Pekka Hirvonen and Timo Lauttamus

Code-switching and Language Attrition: Evidence from American Finnish Interview Speech[1]

1. Introduction

This article deals with code-switching and borrowing between Finnish, an agglutinative language that relies heavily on inflectional morphology, and English, a highly analytic language. The focus of attention in the present article is on code-switching and borrowing behaviour across three generations of Finnish Americans, and we shall attempt to discover if code-switching in particular can be used as an indicator of language attrition and language shift among these three generations of speakers (cf. Halmari 1993a). Since the typological distance between the two languages has an impact on the nature of code-switching and borrowing in a contact situation, it needs to be taken into account in defining the distinctions between different types of code-switching and between code-switching and borrowing (Lauttamus 1990:48).

The present work is based on four theoretical assumptions. First, borrowing and code-switching as language contact phenomena can only be accounted for in terms of a holistic model which incorporates not only structural linguistic factors but also various psycholinguistic, sociolinguistic and pragmatic factors (cf. Romaine 1995: 121-122). Second, code-switching and borrowing should be seen as two opposite poles on a *structural* linguistic continuum, although they can be regarded as different processes from the functional point of view (Lauttamus 1999: 87). In particular, their structural realizations should be described as gradient categories rather than as discrete ones from the synchronic point of view (Lauttamus 1990; 1991; 1992; 1999; Andersson 1993). Third,

[1] The authors gratefully acknowledge the contribution of Dr. Lisa Lena Opas-Hänninen of the University of Joensuu towards the completion of this article. Not only did she help and advice us in all the statistical procedures employed but she carried out the boxplot analyses, produced the boxplots (Figures 1-5) and explained to us how to interpret them.

Thomason & Kaufman's (1988) fundamental division of language contact situations into *language maintenance* and *language shift*, characterized by different kinds of contact-induced change, viz. interference through borrowing and interference through shift, is also synchronically applicable to the contact situations of the different generations of Finnish Americans. Fourth, the choice of code-switching strategies is associated with the bilingual speaker's proficiency in the two languages and thus reflects the speaker's stage in language attrition (Halmari 1993a). We shall briefly discuss each of the four assumptions.

2. Code-switching and borrowing

Given that the structural features and degree of integration into the recipient language are used as critical parameters in the analysis of code-switching and borrowing, evidence from Finnish-English bilingualism in North America supports the division of the corresponding switch and loan types into four (non-discrete) categories: (a) *code-change* and (b) *code-mix* on the one hand, and (c) *integrated loan* and (d) *adapted loan* on the other hand. We will also suggest that these switch and loan types are best regarded as representing categories (prototypes) which have (more or less) invariant cores but indeterminate, or "fuzzy", boundaries.

Adapting the model proposed in Lauttamus (1990; 1991; 1999), table 1 shows how the two processes, code-switching and borrowing, should be regarded as the opposite poles on a (structural) linguistic gradient running from *code-changes* to fully *adapted loans*. On the one hand, code-switching, as Poplack (1980: 583) suggests, is "the alternation of two languages within a single discourse, sentence or constituent". This definition implies that code-switching can take place not only intersententially or intrasententially but also within a single constituent. In addition, it suggests that there are two grammars **sequentially** operational on a given structure. On the other hand, borrowing refers to a process whereby "some lexical and/or structural property is integrated into a language (RL) from another language (SL)" (Lauttamus 1991: 40). The term *loan* is here used to refer to those lexical items where both form and meaning are borrowed from the source language (SL) with at least some integration into or adaptation to the morphosyntactic and phonological system of the recipient language (RL). Within Van Coetsem's (1988: 9) framework *adaptation* should not be confused with *integration*: "adaptation is an adjustment to the native *rl* which does not modify that language",

whereas integration is "incorporation into the native *rl* of something that modifies that language". This distinction can be exemplified by the English word *stove*, which has (at least) two variants in American Finnish: *toovi* and *stouvi*. The former follows the phonological pattern of the Finnish vernacular in that it does not allow consonant clusters in native words in initial position (adaptation), whereas the latter modifies the RL phonological pattern by retaining the SL consonant cluster. Van Coetsem (1995: 79) also points out that even *integratedness* is a continuum: a less integrated element (such as *stouvi*) may become a more integrated one (such as *toovi*), closing the gap between an integrated loan and an adapted one. With adaptation the RL thus preserves its existing phonological structure, and it is on this distinction that we have based our two (non-discrete) categories of *loan* (Lauttamus 1999: 91). Table 1 also illustrates how the notion of *operational grammar* can be used to describe which of the two grammars, the SL or the RL grammar, is operational on each linguistic category.

Table 1. A model for the description of code-switching and borrowing (Lauttamus 1999: 94).

CODE-SWITCHING		***BORROWING***	
OPERATIONAL GRAMMAR			
CODE-CHANGE	CODE-MIX	INTEGRATED LOAN	ADAPTED LOAN
SL	SL ~ RL	RL ~ SL	RL

As table 1 suggests, both code-switching and borrowing are used as "cover" terms for *code-change/code-mix* and *integrated loan/adapted loan* respectively. The section at the bottom of table 1 shows how the intermediate space, covering the categories *code-mix* and *integrated loan*, is characterized by interaction of the two grammars: in code-mixes it is mainly the SL grammar that is operational on the mixed item, while in integrated ('nonce') loans it is mainly the RL grammar which operates. It seems that in the Finnish-English bilingual setting, *morphology* is the most universal indicator of the degree of grammatical integration as far as code-

mixing and nonce borrowing[2] are concerned. Morphological integration is not, however, recognized as a good criterion for distinguishing borrowing from code-switching by all researchers (cf. Romaine 1995: 144). Halmari (1997: 70), for one, regards an example such as (1) as a *code-switch*:

(1) Mää oon sii+nä *green costum*+i+ssa.
I am it+INE +i+INE
'I am in that green costume.'

The SL (English) phonology operates on the "switched" elements (in italics), apart from the Finnish stem formant /i/ -- which facilitates pronunciation -- and the Finnish case (inessive) morpheme {ssA}, which are added to the otherwise unintegrated English stem {costum-}. Halmari (1997: 179) also suggests another reason for considering example (1) a switch rather than a loan: it shows phonological **un**assimilation. Cases such as these, which clearly lend support to the idea of a structural linguistic continuum, are treated as *integrated* ('nonce') *loans* in the present article. Both grammars are operational in an integrated loan, RL Finnish morphology and SL English phonology, contributing to the final product *siinä green costumissa* (NP), as proposed in table 1. The fact that the premodifier *green* is not inflected, as it would be in Standard Finnish (*greenissä*), suggests that *green costumissa* may have been processed as a single unit.

