ML. This is especially evident in mixed NPs. Example (11) includes such an NP (el half chicken) with a determiner from the ML followed by an EL noun.

(10) Palestinian-Arabic-English (Okasha 1999: 156)
?aloo lazim innu el engine locked up fi hadiik el-lahda …
3P.say.PERF must that.3S DET engine locked up in that the moment
‘They said that the engine must have locked up at the moment ….’

¿Qué ha dicho, que el half chicken es mucho, no?
what 3S.PSTPERF AUX say.PSTPERF that DET.M half chicken be.S.PRES a lot no
>What did he say, that half a chicken is too much, isn’t it?=

3.3.2. In contrast to conceptually-activated morphemes, structurally-assigned SMs carry out language-specific indexing roles in large phrases and the full clause. Even though they are represented by lemmas in the mental lexicon, they are called “late” SMs because the divisions the 4-M model makes implies a hypothesis that late SMs only become activated when larger structures are assembled. This occurs at the level of the formulator in the language production model that the 4-M model follows. The formulator puts together the syntactic strings that follow language-specific well-formedness requirements. Morphological realization patterns and phonology produce a surface structure at the next level.
3.3.3. Two types of SMs are included under the category of late SMs, bridges and outsiders. Bridges differ from outsiders in at least two ways. First, they are invariant in form while outsiders are members of a set of variants (e.g., subject-verb agreement morphemes or a declension of case-assigning morphemes). Second, bridges occur to satisfy the well-formedness conditions within the constituent in which they occur, joining together phrases to form a larger constituent. Instead, outsiders depend on directions originating outside where they occur, as will become clearer.

The prototypical example of a bridge is the morpheme that joins two phrases in a construction encoding possession or association in some languages. For example, in English *of* is a bridge between two NPs in the phrase, *a friend of the family*. In CS corpora there are often examples in which both NPs in the larger phrasal constituent come from the EL, with the bridge morpheme from the ML, as illustrated in examples (5) and (6) in section 3.1.

Example (12) illustrates another bridge SM. This example from Arabic-English CS shows two clauses linked by a complementizer. The first clause is entirely in Arabic, followed by the Arabic complementizer *?innu* ‘that’, even though it heads the second clause that is otherwise in English. Myers-Scotton & Jake (2008) argue that such complementizers as *?innu*, along with *that*-complementizers in English (and similar forms in many other languages), are bridge SMs. They meet the definition of bridges because they link two larger constituents.

(12) Palestinian Arabic-English (Okasha 1999:92)

<table>
<thead>
<tr>
<th>?aalat</th>
<th>hiyyi ?innu</th>
</tr>
</thead>
<tbody>
<tr>
<td>PERF.say.3F</td>
<td>she that</td>
</tr>
<tr>
<td></td>
<td>she did not like that</td>
</tr>
</tbody>
</table>

‘She said that she did not like that.’

As in the case of early SMs, there is no requirement in the MLF model that bridges must come from the ML. And yet, they generally do. Again, this finding can be explained by the Uniform Structure Principle, which preferences ML structure.

3.3.4. Most obviously, outsider morphemes contrast with bridges because outsiders look outside the constituent in which they appear for information about their form; hence their name. Outsiders also differ in their relationship with other elements in the clause. That is, while bridges join constituents, outsiders co-index relations. Most specifically, as noted in section 2.1, they relate an argument to a syntactic position. Examples (1) through (4) in section 2.0 show how outsiders in mixed constituents in CS come from the ML. Outsiders include case markers and Agreement (AGR) markers, such as clitics, and certain elements in the verbal assembly as becomes clear in section 5.

Example (13) illustrates a configuration of CS involving Arabic and relations across clauses that requires extra discussion. When the related clauses in question are instances of CS, and Arabic is the language structuring the first clause, the domain of Arabic, as the ML, extends into the second clause even when that second clause is largely or entirely in the EL. This is the case in example (12). However, the picture becomes more complicated for those Arabic complementizers/subordinators that include a suffix that is co-indexed with the subject of the second clause. The subordinator *li?anuhm* ‘because’ in example (13) includes such a suffix. These suffixes are outsider SMs, of course, because they co-index the content morpheme head of the subject of the Infl Phrase (IP). The subordinator *li?anuhm* itself is a content morpheme, although complementizer that take suffixes are bridge SMs. (Myers-Scotton & Jake 2008 make this distinction between subordinators and complementizers.)
A side issue is that information about the nature of the IP from the EL must be available before the IP itself is produced in such examples as (13). This information is needed so that the suffix on the ML complementizer or subordinator can match the person and number of the subject of the IP. The implication is that the EL is activated throughout the entire process. Recent psycholinguistic experimentation also supports the notion that both languages are always activated, even if at different levels, when bilingual data is being produced (cf. Kroll 2008). Still, the insight to take away from this example is that even though the EL supplies the necessary information for co-indexing, the ML builds the structure so that uniformity favored by the Uniform Structure Principle is maintained.

