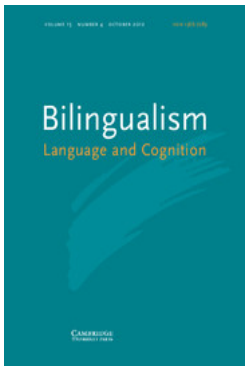

Early emergence of structural constraints on code-mixing: evidence from French–English bilingual children

Johanne Paradis, Elena Nicoladis and Fred Genesee

Bilingualism: Language and Cognition / Volume 3 / Issue 03 / December 2000, pp 245 - 261
DOI: null, Published online: 11 January 2001

Early emergence of structural constraints on code-mixing: evidence from French–English bilingual children*

JOHANNE PARADIS,
ELENA NICOLADIS
University of Alberta
FRED GENESEE
McGill University

Does young bilingual children's code-mixing obey the same structural constraints as bilingual adults' code-mixing? The present study addresses this question using code-mixing data from 15 French–English bilingual children filmed in conversation with both parents at six-month intervals from the age of 2;0 to 3;6. The children's code-mixed utterances were examined for violations of the principles set out in the Matrix-Language Frame model (e.g. Myers-Scotton, 1993, 1997). The results show that the children obeyed all the constraints set out in the Matrix Language Frame model the majority of the time. With respect to the Morpheme Order Principle and to the interaction of Congruence and Matrix Language Blocking, they demonstrated consistent adherence with only marginal violations from the outset. In contrast, the children produced comparatively more frequent violations of the System Morpheme Principle and showed increasing adherence to this principle over time. We discuss possible explanations for the contrast between the children's performance on the System Morpheme Principle and the other constraints, which include the unequal emergence of INFL in the acquisition of French and English.

The mixing of elements from two languages together in one utterance (intra-sentential code-mixing) has been the basis for much speculation on the nature of the developing linguistic representation(s) of simultaneous bilinguals. For example, some researchers have proposed that early code-mixing constitutes evidence for the claim that bilingual children initially have a unitary linguistic representation for their two languages, with differentiation of the two systems occurring later in the preschool years (Leopold, 1949/71; Volterra and Taeschner, 1978; Redlinger and Park, 1980, for example). However, the unitary language system (ULS) hypothesis has been criticized on both methodological and empirical grounds (Genesee, 1989; Lanza, 1997b). Furthermore, substantial counter-evidence to the ULS hypothesis from bilingual children's pragmatic, syntactic, lexical and phonological development strongly argues that a dual linguistic representation is likely established from the earliest stages of acquisition studied (Meisel, 1989; Lanza, 1992; Genesee, Nicoladis and Paradis, 1995; Pearson, Fernandez and Oller, 1995; Quay, 1995; Nicoladis and Genesee, 1996a; Paradis, 1996; Paradis and Genesee, 1996; Lanza, 1997b; Johnson and Lancaster, 1998; Paradis, 1998a, in press-a, in press-b, among others).

Even if we do not consider code-mixing to be the outcome of fused linguistic representations, the function and form of bilingual children's code-mixing can still be considered informative of their developing linguistic knowledge. Research on the form and function of code-mixing (usually referred to as code-switching) in adult bilinguals reveals that this is a rule-governed linguistic behaviour, both socio-pragmatically and grammatically (for overviews, see e.g. Myers-Scotton, 1993; Milroy and Muysken, 1995; Grosjean, 1997). Regarding form in particular, it has been suggested that adult code-mixing is guided by a specific set of structural constraints that form part of a speaker's fundamental linguistic competence (Pfaff, 1979; Poplack, 1980, 1981; di Sciullo, Muysken and Singh, 1986; Myers-Scotton, 1993; Belazi, Rubin and Toribio, 1994; Bhatt, 1997, for example). Structural constraints refer to restrictions on what elements from language[a] can be inserted, and where they can be inserted, into a sentence in language[b], and thus refer to intra-sentential code-mixing and not to the switching of single-language utterances between conversational

* We would like to thank the parents and children for their participation in our research programme, Isabelle Boivin for her assistance in the data collection and Carol Myers-Scotton and two anonymous reviewers for helpful critical comments on an earlier version of this paper. This research was supported by grants from the Social Sciences and Humanities Research Council of Canada to Fred Genesee (410–99–0933 and 410–95–0726), for which we are grateful.

Address for correspondence
Johanne Paradis, Department of Linguistics, 4–46 Assiniboia Hall. University of Alberta, Edmonton, AB, T6G 2E7, CANADA
E-mail: *johanne.paradis@ualberta.ca*

turns (inter-sentential code-mixing). Assuming that young simultaneous bilinguals will become mature adult bilinguals, it is pertinent to ask whether the structural "rules" of code-mixing are subject to developmental change, so that bilingual children gradually approximate the adult rule system, or whether bilingual children demonstrate their obedience to these code-mixing rules as soon as multiword sentences emerge in their speech. This issue of whether there is a developmental shift in the structural organization of code-mixing in bilingual first language acquisition is the focus of the present study. In examining the emergence of organized code-mixing, we are also examining the overall emergence of grammatical knowledge, since it is assumed that rules governing code-mixing are part of that knowledge.

Most investigations of the structural properties of bilingual children's code-mixing and how these relate to adult patterns have looked at the frequencies of different syntactic categories occurring as single-item insertions in mixed utterances (Vihman, 1985; Lanza, 1992, 1997b; Meisel, 1994; Köppe and Meisel, 1995; Köppe, 1996; Deuchar and Quay, 1998). This kind of investigation is related, albeit indirectly, to models of structural constraints because most models include differential restrictions based on syntactic category. For example, grammatical morphemes are usually more restricted in where they can be mixed in a clause than content morphemes like nouns (Poplack, 1980, 1981; Myers-Scotton, 1993; Belazi et al., 1994, for example). Vihman (1985) examined the mixed utterances of an Estonian–English bilingual child, aged 1;8–2;0, and argues that the proportional frequency of different syntactic categories appearing in mixed utterances differs from what has been found for older bilinguals. Mixed words were assigned to three categories: nouns, verbs and function words. In contrast to bilingual adults and older bilingual children, function words and not nouns comprised the largest number of types and tokens mixed. In a subsequent study including data from the same child, Vihman (1998) found that from the age of 2;8 the more prevalent mixing categories tended to be lexical, such as nouns and verbs. Meisel (1994), Köppe and Meisel (1995) and Köppe (1996) also found high rates of function word mixes (as defined by Vihman, 1985) in the early code-mixing of German–French bilingual children, as well as a shift over time to a predominance of nouns as single category insertions. Also employing a functor/contentive distinction, Lanza (1997b) found that for one of her Norwegian–English bilingual subjects, the trend was to mix more functors than contentives; however, the opposite pattern prevailed for the other child studied.

Based on these findings, one might conclude that the structure of code-mixed utterances undergoes developmental change from the immature to the mature bilingual speaker. There are a number of problems with such an interpretation. First, Nicoladis and Genesee (1998), using the same definition of functors and contentives as Lanza (1992), found that five French–English bilingual children had approximately equal rates of mixing for functors and contentives. Furthermore, they found that some children used more grammatical mixing and others more lexical mixing. A second and more important problem, however, is that it is possible this early pattern is not truly evidence for a developmental shift, but rather an artefact of category definition. Vihman (1985) defines function words as an essentially "catch all" category of elements that are not nouns, verbs or adjectives. As such, her category of function words includes items that do not really share syntactic or semantic characteristics. For example, pronouns are classified as function words in Vihman's analysis (see also Lanza, 1992, in which the child's "grammatical" code-mixing is almost entirely due to mixed pronouns), but in recent versions of syntactic theory, pronouns in languages like English fill lexical categories like nouns (Marantz, 1995). Furthermore, as pointed out in Meisel (1994), Vihman's category of function words would include words like "yes/no" and "bye-bye" which play no central syntactic role in a clause and, because of their peripheral status, do not bear on most proposed structural constraints on code-mixing. Meisel (1994) argues that constructions with these kinds of words as mixed items are more properly classified as tag-switches rather than true intra-sentential code-mixes. Therefore, in our view, resolving the issue of whether child and adult bilinguals show similar trends for mixing by syntactic category requires an investigation based on a more motivated taxonomy of syntactic categories.

Although taxonomies offer some basis of comparison between child and adult bilinguals, most proposals of structural constraints on adult code-mixing focus on restrictions on the placement of items within the syntactic string. Thus, investigations of the syntactic structure of bilingual children's code-mixing are more directly relevant to the question of whether bilingual children obey the same constraints as adults or not. For example, Lederberg and Morales (1985) conducted a kind of grammaticality judgment task where older children (8–10 years old) and adults indicated their acceptance of different kinds of code-mixed constituents. They found that the children accepted certain structures the adults did not; however, these results may not indicate develop-

mental changes in knowledge of code-mixing constraints. The authors suggest that the child–adult differences they found could be "due to changes in the ability to make acceptability judgments rather than changes in the grammar of code switching" (Lederberg and Morales, 1985, 134).

