CODE-SWITCHING IN EMERGENT GRAMMARS: VERB MARKING IN BILINGUAL CHILDREN'S SPEECH[1]

Virve-Anneli Vihman
University of Tartu

Abstract. This paper examines the code-switching of verbs in the speech of two children bilingual in Estonian and English (aged 3 to 7). Verbs typically have lower rates of code-switching than nouns, due to their central role in argument structure, lower semantic specificity, and greater morphological complexity. The data examined here show various types of morphological mixing, and include examples which violate the prediction from the literature that only finite verbs bear inflectional morphology from the other language, suggesting that children do not adhere to the same constraints as adults when code-switching.

Keywords: code-switching, acquisition, bilingual children, verb morphology

1. Introduction

The ways in which the cognitive, social and structural aspects of language interact are made particularly evident when two or more languages are at play. For more than one language to be used in a single social situation, the languages must also coexist in the speakers' minds, and the discourse often reflects this. When code-switching

takes place within a clause, and more strikingly, across morphemes within a word, some of the (usually hidden) mechanisms underlying production are made visible. In first language acquisition research, this may be one route to evidence regarding children's knowledge and understanding of the structure of the languages they are acquiring.

In a prevalent view of code-switching, bilingual clauses are always structured according to one of the participating languages, meaning that the languages are clearly distinguishable and remain distinct in bilingual discourse. The language which imposes the basic structure of the clause is called the Matrix Language, and this approach has been best formalised and is best known as the Matrix Language Frame, or MLF, model (Myers-Scotton 2002, 2005; Myers-Scotton & Jake 2000). This has been further developed to incorporate evidence that morphemes are not all alike, as discussed below. One strength of this model is that it provides a principled framework for analysing code-switched utterances and for making predictions regarding code-switching. Myers-Scotton's approach focuses on code-switching within the clause, which is also the focus of this paper. She claims that „it is only in the bilingual clause that the grammars of both languages are in contact and [...] the basic hierarchical opposition [...] between the matrix language and the embedded language makes any sense" (2005: 329).

An alternative view, which may also be construed as complementing and refining the MLF view, is that the two languages interact. Once two languages are used together in bilingual conversation, they may influence each other and they are both likely to show effects of mutual interaction, even when one is dominant or one is identified as the Matrix Language. As linguistic knowledge is not a steady state, but always dynamic, a bilingual language production model will allow interaction and mutual influence even in the production process (De Bot 2004, Vihman 2016). In addition, it has been shown that bilingual conversation is not always best analysed through a matrix language approach (Backus 2014, Johanson 2002, Verschik

2007), and that it is not always possible to identify a Matrix Language (Auer & Muhamedova 2005). The Matrix Language Frame model has also been criticised from a different perspective: e.g. MacSwan (2000, 2005) argues that some of the concepts are vaguely defined and overly rely on code-switching data rather than being independently motivated. We leave this open, and assume that the MLF framework is clear enough to be tested using the data we present.

This paper examines a dataset from two bilingual Estonian and English-speaking children, and focuses on utterances with code-switched verbs in order to explore predictions which have been made regarding morphology in code-switching. In particular, the study aims to investigate the following questions: Can children's code-switching data shed light on how much they generalise abstract knowledge in online production? Do children follow the same constraints proposed in the literature on code-switching more generally? What can the code-switching data reveal about the emergent knowledge of the bilingual child?

2. Background

Before turning to the data, this section introduces background relevant to the study. To begin with, section 2.1 discusses why it is important to investigate bilingual children's productions in addition to adults' discourse; following this, code-switching with verbs is discussed; an overview is given of relevant details of the morphosyntax of the two languages in this study in 2.3.

2.1 BILINGUAL CHILDREN'S SPEECH

Much of the research on childhood bilingualism has focussed on comparing groups of bilingual children with their monolingual peers, and asking whether we find quantitative or qualitative differences, and what factors affect any observed differences. Overall, findings suggest that bilingual acquisition proceeds along similar

developmental paths as monolingual acquisition, and the differences we find tend to be quantitative rather than qualitative, reflecting inevitable differences in amount of exposure to each language in comparison with monolinguals (see Unsworth 2013, Yip 2013). Additionally, much attention has been paid to the question of whether and to what degree the languages interact in bilingual children, with agreement that some interaction may occur, but only under certain conditions (see Hulk & Müller 2000, Müller & Hulk 2001). Interaction has been shown to take place more broadly (see Argyri & Sorace 2007, Pérez-Leroux et al. 2011), but the general consensus is that children separate their languages fairly early (Paradis & Genesee 1996). Even in the literature on adults, disagreement prevails as to what degree of interaction exists between languages in bilingual discourse (Backus & Verschik 2012, De Bot 2004, Johanson 2002). Young children in the process of acquiring their linguistic systems are still discovering what sorts of regularities to expect in language. They make predictions, discover co-occurrences and learn patterns in the languages they are exposed to. We may ask what mechanisms they use in online production to select lexical items from the appropriate language as well as how they structure utterances according to their emerging grammars.