3.4. Summing up the Distribution of Morpheme Types in Codeswitching

This discussion shows that asymmetry and uniformity are two sides of the same coin. The asymmetry between the ML and the EL that the MLF model claims characterizes CS becomes obvious when a CS corpus is held up to the mirror of the 4-M model. The ML regularly supplies all four types of morpheme, but the EL’s contribution is largely limited to content morphemes, as well as occasional EL islands. What seems most essential in bilingual speech is that uniform structure should prevail; this is accomplished by reserving the task of building most or all of the grammatical structure to the morphemes of only one of the participating languages (the ML).

4. The Differential Access Hypothesis and Transfer across Languages

Even though the MLF model’s recognition of the asymmetries in CS leads to a synthetic analysis of morpheme distributions in CS, the model does not explain why its System Morpheme Principle applies to only one type of system morpheme. Also, even though the 4-M model offers a useful classification of morpheme type, again, it does not explain why any of the morpheme types show the distributions they do, not just in CS, but across contact phenomena in general.

The 4-M model does imply a hypothesis about language production that offers an explanation. The Differential Access Hypothesis is one clue as to the extent to which different morpheme types are available for transfer to, or reconfiguration in, another language.

The Differential Access Hypothesis follows:

The different types of morpheme under the 4-M model are differentially accessed in the abstract levels of the production process. Specifically, content morphemes and early system morphemes are accessed at the level of the mental lexicon, but late system morphemes do not become salient until the level of the formulator (cf. Myers-Scotton 2002; 2005a).

The hypothesis is motivated by the extent to which the System Morpheme Principle is supported in classic CS, but also by how distributions differ by morpheme type in other linguistic phenomena. Reported evidence comes from speech errors, the speech of Broca’s aphasics, second language learners, as well as from other language contact phenomena (cf. Myers-Scotton & Jake 2000a; 2000b; Myers-Scotton 2001; 2002; 2003; Wei 2000). The diversity and extent of these data suggest that more than incidental factors motivate this hypothesis.

The Differential Access Hypothesis suggests the following scenario for how morphemes are accessed. Lemmas in the mental lexicon underlying content and early SMs build some phrases, but
also send language-specific directions to the formulator to build larger linguistic units. Some more general directions may also come from an autonomous syntactic component, as well as some from an additional tier such as the Grammatical Function-tier of Culicover & Jackendoff. These directions activate the bridge and outsider late SMs that are the structure-building elements that put together the larger constituents that make up a clause. The basic point is that late SMs are not salient until the level of the formulator.

5. Predicting Externally-Motivated Changes

The Differential Access Hypothesis implies two general predictions about the type of changes to expect when bilingual speakers are in an intense or long-standing contact situation. First, the hypothesis predicts that those elements and patterns that arise because of the speaker’s intentions regarding semantic and pragmatic messages are most likely to show convergence and other types of changes that involve substitution, reconstitution, or loss. This means that the most available morphemes for change in contact situations are content morphemes, followed by early SMs. As already noted, content morphemes are directly elected by speaker intentions and early SMs are indirectly elected by content morpheme heads. They become salient at the level of the mental lexicon, according to the hypothesis. Also, structures at the semantic-pragmatic interface with syntax are predicted to be relatively open to change. In contrast, late SMs are less available, if at all, for changes in contact situations. This follows from the way they are accessed under the hypothesis, as is outlined in section 4.

5.1 Early System Morphemes

In light of the Differential Access Hypothesis, some ways that early SMs participate in classic CS can be explained. Because of their status in both phrase structure and their suggested “early” access in language production, the prediction is that early SMs will differ in their distribution in contact phenomena from late SMs.