Regarding research on younger children, Lanza (1997a) and Vihman (1998) examined bilingual children's code-mixed utterances in reference to the Matrix Language Frame (MLF) model of code-switching constraints (e.g. Myers-Scotton, 1993, 1997). Lanza (1997a) concludes that code-mixing of younger and older bilinguals is of the same kind. However, she concentrates more on the relationship between dominance and the matrix language *vis à vis* overall code-mixing patterns, rather than providing a detailed analysis of whether switch sites and inserted elements in a mixed utterance conform to MLF model constraints. In contrast, Vihman (1998) examines the structural properties of bilingual children's code-mixed utterances with respect to the violation of specific constraints set out in the MLF model, and concludes that the structure of the children's code-mixes follows the predictions of this adult model. Vihman's subjects, Raivo and Virve, range in age from 2;8 to 6;7 and 5;11 to 9;10 and thus, only Raivo's earliest data fall into the same age range that concerns us in our study (2;0–3;6). Because we are interested in determining whether adherence to constraints develops over time, it is possible that children younger than Vihman's subjects show greater evidence of constraint violations.

Meisel (1994) examined the code-mixed utterances of two German–French bilingual children similar in age to our subjects, approximately 1;6–3;0. Meisel (1994) sought to determine whether the structure of the children's intra-sentential code-mixes complied with the Government Constraint proposed for adult bilinguals (di Sciullo et al., 1986). Similarly to Lanza (1997a) and Vihman (1998), Meisel concludes that the children's mixed utterances adhere to a modified version of this constraint. However, the Government Constraint refers to a certain syntactic configuration only and thus is not a comprehensive and integrated set of constraints like the MLF model. In addition, Meisel argues that examination of structural constraints on code-mixing is irrelevant at the early stages of grammatical development and consequently examines only a subset of his data. Specifically, he puts forth a *grammatical deficiency hypothesis*, which asserts that there is a stage in children's grammatical development where their word combinations are not constrained by linguistic principles in general, including the particular principles related to code-mixing. This is the stage before children show productive use of the morphosyntax associated with functional categories. Once there is evidence that functional categories are instantiated in the children's grammars, it is expected that both their single-language and mixed-language utterances conform to linguistic principles. Thus, Meisel (1994) suggests that bilingual children are sensitive to the adult rules of code combination in sentences only when they produce sentences of sufficient grammatical complexity for the rules to apply. By implication, then, there may be an early stage where code-mixing is unconstrained.

This proposal of two stages in grammatical acquisition is pertinent to our investigation for the following reasons. First, we examine code-mixing in bilingual children from the age of 2;0 to 3;6, hence at the younger ages our children should be in Meisel's pre-grammatical/pre-functional category/pre-structural constraints stage. Second, the children in our study are acquiring French and English simultaneously. Previous research on the emergence of functional categories in the monolingual and bilingual acquisition of French and English shows that the functional category INFL, which is mainly associated with tense and agreement markers, is acquired at different rates in the two languages (Pierce, 1992; Ferdinand, 1996; Paradis and Genesee, 1996, 1997; Rice, Wexler and Hershberger, 1998; Paradis and Crago, 1999, 2000). While children tend to master the use of tense and agreement markers well before the age of three years in French, English-learning children demonstrate mastery of these morphemes between the ages of three and four years. In other words, the children in our study would be in a pre-functional stage in English while in a functional stage in French at the same time. Therefore, although the general research question that concerns the present study is whether or not young bilingual children's code-mixing obeys the same constraints as have been proposed for adults, our examination of French–English bilingual children raises additional, more particular questions: Would bilingual children demonstrate sensitivity to code-mixing constraints when the functional category INFL is overtly instantiated in one language before the other? Would their code-mixing patterns demonstrate sensitivity to constraints that demand knowledge of language-specific properties of INFL? If we find that young French–English bilingual children demonstrate such sensitivity, the results of this study will be relevant not only to issues in bilingual development, but also to a broader issue in grammatical development, namely whether children demonstrate possession of complex syntactic knowledge before they consistently use grammatical morphemes associated with functional

categories in their language production (for discussion of this issue, see Meisel, 1992; Wexler, 1994, 1996, 1998; Paradis and Genesee, 1997).

In order to address both the general and specific questions posed, the present study examines French–English bilingual children's adherence to the structural constraints on intra-sentential code-mixing as set out in the MLF model. We examine longitudinal data from the early stages of word combinations to determine if a developmental shift in sensitivity to the proposed constraints occurs. In addition, we focus certain analyses on aspects of code combination that require knowledge of the language-specific properties of INFL in French and English in order to adhere to the MLF model's constraints. In so doing, we expect to show whether sensitivity to structural constraints changes over time and to what extent it is linked to the acquisition of INFL.

The Matrix-Language Frame Model

We chose the MLF model to examine children's adherence to structural constraints on code-mixing because it offers an integrated, comprehensive set of constraints, rather than a single restriction on certain configurations, like the Government Constraint (di Sciullo et al., 1986) or the Functional Head Constraint (Belazi et al., 1994). However, since the various proposals for structural constraints differ in their claims, it is important to point out that our findings may not generalize to other models.

We adopt the MLF model as explicated in Myers-Scotton, (1993), Myers-Scotton (1995), Myers-Scotton and Jake (1995) and (1997), Jake and Myers-Scotton, (1997a, b), and Myers-Scotton and Jake (2000). There are two central components to this model: (1) The distinction between a matrix and an embedded language; and (2) the distinction between system and content morphemes. The matrix/embedded language distinction indicates that the two languages that participate in a code-mixed utterance do not play equal roles and are not subject to the same restrictions. The matrix language (ML) is the base language[1] of the bilingual utterance or CP (also called the "host" or "recipient" language), and the embedded language (EL) is the one whose elements are inserted into the matrix-language frame. We discuss how to determine which language is the ML in our Method section. The ML plays the dominant role in that the morphosyntactic frame for the sentence is set by this language. More specifically, certain system morphemes must be from the ML and the order of all morphemes follows the rules of the ML (System Morpheme Principle and Morpheme Order Principle, defined below). The EL contributes only content morphemes and/or a special set of system morphemes to the sentence with the exception of EL Islands (EL Island Trigger Hypothesis, defined below). Insertion of EL morphemes is further constrained by congruency between the two languages (ML Blocking Hypothesis, defined below). Hence, a mixed utterance can consist of three types of constituents: ML Islands, ML + EL constituents and EL Islands. ML system morphemes and morpheme order prevails in ML islands and in ML + EL constituents. Constituents can be NPs, PPs, APs (adjective and adverb), VPs and IPs.

The second important distinction is between system and content morphemes. System morphemes can be bound or free morphemes. The system/content distinction corresponds approximately but not exactly to the traditional classification of closed-class versus open-class morphemes or to the functional and lexical category distinction in recent syntactic theory. The essential difference between system and content morphemes lies in participation in thematic role assignment and quantification operations. Nominal arguments, like the subject and object of the verb, receive theta-roles like agent of action or patient of action and lexical verbs assign these roles to their nominal arguments. Thus, nouns and verbs are archetypal content morphemes. Examples of system morphemes include determiners, adverbs, negative operators, verb and noun inflections, copulas and auxiliary verbs. More precise indications of how the French and English morphemes used by the children in this study fall into the system versus content split are provided in the Results and Discussion section.

In an extension of the MLF model to the 4–M MLF model, Myers-Scotton and Jake (2000) break down the category of system morphemes into early and late system morphemes. This breakdown has consequences for how system morphemes can participate in ML + EL constituents. Early system mor-

[1] Jake and Myers-Scotton (1997b), Bolonayai (1998) and Myers-Scotton and Jake (2000) argue that the ML can be a composite of EL and ML properties in certain contexts, for example, the inter-language context of L2 learners. In the case of intermediate L2 learners, the L1 influence can be present at the abstract grammatical level in L2–ML constructions such that the ML cannot exclusively be the L2, but instead both the L1 and the L2 set the sentential frame of the mixed CP. However, this kind of composite ML is not directly applicable to simultaneous bilingual acquisition where there is no sequence, hence no L1 and L2. Most important for our study, possible EL influence in the composite ML does not include the presence of overt EL late system morphemes in EL + ML constituents, and thus a composite ML is not an "escape hatch" for violations of the System Morpheme Principle.

phemes are activated by a content morpheme and their form and appearance are determined within the maximal projection of that content morpheme. For example, plural morphemes on the noun are determined by the noun and are within the NP, and so are considered early system morphemes. Also, grammatical features encoded in irregular forms, like the past tense in the verb "ran", are considered early system morphemes because in the language production model Myers-Scotton and Jake (in press) assume, irregular forms have a single lemma, meaning that the past tense feature is stored with the verb stem for irregular forms. In contrast, late system morphemes are activated by positional or functional relations outside of their maximal projections. For example, regular verb inflection for agreement involves late system morphemes because the verb looks beyond the VP to the subject of the CP for the form of the agreement morphemes. Importantly, in the 4-M model, determiners are analysed not as heads of DP, but as NP-internal morphemes. Thus, determiners are early system morphemes.

