A dual-structure analysis of morphosyntactic doubling in code switching[*]

Caleb Hicks

University of North Carolina at Chapel Hill
linghix@live.unc.edu

Code switches in which a single morphosyntactic element is produced twice have been noted in code switching literature (Poplack et al. 1989, Sankoff et al. 1990, Myers-Scotton 1993, Nishimura 1995), but explication of their syntactic structure has remained elusive. In this paper, I present a structural analysis of morphosyntactic doubling in code switching inspired by Sadock's (1991) dual-structure model of the morphosyntactic interface. I argue that certain code switches have an "upper structure," representing one source language, and a "lower structure," representing the other. Elements in the corresponding trees which are not shared by both structures are doubled.

1. Introduction

Intrasentential code switching is a language contact phenomenon in which a speaker utters a sentence containing syntactic and/or morphological material from more than one language.[1] When two languages involved in a switch, hereafter called *source languages*, are very different typologically, the resulting code switches may adhere to the grammatical properties of one language or the other, and much of the better known code switching literature has focused on determining which of the two source languages will dictate the grammar of the utterance as a whole (e.g. Poplack 1980, Myers-Scotton 1993).

The typological parameter relevant to the present discussion is that of

---

[1] It is not my intention here to distinguish between code switching, code mixing, and code blending, or otherwise to deal with the general terminological issues in which the label *code switching* has become entrenched. For more, see Kachru (1978:28), Pfaff (1979:296), Poplack (1980:225-6), Clyne (1987:258), Bokamba (1988:24), Kamwangamalu (1989:321), Myers-Scotton (1993:163), Myers-Scotton & Jake (1995:282), Mashiri (2002:246), and Porte (2003:105), to name a few.

constituent order. When two source languages have different constituent orders, most code switches resulting from these languages will adhere to the constituent order of one source language or the other. In other words, it is most often the case that the constituent order of the code switch complies with that of one source language, at the expense of the order of the other. A rare type of switch, however, occurs when *both* constituent orders are (at least partially) adopted. When this happens, the resulting code switch may contain at least one morphosyntactic element (a phrase, word, or morpheme) which is produced twice: the first occurrence of the element appears in the unmarked position for that element in one source language, while the second occurrence appears in the unmarked position in the other source language.

The double-occurrence of such an element in a code switch, where the source languages have different constituent orders, is a phenomenon for which I will use the label *morphosyntactic doubling*, and I will interchange this term with *doubling*.[2] Code switches containing these doubles have variously been called *portmanteau sentences* (Nishimura 1986, 1995), *copy translation constructions* (Poplack et al. 1989:396), *palindromic switches* (Sankoff et al. 1990:92), and *repetition translation* or *repeat translation constructions* (Sankoff 1998:15-17); almost all authors have noted the remarkable rarity of these types of code switches.

Several authors have proposed explanations for why morphosyntactic doubles occur in the first place (Poplack et al. 1989, Sankoff et al. 1990, Azuma 1993, Myers-Scotton 1993, Nishimura 1995, Sankoff 1998), but to my knowledge, no one has yet provided a structural analysis of doubling of this type. The aim of this paper is to suggest one way in which the structure of code switches which contain morphosyntactic doubles can be analyzed. In brief, the basis of my proposal is that such code switches can be represented by the integration of two (morpho)syntactic trees: one representing the first source language and descending, the other representing the second and ascending; the two trees meet in the middle by sharing certain constituents common to both structures. I show that this analysis successfully models all of the examples of morphosyntactic doubling I have found in the published literature and makes predictions

---

[2] As a more general phenomenon, *morphosyntactic doubling* can also be applied to non-code switched languages (Glaser & Frey 2006, Kortmann & Szmrecsanyi 2006, and see a compendium of papers in Barbiers et al. (eds.) 2008), but in this paper, I am restricting my use of the term to its occurrence in code switching.

about the range of morphosyntactic doubling possible in code switches involving other source languages. In section 2, I will show a variety of examples of the phenomenon of morphosyntactic doubling, and will present the details of the analysis in section 3. Section 4 will close the paper with concluding remarks.

