Introduction:

How Languages Fit Together in Codemixing*

Shana Poplack and Marjory Meechan
University of Ottawa

Acknowledgements*

This project was supported by the Social Sciences and Humanities Research Council of Canada in the form of research grants to Shana Poplack. We are grateful to Dawn Harvie and James Walker for their invaluable assistance with the preparation of this issue.

In virtually all bilingual corpora empirically studied, mixed discourse is overwhelmingly constituted of lone elements, usually major-class content words, of one language embedded in the syntax of another (Berk-Seligson, 1988; Norrier, 1989; Poplack, Sankoff, & Miller, 1988; Treffers-Daller, 1994; to name but a few). The status of these items is notoriously ambiguous. They may be codeswitches or borrowings or yet another type of language mixture. They are at the heart of a fundamental disagreement among researchers about data, not so much how they should be gathered (although this too is contentious), but rather how types of language mixture should be classified. Are these lone items integrated into the grammatical system in which they are embedded (i.e., borrowed) or conditioned by some interaction between grammars (codeswitches)?

Empirical study has confirmed early claims (e.g., Haugen, 1950) that major-class content words such as nouns, verbs, and adjectives are the most likely to be borrowed (e.g., Poplack, et al., 1988; van Hout & Muysken, 1994). The papers in this issue of the International Journal of Bilingualism are consecrated to the study of the behavior of such items in bilingual discourse. Though other linguistic elements are also borrowed occasionally (e.g., tags, discourse markers, or grammatical function words), these generally are instantiations of other products of language contact (e.g., de Rooij, 1995; Poplack 1980). The conclusions of the papers assembled here thus apply only to lone other-language content words.

Classic indications of loanword integration have had mediocre results in disambiguating the status of lone content items because different measures produce conflicting results. While these items are usually syntactically positioned according to the language in which they are embedded, they often appear to retain the phonetic properties of the donor language and sometimes appear to conserve donor-language morphology. Disagreement over the appropriate classification of lone items has resulted in a range of codeswitching theories. At one end of the spectrum, where lone items are defined as codeswitches, researchers tend to consider the relationship between languages in bilingual discourse asymmetrical, with one grammar clearly predominating (e.g., Joshi, 1985;
Myers-Scott, 1993; Rivais, 1981). Where lone items are classified as borrowings and
bilingual interactions, where other-language constructions were used copiously and sponta-
evidence for codeswitching is inferable only from multiword fragments. Both languages are
eously, as is characteristic of a bilingual discourse mode.
postulated to play a role in constraining codeswitching (e.g., Belazi, Rubin, & Toribio,
Any hope of rapprochement between these two views resides in a theory-independent
classification of lone other-language items. If for no other reason, their sheer volume in
documented contact situations enjoins us to determine their status. In their focus on lone
other-language content items in five languages in contact with English—Turkish, Persian, French,
Ukrainian and Igbo—the papers in this issue all illustrate how this can be accomplished.

While disambiguation of language contact phenomena on a case-by-case basis may
remain elusive, it is now clear that classes of items in individual data sets can be quantita-
tively characterized as preponderantly codeswitches or borrowings. Though the language
pairs treated here range from typologically similar to dramatically different on
methodological imperatives

In stable bilingual communities, speakers conventionally make use of both languages with
the same interlocutors, in the same domains, and within the same conversational topic.
Our objects of inquiry are the linguistic norms, or conventions of the community, as opposed to
the idiolect of any isolated individual living in it. Such norms are revealed in situations in
which spontaneous codeswitching (any use of two or more languages in the same discourse)
forms part of a discourse mode.

The study of codeswitching as a discourse mode of a bilingual community dictates two
methodological requirements for any scientific study of the topic. First, we must identify a
totally of any isolated individual living in it. Such norms are revealed in situations in
which spontaneous codeswitching (any use of two or more languages in the same discourse)
forms part of a discourse mode.

