

Distinguishing language contact phenomena: evidence from Finnish–English bilingualism

SHANA POPLACK,* SUSAN WHEELER* and ANNELI WESTWOOD†

ABSTRACT: Using a variationist, i.e. accountable, approach to the study of code-switching, this paper attempts to validate the *equivalence constraint* on intrasentential code-switching on the basis of natural speech data from two typologically different languages, Finnish and English. Though the data include many apparent exceptions to the constraint, the authors show how these are resolved by distinguishing on the one hand between code-switching and borrowing, and on the other between the different functions code-switching fulfills in the discourse.

INTRODUCTION

The bilingual behaviour which has provoked the most controversy in linguistics is undoubtedly intrasentential code-switching. When two languages are to be used in a single sentence, various problems of incompatibility may arise. The most obvious derives from word-order differences—if a switch occurs at a boundary between two constituents which are ordered differently in the two languages, the resulting configuration will be ungrammatical by the standards of at least one. Another type of difficulty involves morphological disparity, as when a noun in one language must be inflected for case, where the other uses alternative means of accomplishing the same function. And there are many other problems having to do with subcategorization patterns, semantic differences, idiomatic constructions, etc.

It has been observed in systematic studies of bilingual communities that speakers tend to avoid these difficulties by eschewing switches at sites which would result in monolingually ungrammatical fragments. How is this accomplished? In earlier studies of Spanish–English bilingualism among Puerto Ricans in New York (Poplack, 1980, 1981; Sankoff and Poplack, 1981) we postulated the *equivalence constraint*, whereby switching is free to occur between any two sentence elements if they are normally ordered in the same way by the grammars of both languages involved, while prohibited elsewhere, as illustrated in (1).

(1) English:	DET + N	
Spanish:	DET + N	
CS:	ENG DET + SP N	
	SP DET + ENG N	
<hr/>		
English:	ADJ + N	
Spanish:	N + ADJ	
CS:	*ENG ADJ + SP N	} 'crossovers'
	*SP ADJ + ENG N	
	*ENG N + SP ADJ	
	*SP N + ENG ADJ	

Thus, for Puerto Rican bilinguals switching is permitted between a determiner of either language and a noun in the other but prohibited between a noun and an adjective because

*Department of Linguistics, University of Ottawa, 78 Laurier E., Ottawa, Canada K1N 6N5.

†Ottawa, Canada.

the result would be contrary to English word order, or between (most) adjectives and the noun, because this would be contrary to Spanish order.

Much has been written about the validity of this constraint, be it for Spanish–English or for other language pairs. A good deal of the work on bilingual discourse has made use of standard linguistic methodologies (e.g. informant elicitation, subjective reaction tests, introspection), yielding data which reflect indirectly (at best) actual language use in real-life situations. This is due in part to the unreliability of grammatical intuitions about bilingual syntax—a problem which is exacerbated in situations of social stigma—and in part to the fact that the study of isolated examples cannot distinguish between the systematic recurrent patterns of everyday interaction and other structures, albeit ‘acceptable’ in some sense, which never or rarely occur.

The work we discuss here, in contrast, has been carried out using a sociolinguistic, or more specifically, *variationist* approach to the study of code-switching, conforming to a number of principles which have come to be associated with this paradigm. These include: (1) the use of appropriate data, resulting from the study of language use in its natural context; (2) the principle of accountable reporting, which implies analysis of *all* of the relevant data; (3) the selection of informants to ensure representativeness, and the knowledge of what they represent; and (4) circumscription of the ‘variable context’, i.e. defining the object of study [cf., among others, Labov (1969, 1971), and G. Sankoff (1974)].

It is not our claim that the equivalence constraint is uniformly pertinent to every bilingual community, even to those in which mixing of the two codes is frequent at the intrasentential level. A key aspect of New York Puerto Rican linguistic behaviour is that intrasentential switching occurs in a smooth manner. Here, switches are usually *not* preceded or followed by hesitations or pauses, nor are switched items translations or repetitions of what went before. Contrary to some claims, no special local rhetorical effect is produced by the overwhelming majority of particular switches. The speaker does not draw attention to the fact that a language change has occurred, nor does the interlocutor have to acknowledge it, as illustrated in the examples in (2).

- (2) (a) So you *todavía* haven’t decided *lo que vas a hacer* next week (01/135)¹
 ‘So you still haven’t decided what you’re going to do next week.’
 (b) *Si tu eres puertorriqueño*, your father’s a Puerto Rican, you should at least *de vez en cuando*, you know, *hablar español*. (34/25)
 ‘If you’re Puerto Rican, your father’s a Puerto Rican, you should at least sometimes, you know, speak Spanish.’

While such smooth intrasentential switching is not confined to Puerto Ricans—indeed, it is also characteristic of fluently bilingual Mexican Americans (Pfaff, 1979)—it may very well be relatively rare in other communities which are equally bilingual. For example, in a more recent, large-scale study (Poplack, 1985/87), we examined the code-switching behaviour of 120 bilingual speakers of French and English in the Ottawa–Hull region of Canada. Since French is typologically very similar to Spanish, patterns of code-switching in the two communities could be expected to be similar if not identical, if the structural properties of the languages involved were a key determining factor. However, the type of code-switching used in Ottawa–Hull is dramatically different from that attested among the Puerto Ricans. Only a minuscule proportion of the French–English switches are genuinely intrasentential. Instead of weaving the two languages smoothly together at imperceptible switch points, French speakers draw attention to the switch by any one of a number of discourse devices—metalinguistic commentary, repetition or translation, English bracketing. Virtually every switch serves a rhetorical purpose, whereas, for the

Puerto Ricans, the smooth switching style is itself a speech mode emblematic of community identity. Indeed, in order for the switch to accomplish its purpose in the French community, it must be highlighted, i.e. salient, and should not pass unnoticed, as can be observed from the examples in (3). One by-product of this flagging is the interruption of the speech flow at the code-switch point, effectively circumventing a grammaticality requirement, or rendering one unnecessary.