2. Morphosyntactic doubling in code switching

In this section, I provide examples of morphosyntactic doubles in code switched sentences where the two source languages have different unmarked constituent orders. The examples are organized by the type of constituent doubled. Each instance of a doubled element is italicized for ease of identification. The first set below shows doubled verbs and auxiliaries.

(English-Hindi, Pandit 1986:41)
(1) she will not come to me because the hindu system *is* tarah kaa *hai*
    she will not come to me because the hindu system is that   of is
    'She will not come to me because the Hindu system is like that.'

(English-Tamil, Sankoff 1990:93)
(2) they *make* candai *pooḍaraanga*
    they make fight   make.3.PL.PRES.
    'They are fighting.'

(English-Tamil, Sankoff 1990:93)
(3) they *gave* me a research grant *koḍutaa*
    they gave me a research grant give.3.PL.PAST
    'They gave me a research grant.'

(English-Japanese, Nishimura 1986:166)
(4) let's *become* kechi ni *naroo*
    let's become tight   become
    'Let's become tight.'

In each of the above examples, English, which is SVO, acts as one source language for the code switch. All of the sentences begin in English and at some point switch into a language with a different basic word order; here, all of the other source languages (Hindi, Tamil, and Japanese) have a basic word order of SOV. In each code switched sentence, some element is produced twice, either a verb or auxiliary.

Two properties of these examples deserve special mention at this point, and I will show that these properties hold for all of the examples of doubling I have found. First, notice that each double is realized in the language whose word order is obeyed at that point the sentence. In (1), the first realization of the auxiliary is in English as *is*, and this is the unmarked position for English auxiliaries (immediately following the subject). The second realization of the auxiliary is in Hindi as *hai*, and auxiliaries are clause final in unmarked Hindi word order. The same is true for the other examples: the position of *make* in (2) corresponds to an SVO order, while the the position of *pooḍaraanga* corresponds to SOV; the situation is isomorphic in (3) and (4).

Second, in each example, there is some shared element. This property was described explicitly by Nishimura (1986) and Azuma (1993) for Japanese, and I have found it be the case in every instance of code switched doubling. By *shared element* I mean some constituent which acts as a syntactic complement to two different heads. As an example, consider (3). The constituent *a research grant* is the object of *gave* and also of *koḍutaa.* The two realizations of this verb share a complement. This property will be elaborated on in section 3.

Now consider a second set of examples below. These sentences are code switches in which an adposition is doubled.

(English-Japanese, Nishimura 1995:139)
(5) We *bought about* two pounds *gurai kattekita* no
  We bought about two pounds about bought
  'We bought about two pounds.'

(English-Japanese, Nishimura 1986:140)
(6) look at the things she buys *for* Sean *ni*
  look at the things she buys for Sean for
  'Look at the things she buys for Sean.'

(English-Finnish, Poplack et al. 1987:404)
(7) mutta se oli kidney-sta *to* aorta-*an*
  but it was kidney-from to aorta-to
  'But it was from the kidney to the aorta.'

(English-Marathi, Hicks 2010:45)

(8) I could run every you know *in* thirty minutes *madhe*   once a day
    I could run every you know in thirty minutes in   once a day
    'I could run every, you know, in thirty minutes, once a day.'

As in (1)-(4), the sentences in (5)-(8) each start out in English and then switch to another source language later in the utterance. In these code switches, an adposition is doubled. In (5), the adposition is realized as the English preposition *about* and then again as the Japanese postposition *gurai* (note here that the verb is doubled as well). Sentence (6), again showing an English-Japanese switch, contains the English preposition *for* and the Japanese postposition *ni*. The English-Finnish code switch in (7) shows the English preposition *to* as well as the Finnish illative case suffix *-an*.[3] In (8), the English preposition *in* is doubled by the Marathi postposition *madhe*. In these code switches, constituent order again plays a role. Since English employs prepositions, the unmarked order of PPs is adposition-noun. Japanese and Marathi are postpositional languages, so their basic order is noun-adposition.

Next, consider the following code switches showing a doubled complementizer.