Much recent research has demonstrated how implicit, statistical learning underlies first language acquisition, and operates on various levels of language (Boyle et al. 2013, Chang et al. 2012, Rebuschat & Williams 2012, Perruchet & Pacton 2006). Statistical learning, the „unconscious process of inducing structure and regularity from exposure to repeated exemplars" (Boyle et al. 2013), has been applied to phonology, morphology and syntactic dependencies in first and second language acquisition. Typically, we think of first language learners as mostly operating via distributional, implicit learning, supplemented by a number of explicit learning processes. Generalisation over exemplars must take place, as learning of linear strings alone would mean that the productive output reflected only structures identical to what is heard in the input. However, the

attribution of pattern generalisation sometimes involves an inferential leap on the part of the analyst, since access to a child's input is limited, nor can we usually see the process underlying the output. This is related to the first research question: Does a particular utterance merely reflect knowledge of co-occurrences based directly on heard utterances, or does it derive from a more abstract level of knowledge, reflecting morphosyntactic generalisations?

Bilingual speech is a good locus for evidence of what knowledge underlies an utterance. With bilingual acquisition, the abstraction of grammatical and collocational regularities may be seen in innovative constructions on the part of the child. Constructions which the child has not heard in the ambient language may reveal something of the grammatical process underlying the production of an utterance. Code-switching is, of course, not limited to children, but the code-switching of young children, whose mental grammars are not yet fully formed, may provide insight into several questions of interest. First, does children's code-switching differ in quantity or quality from that of adults? The data reported on in this study come from the author's own diary notes, which neither provide information on frequency nor include a controlled study of the input speech. Nevertheless, we can assume for the purposes of this study that the intuitions are relatively accurate, regarding the types of code-switching which were present in the child-directed (the author's own and other familiar adults') speech and those which weren't. On the assumption, then, that the children in this study do indeed show evidence of combining their languages in different ways than adults do, we can ask these further research questions: Do the data reflect constraints proposed in the literature on code-switching; and what can these tell us about the emergent linguistic knowledge of the bilingual child?

2.2 VERBS IN CODE-SWITCHING

It has been noted in the code-switching literature that verbs show lower rates of code-switching than nouns (Bolonyai 2005, Cantone 2007:173, Myers-Scotton 2005: 330). This may be largely due to their central role in argument structure: because the verb forms both the semantic and syntactic core of the predicate, it can be seen to provide the core structure of a clause. Verbs also tend to carry more functional information and morphological complexity than other constituents, and to be less semantically specific, which may reduce the motivation to code-switch based on contextually specific lexical items (cf. Backus 2001, Isurin et al. 2009, Backus & Verschik 2012).

For examining the interplay between grammatical structure and lexical units, then, verbs are a useful category to explore. The languages included here both require the expression of certain information via morphological marking when a verb is utilised. Verbs are intimately connected not only to the argument structure of a clause, but also to the tense, aspect and mood of an utterance.

Additionally, clear predictions have been made in the Matrix Language Frame and 4M model of Myers-Scotton (Myers-Scotton 2002, 2005, Myers-Scotton & Jake 2000) regarding the morphosyntactic integration of code-switched verbs.

This model claims, first, that an asymmetry always exists between the participating languages within a clause. The Matrix Language is that which structures the bilingual clause, in that it supplies the morphosyntactic frame.[2] Crucially, the model predicts that only Matrix Language morphemes will indicate grammatical relations within mixed constituents.

The 4M model refines this to differentiate between four types of morphemes. First, content morphemes, lemmas with semantic

---

[2]   A possible composite Matrix Language Frame has been proposed (e.g. Bolonyai 2002, Myers-Scotton 2002), but has not been fleshed out, nor have the implications been fully explored.

content associated with assigning or receiving thematic roles, are set apart from system morphemes. System morphemes are further subdivided into three types, early system morphemes and two late system morpheme types. Both the (thematic role-assigning) content morphemes and (non-assigning) early system morphemes (e.g. determiners and derivational affixes) are conceptually activated, and accessible in the mental lexicon. Late system morphemes, on the other hand, express grammatical information, are structurally assigned and hence accessed later in the production process; they are claimed to be available at the level of the formulator (Levelt 1989) and can therefore participate in code-switching. They mark relations between constituents rather than within a single constituent. These are further subdivided into bridge system morphemes and late outsider morphemes. Person and number in a noun phrase can be accessed within the noun phrase, and are therefore marked with early system morphemes. Person and number marking on a verb, however, is marked with late system morphemes, as it relies on information outside the verb phrase. Case-marking on a noun also derives from information outside the constituent; therefore case markers are also classed as late outsider system morphemes.

According to the 4M model, then, only finite verbs are predicted to bear inflectional morphology from the other language: non-finite verb forms are predicted to be accessed holistically and early, and hence ought not to participate in code-switching. Myers-Scotton & Jake (2000) note that all examples of content words from the embedded language with matrix language inflections[3] support the notion that lexical words with regular morphology are constructed online. We return to this notion after discussing the data.