- (3) (a) Je m'adresse en français, pis s'il dit "I'm sorry", ben là je recommence en anglais. (MM/3254)
'I begin in French and if he says, "I'm sorry", well then I start over in English.'
- (b) Mais je te gage par exemple que . . . excuse mon anglais, mais les odds sont là. (CD/716)
'But I bet you that . . . excuse my English, but the odds are there.'
- (c) Je suis un peu trop anglicisé, anglifié, *anglicized*. (GF/1361)
'I'm a little too anglicized, anglified, anglicized.'

Thus, while the linguistic configuration involved is typologically very similar in the two communities, as indeed are many aspects of their sociological situations, and, while they both make plentiful use of code-switching, the syntactic constraint obtaining in one is largely irrelevant to the other, because of the different discourse functions that code-switching serves in each.

It is thus apparent that not only is code-switching structured differently from community to community, but not all data on switching constitutes acceptable evidence regarding the operation of syntactic constraints. Ungrammaticality is indeed avoided in the French–English context, but this is only a trivial consequence of the flagging of switches. We return to the distinction between smooth and flagged code-switching in our discussion of Finnish–English bilingualism below.

TYOLOGICAL DIFFERENCES BETWEEN LANGUAGE PAIRS

The studies of Puerto Ricans and Franco-Ontarians (as, indeed, the majority of work on code-switching) have involved pairs of typologically similar languages. It has often been claimed that though the equivalence constraint is an intuitively natural way to account for code-switching behaviour in languages which already share many surface similarities, it founders in typologically different structures [e.g. Bentahila and Davies (1983), Prince and Pintzuk (1983), Bokamba (1988), and Nishimura (1985)]. The research we report here is part of a larger project, directed by D. Sankoff, examining code-switching phenomena in communities specifically chosen according to predetermined typological differences in the code-switched pairs.

At one level, the propensity toward code-switching (vs borrowing, for example) is typologically determined. We noted earlier that the Spanish–English case may be an extreme one. Code-switching is copious, transitions are smooth, and it occurs at all permissible code-switch boundaries, of which there are of course many, given the similarities between the languages. Major code-switching sites are between subject NP + VP, V and object NP, preposition + NP, internal to the NP, internal to the PP, around coordinate and subordinate conjunctions, etc. In typologically *different* languages, word order incongruence renders code-switching more problematic because the resulting code-switched sentences risk violating the patterns of one or both languages. In moving from the Spanish–English situation to one where English co-exists with a strongly SOV and/or case-marked language, for example, many permissible switch boundaries are lost.

We propose in this paper to test the validity of the equivalence constraint on the basis

of data on Finnish–English bilingualism, comparing a postpositional with a prepositional language, but where no other major word order difference obtains, as illustrated in (4).²

(4) Spanish (French)	English
SVO	SVO
PREP	PREP
N+ADJ	ADJ+N
Finnish	English
SVO	SVO
POST., CASE	PREP
ADJ+N	ADJ+N

Finnish, unlike English, has a rich system of obligatory morphological case-marking for each of 15 nominal cases. These are used to fulfill subject, direct object, genitive and partitive functions, and in expressions of location and direction, time, instrument, and manner [e.g. Karlsson (1983)], functions which are carried out in English by prepositions or word order. Given the analytic nature of English, the equivalence constraint makes strong predictions about where intrasentential code-switching within this language pair should not occur. If we nonetheless find English-origin material in the ‘prohibited’ switch sites, we may conclude that: (1) either the proposed constraint does not account adequately for the data, or (2) the data involved are not code-switches.

CIRCUMSCRIBING THE VARIABLE CONTEXT

Once we have established that speakers are indeed alternating between languages in a smooth, unflagged way, we must circumscribe the variable context, i.e. determine whether the other-language material under investigation in fact constitutes a code-switch, or, rather, represents some other manifestation of language contact. On the synchronic level, it is often impossible, in a given sentence, to tell whether a genuine switch has taken place, particularly at the level of single-word incorporations, which may be ambiguous as to their status as code-switch or loanword. Borrowing as a process differs radically from code-switching, and failure to separate data on the two phenomena can only obscure the conditioning of each.

The traditional characterization of an established loanword is an L₂ item, phonologically, morphologically and syntactically integrated into host-language discourse, which is both recurrent in the speech of an individual and widespread in the community. Thus, for the same Puerto Rican data, our working hypothesis was that an English-origin word in an otherwise Spanish context that did not satisfy these criteria would only occur in English monolingual discourse or in code-switches from Spanish to English. In general, however, borrowing is a much more productive process and is not bound by all of these constraints. In particular, the social characteristics of recurrence and dispersion need not be satisfied. This type of borrowing has been called *nonce* borrowing (Weinreich, 1953) [cf. also Poplack *et al.* (1988)]. Distinguishing loanwords from code-switches when this process is prevalent is even more delicate.

In this paper we address this problem by proposing and testing a consistent framework for distinguishing code-switching from borrowing, making use of quantitative distributional methods. We will show that the morphological and syntactic role of a nonce borrowing is equivalent to that of an established loanword, which is, in turn, identical to that of its host-language counterpart, and, in this, the two contrast with code-switching. It will be our claim that nonce borrowing is particularly productive in Finnish–English bilingualism,

outweighing code-switching by a factor of at least 5:1, and this may well be the case for typologically different language pairs in general.

DATA AND METHODS

In keeping with the variationist quest for appropriate data, the materials on which this study is based consist of spontaneous tape-recorded conversations between eight first-generation Finnish women and Westwood, a core in-group member of the local Finnish community. Data were collected using standard social network techniques. All informants are fluent native speakers of Finnish who migrated to Canada as adults, although their time of residence in the country varies from 13 to 55 years. Six are also highly proficient in English; indeed, for all but one, English is currently the preferred language in most domains of social interaction, with the notable exception of the friendship network in which Westwood is linked to them, whose meetings are at least partially motivated by the desire to speak Finnish and participate in Finnish culture.