(English-Tamil, Sankoff et al. 1990:93)

(9) just *because* avaa innoru   color and race *engindratunaale*
    just because they   different   color and race of-because
    'Just because they are of a different color and race.'

(English-Japanese, Azuma 1990:199)

(10)   *if* it goes three rounds datta   *ra* ne
    if it goes three rounds was   if   TAG
    'If it goes three rounds.'

---

[3] Finnish does employ postpositions as free morphemes; for a discussion of the formal distinction between semantic cases and adpositions in Finnish, see Nikanne (1993).

(English-Korean, Chan 2008:800)

(11)   everybody *think that*   nay-ka yenge-lul
       everybody think C   I-NOM English-ACC

       cal   hanta-*ko*   *sayngkakhayyo*
       well   do-C   think
       'Everybody thinks that I'm a good English speaker.'

In (9), the complementizer *because* appears first the utterance, while the Tamil equivalent *engindratunaale* comes at the end. The code switch in (10) contains English *if* and then Japanese *ra*. The English-Korean code switch in (11) shows both a doubled complementizer and a doubled verb. The matrix verb, *think* in English and *sayngkakhayyo* in Korean, appears in its unmarked position for each respective language, and its CP complement, headed by English *that* and Korean *-ko*, also appears in the usual spots for both languages.

The examples of doubling discussed so far have each involved English as one source language, and in every instance, English has been the first language of the utterance. This has structural consequences which I will discuss later, but it should be noted that English-first code switches are not the only cases of doubling. Below are examples of code switches in which English is either not a source language at all, or is the second language in the switch.

(Spanish-Aymara, Stolz 1996:10)

(12)   *pero*   sorro-*sti*   wali   astuturi-tajna...
       but   fox-COO   very   keen-3.SG.PRT.EVI
       'But the fox was very keen.'

(Lingala-French, Bokamba 1988:37, Parse: Myers-Scotton & Jake 1995:298)

(13)   ...ba-jeune-*s*   *ko*-comprend-*re*   avenir te
       ...CL2-young-PL   INF-understand-INF   future not
       ...young people did not understand.'

(Lingala-French, Myers-Scotton 1993:133)

(14)   *ba*-parent-*s*
       CL2.PL-parent-PL
       'parents'

(Shona-English, Myers-Scotton 1993:132)

(15)   *ma*-game-*s*
       CL6.PL-game-PL
       'games'

The Spanish-Aymara code switch in (12) shows a doubled conjunction: first in Spanish as *pero* 'but' and again in Aymara as the post-nominal suffix -*sti*, glossed as COO (coordinator). In (13), a doubled infinitive morpheme is present preverbally as Lingala *ko*- and also post-verbally as French -*re*. A plural morpheme is doubled in (14) and (15), where the first realization is prenominal in Lingala and Shona, respectively, and then again post-nominally in French and English, respectively.

In this section, I have shown a number of examples of what I am calling *morphosyntactic doubling* in code switches: some morphosyntactic element is realized twice in each utterance, and each realization occurs in a different source language. Furthermore, the position in which each realization occurs adheres to the constituent order properties of the source language from which it is drawn. Doubled verbs and auxiliaries appeared in (1)-(5) as well as in (11), doubled adpositions appeared in (5)-(8), doubled complementizers in (9)-(11), and doubles where at least one realization is a morphological affix appeared in (12)-(15). In section 3, I will present a syntactic analysis of morphosyntactic doubles.

3. The dual syntactic structure of morphosyntactic doubles

In this section, I argue that morphosyntactic doubles in code switches may be analyzed structurally by exploiting a version of a dual structure model (Sadock 1985, 1991). Such an analysis treats these doubles as consisting of two syntactic structures, each with distinct heads and projections, which are joined by some shared constituent(s).