This prediction is supported in CS corpora where some early SMs do occur. That is, EL plural affixes and, less often, other EL affixes, may appear along with their content morpheme heads. How to explain these occasional appearances? Consider that their form and their very appearance depend on the appearance of their content heads. Also consider that under the Differential Access Hypothesis they are available to be accessed at the same level as content morphemes (the level of the mental lexicon). Taking these two propositions together, the possibility that they are accessed through a “production-based mistiming” seems a plausible explanation, especially because their occurrence is not regular.

This explanation is reinforced by the fact that sometimes these EL early SMs occur—even if the ML also supplies an affix for the intended semantic content. The most common examples of this “double morphology” consists of an EL plural affix along with an ML plural affix (e.g., ma-lesson-s in a Shona-English corpus, with the Shona prefix, ma- for noun class six and the English suffix –s both marking plurality on the English noun). Such examples of double morphology are not limited to one group of languages; for example, lu-civilian-s ‘civilians’ illustrates this phenomenon in CS with a Nilotic language, Acholi, and there are examples in Turkish-Dutch CS. Sometimes other EL early SMs are doubled as well (e.g., French infinitives in Congo Swahili-French CS, e.g. ku-renvoy-er ‘send back’). These occasional happenstances give rise to an Early System Morpheme Hypothesis that distinguishes early SM patterns from those of late SMs. The hypothesis is “Only early system morphemes may be doubled in classic codeswitching”. To date, the hypothesis is supported by all published corpora and this finding makes the notion of “mistiming” even more plausible (cf. Myers-Scotton 2002:92-3).

The status of early SMs in classic CS leads to the prediction that, of the three types of SMs, earlies are the most likely to appear in other types of contact phenomena. The structure of the
split(mixed) language Michif, which includes both elements of French and Cree, supports this prediction. The French grammatical component consists almost entirely of NPs composed of a determiner and a noun. These determiners are early SMs (e.g., un vieu in Bakker 1997). Much of this discussion in this section appeared in Myers-Scotton (2002).

Still, it is only content morphemes that are readily open to movement from one language to another. Even then, they often are modified in some way along the lines of the Abstract Level model (cf. Myers-Scotton & Jake 2001; Myers-Scotton 2002). Under this model, the levels of abstract semantic and structural structure in a lemma underlying a morpheme (largely content morphemes), can be split and recombined with levels from a morpheme from another language variety.

5.2. Bridges

Because the Differential Hypothesis suggests that all late SMs (bridges and outsiders) only become salient in language production at the level of the formulator, the prediction is that bridges are far less available than content morphemes and early SMs for transfer in contact situations. And empirical evidence supports this prediction.

However, the same evidence indicates that bridges are more susceptible to transfer than outsiders. For example, “that-like” complementizers from one language sometimes appear in another language in contact situations. For example, in a Xhosa-English corpus, there are examples of the English complementizer that joining two clauses in a sentence that otherwise includes only Xhosa elements (Myers-Scotton 2005b).

Also, even in classic CS, the EL sometimes supplies a bridge SM. This is only reported to date in some language pairs, such as Moroccan Arabic and French. In these cases, when French is the ML, Arabic dyal ‘of ‘does occur occasionally. Again in Moroccan Arabic-French CS, dyal in possessive constructions can receive an affix that is co-indexed with a content morpheme (typically a pronoun), with the dyal + affix serving as a possessive pronoun. That is, the affix is an outsider SM. (e.g., on étudie la société dyal-na hors du context religieux ‘we study our society out of religious context’ (Ziamari 2003). The suffix -na encodes 1PL and translates here as ‘our’. This, of course, constitutes a counter-example to the System Morpheme Principle, but remember that it is only a single example.

One explanation why even a few bridge SMs seem more available for transfer in contact situations or may come from the EL in ML-framed CS than outsider SMs is their form. As noted above, bridges are morphemes that are not part of a declension or conjugation, as are outsiders. Also, they seem to consist of a single allomorph.

Compare bridges with the resistance to transfer/substitution of the bound morphemes in verbal assemblies of many agglutinative languages. For example, such morphemes in the Dravidian languages have proven to be highly resistant to external influence. Not all of these morphemes need be outsiders, although often some of them are. Certainly, in some case, original structures may be maintained because some of the affixes in the verbal assembly interact with AGR features; that is, their form depends on information from outside the phrase in which they occur and therefore they are outsider SMs. Other affixes that are not likely candidates as outsiders may still be interdependent with the affixes that are outsiders. The result is that the entire verbal assembly resists change.