The placement of system (early and late) and content morphemes from the ML and the EL are regulated by the following interacting set of constraints:

(1) The System Morpheme Principle (SMP): "All syntactically or externally relevant system morphemes come only from the ML in ML + EL constituents" (Myers-Scotton, 1995, p. 239). In terms of the 4-M model, "externally relevant" refers to late system morphemes. So, early system morphemes are excluded from this restriction.
(2) The Morpheme Order Principle (MOP): "Surface morpheme order will be that of the ML in ML + EL constituents" (Myers-Scotton, 1995, p. 239). If a content morpheme from the EL is placed in a different position in the EL than in the ML, the order followed for the ML + EL constituent is that of the ML. For example, if language[a] places adjectives post-nominally and language[b] places them pre-nominally, when language[a] is the ML, an EL adjective from language[b] should be placed post-nominally.
(3) The EL Island Trigger Hypothesis: If an entire EL constituent is mixed as opposed to a single item insertion, then late system morphemes from the EL may appear within that constituent. EL islands can be NPs, VPs, IPs or adjunct constituents like adverb phrases.
(4) The ML Blocking Hypothesis: A filter blocks the insertion of an EL content morpheme not congruent with its corresponding ML morpheme. Congruency refers to whether the morpheme is system or content in the two languages. If an EL content morpheme has a system morpheme counterpart in the ML, insertion is blocked because there is no congruency. For example, if language[a] pronominals are system morphemes like clitics, and language[b] pronominals are content morphemes, in constructions where language[a] is the ML, language[b] pronominals cannot be mixed as EL items even though they are content morphemes.

When adherence to constraints is complicated by absence of congruence or other factors, certain compromise strategies are often employed by bilingual speakers (see especially Jake and Myers-Scotton, 1997a). For example, prevalent use of EL Islands might occur in code-mixing between a language pair when there is an absence of congruence for many structures and morphemes. Also, the insertion of bare verb stems from the EL with a dummy verb from the ML to attach verbal inflections could arise when inflectional procedures are very different between the two languages. What is pertinent about compromise strategies to the present study is that when used consistently they indicate the presence of systematic syntactic knowledge of both languages. We present evidence below for the children's possible use of such compromise strategies for a number of structures, but for the mixing of pronominals in particular.

Method

Participants

The data for this study were drawn from a corpus collected from children who participated in studies examining other aspects of bilingual language acquisition (Genesee, Nicoladis and Paradis, 1995; Genesee, Boivin and Nicoladis, 1996; Nicoladis and Genesee, 1996a, b, 1997, 1998; Paradis and Genesee, 1996, 1997; Paradis, in press-a). Because these prior studies employed similar methods of data collection, it was deemed reasonable to compile the data together for the present study. In addition, one bilingual child (Jason) who had not been included in previous studies was observed using the same methodology as previous studies at the age of 1;11. This child was living in the USA at the time he was filmed, but the family had recently moved from Montreal and it was thought that the structural constraints of his code-mixing would not be changed from the recent move.

We examined data from fifteen French–English bilingual children for this study. Each child was first-

Table 1. *Children's across-Period dominance and ages at observation Periods I to IV*

Children	Dominance	Period 1: 2;0	Period II: 2;6	Period III: 3;0	Period IV: 3;6
Mathieu	English	1;11	2;3	2;11	3;6
Nicholas	English	1;11	2;3	3;1	3;8
Olivier	French	1;11	2;3	2;10	3;6
Stefan	Bal/Eng[a]	2;0	2;7	3;1	3;5
Yann	Balanced	2;0	2;5	3;1	3;8
William	English	2;2	2;10	3;3	
Gene	Balanced	1;10	2;7	3;0	3;7
Elise	English	1;9			
Tanya	English	2;1			
Brigitte	English	1;11			
Jennifer	French	2;1			
Jessica	Balanced	1;11			
Jason	English	1;11			
Joelle	English		2;4		
Leila	English		2;3		

[a] Bal/Eng = dominance changed over time

born and had no siblings at the outset of data collection. Each child had one French-speaking and one English-speaking parent and all families claimed to be using a "one-parent, one-language" strategy of language use with their children. In other words, according to their own reports, each parent used primarily one language when addressing their child. The children's language samples were collected approximately from the age of 2;0 to 3;6 at six-month intervals. Prior research shows this is the period when overt reflexes of the functional category INFL emerge gradually in the children's speech, and are acquired fully in French in advance of English (Paradis and Genesee, 1996, 1997). As shown in Table 1, the data set for this study is semi-longitudinal, with six of the fifteen children providing samples at each time interval.

Also included in Table 1 is an indication of each child's dominant language, if any, because we refer to language dominance in one of our analyses. The children's dominant language was determined according to comparative vocabulary size and MLU in each language as well as parental report on language exposure. Dominant languages for these children were determined in our previous work, and we have transferred this information to Table 1. For further details on how dominance was measured, see Genesee et al. (1995), Nicoladis and Genesee (1996a, b), 1997) and Genesee et al. (1996).

Procedures and data selection

All children were visited in their homes by one or two experimenters and audio- and video-taped in naturalistic play activities with their parents. Each session lasted from 45 minutes to one hour. All children were taped in separate sessions playing and speaking with their mother alone and with their father alone. All children except Jason were also filmed in a session with both parents present. Thus, for each of the age intervals given in Table 1, three samples of spontaneous speech were collected from each child (except for Jason from whom two samples were collected).

Approximately 20 minutes of video-tape from each session were transcribed in accordance with the CHAT system (MacWhinney, 1991) and the children's utterances were coded for language of the utterance (French-only, English-only or Mixed) and addressee (Mother, Father, or Both parents). The sub-corpus used for the present study consists of all the mixed utterances that are constraint relevant with respect to the MLF model (Total = 371). Mixed utterances that are constraint neutral were excluded from our analyses (Total = 109). Our rationale for exclusion is that the latter kind of utterance provides no challenges to any constraints. Constraint-relevant utterances are those that contain any system morpheme, and/or content morphemes which take different word orders in French and English, and/or pronominals. Utterances we categorized as constraint neutral were of one of the following types. The first type are utterances consisting only of content morphemes where no morpheme order or congruency differences between French and English arise, for example, *oiseau sing* "bird sing". This utterance has a bare noun and verb stem in a root clause, so has no system morphemes like tense/agreement/aspect markers or determiners. Since both French and

English are SVO, there are no morpheme order differences. Also, this utterance has no pronominals, which pose congruence differences between French and English. The second type of constraint-neutral mixed utterances are those where the mixed element is on the periphery of the sentence or CP, similar to tag-switching discussed above. Examples are sentences entirely in one language with "yes/no"– *oui/non*, "please/thank you" – *s'il vous plaît/merci*, "hello/byebye" – *bonjour/byebye*, or the French discourse marker *là* attached at the beginning or end. For utterances with "no" – *non* on the periphery, we are referring to anaphoric usage and not to negation within the clause, the latter being included in the constraint-relevant category. The third type of constraint-neutral utterance consists of juxtaposed translation equivalents, like "sleeping *dodo*". The fourth type are counting sequences.

It is noteworthy that by limiting our analyses to constraint-relevant utterances only, the denominators used to calculate per cent violations are smaller and thus the proportion of violations might be higher than if they were calculated for the entire corpus. In the Conclusion section we provide a calculation of total per cent violations out of the entire set of mixed utterances. The advantage of excluding constraint-neutral utterances is that we only perform analyses on utterances that directly pertain to our research question of whether or not young bilingual children's code-mixing obeys the same structural constraints as that of adult bilinguals.

Determining the matrix language

According to Myers-Scotton (1993, Chapter 3), the determination of the ML should be based on both sociolinguistic and psycholinguistic factors for a stretch of discourse (not for an individual utterance). Regarding sociolinguistic factors, the ML is the expected or typical language for the type of interaction in the discourse sample. In the case of the young bilingual children in our study, the language of the adult interlocutor is most likely the expected language since the "one-parent, one-language" strategy is employed in their homes (see also Lanza, 1997a). Regarding psycholinguistic factors, the ML is the language from which the majority of morphemes in the discourse sample are taken. For child–adult discourse samples, the majority of the morphemes contributed to the sample always come from the parent because they are more proficient with language in general. Thus, the application of this psycholinguistic criterion would always yield the same ML as the sociolinguistic criterion. However, most of the children in this study are dominant in one language, and they sometimes use that language when speaking to the parent who speaks their non-dominant language. If the children use more of their dominant than non-dominant language in a session, it is a fair assumption that their dominant language is the ML in the sense of being the morphosyntactic frame for their sentence production. Therefore, a more developmentally appropriate psycholinguistic criterion would be that the ML is the language from which the majority of the child's morphemes come in a stretch of discourse.