Sadock (1985, 1991) developed a theory of autolexical syntax in which tree structure diagrams represent various components or *modules* of the grammar. In this framework, the morphological, syntactic, and semantic modules are each representable as trees, and the trees interface with each other by linking together elements in each tree. An example of Sadock's morphosyntactic interface is shown in diagram 1.
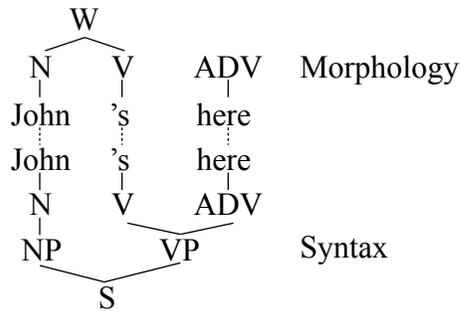
```
              W
            /   \
       N     V     ADV      Morphology
       |     |      |
      John   's    here
      John   's    here
       |     |      |
       N     V     ADV
       |      \    /
      NP        VP           Syntax
        \      /
           S
```

Diagram 1. The morphosyntactic interface in autolexical syntax (Sadock 1985:385)

In diagram 1, the morphological module is represented by the upper structure, while the syntactic module is represented by the lower structure. Where the structures meet in the middle is representative of the interface between the two modules. Sadock advanced this model to account for phenomena such as cliticization and noun incorporation. The very brief sketch of autolexical syntax provided here does not portray its details and implications, but the basic dual-structure nature of the model is shown because it is the inspiration for the analysis to follow[4].

Constructions containing morphosyntactic doubles can be represented by two integrated tree structures, one above and descending and the other below and ascending. As a first example, consider the English-Japanese code switch in (6). An adposition is doubled, first in English as *for* and then in Japanese as *ni*. The NP *Sean* is a shared complement of both of the Ps. In the structure in diagram 2, the English PP is shown in the upper structure, while the Japanese PP is in the lower structure. The two PPs meet in the middle at the shared NP complement.

Notice in diagram 2 that that the NP *Sean* is the only shared element in the structure as a whole. Notice also that the Japanese PP is not contained in a VP or any higher projection as is the English PP. This is because there is no lexical material to license higher-level phrases in the lower structure. There is no Japanese verb or any other item which requires placement in the tree. This situation may be contrasted with the English-Tamil code switch in (2), in which a verb is doubled. Since the verbs also contain tense information, a projection at least as high as the IP must be present in

---

[4] It is worthwhile stressing that mine is not an autolexical analysis; my concern is not the morphosyntactic or syntactic-semantic interfaces. Unlike autolexical theory, both components of the dual structure in my account represent the same "module," be it syntax or morphology.

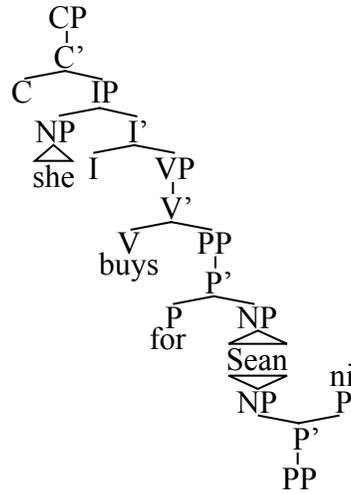both the upper and lower structures, as shown in diagram 3.

Diagram 2. Dual-structure of an English-Japanese code switch with a doubled adposition
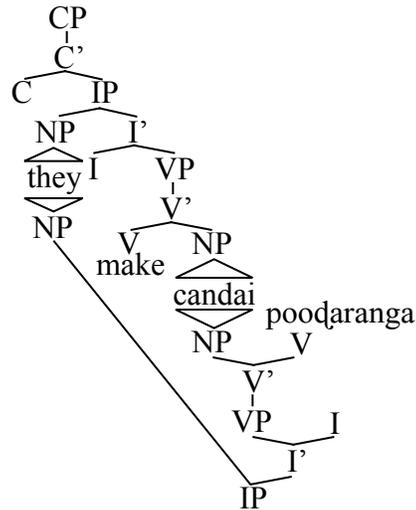
Diagram 3. Dual-structure of an English-Tamil code switch with a doubled verb

Operating under the assumption that in languages like English and Tamil, subjects raise to Spec,IP in order to receive case, an additional feature of the structure in diagram 3 is immediately apparent. The upper and lower trees are able to share a subject in addition to the verbal NP complement already shared. As I will discuss shortly, this has implications for doubling

in code switches whose contributor languages have a subject-final constituent order.