The base language of these conversations is Finnish, but they contain liberal incorporations from English. All of the latter were extracted from the corpus and initially categorized according to whether or not they represented unambiguous code-switches, i.e. multi-word fragments which are lexically, syntactically and morphologically English, as in the italicized portions in (5).

- (5) Niin siellä oli tuota, kättilö joka oli *head of the district who has*
 so there was um midwife-n³ who was
not practiced for twenty years, and there she was sillä oli se vauva kädessä. (19b/326)
 it-ad. was it baby-n. arms-in.

'So there was um, the midwife who was head of the district who has not practiced for twenty years, and there she was, she had the baby in her arms.'

The overwhelming majority of English-origin material in the corpus, however, consists of single nouns, which we have said may equally well represent instances of borrowing as of code-switching. We may make use of the morphological criterion of obligatory case-marking in Finnish⁴ to determine whether these items are in fact behaving like their counterparts in that language, as would be expected if they were loanwords.⁵ Accordingly, we noted for each the case it would require were the sentence entirely in Finnish, and whether that case was in fact marked, as in (6). Not all of the nouns requiring inflection by Finnish rules actually received it, as may be seen in (7), and these were noted as well.

- (6) Misis K. oli *housekeeperina*. (9a/126)
 was -ess.

'Mrs. K. was the housekeeper'.

- (7) Mä laitoin oikein ison semmosen *aluminum pan-ø* lihapullia. (10a/368)
 I made-1p. really big-g. such-g. [g.] meatballs-p.

'I made a really big, like aluminum pan of meatballs.'

In keeping with the distinction drawn above between flagged and smooth code-switching, we also noted the presence of various discourse phenomena in the environment of the English-origin items which might indicate poor integration into host-language discourse. Included here were perceptible pauses, as in (8), and false starts, as in (9), preceding or following the noun in question, and the presence of material repeating, translating or explaining the English-origin form, as in (10).

- (8) Mitä sä haluat se . . . *lunchiks?* (16a/096)
 what you want-2p. it -tr.

'What do you want for . . . lunch?'

- (9) Kaksykyviis dollaria viikko **ruoka-** *room and board-ø.* (10a/204)
 Twenty five dollar-p. week-n. food-
 'Twenty-five dollars a week for food- room and board.'
- (10) Ne lapset tuli ja pyys **kirjoja,** *books.* (26a/409)
 they children-n. came and begged books-p.
 'The children came and begged for books, books.'

We also noted the occurrence of what we refer to as *flags*, forms preceding the borrowed word, which in some sense bracket or highlight it, thereby calling attention to its presence. Forms functioning in this way include *sellainen, semmo(i)nen, tuollainen, t(u)ommo(i)nen* 'such, like that', *tällainen, tämmö(i)nen* 'like this, of this sort', *niin kuin, kuin* 'like', as well as English *you know*.

Throughout the discourse the interlocutor frequently appeared to be ratifying the fact that the speaker had switched languages: to signal either comprehension, acknowledgement or acceptance of the fact that a loanword had been used. Interjections expressing agreement such as *joo, ya, niin, uh huh* were considered possible ratification markers.

Finally, although Finnish does not make use of articles *per se*, initial examination of our data suggested that demonstratives (e.g. *tämä* 'this', *tuo* 'that', *nämä* 'these', *nuo* 'those', *se* 'it', 'that', *ne* 'they'), declinable indefinite pronouns (e.g. *joku* 'someone' and *jokin* 'something') and the numeral *yksi* 'one' were being used much like English determiners, possibly to carry the inflection missing from the loanword, or to otherwise signal English usage [cf. also Lehtinen (1966: 175)]. An example of this use of Finnish 'determiners' may be found in (11).

- (11) Ne aloitti **sen** *union-ø.* (2a/231)
 They started it-g.
 'They started the union.'

Of course all of these markers can in principle occur just as well in monolingual Finnish discourse as in the environment of English-origin incorporations. To assess the extent to which the latter were actually behaving like their Finnish counterparts, we compare rates of inflection, flagging and determiner usage with those for an equal number ($N=803$) of native Finnish nouns in the corpus.

AN ACCOUNTABLE REPORT ON THE DATA

Table 1 summarizes the distribution of incorporations from English in the corpus.

Table 1. Distribution of English incorporations in Finnish discourse

Intrasentential code-switches	Tags/interjections	Single nouns ⁶
12.9% (154) ⁷	18.8% (225)	68.2% (813)

We note first that less than 20% of the incorporations from English are made up by tags, a freely movable category with no syntactic relation to the rest of the sentence, as in (12).

- (12) (a) Mutta en mä viittinyt, **no way!** (9b/134)
 But not-1p. I bothered
 'But I'm not bothered, no way!'
- (b) Jäi pois työstä. **Thank God!** (2a/232)
 left-she away job-el.
 'She left the job. Thank God!'

This is a relatively small number when compared with, for example, the New York Puerto Rican community, among whom tag-switching was found to be a device used by non-fluent bilinguals, allowing them to engage in the code-switching mode without having to respect a grammaticality requirement (Poplack, 1980). Here, however, tag-switching is by no means the domain of non-fluent bilinguals. For the least English-proficient speakers in the sample, such switches do not exceed 6% of the incorporations from English, proportions which are comparable to those found for the most proficient in that language (Wheeler, 1987). Indeed, it is clear that tag-switching is simply not a common strategy among these speakers.

We mentioned earlier that the equivalence constraint makes strong predictions as to where intrasentential code-switching between Finnish and English should *not* occur. Finnish is postpositional, case-marked, and makes little or no use of determiners. When juxtaposed with English, potential switch sites within what corresponds to an English PP, and, in most cases, to an NP (both found elsewhere to be favoured switch sites) are forfeited. Thus, it is not surprising that intrasentential code-switching is also infrequent, not exceeding 13% of the data. We first examine the relatively clear cases of switching involving multi-word fragments of English, including English function words where appropriate and never containing Finnish function words or morphology (Table 2).