---

[3]   This is applied to outsider late system morphemes (Myers-Scotton 1997), but it is not always immediately clear how this plays out in languages other than the ones discussed. Bolonyai (2000), for instance, makes a case for Hungarian preverbs to be treated as both early and late morphemes, partly on the basis of code-switching data which indicate differential behavior.

2.3 ESTONIAN AND ENGLISH

The two languages in this study, one Finnic and one Germanic, have both congruences and dissimilarities which create the potential for structural conflict as well as innovative combinations of structures. English, as is well known, marks argument structure primarily through word order, with a very limited range of (pronominal) case-marking. Estonian, on the other hand, uses a complex system of morphological case-marking to signal grammatical relations, with more flexible word order. English is analytic, whereas Estonian is fusional-agglutinative (Erelt 2003), and uses a combination of post-positions, prepositions, and cases to signal the adverbial meanings usually encoded by prepositions in English.

The difference in morphological richness can be easily demonstrated by the verb paradigms. In English, a total of three overt morphemes are used in regular, finite verb conjugation, as well as frequent zero marking, and only one of the overt affixes signals person/number distinctions. In general, verbal marking signals tense or aspect, with present tense given null inflection in all but one cell, and an overt morpheme (-*ed*) marking past tense. The gerund (-*ing*) is a frequently used verb form which neutralises other grammatical information in the lexical verb, leaving person, tense and number marking to the auxiliary.

In Estonian, on the other hand, each cell in the regular verb paradigm for present and past tense is distinct (with one exception, the isomorphism of 2sg past and 3pl past), and each is marked by overt morphology. As shown in Table 1, the system is nearly agglutinative, with one affix marking past tense and a similar set of affixes marking person/number in both present and past tense: the first and second person morphemes are identical in present and past tense, whereas third person marking is unique in each cell. A total of eleven distinct verb endings are used for regular verbs in present and past tense. This is summarised in Table 1.

Table 1. Regular verb paradigms in English and Estonian

<table>
<tr><td colspan="3"></td><td>1st person</td><td>2nd person</td><td>3rd person</td></tr>
</table>

			1st person	2nd person	3rd person
English	Present	SINGULAR			-s
		PLURAL	-ø		
	Past	SINGULAR	-ed		
		PLURAL			
	Progressive		-ing		
Estonian	Present	SINGULAR	-n	-d	-b
		PLURAL	-me	-te	-vad
	Past	SINGULAR	-si-n	-si-d	-s-ø
		PLURAL	-si-me	-si-te	-si-d

An additional complicating factor is that Estonian verb paradigms often involve stem changes. For instance, the verb *rääki-ma* 'to talk' has two stems (*räägi-* and *rääki-*), used in different parts of the paradigm. Present tense forms use the former (e.g. 1sg *räägi-n,* 'I speak'), whereas the past tense is formed from the latter (e.g. 1sg *rääki-si-n,* 'I spoke'). Participial forms of one verb may also derive from different stems in a paradigm, e.g. *katma* 'to cover' > passive participle: *on kae-tud* 'is covered' vs. active participle *on kat-nud* 'has covered'). There is no room here to go into detail (instead see Blevins 2007, Erelt et al. 1995, Viks 1992), but the above will be relevant for some of the examples discussed in section 4.

In summary, it is crucial that the use of finite verbs in Estonian necessarily involves inflectional endings. When finite verbs are embedded in code-switched utterances framed by the other language, the speaker is forced to make a choice of how to mark the verb, according to the source or target language.

3. Data and Method

The data examined in this study derive from a diary kept by myself of my daughters' utterances. The children are bilingual sisters, aged 6;6–7;11[4] (M) and 2;10–4;3 (K) during the time of the study. The family was living in Estonia. The father, a native Estonian speaker, spoke Estonian with all family members; the mother, raised bilingually in the United States, spoke English with the children and Estonian with the father.

The family, thus, represents a one-parent, one-language model, yet did not strictly observe this model, as when the family setting naturally inclined one or the other parent to join in a discussion in the other language, or when other speakers were present. Estonian was the common parental language, as well as the main social language spoken by most of the family's interlocutors outside the home. The children were attending full-time day-care and school, which took place entirely in Estonian.

Some friends and relatives spoke English with the family. The children visited their grandmother and other relatives and friends at least once a year in English-speaking settings. Additionally, although television was not a constant fixture, it must be said that the abundance and popularity of English-language entertainment and educational media may have had an effect on the children's language skills and preferences. Although the children were born and raised in Estonia, they spent much of their free time (outside educational institutions) with their mother, thus ensuring sufficient English exposure and identity to lead to a preference for English even with each other. During the year when they attended the same day-care center ('pre-school'), they were reported to often speak English with each other on school premises, either as a sign of solidarity and

---

[4]   This notation is commonly used in developmental literature to indicate age: it is to be read as x;y.z = x years, y months, z days. I will continue to use this notation following examples to note age of speaker.

belonging, or else to distinguish themselves from others. Neverthe-
less, their language dominance is not easy to assess, and having used
no formal methods of assessment, I will refrain from saying any-
thing further on the matter.