In some code switches, such as (5) and (11), more than one item is doubled. A dual structure accounts for these instances straightforwardly by representing *multiple* respective positions in the upper and lower trees as being filled with overt material. The English-Korean code switch in (11), in which a verb and a complementizer are both doubled, is shown below in diagram 4.
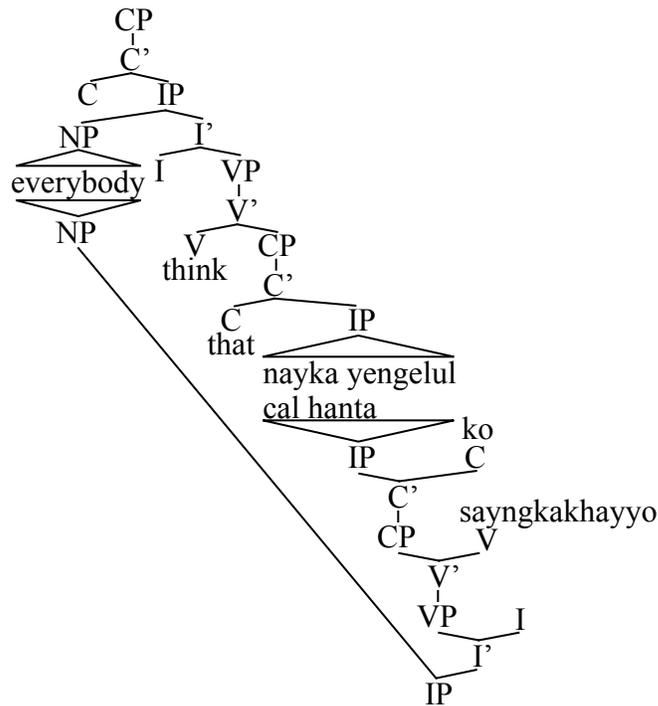

Diagram 4. Dual-structure of an English-Korean code switch with a doubled complementizer and verb

The treatment of morphosyntactic doubles as dual structures has implications about the types of doubling which are probable in code switches. In the examples provided so far, it has only been possible for a subject to be shared between the upper and lower structures (c.f. diagrams 3 and 4) because both contributor languages are assumed to have left-branching specifiers of IP. If a subject-initial language is one contributor language but a subject-final language is the other, and if the lexical content of the code switch licenses a subject for each verb, then sharing a subject is improbable.

Interestingly, a code switch such as this might make a *doubled* (rather than *shared*) subject especially likely because two subjects might exist in IPs which share the same VP or V. As examples, consider two hypothetical code switches: one between Turkish (SOV) and Hixkaryana (OVS), shown in diagram 5; the other between Hindi (SOV) and Lealao Chinantec (VOS), shown in diagram 6.
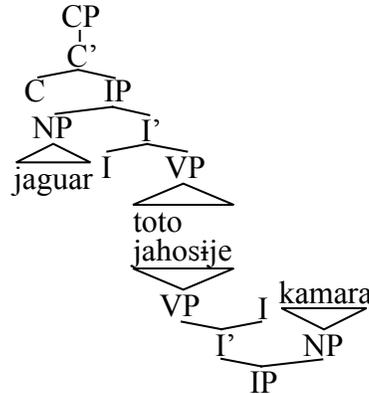
```
                CP
                |
                C'
              /   \
             C     IP
                  /  \
                NP    I'
               /__\  / \
              jaguar I   VP
                        /__\
                        toto
                       jahosɨje
                      /_____\
                    VP      I   kamara
                    / \        /__\
                   I'      NP
                      \    /
                       IP
```