In terms of the approach advocated in the present article, the distinctions can be made as exemplified below. On the one hand, (i) **code-change**, as in (2) to (3), can be distinguished from (ii) **code-mix**, as in (4) to (5), the symbol '+' indicating an audible pause:

(2) Ja se oli se nuorempiki veli +
And it was it+DET younger+CLIT *kin* brother

ei kaikista nuorin veli + se asuu tuol'
not all+ELA youngest brother he lives that+ADE

leikilla kans'+ Heart Lake'lla, joka onää + niinku
lake+ADE also Heart Lake+ADE who is er like
half-brother they sa- say or whatever you want call it+
'And there was also this younger brother, not the youngest of them all,

[2] 'Nonce' borrowing usually "involves the use of single lexical items which are syntactically and morphologically, but not always phonologically integrated" into the RL (Romaine 1995: 153)

he also lives by that lake, Heart Lake, who is er like ...'

(3) ja sitte saksaa, kun Saksassa oltii
and then German+PAR when (we) Germany+INE were
army of occupation after the war was over. Joo.
'and then German, when we were in Germany as an ... Yes.'

Examples such as these clearly support the view that code-changes are, in general, multi-word fragments (mostly clauses or whole phrases), which follow the lexical, phonological and morphosyntactic rules of the SL. Accordingly, code-changes are not integrated into the RL but the *sl* grammar operates on them. Examples (4) and (5) represent the category of **code-mix** (a '0' indicates that the item violates the Finnish case-assignment rule):

(4) Ja suomalaiset, niil' oli paha nimi [laughs] siinä,
And the Finns they+ADE was a bad name it+INE

siihen aikaan, ne joutu *black list*+0
it+DET, ILL time+ILL they got black list
'And the Finns, they had a bad reputation in, at that time, they were "blacklisted."'

(5) Mutta sitte, kun jäin *pension*+0, niin ny, tuli
But then when I+got pension so now became

niin *busy* että se piti myydä.
so busy that it had to be sold
'But then, after I got on *pension,* I became so busy that I had to sell it.']

It should be noted that Poplack *et al.* (1987: 51) regard cases such as (5) as code-switches, characterized by a "total lack of inflection on nouns". The NP *black list* in (4) and *pension* in (5), as well as the AP head *busy* in (5), would normally require inflection in Finnish – the allative of *black list* 'mustalle listalle' and *pension* 'eläkkeelle', instead of the nominative 'musta lista' and 'eläke'. However, the case with the adjective *busy* is more controversial. We normally use the nominative case of the noun *kiire* 'hurry' instead of the adjective *kiireinen* (eg. *meille tuli kiire* 'to us came a hurry'), so the base form *busy* is what we can actually expect. The item *busy* could therefore be placed on the border-line between a code-mix and an integrated loan. The lack of obligatory morphological inflection for *black list* and *pension* indicates that these items are not in agreement with

the Finnish case-assignment rule and should therefore be considered code-mixes rather than integrated loans. The evidence reported in Poplack *et al.* (1987), Pietilä (1989) and Lauttamus (1990) shows that most cases of code-mix (and those of integrated loan for that matter) involve single lexical items (nouns).

There are also some other cases which clearly support the structural continuum. This is exemplified in (6), which represents one single turn (case ending indicated in bold type):

(6) Joo, sillon ei ollu vielä *freeway*+0 niin, sitte,
Yeah, then not was yet freeway well then

sen jälkeen, sillon ne rupes rakentaan *freeway**tä***.
it+GEN after then they began build+INF freeway
'Yeah, then there was no freeway, yeah, then, after that, then they began building the freeway.'

The first item *freeway* (6) shows no obligatory case-marking (PAR), whereas the second item *freewaytä* follows the Finnish morphological rule. In the approach proposed here the first occurrence is analyzed as a code-mix and the second one as an integrated loan. It seems that the speaker is able to move along on the switching – borrowing scale until the item gradually consolidates itself. Halmari (1997: 49) regards a word such as *freeway* as a borrowing, because it has no good Finnish counterpart. She argues, however, that otherwise the determining factor which differentiates a code-switch from a loan is, in fact, phonological unassimilation, instead of morphological unassimilation (p. 179). In this view, both items in (6) should be regarded as 'switches', because neither of them are phonologically assimilated to the RL. Given that our theory involves gradience, we consider it conceptually natural to analyse the two items in (6) in two different ways.

Examples (7) and (8) represent the category of **integrated loan**, while (9) and (10) exemplify that of **adapted loan**:

(7) Ja ne asu miesten *dormitory**ssa***,
And they lived men+GEN dormitory+INE

mutta o, koulun miehiä ei ollu.
but oh school+GEN men+PAR not be+PAST TENSE
'And they lived in men's *dormitory* but there were no men of schooling.'

(8) Ja kymmenen tuntia olla siellä, niin pikata niin ee,
And ten hours be there er pick+INF er er

rakkaa pois siitä *ooripile**sta**,* kun oo iso *stockpail**i***
rock+PAR off it+ELA ore pile+ELA as oh big stockpile+NOM

oli. Ja niin, *wheelpaar**ii**,* ja sitte *wheelpaar**ata***
was. And er wheelbarrow+ILL? and then wheelbarrow+INF

sitä sinne toiseen paikkaan, ja kyllä se päivä oli niin pitkä
it+PAR there the other place+ILL and yes it+DET day was so long

että oo, 'bout ten hours, joo.
that er 'bout ten hours, yeah
'And be there for ten hours, picking rock out of that *ore pile,* as there was a big *stock pile.* And, er, *wheelbarrow,* then *wheelbarrowing* it into the other place, and it sure was a long day, er, 'bout ten hours, yeah.'

(9) You know, niinku *räntti**ä**,* muute' me
You know like rent+PAR by the way (?) we

viistoista *taala**a*** maksamma *kuuränttyä*
fifteen buck+PAR pay monthly rent+PAR
'You know, like *rent*, by the way we pay fifteen dollars monthly rent.'

(10) Ja sitte tuo, joka oli *petiruuma**na***
And then that which was bedroom+ESS

tuolla no, sitte ku Rälfi tuli vanhemmaksi
there well then when Ralph became older

me laitimme sille *petiruuma**n**,* se oli *kitsi**nä*** ennen
we made him bedroom+ACC it was kitchen+ESS before
'And then that which was the *bedroom* there, well then when Ralph became older we made him a *bedroom,* it used to be the *kitchen.*'

Examples (7) and (8) show that integrated loans are both morphologically and syntactically (but not fully phonologically) integrated into the RL, whereas adapted loans, as in (9) and (10), are also phonologically fully integrated into, and in most cases adapted to, the RL.

3. Characterizing the Finnish-English language contact in the United States

The characterization of the types of language contact among Finnish Americans is not an easy task because Finnish Americans are in general quite heterogeneous in their bilingualism (cf. Martin 1988). We shall start our characterization by a discussion of the contact situation of the immigrant (ie. 1st) generation. The following generalizations can be made on the basis of Lauttamus & Hirvonen's (1995: 57) description (based on Karttunen 1977).