As a second imperative, we must obtain a sufficiently large sample of sustained
discourse representative of the bilingual mode in order to carry out the study. It is in these
steps, prior to any linguistic analysis, that social and demographic knowledge of
the community are pertinent. While much work on bilingual discourse has made use of data
collection methods such as informant elicitation, subjective reaction tests, even
introspection by the bilingual linguist, these can only be taken as very indirect evidence of
how bilinguals actually use both languages in real-life situations.

Each paper in this issue exemplifies these two methodological imperatives. All are
carried out in well-defined bilingual speech communities using standard social network
methodology. Most of the authors, as bona fide community members, collected data from
interactions with their own close contacts (Adalar, Lze, Ghafar Samar). The others used
their status as members of the wider bilingual/ethnic community to gain entry and
acceptance. Turpin enlisted a community member to collect the Acadian French-English
materials she analyzes, while Budzak-Jones engaged in long-term participant observation
in the Ukrainian community of Lelington to gather her data. The achievement of an in-
group relationship between interviewer and community is surely responsible for the
extraordinary data on which these articles are based: five sets of unusually rich vernacular

Identification of patterns

How can we infer which of these processes (if any) gave rise to the instances of codeswitching
found in a discourse sample? A discourse mode consists of recurrent patterns of speech
behavior, and analysis of these patterns can reveal the system membership of the mixed
elements. The major analytical tool underlying the articles in this issue is the identification
of such patterns in the data. These are defined as:

- a series of parallel occurrences (established according to structural and/or functional
criteria) occurring at a non-negligible rate in a corpus of language use.

Patterns studied here include the variable marking of spatial reference, plurality, and
case agreement, among others.

Patterns can only be discerned by systematic and exhaustive quantitative analysis of
the data, in conformity with Labov’s principle of accountability (Labov, 1972; Poplack,
1993), which specifies that values must be reported for “every case where the variable
element occurs in the relevant environments as we have defined them” (Labov, 1972, p. 22).
emphasize ours). This permits us to discard imagined examples and isolated tokens, which
are not representative nor recurrent, in favor of the regular tendencies that characterize
natural exchanges in the community.
A variationist approach

The systematic evaluation of corpus materials and the quantitative comparison of monolingual and codemixed vernaculars are most naturally integrated within the conceptual and analytical framework of variation theory (e.g., Labov, 1966/1982; Sankoff, 1974; etc.), with its statistical tools for measuring rates, co-occurrences and environmental conditioning. The assumptions of variation theory as they apply to the disambiguation of language contact phenomena are shown in (1) (Poplack, 1997).

(1) Variationist assumptions as applied to language contact:

a. To some extent the rates, but especially the conditioning of linguistic variability are language-specific; once determined, they may be helpful as diagnostics of language membership.

b. If a set of donor-language (L_d) items embedded in otherwise recipient-language (L_r) discourse features the same hierarchy of constraints as their L_d counterparts (while simultaneously differing from those conditioning L_d items in L_d discourse), we may conclude that the grammar constraining the L_d items in otherwise L_r discourse is that of L_d.

c. If a set of L_d items embedded in otherwise L_r discourse features the same hierarchy of constraints as other L_d items, while simultaneously differing from those conditioning their L_d counterparts in L_r discourse, we may conclude that the grammar constraining the L_d items is that of L_d.

This leads to the hypothesis on loanword integration reproduced in (2):

(2) "If the constraints on variability of L_d-origin forms are parallel to those constraining their L_r counterparts, the former can only be borrowings." (Ibid.)

The papers in this issue each represent an empirical test of the hypothesis that lone other-language elements pattern according to the grammar of L_r, and therefore are borrowed. They also test the alternative hypothesis, that is, that they pattern according to the rules of the donor-language grammar, and hence are codeswitched. They employ different diagnostics and language pairs, but all make use of the same comparative method.