Diagram 5. Dual-structure of a hypothetical Turkish-Hixkaryana code switch with a doubled subject

```
                CP
                |
                C'
              /   \
             C     IP
                  /  \
                NP    I'
               /__\  / \
               vo   I   VP
              aadmi     |
                        V'
                      /    \
                    NP       V
                   /__\     / \
                  meraa sar kaᴸkiu'ʔᴹ   miⱽᴴziiᴸi
                            V           /__\
                                       NP
                                       V'
                                       |
                                      VP    I   zaᴹniᴹ
                                      / \      /__\
                                     I'     NP
                                        \   /
                                        IP
```
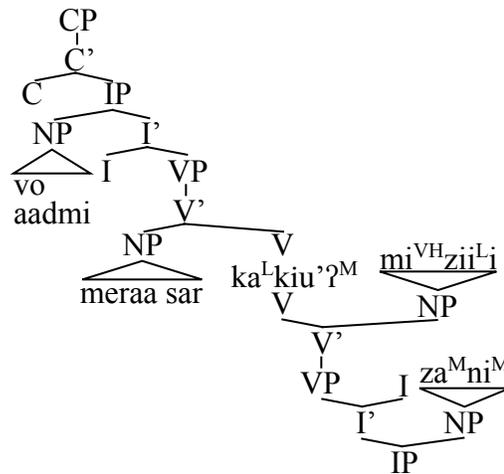
Diagram 6. Dual-structure of a hypothetical Hindi-Lealao Chinantec code switch with a doubled object and subject

Of course, the structures in diagrams 5 and 6 are speculative. I know of no doubles of any kind in code switches whose contributor languages are subject-final, and any code switches from such language pairs must be

very rare, owing to the typological rarity of subject-final languages.

There are further questions raised by the story presented here. There are many other typological parameters in which languages in a code switch can differ, and I have not explored what implications such pairs would have for a dual structure. If, for instance, two contributor languages differ such that one requires subjects to raise in order to receive case while the other allows subjects to receive case *in situ*, it is not clear whether the respective subjects could be shared in a dual structure. Similar questions arise regarding verb raising and other types of movement which may be considered parametric or language-specific. In addition, while the English-Finnish code switch in (7) shows an English preposition and a Finnish case suffix, it remains to be made explicit how different morphosyntactic representations of the same semantic element in the two trees are integrated in a dual structure[5]. Further investigation into morphosyntactic doubling is spurred by such questions, and would benefit tremendously from additional code switching data from other language pairs.

In this section, I presented a treatment of morphosyntactic doubling which promotes a dual structure for code switches in which doubling occurs. Using the dual-structure model of Sadock (1985, 1991) as a springboard, this analysis treats these code switches as having an upper structure and a lower structure which are joined at the constituent(s) shared by both structures.

4. Conclusion

In this paper, I have argued for a dual-structure analysis of morphosyntactic doubling in code switching. As described in section 1, some code switches involve two languages with different constituent orders. These switches have the potential of containing a double of some morphosyntactic element; that is, some item may be produced twice, once in each language.

In section 2, I showed examples of morphosyntactic doubling, as represented in the published literature. The range of constituent types which are doubled in different code switches is quite large: verbs, auxiliaries, adpositions, complementizers, conjunctions, infinitive

---

[5] Thank you to Randall Hendrick for bringing these issues to my attention.

morphemes, and plural morphemes. In addition, the specific languages involved in such switches are quite diverse.

I presented a structural analysis of morphosyntactic doubles in section 3. In this analysis, the code switches shown in section 2 were treated as consisting of two syntactic trees which are integrated into a single structure. The trees are arranged such that one is above and works its way down, while the other is below and works its way up. The trees meet in the middle of the structure by sharing certain lexical material. Material which is present in both trees but not shared is doubled. This analysis presents a straightforward approach to a significant obstacle in code switching syntax: how to account for constituents which are produced twice, occur in different sentential positions, and are realized in different source languages.