On the one hand, the first-generation Finnish Americans can be seen as monolinguals. As Lauttamus & Hirvonen (1995: 57) point out, this immigrant generation "will typically go on speaking their old-country language at home as long as they live, and carry on most of their social life in that language". On the other hand, they can also be seen as "marginally bilingual, as most of them can communicate successfully in English in some situations at least", although "Finnish is clearly their dominant language". In general, these speakers of English can therefore be regarded as *non-fluent* bilinguals with a considerable degree of L2 (English) fossilization, and as L2 learners with varying success in learning English (cf. Hirvonen 1982; 1988; 1993; Pietilä 1989).

The characterization of the language contact described above also implies that Finnish is *linguistically dominant* over English, whereas English is socially dominant over Finnish, at least "in some situations". Characteristic for this transfer situation is *lexical borrowing,* whereby loan words are phonologically and morphologically adapted to the patterns of the RL. The domains of phonology, morphology and syntax ('morpho-syntax') of American Finnish spoken by the old immigrant generation seem to be in general resistant to interference from American English (Martin 1988; Virtaranta 1992; cf. Thomason and Kaufman 1988). All this is entirely expected because vocabulary, which is the least stable component of the RL, is affected, whereas the more stable components of RL grammar (e.g. phonology) are usually left intact (Van Coetsem 1988: 36). The crucial feature is, however, that the first-generation Finnish Americans still maintain their own native language.

With the emphasis on the linguistic *outcome* of the contact, the kind of transfer type which prevails among the first-generation Finnish-born Americans is therefore best described as a type of language *maintenance* whereby foreign elements or features are incorporated into a group's

(linguistically dominant) native language (RL) by speakers of that language. The outcome of the incorporation of foreign elements is that "the native language is maintained but it is changed by the addition of the incorporated features" (Thomason & Kaufman 1988: 37).

In contrast to lexical borrowing, which is typical for language maintenance, the interference *from Finnish into the English* spoken by the first-generation Finnish Americans does not begin with vocabulary but with sounds (phonology) and (morpho)syntax. This pattern of interference from *SL* to *rl* is characteristic of language *shift*. As Thomason and Kaufman (1988: 145) suggest, (interference through) *shift* can also be used to refer to situations involving second language acquisition where learners demonstrate imperfect learning as they study a second language, although "they may not actually shift to the TL [i.e. *rl*]". The authors further state that learners' errors are to a considerable degree comparable to "shift-induced language change". Evidence from the English spoken by the first-generation Finnish Americans demonstrates that the phonological and morphosyntactic patterns often deviate from standard (American) English in the manner typical for 'learner language' or *interlanguage* (cf. Pietilä 1989: 152-189; Hirvonen 1988; 1995). This corroborates the view that the immigrant generation can also be regarded as English learners in a naturalistic setting.

Table 2. The two transfer types and the linguistic levels predicted to be affected by interference in the (American) Finnish – (American) English language contact among the first generation (Lauttamus & Hirvonen 1995: 59).

	English (L2) —> Finnish (L1) sl —> RL MAINTENANCE	Finnish (L1)—>English (L2) SL —> rl SHIFT
lexicon	+	-
phonology	-	+
morphosyntax	-	±

Symbols used: '+' = strong, '±' = moderate *or* unclear, '-' = weak interference. RL, SL, as opposed to rl, sl, indicates linguistic dominance.

The transfer types characteristic of the first generation are depicted in table 2. The section under *maintenance* represents the levels affected by interference from English in Finnish. As noted above, it is primarily the

lexical level that is affected in the transfer situation described as *sl —> RL*. In contrast, the section under *shift* represents the levels affected by interference from Finnish in English. The English spoken by the first-generation Finnish Americans is primarily affected in its phonology [+], to a lesser extent in its morphosyntax [±], while lexical interference is only weak [-] (cf. Pietilä 1989: 135, 190-201; Lauttamus 1990: 36-44; 1991: 35). That lexical interference from Finnish into English is weak could be explained as follows. The restricted variety of English spoken by the immigrant generation is almost invariably used for out-group communication only. Given that (American) English is socially (but not linguistically) dominant over Finnish, massive lexical interference from Finnish would therefore be less desirable for successful communication with monolingual English speakers. The direction of lexical interference is thus from the socially dominant language into the socially subordinate one (Lauttamus & Hirvonen 1995: 60).

In an attempt to apply Thomason and Kaufman's contact typology to the three generations' use of American Finnish, Lauttamus and Hirvonen (1995) came up with different predictions for each generation concerning the interference from English in different language domains. These predictions are presented in table 3.

Table 3. The transfer situations and the language levels predicted to be affected by interference from English in Finnish across the three generations. '+' = strong, '±' = moderate or unclear, '-' = weak interference. SL/sl = linguistically dominant/subordinate source language; RL/rl = linguistically dominant/subordinate recipient language.

	English (L2) -> Finnish (L1) *sl -> RL* 1st GEN	English (L?) -> Finnish (L?) *SL ->RL* 2nd GEN	English (L1) -> Finnish (L2) *SL -> rl* 3rd GEN
lexicon	+	+	±
phonology	-	±	+
morphosyntax	-	±	±

The first generation's transfer situation was already presented under '*maintenance*' in table 2: strong lexical interference and weak or no phonological and morphosyntactic interference were predicted.

The second generation usually learn the ethnic tongue as their first language from their immigrant parents. The oldest child in particular will often not learn any English until he or she goes to school. At any rate, by their teen years at least the second-generation children become fluent bilinguals. Their bilingualism is usually English-dominant: they prefer to speak English to each other, and it is sometimes even difficult to detect any foreign features in their English. But their parents will typically continue to speak the old-country language to them, and they understand it. Whether they answer in the old-country language or in English varies from family to family. As they grow older and move out of the ethnic communities, their old-country language starts to deteriorate from lack of regular reinforcement.

The second generation, too, can be predicted to show strong lexical interference because with them English is even more dominant socially. But since English is typically also linguistically dominant with them, they are also predicted to show stronger phonological [±] and morphosyntactic [±] interference than is shown by the first generation.

Even though second-generation marriages are often within the ethnic group, the spouses are usually not comfortable enough in the ethnic language to use it at home. So *the third generation* will learn English as their first language. From their grandparents they will often learn to understand the old-country language to some extent, but their own productive knowledge of the language is typically quite limited. It is only by considerable stretching of the term that they can be called bilingual.