Comparative method

Insofar as codeswitching and borrowing are based on some principled combination of elements of the monolingual, that is, unmixxed, vernaculars of the bilingual community, it is important to have as explicit an idea as possible of the nature of these vernaculars before concluding that a codeswitched element takes aspects from one or the other or both. In particular, we cannot rely solely on prescriptive or nonquantitative descriptions of more or less related varieties of either language, but must compare and contrast the bilingual structures we find with the unmixed patterns in the same corpus. We thus follow Wentsch (1953/1968, p. 1) in defining the "language-using individuals, as the locus of the contact."

The key to the method employed in these studies is crosslinguistic comparison of patterning in the speakers' vernacular varieties of the source languages. The papers in this issue are all based on detailed analysis of one or more of the source vernaculars in contact. Each compares the quantitative patterning of (at least) lone English-origin nouns with that of Turkish, Persian, Ukrainian, French, or Igbo nouns in the respective recipient languages. Nouns have been widely documented as the category figuring most frequently in codemixing, and the corpora examined here are no exception. Adalar and Taglimonte, and Ghafer Samar and Meechan, extend this comparison to English-origin nouns already attested as established loanwords in Turkish and Persian. If the contentious lone English-origin nouns pattern like attested loanwords, this can be taken as further confirmation that 1) the former are borrowings, if only for the nonce (Sankoff, Poplack, & Vanniaraj, 1999), Weinreich, 1953/1968) and 2) nonce borrowings and dictionary-attested loanwords are not distinct linguistically.

Most of the articles further contrast these patterns with those of nouns in both unmixed English and unambiguous (multiword) intrasential codeswitches to English. Adalar and Taglimonte repeat the exercise in the other direction, comparing lone Turkish-origin nouns in English discourse with their counterparts in unmixed Turkish, unmixed English and unambiguous codeswitches to Turkish. Such multiplex comparisons allow the researcher to discover the associations, if any, between the patterning of the contentious lone English-origin items and that of like items in one of the other language contexts: mixed and/or unmixed—and in this way, disambiguate their status. Thus, as per assumptions (1a) and (1b), if according to some diagnostic criteria, lone English-origin nouns are observed to pattern like their counterparts in the unmixed recipient language, while at the same time differing from patterns of nouns in the donor language and/or in unambiguous codeswitches to that language, we conclude that only the grammar of the recipient language is operative, that is, that they were borrowed. Onl if they pattern with their counterparts in the (unmixed) donor language and (by extension), with nouns in unmixed codeswitches to that language, while at the same time differing from the patterning of nouns in the unmixed recipient language, are we justified in concluding that the lone English-origin nouns result from codeswitches (assumption 1c).

Figure 1 illustrates how this crosslinguistic comparison can be effected in a contact situation featuring a rich complement of language mixing phenomena (cf., Adalar & Taglimonte).
Of course the emphasis on tapping the norms of the community means that not all data sets will provide the opportunity for such exhaustive comparisons. But a corpus lacking comparative data on one or more context(s) can still be used to test for system membership. Where lone English-origin nouns can only be compared to (unmixed) Ukrainian, Budzhab-Jones shows that they replicate not only the categorical, but also the variable, behavior of native Ukrainian nouns: Even the patterning of nonstandard gender marking of both sets of nouns is virtually identical. Such similarities are too detailed to be due to chance.

The conflict site

Comparisons are revealing of system membership to the extent that there are differences in the observed patterns in each context. If two (unmixed) languages each exhibit parallel structures occurring at the same rate (a site of coincidence), bilingual structures showing the same patterns cannot be readily identified as deriving from either language. To unambiguously ascertain the language membership of the bilingual construction, we need to examine areas at which the structures of the language pair do not match, that is, sites where the grammars of the two languages in contact conflict.

The conflict site is another central methodological tool of the papers in this issue. Where two languages differ in distribution of marking of comparable functions, we can evaluate the system membership of lone other-language items: by first determining the rate and conditioning of marking and then comparing them to those of like items in the unmixed source languages. When lone items show not only the grammatical structure of their counterparts in the language in which they are embedded, but also enter into these structures at a rate mirroring that of their unmixed counterparts, we infer that they are full-fledged participants in the grammatical system of the recipient language.