We determined the ML for each video-taped session on the basis of this developmental psycholinguistic criterion. For the single-parent sessions, the determination of the ML was based on the frequency of morphemes produced by the child in his/her single language utterances only (i.e., French-only and English-only utterances). Mixed utterances were excluded from the determination of the ML in order to obtain an independent measure of the ML. For sessions with both parents present, the child's utterances were divided by addressee, essentially creating two mini-sessions in one. On the basis of this division, the ML was determined by the frequency of morphemes used by the child to each of the parents separately. Thus, the ML could alternate in the sessions with both parents. Utterances addressed to both parents were excluded from all analyses because they were too infrequent to serve as the basis for a ML calculation. It is important to note that in spite of our reliance on a psycholinguistic criterion, in the majority of sessions (99 out of the 138 ML calculations; or 72 per cent) the language the child used more frequently was also the parent's language. Thus, our child-centered psycholinguistic criterion produced the same results as a sociolinguistic criterion would have in most cases.

After the ML was determined for each session, and by extension for each utterance in that session, the children's mixed utterances were analysed for their adherence to the key constraints in the MLF model.

It is important to point out that in more recent versions of the MLF model, criteria for determining the ML have been revised such that the ML could be determined on a case-by-case basis for the bilingual CP alone (e.g. Myers-Scotton, 1997; Bolonyai, 1998; Myers-Scotton and Jake, 2000). We chose to adopt the Myers-Scotton (1993) definition of ML for use in our study for the following reasons. First, determining the ML over a stretch of discourse gives predictive power to the analysis of constraint violations within individual utterances. This predictive property is essential for our purposes because we are evaluating children's performance in terms of obedi-

ence to the MLF model constraints. Determining the ML on the basis of each mixed utterance could lead to circularity in our evaluation of obedience to constraints. For example, if an ML + EL constituent contains a late EL system morpheme, it violates the SMP. This violation would be eliminated if we reversed which language we claimed was the ML. Thus, the Myers-Scotton (1993) criteria appear more objective. Second, even if we invoked a different objective criterion for determining the ML, such as considering the ML to be the language contributing the most morphemes to the CP, this criterion would be difficult to implement in our child language data where there are a substantial number of two-morpheme mixed utterances. For two-morpheme utterances, the selection of the ML would be essentially arbitrary.

We acknowledge that use of the Myers-Scotton (1993) criteria means that our analyses are insensitive to possible ML switches within the same stretch of discourse for the sessions with the parents alone. We discuss the implications of our choice of ML determination for our results in our examination of SMP violations.

Results and Discussion

System Morpheme Principle

Recall that according to the SMP, late system morphemes must be in the ML, unless they appear as part of EL Island constituents. In order to test the children's sensitivity to this constraint, all mixed utterances including a system morpheme were analysed for each developmental period for the language of the system morpheme as a function of constituent type, ML+EL, ML Island and EL Island, and system morpheme type, early and late. Acceptable or correct mixes consisted of three types: (1) An early or late ML system morpheme in an ML island or ML+EL constituent; (2) An early or late EL system morpheme in an EL island; or (3) An early EL system morpheme in an ML+ EL constitutent. Incorrect mixes or violations consisted of late EL system morphemes in an ML+EL constituent. Late system morphemes in these data were the following: quantifiers like "some" and "any", tense and agreement inflections, infinitival "to" in English (but not the infinitival inflection in French, see Myers-Scotton and Jake, 2000), auxiliary verbs, modal auxiliaries, copulas, do-support DO, negative operators[2] and pronominal clitics.

[2] The children often used negative markers followed by a nominal to indicate refusal or denial, for example "no *jus*" to mean "I don't want juice" or "no *loup*" to mean "that's not a wolf". These kinds of constructions are characteristic of child language only. Thus, we thought it might be justified to eliminate utterances with non-sentential negation from our analysis of system morpheme distribution because it is uncertain how the formulation of these non-adult-like constructions takes place and whether the negative operators are subject to the same constraints as when they appear in sentential negation.

Table 2. *Incorrect system morpheme mixes (SMP violations) for Periods I to IV*

	SMP Violations
Period I	4.2% (2/48)
Period II	31% (27/87)
Period III	20.9% (29/139)
Period IV	8.1% (7/86)

SM Incorrect = EL late system morpheme in ML + EL constituent.
$X^2 = 22.619, p < .0001$

Results of this analysis are presented in Table 2. The overall rate of SMP violations across the four intervals is 18.1 per cent. Although correct mixes comprise the majority at each time interval, there is a substantial number of violations at Periods II and III (30.1 per cent and 20.9 per cent respectively), followed by a shift towards fewer violations at Period IV (8.1 per cent). The smaller proportion of violations at Period I, as compared with Period II, might be a reflection of the children's less-advanced stage in language development. They may have had fewer system morphemes acquired in their lexicons at this point, and thus fewer opportunities for violations. A Chi-Square analysis of correct and incorrect system morpheme use over Periods I to IV confirms the presence of an interaction between use and time ($X^2 = 22.619$, $p < .0001$). In sum, if we take the 90 per cent-use-in-obligatory-context criterion to indicate mastery in developmental data (Brown, 1973), the 18.1 per cent overall violations of the SMP falls below this limit. In addition, it appears as if adherence to this constraint is subject to developmental change such that mastery is not achieved until the age of 3;6. Let us consider some possible explanations for the substantial number of violations to the SMP.

One consideration might be whether the violations of the SMP are reasonably spread out across the corpora of the individual children, or whether they cluster in the corpus of one child. On the one hand, all the children except four produced utterances in violation of the SMP. The four children who did not, Tanya, Elise, Leila and Nicholas, contributed very few utterances to the mixed corpus, so the absence of violations could be accidental. Furthermore, children

who produced mixed utterances across more than one time interval produced violations at each time interval. (One exception is William, who produced violations at Period I and III but not at Period II.) On the other hand, when the number of violations per child per period is examined, it is possible that one child is contributing more than the others. More specifically, approximately half of all violations at Periods II and III, which are the Periods containing the greatest number of SMP violations in our data, come from Gene's corups. In fact, Gene produced more intra-sentential code-mixes than the other children as a proportion of his overall language use and determination of the ML in his play sessions was often a closer race between French and English than for the other children. In our prior research, Gene has been identified as being a balanced bilingual from the age of 2;0 and his parents code-mixed more than the other parents of the children in this study (Genesee, Nicoladis and Paradis, 1995). Thus, it is possible that Gene's level of bilingual competence and family language context enabled him to not only code-mix more frequently but also perhaps to initiate switches in the ML throughout a parent-alone play session. If this is the case, because our criteria for determining the ML is not sensitive to within-session switches in the ML for parent-alone sessions, some of the apparent violations of the SMP in Gene's data from Periods II and III may not be true violations. Instead, the increase in violations at these periods may be an artefact of our criteria for determining the ML of a mixed utterance.

The potential skew posed by Gene's data notwithstanding, there are still a substantial number of violations to the SMP in these data. Instead of reconsidering whether these violations of the SMP are valid, let us now consider whether there are developmental constraints on the children's ability to obey the SMP. In other words, let us assume that these violations of the SMP are true violations, but that the children have no choice but to violate the SMP in certain circumstances (in contrast to more mature bilingual speakers). One potential source of children's inability to adhere to the SMP at all times is lexical gaps. Children acquiring two languages simultaneously do not acquire translation equivalents for every morpheme in tandem (Pearson et al., 1995; Quay, 1995; Nicoladis and Genesee, 1996a; Nicoladis and Secco, 1998). Hence, a bilingual child may not have acquired the ML late system morpheme required in a certain utterance and thus may be forced to select an EL system morpheme instead. Similarly, even when children acquire a translation equivalent for a system morpheme, the system morpheme that they have already known for some time might be selected frequently at first because its lexical entry or lemma would be easier to activate. Unfortunately, we cannot directly test the hypothesis that lexical gaps may be responsible for violations of the SMP because complete reports of each child's cumulative vocabulary were not part of the data collection procedure. We cannot conclude that a translation equivalent of a certain system morpheme was unknown to a child on the basis of spontaneous speech samples alone (see Nicoladis and Secco, 1998). But we can investigate the influence of factors possibly responsible for lexical gaps, hence SMP violations; for example, individual dominance and unequal morphosyntactic development between French and English.