Sjoberg (2001) offers a hypothesis that focuses on the surface structures. She suggests, “...the morphology of these languages carries an especially heavy functional load of grammatical meaning, much more so than does the syntax” (p. 382). She also refers to the fact that many morphemes in Dravidian languages are “bound on both sides.” Still, I suggest an argument based on differences in how morphemes are accessed in language production (i.e., the Differential Access
Hypothesis) has even more explanatory power because it is more generalizable than an argument based on specific typological features.

5.3. Multi-morphemic Elements

Also, how morphemes interact with other morphemes seems to have consequences for what can happen in language contact. Quite often, functional elements are multi-morphemic and one of the morphemes they include is an outsider. (Typically, they also may include at least one early SM, such as number or gender.) But because of the outsider SM, the activation of the element must wait until directions for that late SM are sent to the formulator. For example, one can argue that German determiners are not available in language production at the same level (the mental lexicon) as determiners in some other languages whose determiners include only gender and number, for example. Not surprisingly, multi-morphemic units that include an outsider SM are predicted to resist transfer in other contact phenomena.

In CS data sets, the presence of multi-morpheme functional elements in one or both of the participating languages may have consequences for possible CS patterns. Myers-Scotton & Jake (2001) explain distributions in an Italian-Swiss German corpus in terms of the Differential Access Hypothesis. They argue that Swiss German NPs are realized later than Italian NPs (because of the case morpheme included in German determiners, but not in Italian determiners). Thus, it seems that German determiners may not occur at all when Italian is the ML and therefore the source of the morpho-syntactic frame. How is this? Evidence from this corpus indicates that when frame-building directions come from Italian for NPs (i.e., Italian is the ML), only NPs with Italian determiners are available when NPs with determiners are assembled at the level of the mental lexicon (cf. Myers-Scotton & Jake 2001; Jake, Myers-Scotton & Gross 2005).

Such situations are covered by the “drag down” principle, which is this: Any multi-morphemic unit containing an outsider SM shows distribution patterns as if it contained solely an outsider SM. This means that the element is dragged down to the level of the formulator for its access in language production (cf. Myers-Scotton 2002:305, which mentions this principle).

5.4. Identifying Outsider SMs

Any attempt to assemble evidence to test predictions about the transferability of outsider SMs in contact situations must first deal with a problem in identification. The problem is that many contact researchers (almost all) hardly differentiate morphemes beyond separating “meaning” morphemes from “grammatical” morphemes when they discuss their susceptibility to change. For example, in an otherwise comprehensive and excellent overview, when Aikhenvald (2006) refers to grammars in contact, she does not refer to specific morpheme type. Still, she does reproduce a figure that graphically indicates that “inflectional (or core) morphology (form/function)” appears at the top of the hierarchy of features that are “more similar to genetic relatives” as opposed to “more similar to neighbouring languages.” (p. 5).

In Aikhenvald’s chapter and elsewhere, there are some useful hints about differences in the patterning of “core” morphology in contact phenomena, but they remain only hints. For example, somewhat wryly, Johanson (2002:47) notes that “it is frequently claimed that a linguistic element is more difficult to replace, the less conscious the speaker is of it”. He attributes this sentiment to Haugen (1950). Johanson himself, of course, is a justly-respected expert on Turkic languages. Such comments suggest that Johanson would include outsider SMs among the less “conscious”.

But along with most contact researchers, Johanson takes the position that while more abstract elements are less likely to be borrowed, what can be borrowed cannot be predicted, because the “relative attractiveness” of forms vary, as do circumstances. For example, in writing about Turkic languages in specific, he states, “One is hard-pressed to find A structures [in a language A] that
absolutely obviate the copying of certain B [from a language B] structures” (p. 53). In addition, he states that “the validity of the structuredness=stability equation ... remains to be proven” (p. 53). The problem remains that it is hard to find even general claims, let alone empirical evidence, that separates out morpheme types.