The third generation's use of Finnish is best regarded, then, as a (temporary) shift from English to Finnish, an imperfectly learned second language. Their speaking of Finnish is thus rather analogous to the first generation's speaking of English. Consequently, the model would predict strong [+] phonological interference and moderate [±] morphosyntactic interference (cf. Thomason and Kaufman 1988: 50), just as in the first generation's shift situation. The lexical level requires a comment: as we noted in connection with table 2 above, the expected direction of lexical interference is from the *socially dominant* language into the socially subordinate one. Now with third-generation Finnish Americans, English is most definitely socially (and also linguistically) dominant. That leads us to predict, even though this is a shift situation, moderate lexical interference.

4. Code-switching behaviour as an indicator of language attrition

Our fourth theoretical assumption, that the choice of code-switching strategies is associated with the bilingual speaker's proficiency in the two languages and thus reflects the speaker's stage in language attrition, has been persuasively put forward by Halmari (1993a). She studied the Finnish-English code-switching behaviour of her two daughters, aged 9 and 8 at the time of the data collection, who had lived in the US one year and five months at the time. From recorded conversations of the girls, she identified all turn-internal code-switches and classified them into the following four classes:

I) Intersentential code-switching,
II) Lexical/Phrasal Insertion,
III) Language Assignment shift, and
IV) Quotes/Translations.

Her main result was that the older daughter's preferred code-switching was **lexical/phrasal insertions** whereas the younger daughter preferred **language assignment shifts**. Halmari suggests that the younger daughter's preference for language assignment shifts is an indication of beginning or on-going **language loss**, pointing out that the syntactic constraints on intrasentential code-switching dictate the use of complete English structures, ie. language assignment shifts, if the speaker is not in control of Finnish inflectional morphology.

Halmari was able to show a marked difference in the code-switching behavior of two sisters as little as one year apart, with the code-switching of the younger daughter reflecting language loss after only one year and five months' residence in the US. This raised the question how much stronger evidence we might be able to produce of code-switching reflecting language loss in our data, spanning three generations of progressively more attrited speakers of Finnish in the US. We looked into code-switching for evidence of the shift to English going on in the second generation and being completed in the third generation of Finnish Americans.

Any comparison of our findings with Halmari's requires an awareness of the differences in our classification systems. For our purposes, Halmari's main categories are *lexical/phrasal insertion* and *language assignment shift*. The former would be termed **integrated loans** in our classification, ie. on the "borrowing side" of the continuum. The latter corresponds to our

code-change mostly, but some of her examples we would classify as **code-mixes**.

We would assume, then, that if Halmari's suggestion about code-switching behaviour reflecting language loss applies to our data, then our third-generation speakers should clearly prefer code-changes to code-mixes and integrated loans, the beginnings of that development should be seen in the second generation already, and the first generation should prefer integrated and adapted loans to either kind of code-switching.

5. The data and the procedure

The data dealt with in this article are recorded sociolinguistic interviews, collected in northern Minnesota between 1988 and 1992 by the first author, of three generations of speakers of American Finnish. Our analyses of code-switching and borrowing covered the interview speech of 12 first-generation, 16 second-generation, and 16 third-generation speakers. (By the first generation we mean Finnish-born people who immigrated to the US at an adult age – i.e. over 16. The second generation are children and the third generation, grandchildren of the first generation.)

From the taped and transcribed interviews we first counted the numbers of word tokens included in "English incorporations", i.e. code-switches or borrowings (cf. the categorization in Section 2) in proportion to the total numbers of word tokens spoken by the informant. We then classified all "English incorporations", interview by interview, as *code-changes, code-mixes, integrated loans* or *adapted loans* (for the criteria of classification see Section 2 above) and tabulated the frequencies of each category and the number of word tokens involved. We made these findings comparable across the varying-length interviews by calculating the frequencies of each category and the number of word tokens involved *per 2000 words spoken.*

We computed mean scores for each generation's use of code-switching and borrowing and, within the generation-groups, also separately for women and men so as to identify any patterned gender differences in code-switching and borrowing behaviour.

To bring into focus the patterns of differences among the three generations in their use of code-switches and borrowings we used *boxplot,* a plotting device provided in the SPSS statistical package. See, eg., figure 1 in subsection 6.1 below. What *boxplot* does is simply draw the median (the heavy black line), surround it by a box which includes the lower and

the upper quartile, and denote the minimum and maximum by drawing tails from this box to the extremes.

If an individual's score is so far outside the distribution shown by the rest of the group that it does not even fit into the extremes, that individual is designated as an *outlier* in the diagram. For example, informants 6 and 12 of the first generation and informant 15 of the second generation are marked as *outliers* in figure 1. Their scores were therefore excluded from the generation-group averages presented in table 4 and figures 6-9. This procedure should not be regarded as an exercise in cosmetics but a respectable statistical procedure which allowed us to exclude from the group figures those very few individuals who were clearly unrepresentative of the group. It is not that these individuals' results are uninteresting but that they require a separate treatment so as to get to the reasons of their exceptionality.

The generation-group medians and distributions of the English-incorporation percentages are displayed graphically as *boxplots* in figure 1. The group mean percentages, computed also separately for women and men, are presented in table 4. The strength of English influence is discussed in the light of these quantitative results in subsection 6.1.

The generation-group medians and distributions of the frequency of each category are displayed graphically as *boxplots* in figure 2 (adapted loans), figure 3 (integrated loans), figure 4 (code-mixes), and figure 5 (code-changes). The mean frequencies of each category per generation-group are presented both graphically (as histograms) and numerically in figure 6 and the same information broken down by gender in figure 7. These quantitative results are discussed in subsection 6.2. The numbers of word tokens involved in each category of code-switching and borrowing are presented graphically and numerically in figure 8 (generation-group averages) and figure 9 (generation-group averages broken down by gender) and discussed in subsection 6.3.

6. Results and discussion

6.1. The strength of English influence in the interviews in terms of word tokens

We estimated the overall strength of English presence in the interviews by calculating the proportions of English-origin word tokens out of all word

tokens uttered by the informant. The patterns observed are portrayed in figure 1 and table 4 below.

Figure 1. Boxplot of the proportions of English-origin word tokens, plotted as generational distributions showing the median frequency (the heavy black line), the lower and upper quartiles (the gray box), the minimum and maximum (the tails), and the outliers (the numbered stars and circle).