Table 2. Distribution of unambiguous code-switch types

	<i>N</i>	%
COMP + S	95	62
V + (OBJECT) NP (SUBJECT)	25	16
NP + VP	10	6
DEM + N	9	6
ADJ/ADV + N	11	7
V + INFINITIVE	4	3
Total	154	100

The largest fraction consists of English sentences embedded by means of the complementizer *että* within a Finnish matrix sentence, as in (13). The next largest group involves English material in post-verbal NPs in object position, as in (14). The remainder all occur at syntactic boundaries where English and Finnish show congruent word order as well: e.g. between subject NP and VP, as in (15), etc.

- (13) Sano että tulla tänne että *I'm very sick.* (16b/412)
 say that come-inf. here that
 'Tell [them] to come here, that I'm very sick.'

- (14) Mun vanhin on- nyt alkaa *part-time nursing in intensive care.* (10b/237)
 my oldest is now starts
 'My oldest is- is now starting part-time nursing in intensive care.'

- (15) Sen tuota titteli *was abolished.* (2a/236)
 it-g. um title-n.
 'Her title was abolished.'

On the other hand, switches at sites prohibited by the equivalence constraint do not occur. Thus, there are no switches: (1) involving negative particles, (2) following English prepositions, (3) preceding Finnish postpositions, (4) following Finnish preposed, question-marked verbs, or (5) in the vicinity of inverted English WH-interrogative structures. We have seen from Table 1, however, that only a small proportion of the English-origin material

is contained in intrasentential switches. The overwhelming majority (68%) consists of single nouns or compounds within otherwise entirely Finnish utterances, largely occurring in precisely the contexts proscribed by the equivalence constraint. This leads us to question whether they are in fact code-switches, or some other manifestation of language contact.

NOMINALS

We have given an informal account of the equivalence constraint. This constraint is difficult to formalize, but this has recently been done mathematically by Sankoff and Mainville (1986a, b) for a particular class of grammars. What are the consequences of this formalization for the present discussion?

We first illustrate the predictions of the equivalence constraint with potential switches involving an English preposition and a case-marked noun (of either native Finnish or English origin), organized according to the four principles in (16). For the moment we set aside the limited class of Finnish prepositions and ignore the presence or absence of the determiner.

(16) Predictions of principle 1: *no crossovers*

- (a) *E prep + F noun
- (b) *E **noun** + F case-marker
- (c) *F noun + E prep
- (d) *F case-marker + E noun

Predictions of principle 2: *monolingual grammaticality*

- (a) *F case-marker + F noun
- (b) *E noun + E prep

Predictions of principle 3: *no omissions*

- (a) ***lone E noun** (no E prep, no F case-marker)
- (b) *bare F noun (no E prep, no F case-marker)

Predictions of principle 4: *no repetitions*

- (a) *E prep + F noun + F case-marker
- (b) *E prep + E noun + F case-marker

The predictions of principle 1 are that no English preposition should precede or follow a Finnish noun, nor should a Finnish case-marker precede or follow an English noun. Principle 2 predicts the obvious restrictions against a Finnish case-marker followed by a Finnish noun or an English noun followed by an English preposition. Principle 3 proscribes a lone English noun (i.e. with no preposition), or a bare (i.e. non-case-marked) Finnish noun, and principle 4 excludes any combination of English preposition and Finnish case-marker in the vicinity of the noun, i.e. 'portmanteau' constructions, or copy translations.⁸

All of these predictions are verified in our corpus, except for 1(b) and 3(a) [given in bold face in (16)], where we find massive exceptions involving English nouns in ungrammatical constructions according to English rules, either with or without Finnish case-marking. We return to these exceptions below. First we examine another set of constructions corresponding to English NPs which would require an article according to English grammar.⁹ What does the equivalence constraint predict for these structures (17)?

(17) Predictions of principle 1: *no crossovers*

- (a) *E det + F noun
- (b) *E **noun** + F case-marker

Predictions of principle 2: *monolingual grammaticality*

- (a) *E noun + E det
- (b) *F case marker + F noun

Predictions of principle 3: *no omissions*

- (a) ***lone E noun** (no E det, no F case-marker)
- (b) ***bare F noun** (no E det, no F case-marker)

Predictions of principle 4: *no repetitions*

- (a) *E det + F noun + F case-marker
- (b) *E det + E noun + F case-marker

Principle 1 prohibits a switch between an English determiner and a Finnish noun; principle 2, a series of internally ungrammatical combinations; principle 3, a lone English noun (i.e. without a determiner) or a non-case-marked Finnish noun; and principle 4, copy translations.

How do these predictions actually hold up in our Finnish–English bilingual data? Again we find the same sort of massive exceptions as in the previous analysis [1(b) and 3(a), namely English-origin nouns with Finnish case-markers as in (18), and English nouns with no case-marker but without the required article (19).

- (18) Näitä kaks, kolme *bypassia* sillä on. (11a/399)
 these-p. two three -p. he-ad. is
 ‘He has two, three bypasses.’

- (19) Mä sanoin että mä menen *interview-ø*. (10a/134)
 I said-1p. that I go-1p.
 ‘I said that I’m going to an interview.’

In sum, a large proportion of the English-origin material in our corpus apparently cannot be accounted for by the equivalence constraint. The fact that most of these words carry the correct Finnish case-marking suggests, however, that they are not code-switches at all, but result from nonce borrowing, a process which (unlike the relatively restricted set of established borrowings) applies to the entire English nominal lexicon.

Recall, however, that Finnish case-marking is *obligatory*, and indeed, appears categorically on all of the native Finnish nouns in our corpus. Yet a good proportion of the English nouns under discussion (21%) do not carry this marking. A recent analysis of nonce borrowing from English into Tamil (Sankoff *et al.*, 1986), an SOV language which is also postpositional with case-marking, but where case-marking is *optional*, showed that the distribution of case-marking on English-origin nouns quantitatively paralleled that for native Tamil nouns. How can the discrepancy between the Finnish- and English-origin nouns be explained?

