Syntactic frames and single-word code-switching

A case study of Mandarin Chinese – Norwegian bilingualism

Tor A. Åfarli & Fufen Jin
Norwegian University of Science and Technology (NTNU) /
The Norwegian Institute of Public Health (FHI)

Mainstream core grammar theory, still to some extent relying on the idealized speaker in a homogeneous speech community, is ill equipped to handle different kinds of periphery data, like code-switching data and other types of language mixing data. In this paper, we defend a model of grammar that we argue is able to handle different types of both core and periphery data. Empirically, we focus on single-word code-switching data obtained from a small corpus of the Chinese production of a bilingual Mandarin Chinese – Norwegian child. We develop the beginnings of a generative frame model of grammar exploiting insights both from late insertion neo-constructional models and from code-switching theories assuming a matrix – embedded language asymmetry. We will be paying special attention to the lexicon – syntax interface.

1. Introduction

It is common in science to make simplifying assumptions and ignore many fuzzy properties of the empirical data in order to obtain a general and elegant theory or analysis. This kind of idealization is a typical characteristic of generative linguistics, too, where a major aim has been and still is understood as giving an account of the language of an idealized speaker – hearer in a homogeneous speech community (cf. Chomsky 1965: 3). Such an account is typically given in terms of what is often called a core grammar, i.e. a grammar that ignores exceptions, mixing and other linguistic idiosyncracies. Such deviating phenomena are counted as periphery phenomena not to be accounted for by the core grammar (cf. Chomsky 1981: 7 ff.).

The mixing of different languages in the speech of one individual, for example as in various types of code-switching, is found in all speech communities and individuals, and has been the subject of much investigation in e.g. sociolinguistics

and language acquisition research. However, less attention has been paid to such periphery phenomena within the generative paradigm, probably because these phenomena would not be easily accounted for in a streamlined core grammar. Thus, Chomsky (1991:42) remarks that beside the core language phenomena, the mind of the speaker "contains a 'periphery' of marked exceptions such as idioms, irregular verbs and the like, and involves a mixture of systems resulting from the diversity of languages ('dialects' or 'idiolects') that coexist in any real human community," which "lie far beyond our current understanding".

In short, core grammars are not designed to accommodate idiosyncratic periphery data. In fact, as noted by Culicover and Jackendoff (2005:25), "[the] 'periphery' tends to become a tempting dumping ground for any irregularity one's theory cannot at the moment explain." Culicover & Jackendoff argue that the core – periphery distinction is both unreal and counterproductive as a research heuristic, and they argue that it should be replaced by a model of grammar designed to cover both core and periphery phenomena. This is also the basic approach of the present paper. Although we will not follow Culicover and Jackendoff in their particular assumption as to how grammar should be construed and organized, we still assume that there is no principled difference between so-called periphery data and core data, and that both kinds of data are equally important, and therefore we assume as a research heuristic that a model of grammar should accommodate both kinds of data equally well. Actually, the so-called periphery data can often be more interesting than core data, since periphery data may show us more clearly the borderline between what is possible and what is impossible in language, and in that capacity periphery data may be particularly revealing as to the nature of the structure and organization of the internalized grammar.

Poplack (2004) contends that the assumption that bilingual syntax can be explained by general principles of monolingual grammar has not been substantiated. However, we suspect that the reason for this is that the main theories of grammar have not taken code-switching and other contact data sufficiently into account as an empirical base for the development of the framework in the first place. In this paper, we will defend a model of grammar, still fledgling, which we hope can eventually be developed into a more robust model or framework that is able to handle many different types of core and periphery data, i.e. both monolingual and bilingual data. After all, in a parsimonious approach one should expect that both monolingual and bilingual production should be constrained by the same basic principles.

Given the limited space available in this article, our aim is quite modest for the time being. We will sketch a model of grammar of the type that we think is required, and we will test this model against a small set of periphery data of the single-word code-switching type. Single-word code-switching should be carefully

distinguished from the kind of 'mixing' involved in the use of established loanwords. Single-word code-switching is a variety of intra-sentential code-switching involving spontaneous lexical borrowing, i.e. a type of on-the-spot lexical borrowing (*nonce borrowing*) where an individual speaker spontaneously picks a lexical item from the mental lexicon of language B that he masters, where this item is not established as a lexical item of his language A that is used for the occasion. A loanword, on the other hand, is a lexical item originally taken from language B that has eventually been established as a member of the mental lexicon of language A. Notice that there is a discussion whether or not single-word code-switching exists at all, but here we will assume that the notion of single-word code-switching is well founded, see Myers-Scotton (1993, 2002), Åfarli (2013) for discussion and motivation, see also Poplack (2004).

In single-word code-switching, the syntactic structures of one of the languages involved are used, and this language is called the host or matrix language. Thus, the host/matrix language provides the syntactic structures or frames that are used, whereas some of the lexical items are inserted into those frames from a guest or embedded language (see Myers-Scotton 1993, 2002). In our approach, we will seek to model the notions of host/matrix language and guest/embedded language in a generative syntax style grammar, paying particular attention to the way the lexicon – syntax interface should be organised. Empirically, the present paper is a case study of Mandarin Chinese – Norwegian bilingualism, where we study a bilingual child who sometimes uses Norwegian words when speaking Mandarin Chinese.