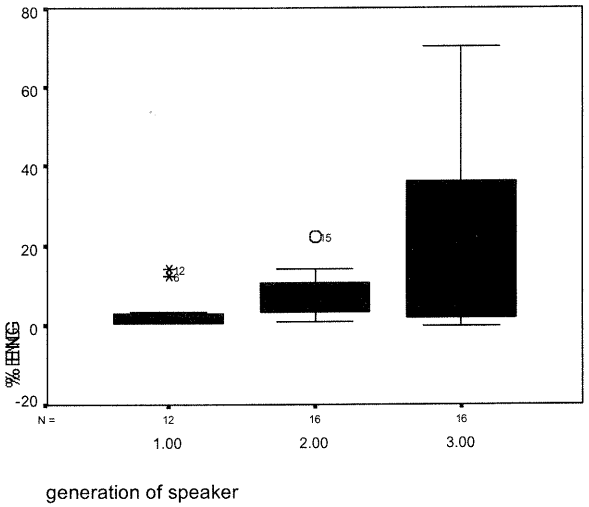


Figure 1 shows, besides the obvious result that there is clearly more English influence in the second-generation speech than the first-generation speech and much more in the third-generation speech, that the first-generation group is quite homogeneous in this respect – with two exceptions; that the second-generation group shows more variation but quite a symmetric distribution, much like a bell curve – with one exception; and that the third-generation group has a great deal of variation and a very skewed distribution of English influence.

Table 4. Average numbers of word tokens spoken by the informants in the interview, numbers of word tokens involved in English incorporations, and percentual shares of English-origin material produced by the speakers: averages per generational group, broken down by gender ("F" and "M").

	ALL WORD TOKENS	ENG WORD TOKENS	% ENG WORD TOKENS
G1F (N=6)	4816	96	1.99%
G1M (N=4)	1568	12	0.77%
G1 ALL	**3517**	**63**	**1.79%**
G2F (N=9)	3833	347	9.05%
G2M (N=6)	2190	134	6.10%
G2 ALL	**3175**	**261**	**8.22%**
G3F (N=8)	1497	510	34.08%
G3M (N=8)	2560	204	7.95%
G3 ALL	**2028**	**357**	**17.59%**

Table 4 shows, first of all, that women's interviews are considerably longer on average except in the third-generation group, where they are considerably shorter. There is a great deal of variability, however, within the groups too: the first-generation range is actually 1062 to 19419 word tokens, the second-generation 618 to 6898 word tokens, and the third-generation one 466 to 4895 word tokens. At least this information shows that it was definitely necessary to ensure the comparability across interviews of the frequencies of code-changes, code-mixes, intergrated loans and adapted loans (discussed in the next subsection) by calculating them as "x instances per 2000 words".

Figure 1 showed graphically how the English influence increases from generation to generation. To this, table 4 adds the average percentages of English incorporations: about 2 % for the first, about 8 % for the second, and about 18 % for the third generation. Also, it is clearly seen that this influence is more pronounced with female than male speakers in every generation and that this difference is particularly great in the third generation.

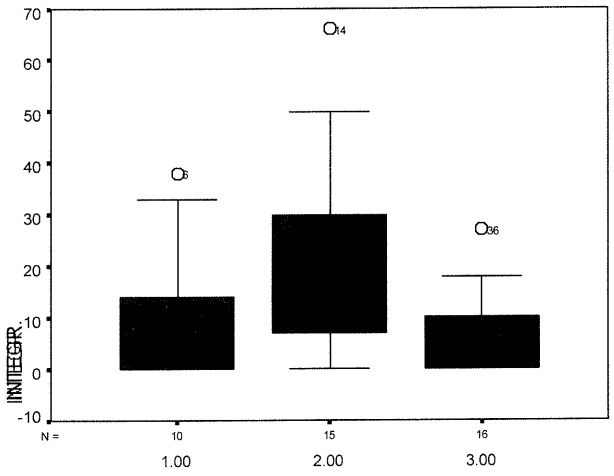
With progressive attrition of Finnish in the later generations, it was entirely to be expected that English incorporations become more frequent from generation to generation. What was perhaps not so obvious at the outset is that this trend should be clearly more pronounced with women than men.

In subsections 6.2 and 6.3 we attempt to locate the English influence more precisely by looking at the distributions of different types of code-switching and borrowing in the three generational groups.

6.2. The frequencies of adapted loans, integrated loans, code-mixes, and code-changes

Let us first examine the generational distributions of the different subtypes of code-switching and borrowing by means of boxplots. They are presented in figures 2 – 5.

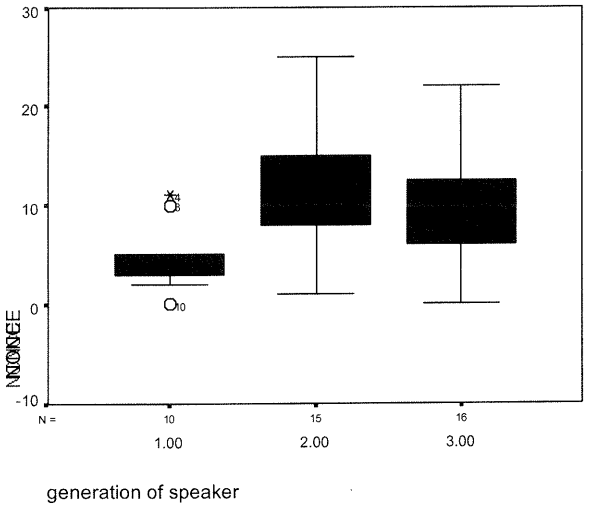
Figure 2. Boxplot of the generational distribution of adapted loans in American Finnish interview speech, showing the median frequency, the lower and upper quartiles, the minimum and maximum, and the outliers.



We see in figure 2 that the first and the third generation tend to have a low frequency of adapted loans, whereas the second generation has the highest median frequency and the widest variation in the distribution. There are three outliers, one in each generation group.

The medians and distributions displayed in figure 2 accord with earlier findings, based on a vocabulary elicitation task done by the same set of informants, that the second generation use by far the most loanwords whereas the figures for the first and the third generation are about the same (eg. Hirvonen 1996; 1998b).

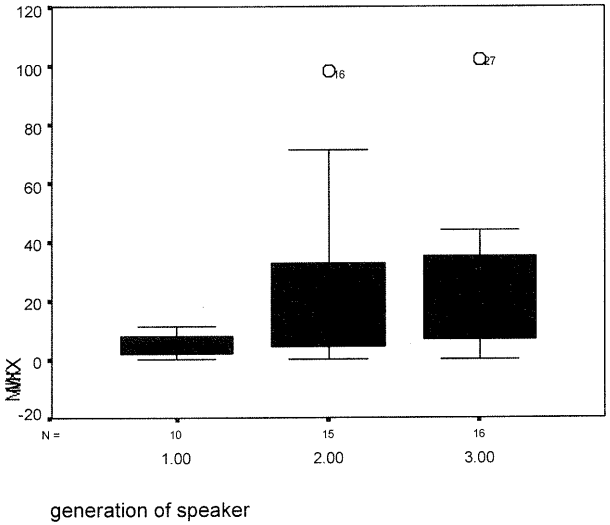
Figure 3. Boxplot of the generational distribution of integrated loans in American Finnish interview speech, showing the median frequency, the lower and upper quartiles, the minimum and maximum, and the outliers.