Variability in case-marking of both Finnish (Lehtinen, 1966; Larmouth, 1974) and English (Lehtinen, 1966) nouns is not unknown amongst transplanted Finns, and has been ascribed to contact and convergence with English. In an attempt to explain this phenomenon in the Finnish of three successive generations of Minnesota-born Finns, Larmouth hypothesized that case-markers were most likely to be deleted in contexts where: (1) they are least strongly bound to the noun, and (2) they have the greatest degree of semantic transparency. Table 3 compares the rates of case-marking on English-origin nouns by case in our corpus with Larmouth’s characterization of their degree of boundness.

Despite sparse data in some of the contexts, the overall distribution of case-marking in our material appears to support his model. Moreover, the partitive, arguably the most semantically transparent of the Finnish cases, can be seen to be marked less than any other case occurring frequently enough to rule out statistical fluctuation. Compare this with the local [marked (L) in Table 3] cases, whose omission would likely involve misinterpretation on the part of the hearer, and which all show moderate to high rates of marker presence. However, when we examine the ranking of case-marking for each individual in the sample,

Table 3. Rates of case-marking of English-origin nouns by case, compared to Larmouth's characterization of their degree of boundness

Case	Degree of boundness (Larmouth, 1974)	% case-marking	N
Nominative	Strong	100 ¹⁰	370
'True' genitive	Strong	94	16
Ablative (L)	Strong	Sparse data	1
Elicative (L)	Moderate	88	17
Illative (L)	Moderate	72	61
Adessive (L)	Moderate	10	20
Inessive (L)	Moderate	68	121
Partitive	Weak	46	134
Allative (L)	Weak	33	6
Accusative	Weak	n/a ¹¹	—
Translative	—	36	11
Essive	—	33	6

we find that no two speakers manifest an identical or even similar ordering. The wide intra-individual variability casts doubt on Larmouth's explanation.

Moreover, Larmouth's hypothesis does not explain why this variability should affect only the *English-origin* nouns in our corpus. Let us now examine these data in more detail. We note first that the presence of English-origin material in Finnish sentences tends to be associated with an unusually high rate of certain discourse phenomena. For example, there is a good deal of what we have referred to as functional flagging, as in the French-English code-switching described earlier. Many English-origin nouns are preceded by discourse markers such as *semmonen* 'such' and *niin kuin* 'like', as in (20), which in some instances seem to be entirely confined to a signalling function.

- (20) (a) Hän ois **niin kuin** *programmeri*. (22b/097)
 he would be like *programmer* -n.
 'He would be like a programmer.'
- (b) Rupesin pitämään **semmosta** *rooming housea*. (15b/136)
 started-1p. to keep such-p. -p.
 'I started to keep a rooming house.'
- (c) Ne on ihan **niin kuin semmosia** *temper tantrum*. (27b/261)
 they is quite like such-pl.-p.
 'They are quite like, like temper tantrums.'

This may be seen in Fig. 1, which shows that English nouns occur far more frequently after flags than do their Finnish counterparts, and this regardless of the case of the noun.

Another discourse phenomenon we analysed was the use of ratification markers (e.g. *yes, uh huh, joo, niin*) on the part of the interlocutor. These were of course quite frequent throughout the conversations amongst all participants, but especially so immediately following English nouns. Figure 2 shows that in all cases, English-origin material occurs more frequently than native Finnish in the environment of these ratification markers.

Indeed, it is apparent from Fig. 3 that, of all the discourse phenomena examined, at least three seem to be clearly associated with the presence of English material: a perceptible preceding pause, a ratification marker on the part of the interlocutor, and a flag.

We now return to the massive exceptions to the predictions of the equivalence constraint outlined above. We first examine those constituents in our data which correspond to English prepositional phrases and in which the noun or noun phrase is of English origin. There

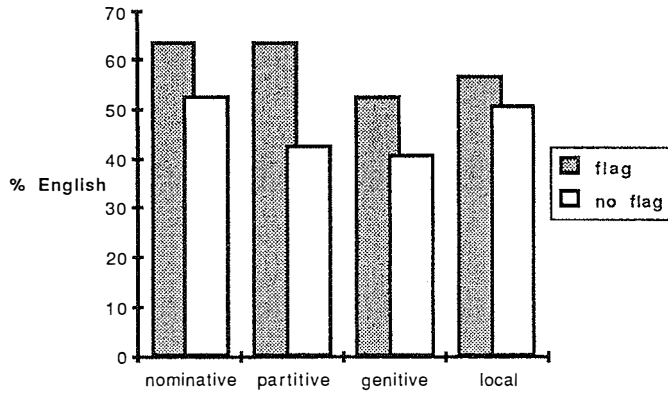


Fig. 1. Percentage of nouns which are English, in the presence or absence of a flag.

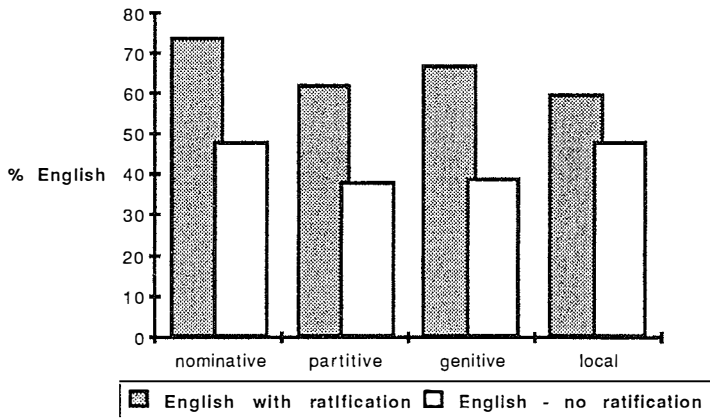


Fig. 2. Percentage of nouns which are marked, by case, in the environment of ratification.