The diary data were gathered during everyday activities by
myself (the children's mother), though at the time of the examples
used in this study this did not reflect all mixed utterances, but
rather those that were particularly noteworthy for any reason (e.g.
they were not typical of adult code-switching), and those which I
was able to record in my notes, often doing so in the course of a
busy work and family schedule. The examples included here come
from a corpus of 600 examples of mixed utterances, of which 85%
come from K (2-4 years old). It may be worth noting that the high
proportion of examples with Estonian insertions into English utter-
ances does not necessarily say anything about actual frequency
of the direction of insertions, but rather reflects the bias created
by the observer's role (or Observer's Paradox, Labov 1972, Lanza
1997) – since I spoke English with the children, the utterances
heard and noted most frequently came from an English discourse
context.

The data were combed for examples with code-switched verbs.
All examples with verb and predicate insertions were collected and
analysed for type of code-switching and amount of inflectional mor-
phology involved in the code-switching. The results are discussed in
the next section.

4. Results

As mentioned above, verb insertions necessarily involve choices
regarding the morphology associated with the verb. The examples in
the dataset were grouped according to what sort of verb morphology
was used. In the following, I discuss the code-switched utterances,
first finite and then non-finite verbs, according to the morphological
make-up of the bilingual clause.

4.1 FINITE VERBS

Numerous verbs in the data involved code-switching. Especially the younger child, K, used various strategies for integrating embedded verbs into clauses in the other language. These included three main combinations: null inflection, embedded language inflection, and matrix language inflection. Each of these are exemplified and discussed below.

4.1.1. FINITE VERBS WITH NULL INFLECTION

Estonian verbs are often used in the data in English-language utterances without any morphological marking. Because many Estonian verbs undergo stem changes in their inflectional paradigms, it is pertinent to ask which stems are selected by the child in the context of insertions into another language. The data include stems which are used as uninflected bare stems in certain contexts in Estonian, such as those in (1–2). Estonian uses these stems in contexts with no affix: for instance, in negative clauses (where the negative particle blocks any person/number marking) and second person singular imperative constructions.

(1)   You **kleebi-ø** [stick] the wings on.   (K, 3;10.2)
(2)   When we do the *õppimise* [learning] thing…
      when we **õpi-ø** [learn] then we don't go upstairs   (K, 3;10.16)

The inserted stem which appears in examples (3–4), on the other hand, is never used alone in Estonian. The verb *jätma* 'to leave.TRANSITIVE' has alternating stems used in different parts of the inflectional paradigm, *jät-* and *jäta-*. The stem ending in a vowel is used with indicative personal endings and contexts without any inflection (2sg imperative, negative), whereas the shorter stem without the vowel is used in past tense forms and some participle formation. However, the CVC form of the shorter stem is more in accord with English phonology, which may be what prompted the forms used in (3–4).

(3) **Jät-ø** [leave] some for me too!   (K, 3;8.24)

(4) I **jät-ø** [leave/left] some for [M...] because she's hungry (K, 3;8.24)

The lack of inflection in many cases can be analysed as following the grammar of English, where typically no inflectional morphology is required (e.g. for present tense 2sg and 1pl in examples 1–2). However, example (4) is intriguing from this point of view, as this utterance denotes a past event and would require past-tense morphology in both English and Estonian. As English past tense verbs usually end in /d/ or /t/, and the lexeme in question corresponds to a one-syllable, irregular verb ending in /t/ in English, the stem *jät* is phonologically felicitous in that it is prosodically similar to the appropriate form, although it does not follow the morphological paradigm of either English or Estonian (though note the existence of similar-sounding, null-marked past-tense verbs in English such as 'set' and 'put').

Hence, examples (1–4) can all be analysed as lacking morphology. Yet, it is not entirely clear which language this should be related to: should it be interpreted as lacking morphology from the perspective of English, which requires null morphology in many of these examples, or from the perspective of Estonian, which would require overt endings. As all these examples have a clear English matrix frame, we would be justified in saying that null morphology is appropriate, yet the form selected for the embedded verb in (4) remains a puzzle. Does this form indicate a misanalysed Estonian stem, or does it reflect a subtle interaction of the morphophonological properties of the lexeme, drawing on knowledge from both languages? Bilingual adults in ordinary code-switching environments might use the form *jäta-* with no ending, but *jät-* is anomalous.

Examples (5–8) all have predicate verbs formed with an auxiliary and lexical verb. The lexical verb comes from language B (Estonian, in these instances) and is inserted with no inflection, appropriately for language A (English, which makes frequent use of auxiliaries). Hence, these are all cases of B-insertions which can be analysed as following A-morphology. In (8), we see another example like (3–4)

above, wherein the selected stem does not appear on its own as a bare stem in Estonian.