The paper is organized as follows. Section 2 outlines the model of grammar that we want to defend, and some general evidence for that model is discussed. Section 3 presents and discusses the Mandarin Chinese – Norwegian bilingual data and shows how the model can accommodate these data. Section 4 concludes the paper.

2. Outline of a model of grammar

The grammar model that we are going to defend belongs to a family of models of grammar that are often called neo-constructional models (see e.g. van Hout 1996; Borer 2005; Åfarli 2007; Ramchand 2008; Lohndal 2012; Nygård 2013; Marantz 2013). Also, the Distributed Morphology framework (Halle & Marantz 1993; Harley & Noyer 1999) may be taken to belong to this family. Notice that the neo-constructional models are generative models, and should be distinguished from (non-generative) models in construction grammar (e.g. Goldberg 1995; Croft 2001). The property of neo-constructional models that is most important for our purposes is separationism, i.e. the assumption that syntactic structure is generated

independently of the words that fill the structure, to put it simply. Thus, the model of grammar that we are going to defend is a separationist model. Embracing separationism, we assume that the grammar generates syntactico-semantic frames and that lexical insertion takes place subsequently. In other words, we assume late lexical insertion.

As for our assumptions about phrase structure, we take a traditional stand, i.e. we assume a basic phrase structural template with head, complement, and specifier, and with possible adjunctions allowed to these basic structures. To be slightly more specific, we assume that the basic operation is a merger of head and complement, and that the existence of a designated specifier position is motivated by the displacement property (of phrases). Thus, every time a head – complement relation is established, a (possible) specifier position is generated to facilitate displacement. We assume that adjunction is reserved for the base generation of phrasal non-arguments (adverbials). Specifically, we assume that movement does not involve adjunction, but substitution into head and specifier positions.

Concerning the actual clausal projections that we assume, we again take a fairly traditional stand. We assume a C-projection and a T-projection headed by abstract operators (see e.g. Rizzi 1997; Kitahara 1997: 8–9; Chomsky 2002: 123), but given our separationist stand, we assume designated open slots (marked with the symbol "¤") for insertion (i.e. substitution) of lexical material, see Nygård (2013: 154 ff.). Instead of vP, we assume a predication phrase (PrP) in all clauses (Bowers 1993; Åfarli & Eide 2000), headed by a predication operator in the Pr head, again with a designated open slot for substitution of lexical material. A proposition cannot be formed unless the predication operator and its Pr-projection are present. As for the verb phrase, we do *not* assume an articulated decomposition structure with various event-related heads (as do e.g. Borer 2005 and Ramchand 2008). We just assume a VP with an abstract V head. Thus, the basic phrase structure of the syntactico-semantic frame of a full clause is as shown in (1).

(1) $[_{CP}$ ¤ $[_{C'}$ $[_{C}$ ¤$]$ $[_{TP}$ ¤ $[_{T'}$ $[_{T}$ ¤$]$ $[_{PrP}$ ¤ $[_{Pr'}$ $[_{Pr}$ ¤$]$ $[_{VP}$ $[_{V}$ ¤$]$ ¤ $]$ $]]$ $]]$ $]]$

Given late lexical insertion, we assume that the V head of the verb phrase is an abstract prototypical 'verb' with an open position where a lexical item (root or stem) is inserted. This is similar to the analysis proposed in e.g. Marantz (1997) or Pylkkänen (2008), where a given root is turned into a verb by being inserted in the complement position of a verbal functional head "v" and then raised to that v. However, we want to adopt a version of this type of analysis where a lexical item is directly inserted into a designated open slot in the V head. As for the abstract prototypical verbs heading the VP, we assume that they are very few in number universally, and that each heads a VP-frame corresponding to a basic

argument structure type. For instance, Åfarli (2007) argues that there are exactly five different basic VP-frames in Norwegian, each corresponding to a prototypical verbal head. The five frames found in Norwegian will be described in some detail below, but first we will give some motivation for this particular approach to the analysis of the VP.

Given our assumptions, a root becomes a verb by being inserted in V. That means that we assume an impoverished lexicon (which is common in neo-constructional approaches) in which lexical items do not contain information about parts of speech or argument structure. Rather, the lexical root becomes a verb with a certain argument structure by being inserted in the V head of one of the five VP-frames. Thus, the syntax that we assume is definitely *not* lexically driven. Interestingly, this stand is also becoming more common in generative syntax outside the initial neo-constructional camp, see e.g. Platzack (2012), Lohndal (2012), and Marantz (2013).

One important question is how the syntactico-semantic frame, i.e. the pre-lexical insertion structure, is generated in the first place. We do not have the space to go into this here, but we tentatively assume that it is formed by composing elements from a very small pool of abstract functional elements (operators) taken from a special "functional lexicon", which also includes the abstract proto-verbs. Further investigation belongs to future work, but see Nygård (2013) for a study of the generation and use of syntactic frames in the analysis of discourse ellipsis phenomena.