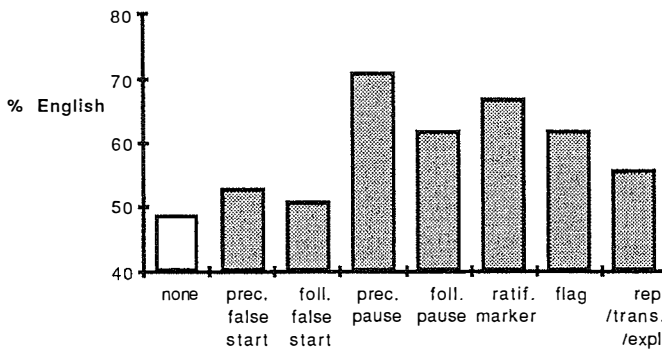


Fig. 3. Percentage of nouns which are English in the environment of various discourse features.

are 246 of these which would require a preposition if the construction were entirely in English. Of these, 10 (4%) were of the type which could be, and indeed were, accompanied by a member of the limited class of Finnish prepositions, as in (21), or a postposition, as in (22).

- (21) Jos mä lasken sen ilman contractia tänne sisälle . . . (11b/172)
 if I let-1p. it-g. without -p here in-al.
 'If I let him in here without a contract . . .'
- (22) Kaurapuuron saa haggiksen sisään työntää. (25b/242)
 porridge-g. can -g. into stuff-inf.
 'Porridge can be stuffed inside the haggis.'

These were all case-marked and can be considered nonce loans. The distribution of the remaining 236 cases is summarized in Table 4.

Table 4. Relationship between flagging and case-marking of English-origin nouns requiring a preposition in English

	+ case marker	- case marker
% flag/pause/ratification marker	40 (63/158)	68 (53/78)

Two-thirds of these (158/236) are case-marked nonce-loans [as in (23)], leaving 78 analytically problematic instances.

- (23) Mä kerran lähetin sen tuonne dry cleaneriin. (29b/397)
 I once sent-1p. it-g. there-al -il.
 'I once sent it to the dry cleaners there.'

Note, however, that these tend to be problematic for the speaker and/or hearer as well, since there is a dramatic increase (to 68%) in the rate of discourse marking of the caseless nouns. Even the 10% of remaining nouns which are neither case-marked nor flagged are mostly accompanied by some other discourse marker than those counted in Table 4 (e.g. a following false start or pause). We may thus treat the set of caseless nouns as either flagged, non-smooth single-word switches, or, in some instances, as poorly integrated nonce borrowings.

While the equivalence constraint correctly predicts the absence of switches after an English preposition, it also predicts that there *should* be some switches before prepositional phrases. In comparing the behaviour of multi-word English sentence fragments (i.e. unambiguous code-switches) with that of nonce loans in this regard (Table 5), we find that among the former, there are 25 instances of constructions which require an initial preposition at the switch point, and this is actually present in two-thirds of the cases. Any missing prepositions tend to be flagged, suggesting some difficulty in constructing the switch boundary, rather than speakers' lack of English competence, as in (24).

- (24) Se oli kai neljä vai oliko se viis vuotta sitta siellä kuin
 She was maybe four or was-int. it five year-p. then there like
 nursing home-ø. (16b/145)
 [in.]
 'She was maybe four or was she five years there like [in a] nursing home.'

Prepositions which are required *internal* to English discourse, in contrast, are virtually always present, while in the vicinity of nonce loans, they are categorically absent.

Table 5. Distribution of English prepositions before nonce loans, at code-switch boundaries and within code-switched fragments

	Before nonce loans	At switch boundary	Within switch
Presence of preposition	0% (0/236)	64% (16/25)	98% (39/490)

We may now examine the nominal constructions which would require a determiner were they entirely in English. No such determiners actually occurred, except as required internal to multi-word English fragments. Table 6 shows the distribution of case-marking and discourse flagging of the single noun items.¹²

Table 6. Relationship between flagging and case-marking of English-origin nouns requiring a determiner in English

	+ case marker	- case marker
% flag/pause/ratification marker	38 (63/168)	66 (50/75)

This enlarged set of nouns shows the same increased discourse flagging of caseless items, confirming our view that most are nonce borrowings. There is some reason to believe that many of the remainder should be considered as code-switches.

Of the small number of nouns which are neither flagged nor case-marked, 68% (17/25) are preceded by what we have called a Finnish 'determiner' [e.g. demonstratives *se* 'it/this', *tämä* 'this', *tuo* 'that', and (in the local cases), the adverbs of place *siellä*, *tuolla* 'there' and *täällä* 'here']. We see from Fig. 4 that these are particularly frequent in the environment of English nouns, where in many instances they may no longer convey a specific marked demonstrative (or locative) value, but simply one of definiteness [cf. also Lehtinen (1966)].

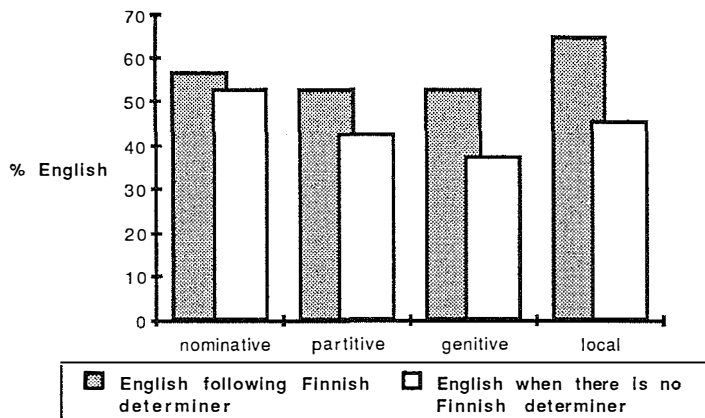


Fig. 4. Use of Finnish determiners preceding English nouns.