For this reason, once the discussion moves away from classic CS, where the MLF model identifies outsiders explicitly, it is difficult to discuss outsiders. As noted above, I define “composite codeswitching” as bilingual data in which more than one of the participating languages contributes abstract structure to the morpho-syntactic frame of a bilingual clause. In truth, this type of data may be more common than classic CS, but most of the CS literature to date deals with classic CS. Muysken (2000) and Clyne (2003) are the most prominent researchers who do discuss data that constitute, in effect, composite CS. But without analyses that separate out morphemes by type, whatever basis is used, it is very difficult to test the hypothesis that outsider SMs are, at the least, at the far end of any borrowability scale. The upshot of this discussion is that many of the “grammatical” morphemes that are reported to be transferred from one language to another are not outsider SMs (or even bridge SMs) under the 4-M classification.

6. Recognizing Other Late SMs

When researchers suggest that most “grammatical” morphemes can be borrowed, are they referring to what the 4-M model identifies as late SMs? As just indicated, this remains unclear in many instances. In this section, three types of “grammatical” morphemes, those for ASPECT, TENSE, and CASE, and the status they have under the 4-M model are discussed. These discussions are quite brief, but their purpose is to give further indications as to how morpheme assignment operates in terms of the 4-M model. As a preliminary, I stress that the counterpart of a form that is an outsider SM in one language is not necessarily an outsider in another language—unless it also meets the general 4-M criteria for outsiders. That is, the definitions apply universally, but like forms across languages do not necessarily both satisfy a given criteria for a same morpheme type.

To begin, following the 4-M model leads to the argument that morphemes (whether self-standing forms or affixes) that mark only ASPECT (and not also TENSE) are best be considered as early SMs. That is, the argument can be made that they add meaning to the verb in a sense that AGR elements, which also occur in a verbal assembly, do not. Certainly, in many cases (all?) aspect markers depend on the verb for their presence as opposed to AGR forms (e.g., subject-verb agreements). AGR forms, as well as TENSE in some languages, depend on information outside the verb.

6.1. Tense as an Outsider SM

In contrast to ASPECT, which always seems to be an early SM, the status of TENSE seems to vary across languages. When it interacts with the AGR paradigm, it is an outsider SM; otherwise, it is an early SM.

TENSE does interact with agreement in many languages, such as German, at least in some instances, and therefore in those instances is an outsider SM. Jake points out (personal communication, 2007). For example, present tense of some German strong verbs is irregular in the 2nd and 3rd person singular. The tensed verbs are co-indexed with the subject. Thus, for the verb essen ‘eat-INF’, TENSE is an outsider. Compare present tense 1st S is ich esse ‘I eat’, but, du isst ‘you eat’, and er, sie, es isst ‘he, she, or it eats’. That is, the vowel of the infinitive ablauts in 2nd and 3rd person. Other strong verbs ablaut in more forms, as in ich weiss (present) ‘I know’ (from wissen ‘know-INF’). This indicates that TENSE is intimately connected and, in many cases, inseparable from agreement in German strong irregular verbs.

Such data suggest that in many languages, TENSE interacts by reference to the AGR system. AGR, of course, is the prime example of an outsider SM. When TENSE interacts with AGR, as it
does in English, it is an outsider, too. That is, English still requires subject-verb agreement in present tense. And, as Jake also points out, “Even if each overt morpheme is not inextricably bound with AGR, it is checked against AGR to see that form of AGR that the system needs because other forms in the same paradigm are bound with AGR” (personal communication, 2007). Another reason to identify TENSE in English as an outsider is that its form is not predictable, except based on information outside of where tense is marked. That is, one must know whether tense interacts with “do support” to know the form that TENSE takes.

Taking the Differential Access Hypothesis together with the Uniform Structure Principle, one can predict that any morphemes (e.g., TENSE) that build structure across the clause are less likely to participate in convergence or attrition than other morphemes. For this reason it is not surprising that Johanson (2002) points out the remarkable stability in Turkic languages of basic aspect-tense categories as well as of the Turkic case system. In Turkic, aspect and tense are one and the same (i.e., in the same morpheme). Further, as indicated above when outsider SMs are exemplified, case marking is a prototypical outsider SM; in many languages, case marking links the basic arguments of subject, object, and beneficiary to syntactic positions. Johanson uses the term “copying” rather than “borrowing” or “transferring” or “switching”. The basic idea is that linguistic elements are copied from language X to language Y. The “copies” then belong to another system than their originals. The emphasis is on both the process and the result. However, one problem with the term, “copying”, is that the copying is not always exact.