Is there any evidence that this is the way the lexicon – syntax interface is organized? Late lexical insertion and the claim that syntax is not lexically driven, but generated independently of the words that fill it, is defended by all proponents of neo-constructional analyses to some degree, see e.g. Borer (2005) for a particularly vigorous defense of these claims. Consider one piece of evidence, having to do with the lexical flexibility predicted by this model, taken from Åfarli (2007). (2a) shows the title of an article by the philosopher David Kaplan (in *The Journal of Philosophy*, Vol. 76, pp. 716–729). (2a) may be paraphrased as (2b, c). The general meaning here is that Kaplan uses an analysis developed by Russell to deal with a problem formulated by Frege and Church.

(2) a. *How to Russell a Frege-Church*
b. *David Kaplan* Russells *a Frege-Church*
c. *David Kaplan* Russelled *a Frege-Church*

As can be seen from this example, the proper name *Russell* is used as a transitive verb effortlessly. How is that possible? According to standard lexically driven analyses, the transitive verb *Russell* should be listed in the lexicon as a verb with

an argument structure specifying an external role and an internal role. Since all proper names (and all nouns, for that matter) can in principle be used as verbs in similar ways (see e.g. Clark & Clark 1979), the standard assumption that syntax is lexically driven amounts to a kind of *reductio ad absurdum* argument, since it would clearly be uneconomical to assume that all nouns are separately listed as verbs in the lexicon.

This example supports the claim that the verbal properties are not projected from the lexicon, but rather that they are imposed by the syntax, i.e. acquired in the syntax. Thus, the sort of lexical flexibility illustrated by (2) is an effect of the syntactic bootstrapping (cf. Gleitman 1990) implied by the model that we assume, whereby any item inserted in a V position becomes a verb. Accordingly, any lexical item (root, stem), e.g. the proper name *Russell*, can in principle become a verb, thus accounting in a straightforward manner for the lexical creativity in examples like (2). The same mechanism readily accounts for the possibility of single-word code-switching, as we shall see later.

Another, but related piece of evidence for a separationist model is what we will call argument structure flexibility. Here we will draw on the basic VP-frames mentioned previously. VP-frames are basic ingredients in all languages. They can be thought of as a particular kind of recurring syntactic patterns, and our claim is that all verbal clauses in a language can be reduced to one of these basic recurring patterns. Here, we will describe the five VP-frames found in Norwegian one by one (cf. Åfarli 2007), briefly commenting on each as we proceed. In each case, we show just the structure from PrP and down.

The first pattern is given by the *intransitive* VP-frame that is shown in (3a). Recall that the symbol "¤" marks the designated open slots that are subsequently to be instantiated by lexical material. In (3b), the frame is instantiated as a clause containing an intransitive verb, or rather, the relevant lexical item is an intransitive verb by virtue of being inserted in the verbal slot in an intransitive frame. In this and in the subsequent examples, the verb raises obligatorily to Pr and is substituted into the designated open slot associated with the predication operator heading the Pr-projection.

(3) a. …[$_{PrP}$ ¤ [$_{Pr'}$ [$_{Pr}$ ¤] [$_{VP}$ [$_{V}$ ¤]]]]
b. …[$_{PrP}$ *babyen* [$_{Pr'}$ [$_{Pr}$ *sparkar*] [$_{VP}$ [$_{V}$ ~~*sparkar*~~]]]]
baby-the kicks
'The baby kicks'

The next frame is called a *transitive* frame, since it accommodates clauses with transitive verbs, as can be seen in the example in (4). Notice, like before, that the lexical item in question becomes a transitive verb by being inserted in the transitive VP-frame.

(4) a. …[$_{PrP}$ ¤ [$_{Pr'}$ [$_{Pr}$ ¤] [$_{VP}$ [$_{V}$ ¤] ¤]]]
b. …[$_{PrP}$ *Ola* [$_{Pr'}$ [$_{Pr}$ *sparkar*] [$_{VP}$ [$_{V}$ ~~*sparkar*~~] *ballen*]]]
Ola kicks ball-the
'Ola kicks the ball.'

The third VP-frame is *ditransitive*, shown in (5).

(5) a. …[$_{PrP}$ ¤ [$_{Pr'}$ [$_{Pr}$ ¤] [$_{VP}$ ¤ [$_{V'}$ [$_{V}$ ¤] ¤]]]]
b. …[$_{PrP}$ *Kari* [$_{Pr'}$ [$_{Pr}$ *sparkar*] [$_{VP}$ *Ola* [$_{V'}$ [$_{V}$ ~~*sparkar*~~] *ballen*]]]]
Kari kicks Ola ball-the
'Kari kicks the ball to Ola.'

The fourth frame is the *resultative* frame, which can be seen as a minor mutation of the transitive frame, i.e. with the lower DP in the transitive frame substituted for by a resultative-propositional PrP (i.e. a small clause). The resultative VP-frame is shown in (6).