Figure 4 shows that, in all cases, English nouns occur more frequently following a Finnish determiner than not. Moreover, the nouns following these determiners are less likely to be marked (in all cases but the nominative) than those which appear without a determiner (Fig. 5).¹³

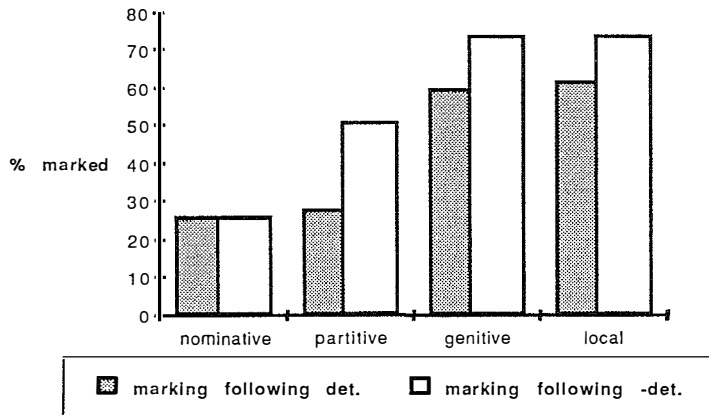


Fig. 5. Rates of case-marking on English-origin nominals following a Finnish determiner.

This leads us to suggest that many of these tokens are in fact code-switches and not borrowings, and that the determiners are serving a flagging or switch-signalling function.

As previously, we now compare this with the behaviour of determiners in unambiguous English sequences, as in Table 7.

Table 7. Distribution of English determiners before nonce loans, switch boundaries and within code-switched fragments

	Before nonce loans	At switch boundary	Within English fragments
Presence of determiner			
Finnish	42% (101/243)	48% (14/29)	0% (0/56)
English	0% (0/243)	17% (5/29)	86% (48/56)

Nonce loans are accompanied by a Finnish determiner less than half the time, with the proportion of determiner presence rising to 66% when a noun requiring a determiner by English rules occurs at a switch boundary. Of course, none of the latter are case-marked. The determiner may either be English (17%) or Finnish (48%), depending on the switch site (between preceding category and noun phrase or between determiner and noun phrase). However, when English noun phrases occur *internal* to English discourse, a full 86% (48/56) of those requiring English determiners have them, in accordance with English grammar.

It must be stressed that all of the empirically studied communities to which we have referred manifest very different patterns of bilingual behaviour, despite the fact that in each we can attest at least some smooth code-switching, at least some nonce borrowing, and at least some flagging. Nevertheless, it would be injudicious to consider the highly-flagged French–English material as the optimum data for the study of smooth code-switching. Nor is the Spanish–English situation, with its ease of switching, the most suitable context in which to analyse nonce borrowing.

However, once we have established that the social role of language mixing is propitious to the smooth integration of elements of both codes, typological considerations are predictive of the types of mixture. Similar typologies are conducive to code-switching, while conflicting typologies are more likely to result in nonce borrowing. And since all these phenomena occur to at least some degree in all the communities, the overwhelming nature of these patterns could never have been distinguished without systematic quantitative analysis of natural performance data.

Acknowledgements—The generous support of the Social Science and Humanities Research Council of Canada to Poplack for the project ‘Sociolinguistic aspects of language contact in the Ottawa–Hull region’, and to D. Sankoff for the project ‘Code-switching constraints in typologically different language pairs’ is gratefully acknowledged. Earlier versions of parts of this material were presented at NWAVE XIII (1984), the IXth Meeting of the American Association for Applied Linguistics (1986), and the VIth International Conference of Nordic and General Linguistics (1986). This paper first appeared in the Proceedings of the latter (*The Nordic Languages and Modern Linguistics* 6. Edited by P. Lilius and M. Saari (1987). Helsinki: Helsinki University Press. pp. 33–56], and in *Bilingualism across the Life Span*. Edited by K. Hyltenstam and L. Obler. Cambridge: Cambridge University Press. pp. 132–154. We thank David Sankoff for his contributions to the analysis and interpretation of the materials, Doug Walker for careful comments on the manuscript, and Marguerite Trudel-Maggiore for her expert preparation of the text.

NOTES

- Codes identify speaker and utterance.
- Word order is actually relatively free in Finnish, variation serving mainly to alter the focus of the sentence. The unmarked order, however, is SVO.
- The following abbreviations are used throughout in the glosses of Finnish cases: n. = nominative, g. = genitive, p. = partitive, in. = inessive, el. = elative, il. = illative, ad. = adessive, ab. = ablative, al. = allative, tr. = translative, ess. = essive, \emptyset = zero morpheme, int. = interrogative, pl. = plural, inf. = infinitive. Grammatical persons are designated 1p., 2p. etc.
- In all cases but the nominative singular, which receives a null mark.
- Readers familiar with the *free morpheme* constraint (Poplack, 1980, 1981), which prohibits a *code-switch* between a bound and a free morpheme of different languages will note that nonce borrowing occurs *without* any change of code on the morphological and syntactic levels.
- The only other single lexical items of English origin found in the corpus were a small number (< 7%) of adjectives and verbs (< 2%), which we do not include in this analysis.
- An additional 163 occurrences originally coded as intrasentential code-switches were in fact examples of reported direct speech, as in (i).
(i) Lääkäri sano että “when you go to Finland, no funerals”, ja mä sanoin “Okay, no funerals”.
(291/065)
doctor-n. said-3p. that and I said-1p.
‘The doctor said “when you go to Finland, no funerals”, and I said, “Okay, no funerals”.’
These all satisfy the equivalence constraint since both languages have the same structure for reporting speech. Switching for this purpose is a well-documented discourse device [e.g. Gumperz (1982) and Poplack (1985)] and, indeed, accounts for the majority of the switches used by some of the speakers in our sample.
- Such examples are in fact attested, but, *contra* Nishimura (1985), they are exceedingly rare. We found two cases in our data:
Mutta se oli *kidneystä to aortaan*. (11a/337)
but it was -el. -il.
‘But it was from the kidney to the aorta (to).’