> (5) Then I **can't *kasta-ø*** [water.TRANS] the flowers when there isn't any rain
>   to go in here.   (K, 3;11.3)
>
> (6) I **don't *mäleta-ø*** [remember] what it was   (K, 3;11.2)
>
> (7) Jakob says that the *soovikivi* [wishing-stone.NOM] **doesn't
>   *täida-ø*** [grant wishes.NOM.PL] *soovid*. (M, 7;6.12)
>
> (8) These **will *kadu-ø*** [get-lost] in here.   (K, 3;11.0)

Also importantly from the point of view of the Matrix Language Frame and the morphology used, two of these examples have direct objects. In (5), only the lexical verb is inserted, with the auxiliary 'can', the negative particle and the direct object all in English and following English matrix structure. Example (7), on the other hand, is a more complicated mix of two grammars. The matrix clause involves reported speech, which is most likely the reason this utterance includes several inserted lexemes: the subordinate clause takes its subject *soovikivi* 'wishing stone', lexical verb *täida* 'fill', and direct object *soov-id* 'wish-NOM.PL' from language B. Moreover, one might assume the phrase 'grant wishes' is taken wholesale from Estonian, but in Estonian, negative polarity requires an object case-marked with partitive case. This is an exceptionless and early-learned syntactic rule in Estonian (Argus 2009, Erelt et al. 1995), but this utterance is an example of two languages interacting in complex ways. Although the lexical verb, and indeed all the content words, are taken from Estonian, the core argument structure derives from English. English structure is marked by the definite article 'the', the negative contracted auxiliary 'doesn't', and the lack of partitive case-marking on the object.

4.1.2. FINITE VERBS WITH MORPHOLOGICAL INFLECTION

Finite verbs also appear in the data with inflection from both languages. In examples (9–11), the embedded language, B, contributes the lexical verb together with its inflectional morphology.

(9) Jüri **kriimusta-s** [scratch-3SG.PAST] my hand when we were going outside. (K, 3;10.16)

(10) When I was there, *Vanaema* [Granny] **korista-s** [clean-3SG.PAST] the *tänav* [street.NOM.SG]. (K, 3;9.21)

(11) I knew daddy's letter but our teacher *ikka* [still] **õpeta-s** [teach-3SG. PAST] **(it to us)**.(K, 3;9.21)

All these examples are third person singular and past tense, and hence would require morphology in either language. Importantly, example (10), like (5) and (7) above, has a direct object which reveals more of the syntactic structure of the clause. *Tänav* 'street' is in nominative case (the unmarked subject case), whereas in Estonian it would require overt, direct object case-marking (partitive or genitive). Additionally, it has an English definite article. Hence, this example contains system morphemes from two different languages, the past tense from Estonian immediately followed by a definite article and nominal case taken from English structure.

Myers-Scotton's 4M model predicts the usage of matrix-language tense and aspect inflection with embedded verbs: B-items with A-inflection. Examples (12–15) follow this pattern, but in the collected data are not distinguished from the examples in (9–11) by context, content or age. The examples are interspersed and show no detectable regularity of usage patterns.

(12) Today we **võimle-d** [exercise-ENG.PAST] at preschool (K, 3;8.15)

(13) Bazoo [=cat] **tagurda-d** [went-backward-ENG.PAST]**.** Like this.

(M, 7;9.17)

(14) When the big boy **karju-s** [yell-s] in the bathroom then my friends can't sleep. (K, 3;5.7)

(15) *Issi kui me käisime   arstis   siis   ma* **choose-*i-si-n***
     Daddy when we went doctor-LOC then I   choose-PAST.1SG
     selle   **sparkly** *konna.*
     this.GEN sparkly   frog.GEN
     'Daddy when we went to the doctor's then I **chose** this **sparkly**
     frog.'   (K, 3;9.14)

Examples (12–13) both involve English past tense inflection on Estonian lexical verb stems. In (14), the verb has an ending which appears in Estonian on verbs, and in fact the form *karjus* means 'yelled'. However, in this context, the subsequent clause disambiguates the -*s* ending as an English 3SG present tense inflection (rather than Estonian 3SG past): 'When the big boy yells, then my friends can't sleep'.

Finally, in (15), an Estonian matrix clause has an inserted English verb. The verb in English is irregular, and would produce a one-syllable past tense with no /d/ ending ('chose'), but here it is given regular Estonian inflection, with a default theme vowel -*i*-, the past tense -*si* and first person singular -*n*. In this case, despite the B-inserted verb and adjective, the direct object is inflected fully according to Estonian object case-marking paradigms, with genitive case.

Example (14), though it can be straightforwardly analysed as an Estonian verb with English inflection, raises the issue of morphology seemingly 'blended' from the two languages, where it is not clear which is the source language. This might be better described, especially in the children's code-switching data, as a means of combining inflectional resources from the available languages.