(6) a. …[$_{PrP}$ ¤ [$_{Pr'}$ [$_{Pr}$ ¤] [$_{VP}$ [$_{V}$ ¤] [$_{PrP}$ ¤ [$_{Pr'}$ [$_{Pr}$ ¤] [$_{XP}$ ¤]]]]]]
b. …[$_{PrP}$ *Kari* [$_{Pr'}$ [$_{Pr}$ *sparkar*] [$_{VP}$ [$_{V}$ ~~*sparkar*~~] [$_{PrP}$ *ballen*
Kari kicks ball-the
[$_{Pr'}$ [$_{Pr}$ ¤] [$_{XP}$ *fillete*]]]]]
tattered
'Kari kicks the ball to pieces.'

The last type is the *ditransitive-resultative* VP-frame, which again can be seen as a mutation of an earlier frame, this time of the ditransitive frame with the lower DP substituted for by a resultative-propositional PrP. The ditransitive-resultative frame is shown in (7). A peculiarity with this frame is that it has obligatory inalienable possession interpretation. Thus, in the example given, the stomach must be interpreted as Ola's stomach.

(7) a. …[$_{PrP}$ ¤ [$_{Pr'}$ [$_{Pr}$ ¤] [$_{VP}$ ¤ [$_{V'}$ [$_{V}$ ¤] [$_{PrP}$ ¤ [$_{Pr'}$ [$_{Pr}$ ¤] [$_{XP}$ ¤]]]]]]]
b. …[$_{PrP}$ *Kari* [$_{Pr'}$[$_{Pr}$ *sparkar*] [$_{VP}$ *Ola* [$_{V'}$[$_{V}$ ~~*sparkar*~~]
Kari kicks Ola
[$_{PrP}$ *ballen* [$_{Pr'}$[$_{Pr}$ ¤] [$_{XP}$ *i magen*]]]]]]]
ball-the in stomach-the
'Kari kicks the ball into Ola's stomach'

VP-frames (and their higher functional structure) are generated prior to lexical insertion, with subsequent late lexical insertion. Thus, the lexical verb that is inserted acquires its argument structure by being inserted into a given VP-frame. Crucially, the same verb is used in all five frames given above. This implies that the verbal item is flexible in terms of which frame it can occur in. This is typical, so it is rather the rule than the exception that a given lexical item that becomes

a verb may occur in more than one frame, and it is not uncommon that a given item may occur in all five frames, as in the examples just given. Thus, argument structure cannot be derived from purported lexical-semantic properties of each particular verb, in which case we would expect each particular verb to head only one VP-frame (the one that corresponds to its lexical-semantic content). Rather, argument structure must be derived from the syntactic frames. The argument structure flexibility exemplified by the Norwegian examples discussed above lends further support to our separationist model (cf. arguments to the same effect in e.g. Goldberg 1995, and Borer 2005).

A question that we cannot go pursue for reasons of space, has to do with possible restrictions on argument structure flexibility, i.e. the seeming resistance that some verbs show against being used in a particular VP-frame. Borer (2005) goes far in suggesting that this problem is something that should not be given a structural explanation (because it has to do with convention and habits of speech), but see Åfarli (2007) for an analysis of this problem based on a notion of harmony/disharmony between the semantics of the frame and the inherent conceptual content of the verb.

The evidence adduced above from lexical creativity and argument structure flexibility has to do with what we could call core phenomena of language (even though such phenomena haven't received much attention in mainstream generative work until quite recently). They have to do with the normal workings of one particular language, without any language mixing or switching being involved. We will now show that the same separationist model that is motivated by these "core" phenomena seems to be just what is needed to handle a common type of language contact phenomenon involving "periphery" code-switching type data.

3. Mandarin Chinese – Norwegian single-word code-switching

We will now demonstrate the usefulness of our model for the analysis of spontaneous single-word code-switching involved in language mixing. To repeat, in single-word code-switching an individual speaker spontaneously picks a lexical item from the mental lexicon of language B (the guest or embedded language) and inserts it into the syntactic frame of language A (the host or matrix language).

We will use data from a six-year-old Chinese child who is growing up in Norway. The Mandarin Chinese – Norwegian bilingual child under study, a boy named Lele, was born into a Chinese immigrant family in Trondheim, Norway. He was first-born, and until he was at the age 5;11, he was the only child in the family. Both his parents are native speakers of Mandarin Chinese, and immigrated to Norway in their late twenty years of age. At the time of recording, they had

good knowledge of English and Norwegian, but the language of communication between the parents was Mandarin. Lele is a so-called second generation immigrant, and like many other second generation immigrants, his input conditions are situation-bound: Mandarin Chinese at home and Norwegian in the wider community. His Mandarin input is mainly from his parents, who were committed to speaking Mandarin with their children. The parents played no role in Lele's Norwegian development, though. His Norwegian input is mainly from the daycare center and subsequently from the school. Lele started to attend the daycare center on a full-day basis (8 hours a day, five days a week) when he was 15 months old. Before that age, his Mandarin input was dominant. After he started attending daycare, his Norwegian input increased and soon became the dominant one.