As shown in figure 3, it is the second and the third generation, especially the second, which mostly use integrated loans; their frequency in the first-generation data is quite low and the variation quite low (but note that there are three outliers in the first-generation group). Considered together with the information of figure 2, this again confirms that the first and the third generation use loanwords (ie. adapted *and* integrated loans) to about the same extent and much less than the second generation.

Code-mixes (figure 4) are used very little by first-generation speakers but a great deal by the second and the third-generation speakers. This may be a symptom of "beginning or on-going language loss" (Halmari 1993a) where the speaker's faltering control of Finnish inflectional morphology dictates the use of short code-switches rather than, or besides, (adapted *or* integrated) loans.

Figure 4. Boxplot of the generational distribution of code-mixes in American Finnish interview speech, showing the median frequency, the lower and upper quartiles, the minimum and maximum, and the outliers.

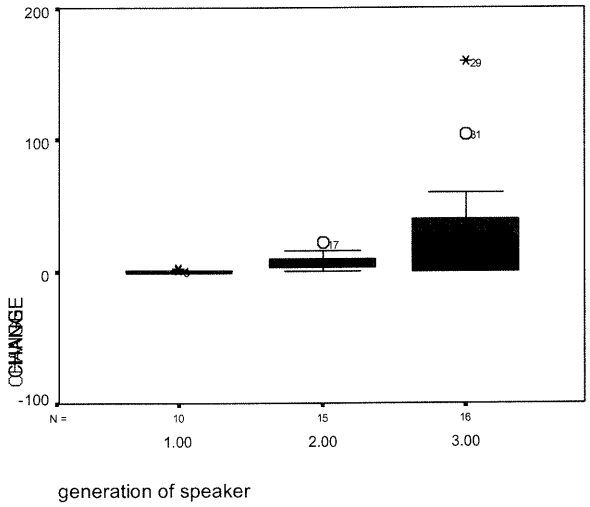


The more extensive subtype of code-switching, code-changes (figure 5), is much favoured by the third generation, quite little by the second generation, and is virtually nonexistent in the first-generation interviews. Code-changing can thus be characterized as a typically third-generation phenomenon. From the point of view of needing to use code-switching as a communication strategy, if extensive use of code-mixes suggests

"beginning or on-going language loss", then in the case of extensive users of code-changes the language loss must have advanced further.

The distributional patterns displayed by figures 2-5 are complemented below with figure 6, which portrays the generational group frequencies of the different types of code-switching and borrowing by means of histograms and also numerically, and with figure 7, which gives the same information separately for women and men speakers.

Figure 5. Boxplot of the generational distribution of code-changes in American Finnish interview speech, showing the median frequency, the lower and upper quartiles, the minimum and maximum, and the outliers.



By bringing all the subtypes of borrowing and code-switching together, figure 6 brings the intergenerational differences into sharp focus. We see that all the subtypes are fairly infrequent with the first generation. They borrow more than code-switch, and when they borrow they favour adapted loans over integrated loans. We can surmise that their limited

command of the source language English may contribute to their tendency to *adapt* their loans to Finnish phonology rather than integrate English phonological features into their Finnish speech.

The second generation use both adapted and integrated loans much more frequently than the first. Being the group which is closest to balanced bilingualism, they can be assumed to have enough control of both Finnish and English phonology to have the option of using either adapted or integrated loans, and as we see in figure 6, they make extensive use of both. Also, they must have learned quite a few adapted loans as just "Finnish" from their parents. Besides borrowing, the second generation also use a great deal of code-switching. The fact that they clearly favour the shorter subtype, code-mixes, which typically involve single lexical items only, suggests that their command of Finnish inflectional morphology allows them most of the time to return to Finnish after just a short code-switch.

The third generation make much less use of loans, both adapted and integrated, than the second generation. Their predominant type of English incorporation is code-switching, where they use the more extensive subtype, code-changes, almost as frequently as code-mixes. This may be largely because of their limited command of the recipient language Finnish. Their low command of Finnish morphology and syntax may dictate the use of complete English structures, ie. continue the code-switch once started to the end of the sentence or even further.

The third generation make much less use of loans, both adapted and integrated, than the second generation. Their predominant type of English incorporation is code-switching, where they use the more extensive subtype, code-changes, almost as frequently as code-mixes. This may be largely because of their limited command of the recipient language Finnish. Their low command of Finnish morphology and syntax may dictate the use of complete English structures, ie. continue the code-switch once started to the end of the sentence or even further.

Examining figure 7, we cannot find a consistent pattern of gender differences in the use of adapted and integrated loans. In the use of code-mixes and code-changes, however, there is a clear pattern: in each generation, the women code-switch more on average; in the second and third generation this difference is rather dramatic.

Figure 6. American Finnish interview speech: Average frequencies, per generation of speakers, of different types of code-switches and lexical borrowings

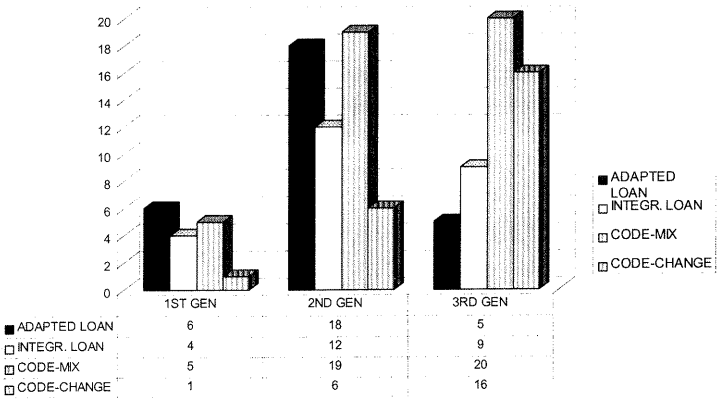
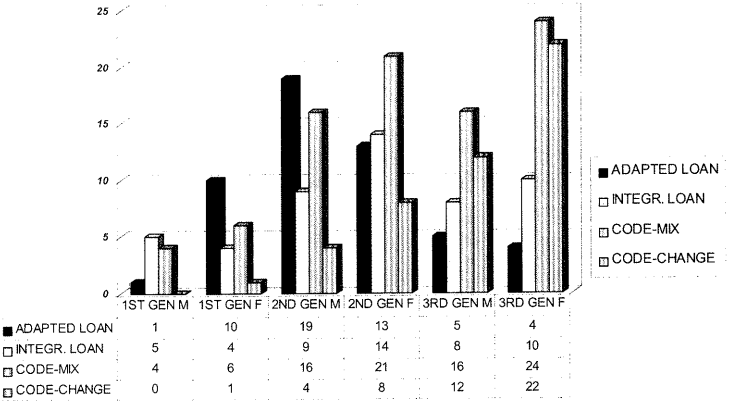


Figure 7. American Finnish interview speech: Average frequencies, per generation and gender of speakers, of different types of code-switches and lexical borrowings



6.3. The numbers of word tokens involved in adapted loans, integrated loans, code-mixes, and code-changes

The patterns observable in figures 6 and 7 become even more graphic when we look at the **numbers of word tokens** involved in the different types of English incorporations. Figure 8 gives this information for the three generational groups by means of histograms and numbers, and figure 9 distinguishes between women and men within each group.