The papers in this issue illustrate a variety of ways in which the conflict site may be employed. All of the papers make use of some kind of morphological criteria, such as noun inflection (e.g., plural marking (Adalar & Tagliamonte; Turpin), gender marking (Budzhab-Jones), reference marking (Adalar & Tagliamonte; Ghafar Samar & Meechan; Turpin) and verb inflection (Eze)). Two papers examine the phonological process of vowel harmony (Adalar & Tagliamonte, Eze). Syntactic criteria include case marking and agreement (Budzhab-Jones) and word order, with respect to the position of the lone element either in the phrase or relative to its modifiers (Adalar & Tagliamonte; Ghafar Samar & Meechan; Eze; Turpin). Regardless of language or diagnostic, the results are remarkably similar, leading to the conclusion that the results of language contact are not monolithic.

Identifying language contact phenomena

Codeswitching is not borrowing
Contra the assumption often made in the literature, codeswitching and borrowing are two distinct phenomena. By comparing unambiguous codeswitches to English (defined as "the juxtaposition of sentences or sentence fragments, each of which is internally consistent with the morphological and syntactic (and optionally, phonological) rules of the language of its provenance" (Poplack, 1993, p.255)) to dictionary-attested English loanwords, Adalar and Tagliamonte demonstrate that loanwords pattern to recipient language (i.e., Turkish) syntax and codeswitches pattern like the donor language (i.e., English). What of the lone other-language items?

Lone other-language items are not necessarily codeswitches
Neither the lone English-origin nouns in otherwise Turkish discourse nor the lone Turkish-origin nouns in otherwise English discourse studied by Adalar and Tagliamonte bear any resemblance to nouns within unambiguous codeswitches than the attested loanwords. This, in conjunction with earlier findings (e.g., Meechan, & Poplack, 1995), disproves the hypothesis that all lone items are codeswitches. This finding is replicated by Eze, who shows that lone English-origin items in otherwise Igbo discourse display quantitative patterns which are clearly at variance with both unmixed English and multilingual codeswitches to English. In no case is there evidence that the language of the noun plays a part in determining its position in discourse. Since codeswitching and borrowing are clearly two distinct categories, and given the evidence from these two language pairs that lone items are not codeswitches, then unless justification can be found for positing another category, we infer that they are borrowings.

Of course, we have not yet ruled out the possibility that these items are constrained by a language system other than the one in which they are embedded. For this, we compare the ambiguous lone other-language items with unambiguous recipient-language items (including longstanding, widespread, dictionary-attested loanwords). If the former cannot be shown to differ from the latter, the borrowing hypothesis is supported.

Lone other-language items pattern like attested loanwords
Where the bilingual corpora include attested (unambiguous) loanwords, lone other language items are generally found to pattern with them. Both Ghafar Samar and Meechan, and Adalar and Tagliamonte, show that lone English-origin nouns are not different from attested loanwords from English in Persian and Turkish. If lone items pattern like attested loanwords, there is no reason to posit another category.

Lone other-language items pattern like their recipient-language counterparts

Morphological measures. The patterning of lone other-language items generally reflects that of like recipient-language items. Budzhab-Jones shows that English-origin nouns in otherwise Ukrainian discourse are inflected with Ukrainian case markers following the same system speakers use to inflect Ukrainian nouns in Ukrainian discourse. The English system is nowhere in evidence. This is equally true of the other language pairs studied here. Ghafar Samar and Meechan show lone English-origin nouns to employ Persian reference markers at the same rate as native Persian nouns. Adalar and Tagliamonte show rates of Turkish plural marking of lone English-origin nouns to be identical to that of native Turkish nouns. Rates and selection of verbal inflections on lone English-origin verbs replicate almost exactly those of Igbo verbs (Eze). In all of these cases, establishment of language membership is facilitated by the presence of overt morphological clues in the recipient language. Elsewhere no such clues may be available.