Lele spoke Mandarin Chinese to his parents and other Chinese-speaking people, and he spoke Norwegian to Norwegian-speaking people. As most Chinese-speaking people around him spoke some Norwegian as well, he might occasionally embed some Norwegian words into his Chinese speech. However, no Norwegian-speaking people around him spoke Mandarin Chinese, so he did not embed Chinese into his Norwegian speech. Lele was perceived as a competent speaker of Norwegian, with not much difference from his monolingual Norwegian peers, by the native Norwegian speakers (e.g. his caretakers in the daycare center, his neighbors, and the parents of his Norwegian friends). He was also perceived as a competent speaker of Mandarin by the Mandarin-speaking people who knew him. His Norwegian production data and Mandarin production data show that he had essentially the same type of abstract linguistic system as his peers (cf. Jin & Eide 2011; Jin 2010). Nevertheless, we believe his Norwegian is slightly dominant, based on the following observations: First, he had longer average exposure to Norwegian than Chinese after attending the Norwegian daycare, and secondly he sometimes had difficulty retrieving Chinese words, but never had difficulty retrieving Norwegian words.

Production data from Lele were collected at Lele's home both in a Mandarin language context and in a Norwegian language context from age 5; 11 until age 6; 11. The interactions in Mandarin consisted of conversations between Lele and his parents. The interactions in Norwegian consisted of conversations between Lele and his Norwegian friends, who were his peers. They were recorded with audio equipment in spontaneous interaction. The Chinese corpus, from which the data in the present study are drawn, contains 7 audio recordings which is 4 hours and 51 minutes long, and has a total of 2147 complete and clear sentences. The Chinese production data indicates that he was indeed a competent speaker of Chinese. The use of Norwegian words in matrix Chinese was not at all frequent.

It is known from other studies of single-word code-switching that words borrowed into a matrix language are practically without exception uninflected lexical

items, with the inflection belonging to the matrix language (Myers-Scotton 1993, 2002; Kamwangamalu 1997). Åfarli (2013) gives the pattern in (8) for the inflection of verbs in early varieties of American Norwegian, the immigrant language spoken by Norwegian immigrants to the USA and their descendants from c. 1850 and onwards. This is a variety of Norwegian in which borrowing from English is relatively common (see Haugen 1953; Hjelde 1992).

(8) a. $\text{stem}_{\text{NORW}}$-$\text{tense}_{\text{NORW}}$
b. stem_{ENG}-$\text{tense}_{\text{NORW}}$
c. *$\text{stem}_{\text{NORW}}$-$\text{tense}_{\text{ENG}}$
d. *stem_{ENG}-$\text{tense}_{\text{ENG}}$

As can be seen from this pattern, both Norwegian and English stems can occur with Norwegian tense inflection, whereas neither Norwegian nor English stems can occur with English tense inflection. In short, English inflections do not occur even though English stems may occur quite frequently. The pattern in (8b), with an English stem borrowed into American Norwegian, is illustrated in (9) (from Haugen 1953). Here the stem is clearly English (*settle*) whereas the inflectional tense affix (*-a*) is clearly Norwegian.

(9) *Då dæ kåm ti detti landi, då setla dæ på…*
when they came to this country-the, then settled they on…
'When they came to this country, they settled on…'

In terms of the model sketched in the previous section, the pattern in (8) indicates that inflection is a property of the matrix language frame. Therefore, the simplest analysis is one where the matrix language determines the syntactico-semantic frame, including the functional properties of the frame (i.e. the abstract operators heading the functional projections, including the T-projection), whereas appropriate lexical items (roots/stems) are taken into that frame and are inserted into the empty slots in the relevant V positions. Thus, the inflectional affixes are the morpho-syntactic spell-outs of the matrix language operators and must belong to the matrix language, whereas the lexical roots/stems can come from any lexicon, in particular from a lexicon belonging to another language, as is the case in single-word code-switching. See Åfarli (2013) for a more detailed analysis of verbs in American Norwegian along these lines. Also, see Myers-Scotton (1993, 2002) for numerous examples from various languages that inflections (almost) invariably belong to the matrix language in cases where the root/stem comes from a guest language.

We will assume a similar type of analysis for the Mandarin Chinese – Norwegian bilingual data produced by Lele, but whereas there is a clear asymmetry between English and Norwegian in the American Norwegian case, in that Norwegian is clearly the "stronger" language since Norwegian is the L1 of the

speakers and English is their often quite fragile L2, Norwegian and Mandarin Chinese in the case at hand are basically equal in strength, both being reasonably considered as the L1s of the child in question. Thus, in this particular bilingual case, there are two reasons why it should be likely that *complete inflected forms* could be borrowed from Norwegian, if it is at all possible to borrow complete inflected forms in principle. First, Mandarin Chinese, being an isolating language, has no (or extremely few) inflected words at all, so Norwegian inflected forms that are borrowed would have no competition from Chinese inflections. Second, this child is bilingual and masters Norwegian inflection completely, so there is no lack of knowledge of the inflectional properties of the guest language. In contrast, in the American Norwegian case, one could imagine as a possible explanation for the lack of English inflections on the verbs that there is a lack of knowledge of the English inflectional paradigms rather than any principled structural inaccessibility to the inflectional properties of the guest language (as we would wish to argue).

Importantly, in spite of the favorable conditions for the borrowing of inflections in the Mandarin Chinese – Norwegian case, still it turns out that Norwegian words borrowed into Lele's Chinese are *uninflected* forms. The data in (10), with noun phrases as direct objects, illustrate this.