6.2. Case Markers as Morpheme Types

It should be clear from the discussion above that case markers that identify the main arguments of a clause are outsider SMs. In case-marking languages, these affixes typically are assigned by verbs or prepositions, but they are realized on other elements, such as nouns, or adjectives, or determiners, as they are in German or Russian and other Slavic languages, for example.

But what about other cases, notably instrumental and locative cases, in languages with extensive case systems? The extraordinary extent of apparently contact-induced change that Heath (1978) reports in Arhem Land in Northern Australia is well known to contact researchers. Some of the apparent borrowings that he reports are in the case system, notably various instrumental affixes. As the summary of a lengthy discussion, Heath writes, “It is thus fairly certain that Nu, Wa, and Enindhilyagwa have borrowed ‘having’ suffix —miri from the Yuungu languages, adopting it as an instrumental suffix to fill a gap left by the absence of a stable instrument or ergative-instrumental affix in the relevant proto-languages" (p. 79).

Although the extent of borrowing in this geographical area is indeed great, under the 4-M model, such case markers as instrumental and locative would be analyzed as early SMs. These case markers encode content more than they have anything to do with the main argument structure in the clause. That is, they are best considered as early SMs that flesh out the meaning of the verb. It is true that the INSTRUMENTAL did develop an ergative function, as Heath reports, and so today one could argue this morpheme functions as an outsider SM because of the role of ergative case in argument structure. But the suggestion is that when it was borrowed, it was—in all likelihood, an instrumental. Heath (personal communication, 2007) concurs with the gist of this argument as a possible development. In other languages, it is possible that INSTRUMENTAL is an outsider SM. But, as argued above, the status of any morpheme type depends on how parameters are set in the language in question.

Other languages with well-known extensive case systems include Hungarian and the Turkic languages. But within these systems, all case markers clearly do not function alike. Under the 4-M model, some are outsider SMs (those that map primary argument structure) and some are early SMs (those that refine the meaning of the verb by indicating different types of location). Bolonyai (2002)
supports the hypothesis that there is such a division in the Hungarian case system.

Bolonyai studied six Hungarian children in the U.S. whose informal conversations showed L1 attrition; English had become their dominant language. Her findings make a distinction between the Hungarian morpheme marking accusative case and “local” morphemes marking various ways to indicate the location of an object. She writes, “These findings demonstrate that the structurally-salient system morpheme (accusative) and the conceptually-salient system morphemes (local cases) do not pattern alike” (p. 17). The two types of morpheme show different patterns of retention and substitution for those Hungarian case elements that can be analyzed as early SMs as opposed to one case marker that is more clearly a late SM (accusative marker). Those Hungarian cases that encode locations are considered early SMs.

Bolonyai’s study is particularly useful because it includes quantitative data and a statistical analysis. In her study, 89% (1216/1366) of their obligatory case markers corresponded to standard Hungarian usage. The accusative case affix, marking the direct object, was the most frequent in the data, but it is also the most accurately applied. Out of 783 accusative contexts, 91.8% (719) are correct. For the local cases, 241/306 or 78.8% are correct. This difference is statistically significant ($p > .0005$). This difference in accuracy indirectly supports the Differential Access Hypothesis presented above; because the hypothesis states that the two types of morpheme are accessed differently, early SMs at the level of the mental lexicon and outsider SMs at a later level in production, at the formulator. The finding also supports my overall claim that late SMs tend to be retained in language contact situations rather than undergo loss or substitution.

The two types of errors that the children make also support the claim that there are two types of morpheme in the Hungarian case system, outsider SMs, such as accusative, and early SMs, such as the locative case markers. Where both English and Hungarian overtly mark thematic roles (by prepositions in English and case markers in Hungarian), the children are less likely to omit the case morpheme when it also contains thematic information. Sometimes they substitute another Hungarian case marker that was closer in meaning to an English counterpart. In regard to accusative case (an outsider SM), which is only overtly marked in Hungarian, the children made some omission errors. Bolonyai notes that in data on adult Hungarian-American immigrants, speaker errors with accusative were also errors of omission. Further, she notes that independent evidence of the different treatment of morpheme types comes from Hungarian speakers with Broca’s aphasia (MacWhinney & Osmán-Sági 1991). Their subjects demonstrated an almost identical pattern of error/substitution on the two types of case.