In terms of numbers of word tokens involved in the various subtypes of borrowing and code-switching, which can be considered a good measure of the strength of English influence in the American Finnish speech of the informants, we can see in figure 8 that the first-generation speakers have next to no need for English in their speech.

The second generation both borrow and code-switch a great deal more, with the emphasis already on the code-switching side as far as the numbers of word tokens are concerned.

The third generation borrow less than the second but more than make up for it by massive code-changing.

Figure 8. American Finnish interview speech: Average numbers of word tokens, per generation of speakers, produced in the different types of code-switches and lexical borrowings

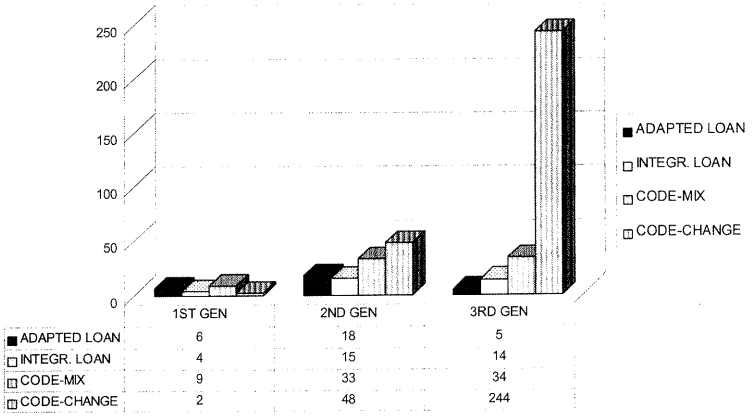
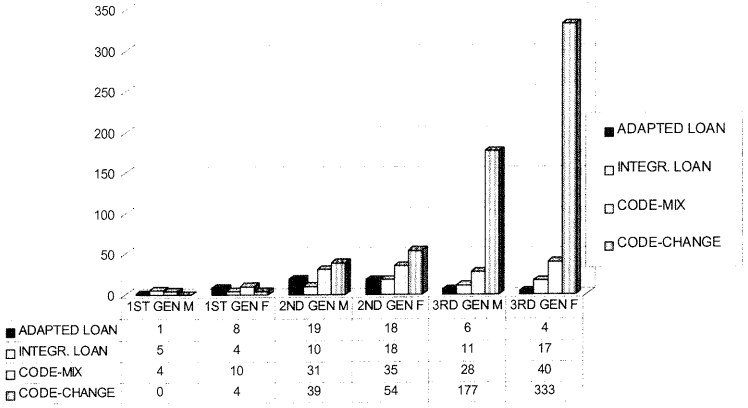


Figure 9. American Finnish interview speech: Average numbers of word tokens, per generation and gender of speakers, produced in the different types of code-switches and lexical borrowings



	1ST GEN M	1ST GEN F	2ND GEN M	2ND GEN F	3RD GEN M	3RD GEN F
ADAPTED LOAN	1	8	19	18	6	4
INTEGR. LOAN	5	4	10	18	11	17
CODE-MIX	4	10	31	35	28	40
CODE-CHANGE	0	4	39	54	177	333

In terms of gender differences, the pattern in figure 9 is the same as in figure 7: no clear pattern observable in the use of adapted and integrated loans, but women code-switch more in each generation, and this difference is rather dramatic in the case of code-changing in the third generation.

7. Conclusion

This may be a good place to remind the reader that our informants and the speech situation of our data collection is rather far removed from what is sometimes referred to as a situation of "classical code-switching" – a situation where fluent bilinguals talk among themselves as members of a bilingual speech community and have a genuine choice of two fluently mastered codes at their disposal. In such a situation, code-switching may well be motivated simply by attempts to find the best way of expressing something. – In our data, only the second generation can be said to approach the condition of being fluent bilinguals between Finnish and English. Our first-generation speakers have great deficiencies in their English and our third generation speaker have great deficiencies in their Finnish, so that in many situations they do not have this free choice of

codes but have to supplement one code with the other, ie. use code-switching – and borrowing – to fill lexical and structural gaps, ie. as strategies of communication. Furthermore, our data come from interviews in which the informants had agreed to be interviewed "in Finnish" not by a member of the Finnish-American community but by a scholar from the old country. These conditions can be assumed to have encouraged the informants' staying in Finnish as much as they were able to, which may have caused deviations from the bilingual community's own internal code-switching norms. This situation seems to call for a more comprehensive view of the possible motivations of code-switching and borrowing and a model of these processes that entail a conception of their intimate relationships and gradient nature.

Accordingly, as far as the present data are concerned, we regard code-switching first and foremost as a communication strategy – filling lexical gaps. This is not to be understood as denying other possible motives for code-switching, however, even in the present data.

In her case study of the code-switching behaviour of her two daughters that was referred to in section 4 above, Halmari (1993a) has this to say about the preferred code-switching type of her older daughter, ie. lexical/phrasal insertion (which we categorize as integrated borrowing): "The Finnish syntactic frame is left intact, and only separate lexical items are inserted from English. The morphology of Finnish prevails..." The younger daughter, whose preferred code-switching type is the language assignment shift, "often clearly prefers to complete an utterance, which she has started in her L1 (Finnish), in her L2 (English)" and "we can detect an obvious difficulty experienced by her to successfully complete the utterance in her L1."

These characterizations also seem to fit our first-generation and third-generation speakers, respectively, rather well. All the English incorporations that the first-generation, or immigrant-generation, speakers need in their Finnish discourse is a few single-word insertions, which are typically integrated morphosyntactically – and often also phonologically – into their Finnish discourse (ie. integrated loans or adapted loans). The second generation needs more English in their Finnish discourse, and since their mastery of the rich inflectional morphology of Finnish is less than perfect, they are not nearly always able to integrate their English incorporations morphosyntactically into their Finnish discourse, which results in code-mixes, and even code-changes, more often than borrowings. And the third generation, whose Finnish is **not** a continuation of the

grandparents' speech but a second language imperfectly learned after childhood, have, besides a restricted vocabulary, so little Finnish grammar at their control that after switching to English, which they have to keep doing all the time, they are seldom able to switch back to Finnish in the same constituent, or the same clause, or even the same speaking turn. To put it in another way, the third-generation speakers are not only English-dominant but but have such a limited mastery of Finnish structure that they are unable to cope in a Finnish conversation without massive code-switching into English.