Bare forms: Even where lone other-language items surface bare, the comparative method enables us to determine their status. Bare forms have figured prominently in the formulation of codemixing theories, where they are frequently cited as examples of exceptional bilingual strategies for incorporating foreign material. In contrast, Eze shows that lone English-origin forms, with respect to both morphological and syntactic criteria, surface only in context where bare forms are permitted in Igbo, and more tellingly, at exactly the same rate. The
appearance of untested bare English-origin forms in Persian discourse is likewise shown by Ghafar Samar and Meechan to be fully consistent with the patterns observed for attested loanwords from English. Turpin finds that bare English-origin nouns in French discourse surface at the same rate as (unmixed) French nouns appearing with null determination. At the same time, they differ diametrically from the rates of zero determiner displayed by English nouns within unambiguous codeswitches to English.

Where the (unmixed) comparison varieties are those spoken by the bilinguals who produced the mixed forms under study, like those contained in the corpora analyzed here, as opposed to some intimated or idealized variety, it becomes clear that bilingual structures that may seem exceptional or odd by comparison to standardized forms of the source languages generally conform closely to one, the other, or both of those languages. The related, though rarely discussed, issue of productivity may also explain speakers' apparent use of unusual morphological strategies to integrate lone other-language items.

Productivity. Such strategies are often invoked as exceptions to recipient-language rules, that is, as evidence of lack of integration into recipient-language grammar. This is nicely illustrated by the primeface exceptional noun class assignment to foreign-origin items in Wolof encountered in earlier research into lone L4 items in L3 discourse (Poplack & Meechan, 1995). Lone French-origin nouns are assigned exclusively to the class marker hr, in contrast to native Wolof nouns, which are purportedly distributed among the set of eight Wolof nominal class markers. However, closer inspection revealed that over half of the unmixed Wolof nouns in the Wolof-French corpus were also classed with hr, some nonrandomly. No other marker appeared in nonstandard construction. Now Wolof noun class distinctions have been neutralized for at least a century (Irving 1974). So, Wolof nouns classified with any marker other than hr are relics of the once productive system. It is thus unsurprising that virtually all French-origin items, which postdate the change, co-occur with hr. For a foreign-origin item to co-occur with a recipient-language inflection, the rule assigning that inflection must be productive in the recipient language at the moment of borrowing. Inherent variability and/or internal change in the recipient language must be ruled out before characterizing nonstandard behavior of foreign-origin items as unusual or exceptional codeswitching strategies.

Inherent variability. Nonstandard case-marking on lone English-origin nouns in Ukrainian discourse is a classic example of the importance of locating inherent variability prior to explaining apparent exceptions. Budzik-Jones shows that the ostensible failure of lone English-origin nouns to integrate simply reflects the variable (and not the prescribed categorical) nature of gender assignment in spoken Ukrainian, rather than any extraordinary morphological process associated with codeswitching.

Related to this is the question of phonetic criteria for determining integration. These are well documented as poor predictors of loanword (or codeswitch) status (e.g., Mouton, 1981; Myers-Scotton, 1993; Poplack et al., 1988). One reason is that bilingual speakers often produce accented forms in their second language even when no codeswitching is taking place. Moreover, the extent to which the phonetic (or underlying) form appearing in mixed environments is determined by active grammatical rules is rarely, if ever, established. We are thus unable to ascertain whether the speaker failed to adapt a lone item to recipient-language phonotactic constraints because it is a codeswitch, or because s/he has borrowed both the word and its phonemes, in much the same way as speakers sometimes treat phonologically single units (e.g., bare French-origin infinitives in a variety of language pairs). Where a productive phonological rule involving alternate forms can be observed to operate in the recipient language, as in Igbo (Eze) and Turkish (Adalat & Tagliamonte) vowel harmony, phonological integration is seen to proceed in the same way as the traditionally preferred indices of morphological and syntactic integration. Borrowings only rarely undergo grammatical processes that do not consistently apply to native items.

These results illustrate the utility not only of examining the unmixed vernaculars entering into the contact situation, but also of using quantitative methodology to reveal inherent variability and ensure that our structural diagnoses of system membership are valid.