The dominant language, if any, of each child in this study based on analyses from our prior research is given in Table 1. It could be hypothesized that children would have acquired a greater number of system morphemes in their dominant language and, thus, would be forced to use them even in violation of the SMP. Other researchers have found dominance to be a predictor of the overall directionality of grammatical morpheme mixing (Petersen, 1988; Lanza, 1997a, 1997b).[3] The bilingual children these researchers studied tended to mix grammatical morphemes from their dominant to their non-dominant language, but not vice versa. We were not interested in examining dominance with respect to overall mixing directions because such an analysis is not pertinent to the questions posed in this study regarding structural constraints. Instead, we investigated whether dominance could predict the directionality of system morpheme mixing only in utterances that violate the SMP. Accordingly, we examined the SMP violations for the seven children in this study who were identified as having a dominant language and who produced SMP violations. SMP violations were categorized as: (1) The use of a dominant language late system morpheme in an utterance where the ML is the non-dominant language, or (2) The use of a non-dominant language late system morpheme in an utterance where the ML is the dominant language. A preponderance of mixing type (1) might imply that dominance plays a role in forcing SMP violations. The results of this calculation are presented in Table 3. Only three children, Mathieu, William and Brigitte, showed a majority of violations consisting of dominant language system morphemes in a non-dominant ML utterance. For Brigitte and Mathieu, only one viola-

[3] Lanza (1997b) appears to consider directionality of mixing to be a determiner of dominance, as well as dominance to be a predictor of directionality. Such circular reasoning with respect to this issue renders the relationship between the two phenomena unclear (see Paradis, 1998b).

Table 3. *Distribution of SMP violations according to dominance*

Child	EL-Dom+ML-Non-Dom	EL-Non-Dom+ML-Dom
Mathieu	1	0
Olivier	1	1
William	9	6
Brigitte	1	0
Jennifer	0	3
Jason	0	1
Joelle	0	1
TOTAL	12	12

EL-Dom + ML-Non-Dom = EL late system morpheme from the dominant language in a non-dominant ML utterance; EL-Non-Dom + ML-Dom = EL late system morpheme from the non-dominant language in a dominant ML utterance

tion occurred, so directionality cannot really be reliably determined. It is worth noting that the overall frequencies of violations in Table 3 are low compared to the total number of SMP violations (25 versus 65), indicating that the majority of violations detected in the entire corpus were contributed by balanced bilingual children, a finding not expected if SMP violations were driven by individual language dominance.

A second developmental factor that could underlie SMP violations is the unequal morphosyntactic development of French and English. As mentioned in the Introduction, late system morphemes associated with the functional category INFL, for example auxiliary verbs, copulas, modals, tense and agreement inflections, emerge later and are mastered later in English than in French. Moreover, Paradis and Genesee (1996, 1997) found that the earlier emergence of INFL in French occurred even for bilingual children who were dominant in English. The presence of unequal development of such system morphemes could result in unequal or directional mixing patterns between French and English, where far more French system morphemes than English system morphemes of this type appear in mixed utterances. Thus, the presence of French EL INFL-related system morphemes in English ML constructions might be greater than English EL INFL-related system morphemes in French ML constructions.

To test this prediction, we calculated the distribution of all SMP violations according to the language of the ML and EL and according to whether the late system morpheme was non-INFL related or INFL related. The results presented in Table 4 provide some support for the directionality prediction. The largest category of SMP violations consisted of ML-English + EL-French INFL system constructions (33/65), and the number of violations involving INFL-related system morphemes for ML-English + EL-French constituents was nearly double that for the ML-French + EL-English constituents (33 versus 17). Therefore, the morphosyntactic gap inherent in the bilingual acquisition of French and English could underlie many SMP violations.

Table 4. *Distribution of mixed constituent and system morpheme type in violations of the System Morpheme Principle*

	Non-INFL	INFL
ML-Fr + EL-Eng	9	17
ML-Eng + EL-Fr	6	33

Non-INFL = quantifiers, adverbs, negative markers; INFL = auxiliary verbs, modals, copulas, clitics, verb inflections

This directional pattern of mixing of INFL items from one language to another has also been attested in a German–English bilingual child (Gawlitzek-Maiwald and Tracy, 1996). The researchers in this case argue that this child's mixing patterns could be viewed as a gap-filling strategy because, like French, INFL-related items emerge earlier in German than in English. Thus, the child was profiting from her bilingualism to increase the communicative complexity of her English sentences, so-called "bilingual bootstrapping". If we view the SMP violations consisting of French INFL items in English ML clauses in the same way as Gawlitzek-Maiwald and Tracy (1996), we could consider all violations of this type to be akin to the compromise strategies used by older bilinguals or, at least, violations of the SMP that serve a developmental purpose.

Morpheme Order Principle

Another key constraint in the MLF model is the MOP. Recall that this constraint stipulates that in ML + EL constituents, the order of the morphemes must be that of the ML. The application of this constraint is visible when the two languages have divergent word orders for morphemes in certain constituents. For French and English, we have identified three constituent types where the morpheme order differs between the two languages: possessor–possessed constructions, adjective–noun constructions and negative-marker–thematic-verb constructions. As will be demonstrated below, violations of the MOP are infrequent in comparison to the SMP because they occur in just 8.8 per cent of all relevant utterances across construction type.

Possessor–possessed constructions

As illustrated in (5a) and (5b), the order of the possessor noun and possessed noun in possessor–possessed constructions differs in French and English, with possessor + possessed as the English order and possessed + possessor as the French order. Note that a construction such as "the house of the teddy-bear" is possible in English, but is not as common as "the teddy-bear's house" and more importantly, the children in this study did not use possessive "of" in English. If the possessor–possessed construction is in an ML + EL constituent, the order of each element will differ depending on the ML. If French is the ML, the order in (5c) should obtain. If the ML is English, the opposite order as in (5d) should obtain. These divergent word orders should occur even if the possessive marker, the system morpheme, is omitted from the NP, which is a common phenomenon in early child language.

(5) a. the teddy-bear's house
b. *la maison du nounours*
the house of the teddy-bear
c. *maison* + teddy-bear
d. teddy-bear + *maison*

There are just eight examples of possessor–possessed constructions in our entire corpus; however, all examples follow the predictions of the MOP without exception. The order differences within possessor–possessed constructions as a function of ML are demonstrated by the examples in (6), all from one child. EL morphemes are in upper case.

(6) a. *MADAME* cookies? (Jessica I; ML-Eng)
"lady's cookies"
b. *MADAME* ball (Jessica I; ML-Eng)
"lady's ball"
c. GUITAR *monsieur* (Jessica I; ML-Fr)
"mister's guitar"
d. FLOWER *madame* (Jessica I; ML-Fr)
"lady's flower"

Adjective – noun constructions

In English, all adjectives precede the noun they are modifying, as shown in (7a) and (7b). In contrast, French has two categories of adjectives, one which precedes the noun and another which follows the noun. An example of each type is given in (7c) and (7d). Based on this difference between French and English, we can predict that in English ML constituents with French adjectives, the adjectives must precede the noun regardless of their category type in French, for example (7e) and (7f). Conversely, if the ML is French, the adjectives should be positioned according to their category, as shown in (7g) and (7h).

(7) a. big dog
b. red dog
c. *grand chien*
big dog
d. *chien rouge*
dog red
e. *grand* dog
f. *rouge* dog
g. *grand* dog
h. DOG *rouge*

We found a total of 20 utterances with mixed adjective–noun constituents in our corpus, some examples of which are presented in (8). The utterances in (8a) and (8b) have English as the ML with a French adjective and English noun, and an English adjective and a French noun, respectively. The adjective *petit* is noun-preceding in French. The utterances in (8c) and (8d) are French ML utterances, one with an English adjective and French noun and one with a French adjective and an English noun, respectively. The example in (8e) shows the use of a French adjective in an English ML constituent where this adjective, *rose*, normally follows the noun in French. In accordance with the MOP, *rose* precedes the noun in this sentence. The example in (8f) is the only violation we found in this set. The French adjective *bon* should precede the noun; however, because the ML of this utterance is English, no adjective should follow the noun. Perhaps this utterance is actually an attempt at "the leg is good" and has an omitted copula. If this is the case, it does not belong in this set.

(8) a. *PETIT* bird (Elise I)
little
b. big *BOBO* (William II)
booboo
c. TWO *pirates* (William III)
d. *des petits* CAR (Yann IV)
some little
e. my *ROSE* bat (Olivier III)
pink
f. leg *BON* (William III)
good

Negative marker – thematic verb constructions

In French, all finite verbs, thematic and non-thematic, are situated to the left of the negator *pas* "not" (Pollack, 1989). This process is shown below in (9a) for thematic verbs in the present tense and for

the non-thematic, auxiliary verb *avoir* "have" in (9b). Note that in (9b) the thematic verb is in a non-finite form, and thus appears to the right of the negator. On the contrary in English, thematic verbs are always placed to the right of the negator. Non-thematic verb forms bear the tense features in negative constructions and appear to the left of the negator. These verb forms include do-support DO, shown in (9c) and auxiliary HAVE, as in (9d). Thus, the [± thematic] status of a verb determines its placement *vis à vis* the negative marker in English, whereas, the [± finite] status of a verb is the crucial distinction for determining verb placement in French.