(10) a. *Da na ge ball*
hit that-GE ball
'Hit that ball.'

b. **Da na ge ball-en*
hit that-GE ball-DEF-SG
'Hit that ball.'

c. **Da na xie ball-ene*
hit those ball-DEF-PL
'Hit that ball.'

In cases where a Norwegian noun (here: *ball* 'ball') is borrowed into the Lele's Chinese, it occurs without any inflection, as in (10a), even though Norwegian agreement inflection would have been natural in the context, as in (10b, c). (The suffix *-en* is the Norwegian sg, def suffix for this masculine noun, and the suffix *-ene* is the Norwegian pl, def suffix.) However, data like the ones in (10b, c) are not attested in our corpus. This fact clearly indicates that the inflections of the guest/embedded language are not available in the matrix language, in principle.

A corresponding pattern is found with the single-word code-switching of verbs, see the data in (11) and (12):

(11) *Xiao meimei bæsje le*
little sister poop LE
'My younger sister has pooped.'

(12) *Wo-de baba jintian bu jobbe*
I-POSS daddy today not work
'My daddy does not work today.'

Notice in particular that examples like (13) and (14), where Norwegian tense inflection is borrowed along with the Norwegian stem, are *not* attested. (Notice that *-et* is the past tense suffix and *-r* is the present tense suffix in Norwegian for the two verbs involved, respectively.)

(13) **Xiao meimei bæsje-et le*
little sister poop-PRET LE
'My younger sister has pooped.'

(14) **Wo-de baba jintian bu jobbe-r*
I-POSS daddy today not work-PRES
'My daddy does not work today.'

Actually, in (11) and (12), it is the infinitival forms of the Norwegian verbs that are used, not the bare roots/stems. This is also observed in the borrowing of Norwegian verbs in Turkish light verb constructions in Turkish – Norwegian code-switching, see Türker (2000: 89 ff.). Also, a similar observation is made in Kamwangamalu (1997: 47) concerning French verbs that are code-switched into African languages. This clearly indicates that it is the infinitival form that is listed in the lexicon, not the bare stem, and probably this is the case universally. Notice, by the way, that English infinitival stems borrowed into American Norwegian, actually get the Norwegian infinitival affix (normally *-e*) added, which indicates that the infinitival property is both a functional property of the frame (a possible value of T), and a property of the encyclopedic entry that is listed in the lexicon. As far as we can see, this double "existence" of the infinitival property should cause no serious problems for the analysis, although it raises a question about what counts as a lexical item (root, stem, listeme), which we will not go into here. No doubt, the status of the infinitival affix clearly deserves further investigation both generally and in the context of the model that we propose. Still, what is significant for present purposes is that present and past tense inflectional affixes from the guest/embedded language are never used, even though the lexical item is taken from the guest/embedded language (Norwegian in the examples above).

Now, to analyse the Mandarin Chinese – Norwegian data in our model, recall first that Mandarin Chinese is the host/matrix language, and Norwegian is the guest/embedded language in the Chinese production of Lele. Thus, the lexical items may belong to the host/matrix language (Mandarin Chinese) or the guest/embedded language (Norwegian), but the functional morphology invariably belongs to the matrix/host language (Mandarin Chinese). In e.g. (11) above, the perfective aspectual particle *le* belongs to the functional morphology, whereas the

borrowed lexeme is the main verb *bæsje* 'poop' which, we assume following the approach described above, is inserted into the open slot in the V position of the Chinese matrix language frame.

We will now sketch a derivation, step by step, to illustrate our analysis. We will use (11) as our example, repeated below. As for the syntactic analysis of Chinese aspectual particles, see e.g. Xu (1997) and Huang, Li and Li (2009).

(11) *Xiao meimei bæsje le*
little sister poop LE
'My younger sister has pooped.'

The first step is the formation of an intransitive frame in Chinese, see (15) (the structure is shown just from AspP and down). Notice that the inflectional aspectual property is already spelled out at this step (signaled by the aspectual marker *le*), since it is part of the matrix frame.

(15) $[_{\text{AspP}}\ ¤\ [_{\text{Asp'}}\ \ [_{\text{Asp}\langle\text{perf}\rangle}\ ¤+le\ ]\ [_{\text{PrP}}\ ¤\ [_{\text{Pr'}}\ \ [_{\text{Pr}}\ ¤\ ]\ \ [_{\text{VP}}\ [_{\text{V}}\ ¤\ ]]\ ]]\ ]]$

The next step consists of the insertion of lexical items or content words into the intransitive frame (in **bold**), see (16). This is the late lexical insertion stage.

(16) $[_{\text{AspP}}\ ¤\ [_{\text{Asp'}}\ [_{\text{Asp}\langle\text{perf}\rangle}\ ¤+le\ ]\ [_{\text{PrP}}$ **xiao meimei** $[_{\text{Pr'}}\ [_{\text{Pr}}\ ¤\ ]\ [_{\text{VP}}\ [_{\text{V}}$ **bæsje** $]]\ ]]\ ]]$

Notice that whereas the lexical item that becomes the verb is picked from the lexicon and inserted in the V-position, the DP *xiao meimei* 'little sister' is formed as a DP in the syntax by items picked from the lexicon. We assume that the formation of the DP takes place in a separate work-space, and that the DP is inserted into its specifier position by a generalized transformation.