Some lone other-language items are not borrowings

The method we have outlined here also reveals that a small minority of lone other-language items may not pattern with recipient-language counterparts or established loanwords. For example, in Poplack and Meechan's (1995) study of Wolof-French, lone French-origin nouns patterned with unmixed Wolof nouns for all forms of nominal modification except indefinite marking. In fact, lone French-origin nouns surfaced with Wolof indefinite markers at a rate exceeding that for Wolof nouns in unmixed Wolof contexts. Now Wolof indefinite markers, which are proposed, are the only Wolof determiners that are structurally equivalent to French determiners. Interestingly, it is only at these sites that lone French-origin terms approach the determination rate shown by their French counterparts. We hypothesized that this intermediate rate of indefinite marking of French-origin nouns was due to the inclusion among them of single-word codeswitches to French (at the equivalence site between Wolof proposed indefinite articles and French nouns). Turpin confirms and builds on this finding with a detailed analysis of plural marking of lone English-origin nouns at French-Wolof equivalence sites. She finds that patterns consistent with English grammar are most common in lone English-origin nouns which, on independent analysis, do not show the extralinguistic characteristics of loanwords (i.e., recurrence, diffusion, and dictionary attestation), and which had in addition been uttered by speakers also showing high rates of unambiguous (multitword) codeswitches to English. These facts, taken together, make it likely that at least some of the lone English-origin nouns in French discourse are also codeswitches.

Most lone other-language items are borrowings

The most recent finding in the papers collected here is that lone major-class content words of one language incorporated in discourse of another are almost always borrowings. Only Turpin's paper just cited could find, on one diagnostic, and for a small subsample of plural English-origin nouns, any evidence to the contrary. This, despite exhaustive quantitative analysis and comparison with multiple source varieties. In fact, our entire research program has uncovered only one other situation where a set of lone items could be unambiguously classified as codeswitches: French adjectives in English discourse (Meechan & Poplack, 1995). Interestingly, we are aware of no theory that predicts codeswitching these lone other-language items from the domain of codeswitching; it is simply a total bilingual behavior that only exceedingly rarely do they function as such.
Summary

This issue of the International Journal of Bilingualism, entirely consecrated to the study of lone other-language items, is testimony to the utility of a consistent scientific method. Lone other-language items have been widely documented as the most prevalent type of code-mixing, and the corpora on which the papers assembled here are based bear this out. Despite their controversial status in the literature, owing to no small part to their inherent ambiguity in isolation, there is little doubt as to the classification of these items as a set. The rigorous quantitative methodology employed by the authors provides strong support for a borrowing analysis and limited evidence for code-switching behavior. Convergence on this result is unanimous, no matter how stringent the comparison (with two, four, or five language contexts) employed. Whatever the linguistic properties of the language pair examined, ranging from typologically distant to nearly identical, and the diagnostic employed—phonological, morphological, or syntactic—lone other-language items overwhelmingly surface with the patterns of the language in which they are incorporated. This is true not only of the grosser linguistic structures, but more remarkably, of the fine details of quantitative conditioning of linguistic variability. This is evidence that they have been borrowed into that language, despite the lack, in some cases, of any dictionary attestation or diffusion within the community.

The fact that, in most situations, lone other-language items can be shown to be borrowings should not obscure the heuristic value of the method we propose. It permits us not only to disambiguate code-mixing phenomena, but also provides valuable insights into the structures they may enter into. Where we have located clear evidence of systematic single-word code-switching, the quantitative results additionally reveal how they are constrained by linguistic structure: the majority appear at permissible (i.e., Equivalence) codeswitch sites, and relatively few at prohibited ones.

The papers in this issue also illustrate the versatility of the variationist method for hypothesis testing. This is because it is not tied to any particular theory of code-switching (cf., also Laks, 1992). As explicitly illustrated by at least two contributions, it is applicable to any theory making testable claims about system membership of particular categories of code-mixing. Ghafar Sanar and Meechan test the fit of the Null Theory of Code-switching (Mehoojian, 1993) to their Persian-English data. Budzik-Jones provides a detailed assessment of the evidence in Russian-English for word-interval codeswitching, an assumption underlying any theory that permits switching between stems and morphemes (e.g., the Matrix Language Frame model or the Null Theory).