(9) a. *Le lion (ne) voit pas l'éléphant.*
the lion see-(pres) not the elephant
"the lion does not see the elephant"
b. *Le lion (n') a pas*
the lion have-(aux-past) not
vu l'éléphant.
see-(past part.) the elephant
"the lion did not see the elephant"
c. The lion does not see the elephant.
d. The lion has not seen the elephant.

The differences in verb placement between French and English are explained in current syntactic theory through the operation of overt verb movement (Pollack, 1989; Chomsky, 1992; Marantz, 1995). In languages like French, verbal forms that are finite, or bear tense features, move overtly from the VP to INFL, and such movement is surface transparent when a negative marker is present in the syntactic string. This process is represented by the bracketed version of example (9a) presented in (10a). In languages like English, only non-thematic verb forms bearing tense are present in INFL on the surface. The absence of overt thematic verb movement is demonstrated in the bracketed version of the example in (9c) given in (10b).

(10) a. [$_{IP}$ *le lion* [$_{INFL}$ *voit*] [$_{NEGP}$ *pas*] [$_{VP}$ [*l'éléphant*]]]
b. [$_{IP}$ the lion [$_{INFL}$ does] [$_{NEGP}$ not] [$_{VP}$ see [the elephant]]]

Based on this difference in verb placement between French and English, we can make the following predictions regarding mixed utterances with negative markers, in line with the MOP. If the ML is French, then all non-finite verb forms, such as participles, infinitives and bare verb stems (English), should follow the negative marker, while all finite verb forms, whether thematic or non-thematic, should precede the negative marker. If English is the ML, then all thematic verb forms should follow the negative marker, while all non-thematic verb forms should precede the negative marker.

We found 29 utterances with sentential negation in our corpus, 13 with French as the ML and 16 with English as the ML. For French ML utterances, violations of the MOP consist of NEG–V [+finite] and V[−finite] – NEG order combinations. For English ML utterances, violations of the MOP consist of V[+thematic] – NEG and, NEG – V[−thematic] order combinations. Four utterances in the sentential negation set (14 per cent) contained violations of the MOP and these are given in (11). First, the utterance in (11a) has English as the ML, but the negative marker appears to be in an EL Island due to the left-dislocated subject, thus, this may not be a true violation of the MOP. Excluding this utterance, the total proportion of violations becomes 9.7 per cent. Second, note that all the utterances in (11) are also SMP violations because the negative marker, a system morpheme, is from the EL lexicon. In each case, the word order used by the child is in accordance with the language of the negative marker. It is possible that this is a compromise strategy when an EL system morpheme is erroneously accessed. However, it is important to note that two other negative utterances from the greater set that are also SMP violations follow the morpheme order of the ML. Therefore, it seems that in the case of an SMP violation for a negative marker, the children choose either the ML or EL word order. Overall, the key generalization from this analysis is that we found no violations of the MOP for negative constructions that were not also SMP violations.

(11) a. fish, *IL MANGE PAS.* (Gene III)
"fish, he doesn't eat"
b. NO *va.* (Mathieu II)
"no goes"
c. I *AIME PAS MAMAN.* (Gene III)
"I don't love Mommy"
d. I like *PAS* strawberries. (Gene III)
"I don't like strawberries"

Congruence, the SMP and ML Blocking

In addition to the SMP, system morpheme mixes are also constrained by the degree of congruence between the language pair. As mentioned above, if pronominals are content morphemes in one language and system morphemes in the other, then congruence between the two languages does not exist for pronominals and further restrictions such as ML Blocking apply to mixed utterances with pronominals (see also Jake, 1994).

French and English are only partially congruent

with respect to pronominals. In English, pronouns (I, you, he, me, him, etc.) are free-standing content morphemes which occupy argument positions within an NP. In French, there are two types of pronominals. Pronominal subjects (*je* "I", *tu* "you", *il*, "he" etc) are clitics, late system morphemes which attach to a verbal host and, under certain analyses, do not occupy argument positions (Cummins and Roberge, 1993; Kaiser, 1994; Auger, 1995). The second type of pronominals are the "strong pronouns" (*moi* "me", *toi* "you", *lui* "him", etc.), which are free-standing content morphemes that can occupy certain argument positions. The contrast between the two types of pronominals in French is best illustrated in subject-doubled constructions, for example *Moi j'aimerais aller au parc* "(me) I would like to go to the park". In subject-doubled constructions, the strong pronoun *moi* occupies the subject argument position and the clitic *j'/je* is arguably functioning as an agreement marker on the verb. Subject-doubled constructions can be distinguished from left dislocations by the absence of a pause between the pronoun and the clitic and by the prevalence of usage which is approximately 75 per cent of the time in the Québec French dialect (Auger, 1995). Also, children acquiring French either in a monolingual or bilingual context, often go through a stage where they use the strong pronouns alone as full NP subjects, for example *Moi tomber* "me fall" (Pierce, 1992; Ferdinand, 1996; Paradis and Genesee, 1996).

According to a strict interpretation of the ML Blocking Hypothesis, when French is the ML, English pronouns should not be mixed even though they are content morphemes because of the lack of congruency between the two languages with respect to pronoun category status. However, because French also has content morpheme pronominals, and these strong pronouns are used as subjects in child French, we could expect English pronouns to be mixed in a French ML utterance because they are congruent with strong pronouns. However, only strong pronouns and not clitics from French could be mixed in an English ML utterance because of the SMP. The only exception would be an EL Island constituent.

The difference in status between clitics and pronouns intersects with verb movement (discussed above regarding the MOP) to predict further limitations on the mixing of pronominal morphemes between French and English. Because subject clitics behave syntactically like bound morphemes marking person agreement, they are INFL-related items and move with the verb in the syntax in French. Recall that in English, thematic verb movement is not apparent on the surface. These structural differences are illustrated in (12) for the sentence *Moi j'aime l'été* "I love summer". Note that the key difference is the empty versus filled INFL constituent. The consequences of verb movement could influence mixing of clitics in the following way: even if French is the ML, a switch might not occur between a clitic and the verb. In other words, insertion of an English verb, in spite of verbs being content morphemes, could be blocked or avoided because of the lack of congruence in both the status of subject pronominals and the placement of clitic + verb constructions in the sentence.

(12) a. $[_{IP}$ *moi* $[_{INFL}$ *j'aime*$]$ $[_{VP}$ [*l'été*]]]
b. $[_{IP}$ I $[_{INFL}$ $]$ $[_{VP}$ love [summer]]]

All mixed utterances including a pronominal subject from the corpus were examined with respect to our predictions. The distributional frequencies of mixed utterance types are presented in Table 3 and examples of each acceptable mixed utterance type are given in (13). All mixed utterance types are acceptable in the MLF model except the final category, an English ML utterance with a clitic. The data are not divided according to developmental period because we detected no changes over time in the distribution of mixing patterns. Table 3 shows that the vast majority of mixed utterances with pronominals fall under the acceptable categories; violations of constraints with respect to pronominals comprise just 13 per cent of all mixed utterances with pronominals. Furthermore, as predicted, we found no examples of French ML utterances with clitics that included an English verb; whereas, examples (13e) to (13h) show that switching between a pronominal and a verb is possible for other pronominal types.

(13) a. *non, il est* BAD GUY.
(William III; ML-Fr + clitic)
"no, he is bad guy"
b. *je veux aller manger* TOMATO.
(Olivier II; ML-Fr + clitic)
"I want to go eat tomato"
c. I *aime pas ça, moi*. (Gene IV; ML-Fr + pro)
"I don't like that, me"
d. me *cacher loup* YUM+YUM.
(Yann III; ML-Fr + pro)
"me hide wolf yum+yum" (syntactic relationship of yum+yum is uncertain)
e. AND HE *tombe* WITH THE BICYCLE.
(Mathieu IV; ML-Fr + pro)
"and he falls with the bicycle"
f. I *VAS* taxi. (Jason I; ML-Eng + pro)
"I go taxi"
g. he *A POMME*. (Gene II; ML-Eng + pro)
"he has apple"

Table 5. *Distribution of mixed utterance types with pronominal subjects*

Utterance Type	Frequency
ML-Fr + clitic	17
ML-Fr + spro	1
ML-Fr + pro	41
ML-Eng + pro	38
ML-Eng + spro	7
*ML-Eng + clitic	16

clitic = French system morpheme; spro = strong pronoun = French content morpheme; pro = pronoun (nominative, accusative or genitive) = English content morpheme

h. *MOI* do it this, *MOI*.
(William III; ML-Eng + spro)
"me do it this, me"

i. *MOI VEUX* more.
(Joelle II; ML-Eng + spro)
"me want more"

Even though the percentage of violations is low, it is worth noting that the majority of them appear to demonstrate the use of compromise strategies that might qualify their status as true violations. All examples of English ML utterances with clitics are presented in (14). In examples (14a) to (14f), the clitic appears to be situated in an EL Island, which is a permissible constituent for a French system morpheme to appear in. Specifically, in (14a) and (14b), the English word "there" seems to be an adjunct to an otherwise French sentence. In (14c) to (14e), the English subject NP is left-dislocated. Finally, in (14f) an English CP contains an entire IP in French. Furthermore, note that all the examples except (14l) and (14n) include a finite French verb with the clitic. Since a finite French verb resides in INFL, it appears as if an entire INFL constituent, perhaps an EL island,[4] is mixed into an English utterance, rather than just one French system morpheme. Because of the divergence between English and French with respect to verb movement, the children's choice to mix an entire INFL constituent instead of simply a clitic seems to be evidence of the use of a compromise strategy based on sophisticated knowledge of pronominal status and verb movement rules in each language. If we exclude the EL Island constructions as well as the mixed-INFL constructions, this leaves two true or uncompensated violations of MLF constraints, representing 1.7 per cent of all mixed utterances with pronominal subjects. Note that these excluded utterances are also SMP violations, so if they could be considered compromise structures with respect to congruence, they could also be considered non-violations of the SMP.