The last step in the syntactic derivation consists of syntactic processes (like movement), as indicated in (17), where the verb has moved from V to Pr to its empty slot under Asp, and the DP subject has moved to the specifier position of AspP.

(17) $[_{\text{AspP}}$ **xiao meimei** $_{\text{j}}\ [_{\text{Asp'}}\ [_{\text{Asp}\langle\text{perf}\rangle}$ **bæsje**$_{\text{i}}$ $+le\ ]\ [_{\text{PrP}}\ \text{t}_{\text{j}}\ [_{\text{Pr'}}\ [_{\text{Pr}}\ \text{t}_{\text{i}}\ ]\ [_{\text{VP}}\ [_{\text{V}}\ \text{t}_{\text{i}}\ ]]\ ]]\ ]]$

Notice that potential additional functional structure is not taken into account here, as it is not essential for our point.

Next, consider adjectives that are used as verbs. It is well known that some adjectives behave as intransitive verbs in Mandarin Chinese. Thus, aspectual markers (e.g. *le*) can be attached both to verbs and adjectives in Chinese, and an adjective can function as the predicate of the clause without the support of the copula (*shi* 'be'). Given our analysis, it is predicted that Norwegian adjectives can be used as verbs, like in Chinese, which means that Norwegian adjectives may be inserted into Mandarin Chinese VP-frames to get the verbal properties of the matrix language. This is exactly what we find, as illustrated by the example in (18).

(18) *Women liang ge forelsket le*
I-PL two GE in-love LE
'We two are in love.'

Here the Norwegian participal adjective form *forelsket* 'be-in-love' is inserted in the intransitive Mandarin Chinese VP-frame, and what we otherwise know as an adjective becomes a verb. Notice that *ge* is a classifier that is part of the subject noun phrase.

Also, consider the example in (19) where the Norwegian noun *ferie* 'vacation' is code-switched into Chinese and is used as a verb.

(19) *Ta ferie le*
he vacation LE
'He is on vacation.'

This is possibly modeled on the corresponding Chinese verb *xiu jia*, meaning 'take off' or 'taking off'. Notice that *xiu jia* can be used both as a verb and a noun in Chinese, without changing its form. This is a common process in Chinese, but not in Norwegian. We speculate that the code-switching here occurred in the following process: When Lele had the abstract concept of 'xiu jia', and searched his mental lexicon for it, the corresponding Norwegian word *ferie*, which is highly frequent, first popped up due to its meaning association with the nominal *xiu jia*. Then *ferie* was inserted into the relevant verbal slot of the syntactic frame. Thus, a code-switched Norwegian noun becomes a verb by being inserted in the V slot in the intransitive Mandarin Chinese VP-frame (cf. the *Russell* example shown in (1) where a proper noun was similarly turned into a transitive verb).

The intermediate step of the derivation of (19) is sketched in (20).

(20) $[_{\text{AspP}}$ ¤ $[_{\text{Asp'}}$ $[_{\text{Asp⟨perf⟩}}$ ¤+*le*] $[_{\text{PrP}}$ **ta** $[_{\text{Pr'}}$ $[_{\text{Pr}}$ ¤] $[_{\text{VP}}$ $[_{\text{V}}$ **ferie**]]]]]]

To sum up, the data discussed above suggest that the child under consideration used a full-fledged Mandarin Chinese syntactico-semantic matrix frame in his Chinese speech, although Norwegian words were code-switched occasionally. Also, the data suggest that inflectional properties belong to the matrix frame, not to the spontaneously borrowed word.

4. Conclusion

The patterns displayed in the Mandarin Chinese – Norwegian bilingual case are fully consistent with the patterns displayed in the American Norwegian case, and the patterns are accounted for by the same type of separationist model, namely the model outlined in Section 2. Moreover, the very same model also accounts

for cases of lexical creativity and argument structure flexibility in monolingual contexts. None of these facts are easily explained in the lexically driven grammar models of mainstream syntactic theory (outside of the constructional or neo-constructional family of grammar models), as indicated in Section 2. Thus, it seems to us to be a promising hypothesis that a model of grammar that accommodates both "core" type data and "periphery" type data should be a separationist model with late lexical insertion along the lines that we have suggested.

As already suggested, the model that we have argued for is still fledgling, and many problems remain to be solved, e.g. how the syntactico-semantic frames are generated in the first place (Section 2), or how to handle the infinitival suffix (Section 3), but at least we are confident that our model is a promising starting point for the development of a model of grammar that is equally well suited to accommodate both "core" data and "periphery" data, as well as both monolingual and bilingual data.

Acknowledgements

An earlier version of this paper was presented at the International Symposium on Bilingualism 8, 15–18 June 2011, University of Oslo (see Åfarli & Jin 2011). Many thanks to the audience there for useful comments. Also, many thanks to three anonymous reviewers for the present volume for very useful comments and suggestions.

References

Åfarli, Tor A. 2007. Do verbs have argument structure? In *Argument Structure* [Linguistik Aktuell/Linguistics Today 108], Eric Reuland, Tanmoy Bhattacharya & Giorgos Spathas (eds), 1–16. Amsterdam: John Benjamins.