In Croatian, a language with quite a different typology, but also with an extensive case system, case affixes are maintained very well among second generation Croatian young adults raised in Australia, even though English has become a dominant language for them. Hlavac (2003:71) reports that of the morphologically-integrated EL (English-based) forms in his data, 90% (258/280) show standard Croatian case marking. Admittedly, these statistics do include content morphemes that would take zero case marking or do not require overt marking in Croatian. The remaining 22 EL nouns, where overt case marking is required, have no markers or non-standard markers. Most (85%) of these English-based forms in Hlavac’s corpus are nouns, as in example (2). Again, this finding supports the argument that outsider SMs resist replacement in a contact situation.

6.3. Late SMs in Creole Development

More evidence that distributions in other data add up to a reason to distinguish morpheme types comes from data on the form of creoles. Elsewhere, I have argued for a scenario in creole formation that segregates outsider SMs from other lexifier/substrate language morphemes in regard to their participation. (Myers-Scotton 2001, 2002). While lexifier content morphemes are reanalyzed as outsiders and other grammatical elements in the developing creole, lexifier outsider SMs are not
available to become part of the creole. Part of my argument is that in the turbulent social milieu in which creoles developed, the basically structure roles of outsider SMs made them less learnable than content morphemes. But even more important are the reasons related to their nature and production that are discussed in the above sections. How, then, were the needs for argument structure and co-indexing met when creoles developed? Typically, content morphemes from the lexifier (superstrate) were reanalyzed as late system morphemes. I cite just three examples, first, in Hawaiian Creole, the English morpheme *get* serves as both a “that-like” complementizer and as a locative existential meaning ‘there is’. Second, it is well-known by researchers who study Tok Pisin that the future tense marker comes from English *by and by*. Also, in Tok Pisin, at least as spoken on New Ireland, the transitive marker (*-i*) comes from English *he*. This marker resembles the transitive marker in Tigak, the substrate language for Tok Pisin in New Ireland in both form and function (Jenkins 2002). Evidence about the structure of creoles today largely supports this scenario. In sum, outsider SMs from the lexifier do not participate in creole formation; instead, they largely are created from contentful elements.

6.4. Abruptness in Language Shift

As mentioned earlier, outsiders typically are members of sets, conjugations or declension. The case of the East African language, Ma’a is suggestive of what this means for their borrowability. Ma’a’s speakers seem to have moved from an earlier (apparently Cushitic) grammatical system to a largely very different (Bantu) grammatical system. That is, the entire set of agreement morphemes, which largely resemble those in neighboring Bantu languages, characterize Ma’a today. This outcome suggests that if individual outsiders are to be “borrowed”, an entire system of outsiders is borrowed (cf. Mous 2003 *inter alia* on Ma’a).

Also, my own recent research indicates the scenario of Ma’a development may be the general model for what happens when speakers shift their allegiance from one language to another (cf. Myers-Scotton 2007). The study deals with Xhosa L1 speakers and their allegiance to Xhosa in the face of English dominance in the South African public society. In such a scenario, outsider SMs are not replaced in an L1 in a piece-meal fashion, nor are L1 outsiders found in a new dominant language in the context of a language shift.

Specifically, my research tested the hypothesis that when individual bilingual speakers begin to use a community-dominant L2 as their main out-group language instead of their L1, they do not accomplish this shift by incorporating L2 late SMs into their L1. Rather, the hypothesis continues, the shift is abrupt, from one grammatical system to another (cf. Haugen 1950 on Norwegian immigrants in the U.S. and their shift to English).

Xhosa-English bilinguals were studied who are vulnerable to shift in that they live far from their home area (Eastern Cape) in the main industrialized area of South Africa (Gauteng Province). There, English is the main lingua franca of public life. (Various indigenous languages, especially Zulu, also are lingua francas in informal interactions among black South Africans.)

Data were gathered in informal interviews with 48 Xhosa-English bilinguals. The interviewer, also an L1 speaker of Xhosa, was instructed to conduct the interview mainly in Xhosa, switching only occasionally to English words or phrases. Clearly, the bias was for the interviewees to speak Xhosa. Instead, many speakers showed a good deal of CS and many of them produced full clauses in English only. However, even in the group showing the most advanced use of English, the greatest percentage of their clauses had Xhosa grammatical frames. Even among these front runners, 30% was the highest percentage of English-framed clauses.