To summarize, not only are there a marginal number of exceptional English ML with clitic utterances, the majority of these reveal deft use of compromise strategies and hence, complex language-specific grammatical knowledge.

(14) a. *oui, il met* there. (Gene III)
"yes, he puts there"

b. *IL FAIT DODO* there. (Gene III)
"he is sleeping there"

c. fish, *il mange pas*. (Gene III)
"fish, he doesn't eat"

d. flat, *IL A MANGÉ*. (Gene III)
"flat, he ate"

e. *LA* boy, *IL A WOO+WOO*. (William III)
"the (?) boy, he has woo+woo" (woo+woo is onomatopoeic for train in both languages)

f. but *PAPA IL MANGE HOTDOG*. (Gene IV)
"but papa he is eating hotdog" (hotdog is a borrowed word in Quebec French)

g. I *j'aime maman*. (Olivier I)
"I love mommy"

h. *IL A* finish. (Gene II)
"he finished"

i. *IL Y A* my toesie there. (Gene III)
"there are my toes there"

j. *IL MET* tootsie *BOTTE*. (Gene III)
"he puts foot boot"

k. *IL A* put this. (William III)
"he put (past tense) this"

l. *quoi il* want from eating? (Gene IV)
"what he wants from (for?) eating"

m. *POUQUOI IL A LE* lights? (Gene IV)
"why he has the lights"

n. *ON* buy *AUTRE AMI* (Stefan IV)
"we buy other friend"

o. *ELLE COUPE* her hair (Stefan IV)
"she cuts her hair"

p. *ELLE COUPE* her hair (Stefan IV)
"she cuts her hair"

Conclusions

The principal question we sought to address in the present study is whether bilingual children demonstrate a developmental shift in structural properties of code-mixing from a non-adult to an adult-like system. As a whole, our data do not support the developmental shift hypothesis. First, our examina-

[4] It is uncertain whether INFL itself could be an island because of the hierarchical nature of CP construction. An IP constituent dominates VP, and so on a conventional analysis, the entire IP must be in the EL for it to constitute an island.

tion of adherence to the MOP and to the intersection of ML blocking, convergence and the SMP for pronominals revealed only a small number of true violations (those without evidence of compromise strategies), ranging from 0 per cent to 9.7 per cent, and with no discernible changes over time. Based on a 90 per cent-use-in-obligatory-context criterion, mastery of the MOP and the constraint interactions regarding pronominals appeared to be achieved from the outset of word combinations. Second, in spite of the greater number of violations to the SMP, the children obeyed this constraint the majority of the time, 82 per cent overall. Third, because we excluded constraint-neutral utterances from our analysis, our calculations of the children's violations of MLF model constraints are greater than they would have been if all mixed utterances had been considered. If we calculate the total number of utterances with a violation to any constraint (both with and without compromise strategies evident) out of the total number of constraint-relevant mixed utterances, the violation rate for the data set is 24.8 per cent. In contrast, if we calculate the total number of utterances with a violation to any constraint out of the total number of mixed utterances, the overall violation rate for the data set is 17.5 per cent. Thus, taken together, these patterns are not consistent with an across-the-board qualitative shift from no sensitivity to structural constraints to a stage where code-mixing adheres to structural constraints.

However, we did find a developmental trend in violations to the SMP. Violations diminish with time, and the 82 per cent overall adherence rate to this constraint falls far below a 90 per cent criterion. Because this developmental trend pertains to only one constraint, it is relevant to ask whether the children's sensitivity to or awareness of this constraint changes over time, or whether other factors underlie the developmental trend. Among the factors we considered, both the skew from one subject who could have been switching the ML during sessions and the later emergence of INFL in English appeared to play a role in the frequency of SMP violations. Concerning the latter factor, our data suggest that until children have acquired a sufficient number of lexical items associated with tense and agreement in English, they may insert such items from French in English ML utterances in violation of the SMP. Consequently, it may not be the SMP constraint that matures or is learned over time, but instead it is the children's lexicons which must develop in order to give them the tools to adhere more strictly to the constraint (see Nicoladis and Secco, 1998, for a similar argument about pragmatic constraints). Future research should focus on the relationship between children's vocabulary and their violations of the SMP constraint in order to determine whether the tendency shown in these data is truly a robust phenomenon that could underlie the majority of early SMP violations.

Even though the later emergence of INFL-related morphology in English may contribute in part to SMP violations, our data do not seem to support the notion that code-mixing, or grammatical structure in general, is unconstrained before sufficient use of such morphology occurs in both languages, as proposed by Meisel (1994). The children in the present study show evidence from the outset of language-specific, INFL-related grammatical knowledge in their mixing patterns involving sentential negation and pronominal subjects. Moreover, even apparent constraint violations reveal knowledge of language-specific syntactic structures, as in the compromise strategies shown in the mixing of pronominals. Therefore, we need to distinguish between children's language-specific syntactic knowledge associated with INFL, and their acquisition and use of morphemes marking tense and agreement. Our data seem to indicate that the former can be apparent before the latter is mastered in production. In sum, akin to Meisel's (1994) *grammatical deficiency hypothesis*, we found evidence for a relationship between grammatical development and the emergence of structural constraints on code-mixing. However, in contrast to Meisel's proposal, we only found evidence for the impact of a certain aspect of grammatical/lexical development on children's ability to obey a certain constraint, rather than evidence for an overall shift in grammatical organization, including the emergence of organized code combination.

In conclusion, that these children demonstrated general adherence to adult-like structural constraints in most of their code-mixing implies not only that they have complex knowledge of how to fit their two languages together in one utterance during production, but also that they possess language-specific syntactic knowledge even during an early period of development where the use of INFL-related morphosyntax is variable in their two languages, and the mastery of INFL-related morphosyntax lags in one of their two languages.

References

Auger, J. (1995). Les clitiques pronominaux en français parlé informel: une approche morphologique. *Revue Québécoise de Linguistique*, 24, 21–60.

Belazi, H., Rubin, E. & Toribio, J. (1994). Codeswitching and X-bar theory: the functional head constraint. *Linguistic Inquiry*, 24, 221–237.

Bhatt, R. (1997). Code-switching, constraints and optimal grammars. *Lingua*, 102 (4), 223–251.

Bolonyai, A. (1998). In-between languages: language shift/maintenance in childhood bilingualism. *International Journal of Bilingualism*, 2 (1), 21–43.

Brown, R. (1973). *A first language: the early stages*. Cambridge, MA: Harvard University Press.

Chomsky, N. (1992). A minimalist program for linguistic theory. *MIT Occasional Papers in Linguistics*, 1, 1–69.

Cummins, S. & Roberge, Y. (1993). A morphosyntactic analysis of romance clitic constructions. In M. Mazzola (ed.), *Issues and theory in Romance linguistics*, pp. 239–257. Washington, DC: Georgetown University Press.

Deuchar, M. & Quay, S. (1998). One vs. two systems in early bilingual syntax: two versions of this question. *Bilingualism: Language and Cognition*, 1, 231–243.

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

Ferdinand, A. (1996). *The development of functional categories: the acquisition of the subject in French*. Dordrecht: ICG.

Gawlitzek-Maiwald, I. & Tracy, R. (1996). Bilingual bootstrapping. *Linguistics*, 34, 901–926.

Genesee, F. (1989). Early bilingual development: one language or two? *Journal of Child Language*, 6, 161–179.

Genesee, F., Boivin, I. & Nicoladis, E. (1996). Talking with strangers: a study of bilingual children's communicative competence. *Applied Psycholinguistics*, 17, 427–442.

Genesee, F., Nicoladis, N. & Paradis, J. (1995). Language differentiation in early bilingual development. *Journal of Child Language*, 22, 611–631.

Grosjean, F. (1997). Processing mixed languages: issues, findings and mode. In A.M.B. de Groot & J.F. Kroll (eds.), *Tutorials in bilingualism: psycholinguistic perspectives*, pp. 225–254. Mahwah, NJ: Lawrence Erlbaum.