REFERENCES

Azuma, Shoji. 1993. Word order vs. word class: Portmanteau sentences in bilinguals. In Patricia M. Clancy (ed.). *Japanese/Korean Linguistics* 2. Stanford: Center for the Study of Language and Information.
Barbiers, Sjef, Olaf Koeneman, Marika Lekakou & Margreet van der Ham (eds.). 2008. *Syntax and Semantics* 36: *Microvariation in syntactic doubling*. Bingley: Emerald.
Bokamba, E. 1988. Code-mixing, language variation, and linguistic theory: Evidence from Bantu languages. *Lingua* 76. 21-62.
Chan, Brian Hok-Shing. 2008. Code-switching, word order and the lexical/functional category distinction. *Lingua* 118. 777-809.
Clyne, M. 1987. Constraints on code-switching: How universal are they? *Linguistics* 25. 739-64.
Glaser, Elvira & Natascha Frey. 2006. Doubling in Swiss-German dialects, http://www.meertens.knaw.nl/projecten/edisyn/. (May 2010.)
Hicks, Caleb. 2010. *Morphosyntactic Doubling in Code Switching*. Chapel Hill: University of North Carolina MA thesis.
Kachru, Braj B. 1978. Toward structuring code-mixing: An Indian perspective. *International Journal of the Sociology of Language* 16. 27-46.
Kamwangamalu, Nkonko Mudipanu. 1989. Code-mixing and modernization. *World Englishes* 8(3). 321-332.
Kortmann, B. & B. Szmrecsanyi. 2006. Morphosyntactic doubling phenomena in non-standard varieties of English, http://www.meertens.knaw.nl/projecten/edisyn/. (May 2010.)
Mashiri, Pedzisai. 2002. Shona-English code-mixing in the speech of students at the University of Zimbabwe. *South African Linguistics and Applied Language Studies* 20. 245-261.
Myers-Scotton, Carol. 1993. *Duelling languages: Grammatical structure in codeswitching*. Oxford: Clarendon Press.
Myers-Scotton, Carol. & J.L. Jake. 1995. Matching lemmas in a bilingual competence and production model. *Linguistics* 33. 981-1024.

Nikanne, Urpo. 1993. On assigning semantic case in Finnish. In Anders Holmberg & Urpo Nikanne (eds.). *Case and other Functional Categories in Finnish Syntax*. Berlin: Mouton de Gruyter. 75-88.

Nishimura, Miwa. 1986. Intrasentential code-switching: The case of language assignment. In Jyotsna Vaid (ed.). *Language Processing in Bilinguals: Psycholinguistic and neuropsychological perspectives*. Hillsdale: Lawrence Erlbaum Associates.

Nishimura, Miwa. 1995. A functional analysis of Japanese/English code-switching. *Journal of Pragmatics* 23. 157-181.

Pandit, Ira. 1986. *Hindi English Code Switching: Mixed Hindi English*. Delhi: Datta Book Centre.

Pfaff, C. 1979. Constraints on language mixing: Intrasentential code-switching and borrowing in Spanish/English. *Language* 55. 291-318.

Poplack, Shana. 1980. Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of code-switching. *Linguistics* 18. 581-618.

Poplack, Shana, Susan Wheeler & Anneli Westwood. 1989. Distinguishing language-contact phenomena: Evidence from Finnish-English bilingualism. *World Englishes* 8(3). 389-406.

Porte, Graeme. 2003. English from a distance: Code-mixing and blending in the L1 output of long-term resident overseas EFL teachers. In Vivian Cook (ed.). *Effects of the Second Language on the First*. Tonawanda: Multilingual Matters Ltd.

Sadock, Jerrold M. 1985. Autolexical syntax: A proposal for the treatment of noun incorporation and similar phenomena. *Natural Language and Linguistic Theory* 3. 379-439.

Sadock, Jerrold M. 1991. *Autolexical Syntax: A theory of parallel grammatical representations*. Chicago: University of Chicago Press.

Sankoff, David. 1998. The production of code-mixed discourse. *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, August 10-14, 1998*. 8-21.

Sankoff, David, Shana Poplack & Swathi Vanniarajan. 1990. The case of the nonce loan in Tamil. *Language Variation and Change* 2. 71-101.

Stolz, Thomas. 1996. Grammatical hispanisms in Amerindian and Austronesian languages: The other kind of transpacific isoglosses. *AMERINDIA* 21. 1-22.