Ja sitten . . . uh missä hän n- -at yliopistossa otti . . . niin kuin *art history*. (11a/252)

and then where she n- university-in. took-3p. like

'And then . . . uh where did she- at university (at) she took . . . like art history.'

9. Note that these data contain, as a subset, the NPs contained in the structures corresponding to English prepositional phrases discussed above, but include, in addition, many others corresponding to English subjects and direct objects which also require case-marking according to Finnish rules.
10. For the nominative singular, which in standard Finnish takes a null mark, we considered bare English nouns to be 'marked' for case, i.e. to follow native-language patterns. Of course, it has been widely reported [e.g. Lehtinen (1966), Karttunen (1977), Martin (1981), and Virtaranta (1981)] that many consonant-final nouns borrowed into Finnish take a vocalic ending (e.g. *jobi* 'job', *kaara* 'car'), which serves as a stem for further case-marking. However, since consonant-final nouns occur natively in Finnish with no such vowel addition, we are reluctant to consider the 72% of consonant-final English-origin singular nominal stems in our data as failing to show obligatory case-marking, particularly since this rate far exceeds the incidence of case-absence found for any other case (Table 3). Such an inflated rate would be particularly curious in view of the fact that the nominative is by far the most frequently used case in the data. We suggest that the optional presence of the stem vowel serves as a phonological marker of loanword integration. This would explain its relatively sparse occurrence in our data, which contain both established and nonce borrowings.
11. Because our data were analysed according to morphological case rather than syntactic function, the figures for the accusative case are contained in the calculations for the nominative and genitive, and thus cannot readily be brought to bear on Larmouth's claim.
12. Nominative contexts have been excluded from this calculation.
13. 'Marking' for the nominative case in this instance refers to the previously mentioned addition of the stem vowel to consonant-final English nouns. The fact that there is no difference between marking rates in the two contexts, in contrast to the other cases, lends support to our suggestion (note 10) that addition of the stem vowel is a phonological integration device rather than specifically a marker of case.

REFERENCES

- Bentahila, Abdelali and Davies, Eirlys (1983) The syntax of Arabic-French code-switching. *Lingua*, **59**, 301-330.
- Bokamba, Eyamba (1988) Code-mixing, language variation, and linguistic theory: evidence from Bantu languages. *Lingua*, **76**, 21-62.
- Gumperz, John (1982) Conversational code-switching. In *Discourse Strategies*. Edited by J. Gumperz. Cambridge: Cambridge University Press. pp. 59-99.
- Karlsson, Fred (1983) *Finnish Grammar*. Helsinki: Werner Söderström Osakeyhtiö.
- Karttunen, Frances (1977) Finnish in America: a case study in mono-generational language change. in *Sociocultural Dimensions of Language Change*. Edited by B. G. Blount and M. Sanches. New York: Academic Press. pp. 173-184.
- Labov, William (1969) Contraction, deletion and inherent variability of the English copula. *Language*, **45**, 715-762.
- Labov, William (1971) Some principles of linguistic methodology. *LiS*, **1**, 97-120.
- Larmouth, Donald (1974) Differential interference in American Finnish cases. *Language*, **50**, 356-366.
- Lehtinen, Meri (1966) Analysis of a Finnish-English bilingual corpus. Ph.D. dissertation, Indiana University.
- Martin, Marja-Liisa (1981) Finnish as a means of communication in Thunder Bay. In *Finnish Diaspora 1: Canada, South America, Africa, Australia and Sweden. Papers of the Finn Forum Conference, Toronto, Nov. 1979*. Edited by M. G. Karni. Toronto: Multicultural History Society of Ontario. pp. 117-130.
- Nishimura, Miwa (1985) Intrasentential code-switching in Japanese and English. Ph.D. dissertation, University of Pennsylvania.
- Pfaff, Carol W. (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 (1981) Syntactic structure and social function of code-switching. In *Latino Discourse and Communicative Behavior*. Edited by Richard Duran. Princeton, NJ: Ablex. pp. 169-184.
- Poplack, Shana (1985/87) Contrasting patterns of code-switching in two communities. In *Aspects of Multilingualism*. Edited by Erling Wande *et al.* Uppsala: Borgström. pp. 51-77.
- Poplack, Shana, Sankoff, David and Miller, Christopher (1988) The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics*, **26**, 47-104.
- Prince, Ellen F. and Pintzuk, Susan (1983) Bilingual code-switching and the open/closed class distinction. Ms.
- Sankoff, David and Mainville, Sylvie (1986a) Un modèle de l'alternance de langue sous la contrainte d'équivalence. *RQL*, **15**, 233-246.
- Sankoff, David and Mainville Sylvie (1986b) Code-switching of context-free grammars. *TL*, **13**, 75-90.

- Sankoff, David and Poplack, Shana (1981) A formal grammar for code-switching. *Papers in Linguistics: International Journal of Human Communication*, **14**, 3–46.
- Sankoff, David, Poplack, Shana and Vanniarajan, Swathi (1986) The case of the nonce loan in Tamil. Centre de recherche de mathématiques appliquées, Technical Report No. 1348. University of Montreal, Montreal.
- Sankoff, Gillian (1974) A quantitative paradigm for the study of communicative competence. In *Explorations in the Ethnography of Speaking*. Edited by R. Bauman and J. Sherzer. Cambridge: Cambridge University Press. pp. 18–49.
- Virtaranta, Pertti (1981) Finnish dialects in America: some experiences and problems. In *Finnish Diaspora II: United States. Papers of the Finn Forum Conference, Toronto, Nov. 1979*. Edited by M. G. Karni. Toronto: Multicultural History Society of Ontario. pp. 303–316.
- Weinreich, Uriel (1953) *Languages in Contact*. The Hague: Mouton.
- Wheeler, Susan (1987) Code-switching and borrowing in a Finnish–English bilingual situation. M.A. thesis. University of Ottawa.