This leads to the question of analysis of examples such as those in (16–19), which were frequent for a period of nearly six months. In these, a clause typically displaying English matrix clause structure employs an inserted Estonian verb in a predicate in progressive aspect. However, the lexical verb is marked with a -*n* affix. This is simpler phonologically than -*ing*, as well as following the phonological structure of Estonian, which lacks /ŋ/. However, in Estonian, the -*n* inflection on verbs indicates first person singular. In (16), then, the -*n* could be drawing on the Estonian structure not

only phonologically, but also morphologically. It is clear that *-n* is an adapted form of the progressive ending, but it may also incorporate the frequent Estonian verbal *-n* morpheme, to produce a form which draws on grammatical resources and satisfies grammatical requirements of both languages. From examples (17–18), we can see that the *-n* ending is not limited to first person contexts.

(16) I wasn't hitting you, I **was *koputa-n*** [knock-n]   (K, 3;5.24)

(17) Why you**'re *kirjuta-n*** [write-n], Mommy?   (K, 3;5.25)

(18) a. K: This is a bad plate because this guy **is *aja-n tema taga*** [drive-n 3SG.NOM/GEN behind].

b. Mother: What is he doing?

c. K: He **is *aja-ing tema taga*** [driving 3SG.NOM/GEN behind = chasing him].   (K, 3;5.17)

(19) We **are *mängi-n*** [play-n] *peitust* [hide-and-seek.PAR] with the *karu* [bear]   (K, 3;5.10)

In (18a), K comments on a picture of Mr. McGregor chasing Peter Rabbit with a rake, depicted on her dinner plate. The entire matrix clause and beginning of the subordinate clause are in English, until the lexical verb *taga ajama,* 'chase'. This is a challenging word for code-switching, a phrasal verb composed of a light verb *ajama* 'drive/make/push' and adverbial *taga* 'behind', the semantics of which is not compositional, but depends on the whole phrase. Perhaps because the direct object occurs inside the complex verb, the language shifts at this point, and the rest of the clause is in Estonian, but the verb has the same *-n* ending which seems to mark the progressive in combination with the English auxiliary, and the pronominal direct object is not in the expected partitive case, but rather in nominative / genitive (ambiguous with this pronoun). The construction seems so surprising that I ask K to repeat it, upon which she slows down and enunciates the progressive marker, resulting in a more difficult word to pronounce (with two successive non-diphthong vowels from *aja* + *-ing* as well as the /ŋ/ ending), while also demonstrating that she knows and can say the English progressive *-ing*.

Finally, the example in (19) involves the same pattern, this time in first person plural, with an English auxiliary and Estonian verb 'play' with the -*n* ending, and with a direct object following Estonian case-marking patterns.

Examples with the type of 'blended' morphology discussed above and exemplified in (14) and (16–19) were regularly present in K's speech for approximately six months. These always involved Estonian verbs inserted into English utterances, with morphology which can be seen as drawing on both languages. A limited set of three inflections were used: -*d, -s,* and -*n*. These can all be mapped to English verb endings, but they (a) suit the Estonian verbs phonologically and (b) also occur in the Estonian verbal paradigm, though with different meanings, while (c) being selected according to similarity with English morphemes to map to grammatical functions required by English grammatical structure: -*d* (past), -*s* (third person singular present) and -*n* (progressive -*ing*).

4.2 NON-FINITE VERBS

The above examples demonstrate that the child is constructing finite verb forms online, as well as producing flexible argument structure constructions around them. According to Myers-Scotton, the combining of regular inflectional morphemes from one language with content morphemes from another is „strong evidence that regular inflections are supported as individual elements in the mental lexicon" (2005: 333). She claims that: „code switching also provides good evidence that Jackendoff's semi-productive elements are based on single units in the mental lexicon; they are not constructed on line. The evidence is that Embedded Language nonfinite verb forms, especially for the participles, from different languages always appear as holistic units in code switching" (Myers-Scotton, 2005: 333).

In this section, we examine code-switched participles to investigate whether this claim is well founded. Participles do participate in the code-switching data in question. In (20), K inserts a participle,

taken wholesale from Estonian, as predicted by Myers-Scotton, along with the participial morphology and the particle which comprises the phrasal verb 'tied up'.

(20) When your eyes are *kinni* ['closed'] **seo-tud** ['tied'] then you can't see. And then you break piñatas and then you get candy!

(K, 3;4.27)

In (21–22), however, we see that participles are not necessarily retrieved as holistic units, which may cast doubt on the argument above. If participles are always single units in the lexicon, that implies that they cannot be formed through morphological patterning. Nevertheless, examples of children producing erroneous forms such as 'eated' or 'putten' are not uncommon in English, indicating that either the story of mental storage of participles is more complicated than suggested, or else that children store them differently, perhaps not having yet acquired the crucial distinction between holistic units and grammatical forms constructed online. We first add to this question attested examples of participial forms used as insertions in code-switched utterances with erroneous, Estonian-based forms in (21–22). The error in example (21) arises from selection of the incorrect stem, while the error in (22) involves an incorrect allomorph.