Jake, J. (1994). Intrasentential code switching and pronouns: on the categorial status of functional elements. *Linguistics*, 32, 271–298.

Jake, J. & Myers-Scotton, C. (1997a). Codeswitching and compromise strategies: implications for lexical structure. *International Journal of Bilingualism*, 1, 25–39.

Jake, J. & Myers-Scotton, C. (1997b). Relating interlanguage to codeswitching: the composite matrix language. In E. Hughes, M. Hughes & A. Greenhill (eds.), *BUCLD 21 Proceedings*, pp. 319–330. Somerville, MA: Cascadilla Press.

Johnson, C. & Lancaster, P. (1998). The development of more than one phonology: a case study of a Norwegian–English bilingual child. *International Journal of Bilingualism*, 2, 265–300.

Kaiser, G. (1994). More about INFL-ection and agreement: the acquisition of clitic pronouns in French. In J. Meisel (ed.), *Bilingual first language acquisition: French and German grammatical development*, pp. 131–160. Amsterdam/Philadelphia: John Benjamins.

Köppe, R. (1996). Language differentiation in bilingual children: the development of grammatical and pragmatic competence. *Linguistics*, 34, 927–954.

Köppe, R. & Meisel, J. (1995). Code-switching in bilingual first language acquisition. In L. Milroy & P. Muysken (eds.), *One speaker, two languages*, pp. 276–301. Cambridge: Cambridge University Press.

Lanza, E. (1992). Can bilingual two year olds code-switch? *Journal of Child Language*, 19, 633–658.

Lanza, E. (1997a). Language contact in bilingual two-year-olds and code-switching: language encounters of a different kind? *International Journal of Bilingualism*, 1, 135–162.

Lanza, E. (1997b). *Language mixing in infant bilingualism: a sociolinguistic perspective*. Oxford: Clarendon.

Lederberg, A. & Morales, C. (1985). Code switching by bilinguals: evidence against a third grammar. *Journal of Psycholinguistic Research*, 14 (2), 113–136.

Leopold, W. (1949/71). *Speech development of a bilingual child: a linguist's record* (Vols. I–IV). New York: AMS.

MacWhinney, B. (1991). *The CHILDES project: tools for analyzing talk*. Hillsdale, NJ: Erlbaum.

Marantz, A. (1995). The minimalist program. In G. Webelhuth (ed.), *Government and binding theory and the minimalist program*, pp. 349–382. Cambridge: Blackwell.

Meisel, J. (1989). Early differentiation of languages in bilingual children. In K. Hyltenstam & L. Obler (eds.), *Bilingualism across the lifespan: aspects of acquisition, maturity and loss*, pp. 13–40. Cambridge: Cambridge University Press.

Meisel, J. (1992) (ed.). *The acquisition of verb placement: functional categories and V2 phenomena in language acquisition*. Dordrecht: Kluwer.

Meisel, J. (1994). Code-switching in young bilingual children: the acquisition of grammatical constraints. *Studies in Second Language Acquisition*, 16, 413–441.

Milroy, L. & Muysken, P. (1995). *One speaker, two languages*. Cambridge: Cambridge University Press.

Myers-Scotton, C. (1993). *Duelling languages: grammatical structure in codeswitching*. Oxford: Clarendon.

Myers-Scotton, C. (1995). A lexically based model of codeswitching. In L. Milroy & P. Muysken (eds.), *One speaker, two languages*, pp. 233–256. Cambridge: Cambridge University Press.

Myers-Scotton, C. (1997). Afterword. *Duelling languages: Grammatical structure in codeswitching* (Paperback Edition). Oxford: Clarendon.

Myers-Scotton, C. & Jake, J. (1995). Matching lemmas in a bilingual language competence and production model: evidence from intrasentential code switching. *Linguistics*, 33, 981–1024.

Myers-Scotton, C, & Jake, J. (2000). Explaining aspects of codeswitching and their implications. In J. Nicol (ed.), *One mind, two languages: bilingual language processing*. Cambridge: Blackwell.

Nicoladis, E. & Genesee, F. (1996a). A longitudinal study of pragmatic differentiation in young bilingual children. *Language Learning*, 46, 439–464.

Nicoladis, E. & Genesee, F. (1996b). Bilingual communica-

tion strategies and language dominance. In A. Stringfellow et al. (eds.), *BUCLD 20 Proceedings*, pp. 518–527. Somerville, MA: Cascadilla.

Nicoladis, E. & Genesee, F. (1997). The role of parental input and language dominance in bilingual children's code-mixing. In E. Hughes et al. (eds.), *BUCLD 21 Proceedings*, pp. 422–432. Somerville, MA: Cascadilla.

Nicoladis, E. & Genesee, F. (1998). Parental discourse and codemixing in bilingual children. *International Journal of Bilingualism*, 2, 85–99.

Nicoladis, E. & Secco, G. (1998). The role of translation equivalents in a bilingual family's code-mixing. In A. Greenhill et al. (eds.), *BUCLD 22 Proceedings*, pp. 576–585. Somerville, MA: Cascadilla.

Paradis, J. (1996). Phonological differentiation in a bilingual child: Hildegard revisited. In A. Stringfellow et al. (eds.), *BUCLD 20 Proceedings*, pp. 528–539. Somerville, MA: Cascadilla.

Paradis, J. (1998a). Prosodic phonology in bilingual two year olds: one system or two? Paper presented at the International Conference on Infant Studies, Atlanta, GA.

Paradis, J. (1998b). Review of language mixing in infant bilingualism: a sociolinguistic perspective by E. Lanza. *Journal of Child Language*, 25, 728–733.

Paradis, J. (in press-a). Beyond "one system or two": Degrees of separation between the languages of French–English bilingual children. To appear in S. Döpke (ed.), *Cross-linguistic structures in simultaneous bilingualism*. Amsterdam/Philadelphia: John Benjamins.

Paradis, J. (in press-b). Do bilingual two year olds have separate phonological systems? *International Journal of Bilingualism*.

Paradis, J. & Crago, M. (1999). Is French SLI EOI? Paper presented at the 24th Annual Boston University Conference on Language Development, University of Boston, Boston, MA.

Paradis, J. & Crago, M. (2000). Tense and temporality: a comparison between children learning a second language and children with SLI. *Journal of Speech, Language and Hearing Research*, 43 (4), 834–847.

Paradis, J. & Genesee, F. (1996). Syntactic acquisition in bilingual children: autonomous or interdependent? *Studies in Second Language Acquisition*, 18, 1–15.

Paradis, J. & Genesee, F. (1997). On continuity and the emergence of functional categories in bilingual first language acquisition. *Language Acquisition*, 6, 91–124.

Pearson, B., Fernandez, S. & Oller, D.K. (1995). Cross-language synonyms in the lexicons of bilingual infants: one language or two? *Journal of Child Language*, 22, 345–368.

Petersen, J. (1988). Word-internal code-switching constraints in a bilingual child's grammar. *Linguistics*, 26, 479–493.

Pfaff, C.W. (1979). Constraints of language mixing: intersentential code-switching and borrowing in Spanish/English. *Language*, 55, 291–318.

Pierce, A. (1992). *Language acquisition and syntactic theory*. Dordrecht: Kluwer.

Pollack, J-Y. (1989). Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry*, 20, 365–424.

Poplack, S. (1980). "Sometimes I start a sentence in English y termino en Espanol": Toward a typology of code-switching. *Linguistics*, 18, 581–618.

Poplack, S. (1981). Syntactic structure and social function of codeswitching. In R. Duran (ed.), *Latino language and communicative behavior*, pp. 169–184. Norwood, NJ: Ablex.

Quay, S. (1995). The bilingual lexicon: Implications for studies of language choice. *Journal of Child Language*, 22, 369–387.

Redlinger, W.E. & Park, T. (1980). Language mixing in young bilinguals. *Journal of Child Language*, 7, 337–352.

Rice, M., Wexler, K. & Hershberger, S. (1998). Tense over time: the longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language and Hearing Research*, 41, 1412–1431.

Vihman, M. (1985). Language differentiation by the bilingual infant. *Journal of Child Language*, 12, 297–324.

Vihman, M. (1998). A developmental perspective on codeswitching: conversations between a pair of bilingual siblings. *International Journal of Bilingualism*, 2, 45–84.

Volterra, V. & Taeschner, T. (1978). The acquisition and development of language by bilingual children. *Journal of Child Language*, 5, 311–326.

Wexler, K. (1994). Optional infinitives, head movement and the economy of derivations. In D. Lightfoot & N. Hornstein (eds.), *Verb movement*, pp. 305–363. Cambridge: Cambridge University Press.

Wexler, K. (1996). The development of inflection in a biologically based theory of language acquisition. In M. Rice (ed.), *Towards a genetics of language*, pp. 114–144. Mahwah, NJ: Lawrence Erlbaum.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: a new explanation of the optional infinitive stage. *Lingua*, 106, 23–79.

Received September 30, 1999 Revision accepted May 31, 2000