Åfarli, Tor A. 2013. Hybrid verb forms in American Norwegian and the analysis of the syntactic relation between the verb and its tense. Ms, NTNU, Trondheim.

Åfarli, Tor A. & Eide, Kristin M. 2000. Subject requirement and predication. *Nordic Journal of Linguistics* 23: 27–48.

Åfarli, Tor A. & Jin, Fufen. 2011. The Matrix Language Frame Model in a principles and parameters adaption: A case study of Mandarin – Norwegian code-switching. Talk at ISB8, University of Oslo, June 15–18.

Borer, Hagit. 2005. *Structuring Sense,* Vols. I & II. Oxford: OUP.

Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24: 591–656.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge MA: The MIT Press.

Chomsky, Noam. 1981. *Lectures on Government and Binding. The Pisa Lectures*. Dordrecht: Foris.

Chomsky, Noam. 1991. Linguistics and cognitive science: Problems and mysteries. In *The Chomskyan Turn,* Asa Kasher (ed.), 26–53. Cambridge: Blackwell.
Chomsky, Noam. 2002. *On Nature and Language.* Cambridge: CUP.
Clark, Eve & Clark, Herbert. 1979. When nouns surface as verbs. *Language* 55: 767–811.
Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: OUP.
Culicover, Peter & Jackendoff, Ray. 2005. *Simpler Syntax*. Oxford: OUP.
Gleitman, Lila. 1990. The structural sources of verb meanings. *Language Acquisition* 1: 3–55.
Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure.* Chicago IL: The University of Chicago Press.
Halle, Morris & Marantz, Alec. 1993. Distributed morphology and the pieces of inflection. In *The View from Building 20. Essays in Linguistics in honor of Sylvain Bromberger,* Kenneth Hale & Samuel Jay Keyser (eds), 111–179. Cambridge MA: The MIT Press.
Harley Hidi & Noyer, Rolf. 1999. Distributed morphology. *Glot International* 4: 3–9.
Haugen, Einar. 1953. *The Norwegian Language in America: A Study in Bilingual Behavior,* Vol. I & II. Philadelphia PA: University of Pennsylvania Press.
Hjelde, Arnstein. 1992. *Trøndsk talemål i Amerika*. Trondheim: Tapir forlag.
van Hout, Angeliek. 1996. Event Semantics and Verb Frame Alternations [TILDIL Dissertation Series]. Ph.D. dissertation, Tilburg University.
Huang, C.-T. James, Li, Y.-H. Audrey & Li, Yafei. 2009. *The Syntax of Chinese.* Cambridge: CUP.
Jin, Fufen. 2010. Linguistic competence in the minority language. Paper presented at GSFL, NTNU, Trondheim, September 22–26.
Jin, Fufen & Eide, Kristin M. 2011. Contradictory parameter settings in one mind: A case study of a bilingual's acquisition of finiteness and V2 in Norwegian. Talk at ISB8, University of Oslo, June 15–18.
Kamwangamalu, Nkonko M. 1997. Language contact, code-switching, and I-languages: Evidence from Africa. *South African Journal of Linguistics* 15: 45–51.
Kitahara, Hisatsugu. 1997. *Elementary Operations and Optimal Derivations.* Cambridge MA: The MIT Press.
Lohndal, Terje. 2012. Without Specifiers: Phrase Structure and Events. Ph.D. dissertation, University of Maryland.
Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In *Pennsylvania Working Papers in Linguistics* 4(2), Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark & Alexander Williams (eds), 201–225. Philadelphia PA: University of Pennsylvania, Penn Linguistics Club.
Marantz, Alec. 2013. Verbal argument structure: Events and participants. *Lingua* 130: 152–168.
Myers-Scotton, Carol. 1993. *Duelling Languages: Grammatical Structure in Codeswitching*. Oxford: OUP.
Myers-Scotton, Carol. 2002. *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: OUP.
Nygård. Mari. 2013. Situational Ellipses in Spontaneous Spoken Norwegian: Clausal Architecture and Licensing Conditions. Ph.D. dissertation, NTNU, Trondheim.
Platzack, Christer. 2012. Towards a minimal argument structure. Ms, University of Lund.
Poplack, Shana. 2004. Code-switching. In *Soziolinguistik. An International Handbook of the Science of Language,* 2nd edn, Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier & Peter Trudgill (eds). Berlin: Walter de Gruyter.

Pylkkänen, Liina. 2008. *Introducing Arguments.* Cambridge MA: The MIT Press.
Ramchand, Gillian C. 2008. *Verb Meaning and the Lexicon. A First-Phase Syntax.* Cambridge: CUP.
Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar. A Handbook in Generative Syntax*, Liliane Haegeman (ed.), 281–337. Dordrecht: Kluwer.
Türker, Emel. 2000. Turkish – Norwegian codeswitching: Evidence from Intermediate and Second Generation Turkish Immigrants in Norway. Ph.D. dissertation, University of Oslo.
Xu, Ding, 1997. *Functional Categories in Mandarin Chinese*. The Hague: HAG.