These data imply that these speakers have not necessarily shifted to English as a main public language; for example, even the front runners do not “override” the Xhosa interviewer and speak just English. True, when they speak English they often are only integrating English words into a
Xhosa frame (i.e., CS), but these are not English late SMs, but content morphemes. For example, rather than use the English existential construction *there is/are*, which includes an outsider SM, they use the Xhosa counterpart, *ku-kho/a-yi-kho* ‘there is/ there is not’. The Xhosa counterpart appears even in such an English-ladden example as (14).

(14) Xhosa-English (Myers-Scotton 2005a corpus)

| I don’t agree with i-notion ...i-crime i-be-ninzi because ku-kho i-democracy, |
| But i-democracy because ku-kho ne-transparency, |
| ku-kho i-availability ye information ... |

‘... I don’t agree with the notion, crime has increased because there is democracy. But...democracy, it exists, along with transparency, [and] there is [the] availability of information.’

Also, remember that these Xhosa-English bilinguals often produce full clauses in English. Thus, these data about the production of such clauses support the hypothesis that the progress of shift is abrupt as far as the grammatical system is concerned.

Also note the Xhosa bridge SM (*ye*) in the associative constituent (*i-availability ye information*) in example (14).

7. Concluding Remarks

Although this paper devotes a good deal of space to presenting “something old” (i.e., descriptions of the MLF and 4-M models), it also presents “something new”, new arguments about the nature and roles of morpheme types that include insights that are relevant to contact linguistics. Understanding the arguments and recognizing any insights depends on familiarity with the two models.

A major development in this paper is that the new ways of characterizing “outsider” late SMs integrate them more clearly into the general architecture of language. That is, the role of these morphemes in the identifying and positioning of NP arguments in a clause is now emphasized in their characterization. In effect, these morphemes are seen as making the primary arguments in any clause transparent. This insight about their importance in the structuring of clauses helps explain why morphemes of this type are unlikely to be open to transfer or copying in contact situations. They are tied into an agreement system that structures the clause such that undoing the knot is difficult at best.

Another development is the introduction of the Uniform Structure Principle. This simply-worded principle serves to recognize that each constituent in language has a uniform structure that is maintained whenever the constituent appears. In bilingual speech, such as CS, the principle prefers the structure of one of the participating languages. By extension, the principle predicts that there will be asymmetry between the roles of the participating languages in bilingual speech. Overall, the principle can be seen as a resisting force in contact situations. This does not mean that contact-induced changes do not happen frequently, especially in intense contact situations. But it is important to distinguish the type of changes that do happen from those that are much less frequent. Frequent changes are largely to the elements and patterns that make up the semantic-pragmatic interface (e.g., Aikhenvald 2006; Doğruöz & Backus 2007; Heine & Kuteva 2005).
Yet another development is the Differential Access Hypothesis, a hypothesis that is motivated both by the distribution of morpheme types in bilingual data and by the characterization of these morpheme types under the 4-M model. This hypothesis suggests that what are called content morphemes and early SMs under the 4-M model become salient early in the production process (at the level of the mental lexicon). That is, these two types of morpheme are conceptually-activated. In contrast, the two types of structurally-assigned morphemes (late SMs) only are accessed when language-specific directions about constructing clause-level constituents are received at the formulator.

The implication of this hypothesis is that when a morpheme becomes salient in a production model influences the patterns that are open to that morpheme at the surface level.

Indirectly, this hypothesis offers an explanation of why content morphemes and also early system morphemes, although to a much more limited extent, are more open to transfer or copying to another language, or substitution by another morpheme in their original language, than are the late SMs. The suggestion is that morphemes that become “available” at the level of the mental lexicon are that much more available to change in contact situations.

In summary, the paper places a value on recognizing the abstract nature of late SMs, and especially outsider late SMs, both how it regard to their place in the architecture of language and to how they are accessed in language production. Their abstract nature seems to offer explanations for why these morphemes differ from other morphemes in their availability to change in contact situations. When the Uniform Structure Principle is linked to the Differential Access Hypothesis, the prediction is that these morphemes are not easily open to transfer or substitution with elements from another language. This limiting has to do their critical roles in the clause, how they relate elements in the clause to each other.

References