(21) Now Katie's necklace is ***tee-dud*** ['made', pro: *'teh-tud'*]

(K, 3;6.2)

(22) Now it's even more ***mur-dud.*** ['broken' pro: *'mur-tud'*]

(K, 3;11.8)

However, more problematic for Myers-Scotton's argument is the fact that we also find examples of participles with code-switching across morpheme boundaries, a phenomenon which suggests that these forms are not simply stored or retrieved as holistic forms and that they may involve online construction. In (23), the Estonian verb *jätma* is given an English participial ending, resulting in *jät-en,* to express an English phrasal verb 'left over'. In (24), an English verb is inserted into an Estonian matrix clause and combined with an

Estonian participial ending. The affix involves metathesis, presumably as a consequence of the repetition of 'i' in the lexical verb, but the use of a participle here is in accord with Estonian syntax: negative past is formed with the negative particle *ei* and an active past participle, typically ending in (Vowel)-nud.

(23) Kribu [=cat] ate a drumstick that was ***jät*-en** [leave-ENG.PRT] over   (K, 3;6.5)

(24) *Ma jätsin   pildi   mida ma   ei   **finish-onid***
      I   leave.PAST.1SG picture what.PAR 1SG NEG finish-EST.PRT
      lasteaeda
      at-preschool
      'I left a picture that I didn't finish at preschool.'   (K, 3;11.3)

These and other examples of code-switching in nonfinite verb forms across morpheme borders are clear violations of Myers-Scotton's prediction that participles, as holistic units, are embedded holistically in code-switching. They occur rarely enough that it would not seem plausible to claim the verbal stems are borrowings, for instance, but their recurrence in at least some children's productions merits attention in any framework for analysing code-switching data.

5. Summary and Conclusions

The examples discussed here are drawn from diary data from two children simultaneously bilingual in Estonian and English, with fairly balanced competence in both languages. As shown, verb insertions appear in the children's code-switching with various combinations of morphological marking. This includes embedded verbs with:

a. zero marking in constructions which may be either well- or ill-formed in the matrix language of the utterance;
b. morphology from the source language;
c. morphology and function words from the target language;
d. a blend, where both source and target languages constrain the forms used in mixing.

Returning to the research questions posed at the beginning, we claim that children's code-switching provides a fruitful vantage point for analysing their emerging knowledge of grammar and linguistic structure. The data examined here come from only two children and two languages, but suggests strongly that the proposed constraints are either not followed by children or not as generally applicable as has been suggested in the literature. This study looks at only a subset of data involving verbs, but it holds promise for future research on interaction between languages, emergent grammatical systems, and possible combinations in code-switching, drawing on grammatical resources from both languages.

Children, whose grammatical competence is less than fully formed, are in a continuous process of discovering regularities, making predictions, and analogising across forms, in addition to learning new concepts and vocabulary. Some code-switching reflects an imbalance in lexical knowledge between the languages, but the more interesting examples of how those embedded words are grammatically accommodated may shed light on the knowledge underlying children's productions. When they embed a verb from one language in an utterance which begins in a different language, they must immediately make choices regarding the use of grammatical resources and how they interact. Data from code-switching is akin to a naturally occurring novel-word experiment. The degree and type of abstraction of regularities can be clearly seen in some of these innovated, online constructions.

Examples from the data show that both finite and nonfinite forms can occur with morphemes drawn from either the source or target language. It is beyond the scope of this paper to assess whether these constitute evidence for a lack of separation in the emergent morphological inventory, but they raise questions.

The division of lexical items into function and content morphemes may not always be cross-linguistically comparable. In analysing a child's developing grammar, these categories must not be taken for granted, either. Likewise, the question of separation of lexicons

from each language must be revisited, as it is unclear whether bilingual children draw on two distinct lexical inventories, and how the distinctions between the languages emerge. Individual differences compounded with language experience may lead to greater effects of differential inhibitory control, executive function and verbal memory in bilingual children. This multiplied effect may also lead to greater differences among bilingual than monolingual children in the storage and retrieval of lexical items and grammatical structures.

Finally, violations of predicted constraints in children's codeswitching data may mean that the predictions are incorrect, but it may also mean that children do not adhere to the same constraints on code-switching as adults do. This needs to be further explored, as much for analysing code-switching data as for what this implies about the developing cognitive systems, before we assess the merits of one or another set of constraints on code-switching. Regardless, it must be noted that these constraints have been criticised also from the perspective of adult code-switching (e.g. Backus 2014, Auer & Muhamedova 2005). The creative, online code switching of children affords a window onto the mutual influence of the languages in the mind, as well as the child's emergent knowledge of both linguistic systems.

REFERENCES

Argus, Reili 2009. The early development of case and number in Estonian. In M. D. Voeikova & U. Stephany (Eds.) Development of Nominal Inflection in First Language Acquisition: A Cross-Linguistic Perspective, pp. 111–152. Berlin: Mouton De Gruyter.