The results we have obtained also support the view that the code-switching behaviour of the third-generation speakers is ultimately a consequence of the language shift to English which has already taken place in their parents' speech.

References:

Andersson, P. (1993) Finns and Americans in Sweden: Patterns of linguistic incorporation from Swedish". In Guus Extra & Ludo Verhoeven (eds.), *Immigrant Languages in Europe* pp. 249-269. Clevedon, Philadelphia: Multilingual Matters.

Gumperz, J.J. (1982) *Discourse Strategies*. Cambridge: Cambridge University Press.

Halmari, H. (1993a) Code-switching strategies as a mirror of language loss: A case study of two child bilinguals. In D. Staub & C. Delk (eds.), *The Proceedings of the Twelfth Second Language Research Forum. Papers in Applied Linguistics* pp. 200-215. East Lansing, MI: Michigan State University.

Halmari, H. (1993b) Structural relations and Finnish-English code switching. *Linguistics* 31: 1043-1068.

Halmari, H. (1997) *Government and Codeswitching: Explaining American Finnish.* Amsterdam: John Benjamins.

Haugen, E. (1953) *The Norwegian Language in America: A Study in Bilingual Behavior*. Philadelphia: University of Pennsylvania Press. [Reprinted in 1969. Bloomington: Indiana University Press.]

Haugen, E. (1956) *Bilingualism in the Americas: A Bibliography and Research Guide.* Publications of the American Dialect Society, 26.

Hirvonen, P. (1982) Aspects of Fossilized Interlanguage: The English of Finnish Americans. In T. Fretheim & L. Hellan (eds.), *Papers from the Sixth Scandinavian Conference of Linguistics* pp. 260-268. Trondheim: Tapir.

Hirvonen, P. (1988) American Finns as Language Learners. In M. Karni, O. Koivukangas & E. Laine (eds.), *FINNS IN NORTH AMERICA: Proceedings of Finn Forum III, 5-8 September 1984, Turku, Finland* pp. 335-354. Turku: Institute of Migration.

Hirvonen, P. (1993) Return of the Klörkki: More about the Fate of Old-country Words in American Finnish. In M. Suojanen & A. Kulkki-Nieminen (eds.), *19. Kielitieteen päivät Tampereella 8. - 9. toukokuuta 1992* pp. 81-96. Folia Fennistica & Linguistica, 16. Tampere: University of Tampere.

Hirvonen, P. (1995) Phonological and Morphological Aspects of Finnish Language Attrition in the United States. In Wolfgang Viereck (ed.), *ZDL-Beiheft 77: Verhandlungen des Internationalen Dialektologenkongresses Bamberg 1990, Band 4* pp. 181-193. Stuttgart: Franz Steiner.

Hirvonen, P. (1996) Retention and Replacement of Original Finnish Words in American Finnish. In P. Sture Ureland and Iain Clarkson (eds.), *Language contact across the North Atlantic. Linguistische Arbeiten 359* pp. 333-349. Tübingen: Niemeyer.

Hirvonen, P. (1998) The Finnish-American Language Shift. In Jussi Niemi, Terence Odlin and Janne Heikkinen (eds.), *Language Contact, Variation and Change. Studies in Languages 32* pp. 136-150. Joensuu: University of Joensuu, Faculty of Humanities.

Karttunen, F. (1977) Finnish in America: A Case Study in Monogenerational Language Change". In B.G. Blount & M. Sanches (eds.), *Sociocultural Dimensions of Language Change* pp. 173-184. New York: Academic Press.

Lauttamus, T. (1990) *Code-switching and Borrowing in the English of Finnish Americans in an Interview Setting* (= *Studies in Languages*, 20). Joensuu: University of Joensuu.

Lauttamus, T. (1991) Borrowing, Code-switching, and Shift in Language Contact. In Muusa Ojanen & Marjatta Palander (eds.), *Language Contacts East and West* pp. 32-53. (= *Studies in Languages*, 22). Joensuu: University of Joensuu.

Lauttamus, T. (1992) Lainaaminen ja koodinvaihto: havaintoja amerikansuomalaisten kielistä. *Virittäjä* 1/1992: 3-16.

Lauttamus, T. & P. Hirvonen (1995) English Interference in the Lexis of American Finnish. *The New Courant* 3: 55-65. Helsinki: Helsinki University Press.

Lauttamus, T. (1999) Fuzzy Switch and Loan Types in the Languages of Finnish Americans. *SKY Journal of Linguistics* 12: 87-109.

Martin, M. (1988) *Amerikansuomen morfologiaa ja fonologiaa.* An unpublished Licentiate Thesis at the University of Jyväskylä (Finland).

Pietilä, P. (1989) *The English of Finnish Americans with Reference to Social and Psychological Background Factors and with Special Reference to Age*. Turun yliopiston julkaisuja, Sarja B, Osa 188. Turku: Turun yliopisto.

Poplack, S. (1980) Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: Toward a typology of code-switching. *Linguistics* 18: 581-618.

Poplack, S., S. Wheeler & A. Westwood (1987) Distinguishing Language Contact Phenomena: Evidence from Finnish-English Bilingualism. In P. Lilius & M. Saari (eds.), *Proceedings of the Sixth International Conference of Nordic and General Linguistics in Helsinki, August 18-22, 1986* pp. 33-56. The Nordic Languages and Modern Linguistics, 6. Helsinki: Helsinki University Press. (Also *World Englishes* 8 [1989]: 389-406.)

Romaine, S. (1995) *Bilingualism.* 2nd ed. Oxford: Blackwell.

Thomason, S.G. & T. Kaufman (1988) *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
Torres, L. (1989) Code-mixing and Borrowing in a New York Puerto Rican Community: a Cross-generational Study. *World Englishes* 8: 419-432.
Van Coetsem, F. (1988) *Loan Phonology and the Two Transfer Types in Language Contact*. Dordrecht: Foris.
Van Coetsem, F. (1990). Review of Thomason and Kaufman (1988), Lehiste (1988), and Wardhaugh (1987). *Language in Society* 19: 260-268.
Van Coetsem, F. (1995) Outlining a Model of the Transmission Phenomenon in Language Contact. *Leuvense Bijdragen* 84: 63-85.
Virtaranta, P. (1992) *Amerikansuomen sanakirja / A Dictionary of American Finnish.* Turku: Institute of Migration.
Wardhaugh, R. (1992) *An Introduction to Sociolinguistics.* 2nd ed. Oxford: Blackwell.

Contact Address:
Pekka Hirvonen
English Department
Joensuu University
P.O. Box 111
80101 Joensuu, Finland
pekka.hirvonen@joensuu.fi

Timo Lauttamus
English Department
Oulu University
P.O.Box 1000
SF-90014 Oulun yliopisto
timo.lauttamus@oulu.fi