In sum, we agree with Grosjean (1995) who cautions that lone items cannot simply be collapsed with other types of code-mixing into one "global category" (p. 263). Given that they constitute the greatest portion of code-mixing phenomena, to include them in a theory of code-switching (or borrowing) without appropriately determining their status, as is commonly done in the (nonvariationist) literature, may obscure any patterns of true code-switching that exist, particularly since, as the papers assembled here show, they are most likely to be borrowings.

Despite some skepticism in the field as to the feasibility of distinguishing between code-switching and borrowing (e.g., Boyd, Andersson, & Thornell, 1991; Eliasson, 1990; Gardner-Chloros, 1987; Mehoojian, 1993), the papers in this issue all illustrate how this can be accomplished. They converge in confirming, for five different language pairs, that nonce borrowing differs from code-switching, and resembles established borrowing in all but its extralinguistic characteristics of recurrence and diffusion. These findings may be added to the accumulating evidence (Budzik-Jones & Poplack, 1997; Meechan, 1992; Meechan & Poplack, 1995; Meechann, Tagliamonte, & Poplack, 1996; Poplack, 1997; Poplack & Meechan, 1993; Poplack et al., 1988; Sankoff, Monpetit, & Dupont, n.d.; Sankoff et al., 1990) that nonce borrowings pattern exactly like their native counterparts in the (unmixed) recipient language, and not like elements of the language of their etymological origin. This confirms the Nonce Borrowing hypothesis for lone other-language items (Sankoff et al., 1990) and provides further testimony to the amazing productivity of bilingual borrowing.

References

Borrowed Nouns; Bilingual People: The Case of the “Londrali” in Northern Cyprus

Nevin Adalar and Sali Tagliamonte
Eastern Mediterranean University and University of York

Acknowledgments

We thank audiences at the University of York staff/student seminar and the 1997 International Symposium on Bilingualism for their comments. We are grateful to members of the Londrali community for their contribution to this work.

Abstract

In this paper we provide a quantitative analysis of the behavior of nouns in two generations of speakers from a bilingual community in northern Cyprus. Diagnoses from three areas of grammar — phonology, morphology, and syntax — provide corroborating evidence that when a loan noun, either English or Turkish origin, appears in contexts in which it is surrounded by the other language, it patterns systematically in accordance with its counterparts in that other language. On the other hand, when a loan noun, either English or Turkish, appears within a multieword fragment of English or Turkish, it patterns overwhelmingly with the language of its etymology. The strikingly different grammatical patterns reflect the fact that they represent two different types of behavior: borrowing in the former, codeswitching in the latter. These findings demonstrate that empirical investigation can disambiguate the community-specific status of language contact phenomena. Further, contrastive analysis across generations reveals that although the generations differ in the frequency with which they use English and Turkish, they employ the same strategies when they do. These strategies have far more to do with their preferred language of discourse, than with any purely linguistic factor.

Key words

codeswitching
loanword integration
nucleus borrowing
Turkish-English

1. Introduction

While the methods and questions addressed in the literature vary considerably, perhaps the most uncontroversial issue in the study of language contact is that social context influences language use. Research in this area has demonstrated that bilingual speakers differ dramatically depending on social circumstances and a host of other cultural factors (e.g., Poplack, 1985, 1987; Milroy & Li, 1995). But how these extralinguistic effects impinge on the micro-level details of codeswitching and/or borrowing phenomena is relatively little studied (but see, e.g., Poplack, 1980, 1988; Poplack, Sankoff, & Miller, 1988; Tetreau-Daller, 1997; Budzynska-Jones & Poplack, 1997). Such research is particularly relevant to ongoing investigations which attempt to elucidate how codeswitching and borrowing as such can be classified (Eze, 1995a, 1995b; Poplack & Meechan, 1995; Turpin, 1996, Meechan &...