Argyri, Efrosyni; Sorace, Antonella 2007. Crosslinguistic influence and language dominance in older bilingual children. Bilingualism: Language and Cognition, 10(1): 79–99.

Auer, Peter; Muhamedova, Raihan 2005. 'Embedded language' and 'matrix language' insertional language mixing: Some problematic cases. Italian Journal of Linguistics / Rivista di linguistica, 17(1): 35–54 (special issue, ed. G. Berruto).

Backus, Ad 2014. Towards a Usage-Based account of language change: Implications of contact linguistics for linguistic theory. In R. Nicolai (ed.) Questioning Language Contact: Limits of Contact, Contact at its Limits. Leiden: Brill, pp. 91–118.

Backus, Ad; Verschik, Anna 2012. Copiability of (bound) morphology. In L. Johanson & M. Robbeets, Copies versus Cognates in Bound Morphology. Leiden: Brill, pp. 1–32.

Bolonyai, Agnes 2000. „Elective affinities": Language contact in the abstract lexicon and its structural consequences. International Journal of Bilingualism, 4(1): 81–106.

Boyle, Whitney; Lindell, Annukka; Kidd, Evan 2013. Investigating the Role of Verbal Working Memory in Young Children's Sentence Comprehension. Language Learning, 63(2): 211–242.

Chang, Franklin; Janciauska, Marius; Fitz, Hartmut 2012. Language adaptation and learning: Getting explicit about implicit learning. Language and Linguistic Compass, 6(5): 259–278.

De Bot, Kees 2004. The Multilingual Lexicon: Modelling Selection and Control. International Journal of Multilingualism, 1(1):17–32.

Demirçay, Derya; Backus, Ad 2014. Bilingual constructions: Reassessing the typology of code-switching. Dutch Journal of Applied Linguistics, 3(1):30–44.

Erelt, Mati; Kasik, Reet; Metslang, Helle; Rajandi, Henno; Ross, Kristiina; Saari, Henn; Tael, Kaja; Vare, Silvi 1995. Eesti keele grammatika. [=Grammar of Estonian]. Tallinn: Eesti Teaduste Akadeemia Eesti Keele Instituut.

Erelt, Mati 2003. Estonian Language. (Linguistica Uralica, Supplementary Series 1) Tallinn: Eesti Akadeemia.

Hulk, Aafke; Müller, Natascha 2000. Bilingual first language acquisition at the interface between syntax and pragmatics. Bilingualism: Language and Cognition, 3(3): 227–244.

Johanson, Lars 2002. Contact-induced linguistic change in a code-copying framework. In: M.C. Jones & E. Esch (eds.) Language change: The interplay of internal, external and extra-linguistic factors. Berlin: Mouton de Gruyter, pp. 285–313.

Levelt, Willem J. M. 1989. Speaking, from Intention to Articulation. Cambridge, MA: MIT Press.

Müller, Natascha; Hulk, Aafke 2001. Crosslinguistic Influence in Bilingual Language Acquisition: Italian and French as Recipient Languages. Bilingualism: Language and Cognition, 4(1): 1–21.

Myers-Scotton, Carol 2002. Contact linguistics, bilingual encounters and grammatical outcomes. OUP.

Myers-Scotton, Carol 2005. Supporting a Differential Access Hypothesis. In J. Kroll & A. de Groot (eds). Handbook of bilingualism, psycholinguistic approaches, pp. 326–348. New York: OUP.

Myers-Scotton, Carol; Jake, Janice L. 2000. Four types of morpheme: Evidence from aphasia, code switching, and second-language acquisition. Linguistics 38: 1053–1100.

Paradis, Johanne; Genesee, Fred 1996. Syntactic Acquisition in Bilingual Children. Studies in Second Language Acquisition, 18: 1–25.

Pérez-Leroux, Ana Teresa; Cuza, Alejandro; Thomas, Danielle 2011. Clitic placement in Spanish / English bilingual children. Bilingualism: Language and Cognition, 14: 221–232.

Perruchet, Pierre; Pacton, Sébastien 2006. Implicit learning and statistical learning: One phenomenon, two approaches. Trends in Cognitive Sciences, 10, 233–238.

Rebuschat, Patrick; Williams, John N. (eds.). 2012. Statistical learning and language acquisition. Berlin: Mouton de Gruyter.

Unsworth, Sharon 2013. Current issues in multilingual first language acquisition. Annual Review of Applied Linguistics, 33: 21–50. http://dx.doi.org/10.1017/S0267190513000044

Verschik, Anna 2007. Multiple language contact in Tallinn: Transfer B2 > A1 or B1 > A2? The International Journal of Bilingual Education and Bilingualism, 10 (1): 80–103.

Vihman, Marilyn M. 2016. Prosodic structures and templates in bilingual phonological development. Bilingualism: Language and Cognition, 19: 69–88.

Yip, Virginia 2013. Simultaneous language acquisition. In F. Grosjean and L. Ping (eds). The Psycholinguistics of Bilingualism, Wiley-Blackwell, pp. 119–136.