Language in Society **34**, 493–531. Printed in the United States of America
DOI: 10.1017/S0047404505050190

Spelling bilingualism: Script choice in Russian American classified ads and signage

PHILIPP SEBASTIAN ANGERMEYER

Department of Linguistics
New York University
New York, NY 10003-6860
psa208@nyu.edu

ABSTRACT

This article investigates the role of script choice in bilingual writing, drawing on classified advertisements and other texts written for and by Russian-speaking immigrants in New York City. The study focuses on English-origin items that appear in Russian texts, which are found to be written either in roman or Cyrillic script. Through an investigation of categorical and variable constraints on this variation, it is found that script choice relates to the distinction between lexical borrowing and single-item codeswitching. It is argued that writers may, consciously and on a token-by-token basis, choose the Cyrillic script to mark a word as borrowed or the roman script to mark it as foreign. However, they may also avoid this choice, as hybrid forms attest, especially when the use of characters shared by both alphabets allows ambiguous readings. The findings thus have implications for understanding notions of language boundaries in bilingual language use. (Writing systems, Russian, English, codeswitching, borrowing, hybridity.)*

INTRODUCTION

The attribution of lexical items to particular languages is a crucial issue for research on language contact phenomena, even if it is not always acknowledged as such. In fact, there are many instances where it is difficult, if not impossible, to tell which language a given word belongs to. Bilinguals frequently produce forms that diverge from monolingual norms of either language, as well as other forms that conform to both norms at the same time. The classification of such divergent elements and the question of whether they can be classified at all have been controversial among linguists studying bilingual language use, and this topic has proven to be of great theoretical importance for the advancement of research in this field.

In particular, debates in language contact studies have centered on the classification of single lexical items, most commonly nouns, which are perceived as originating in one language but which occur in structures of another language. Have such words become part of the lexicon of the receiving language (i.e., lexical borrowing), or does their use indicate that a speaker has temporarily

switched from one language to another (i.e., single-item codeswitching)? This distinction plays a crucial role in the debate surrounding two conflicting models of codeswitching: on the one hand, those that see codeswitching primarily as an alternation between two grammars (e.g., Poplack 1980), and on the other hand, those that see it as an insertional phenomenon in which elements from one language are inserted into the grammatical structure of another language (e.g., Myers-Scotton 1993). While proponents of both models seek to distinguish between codeswitching and borrowing, they differ markedly in how they attribute tokens to one or the other category. In the alternational view, a word is borrowed if it behaves like a native word with respect to the categorical and variable constraints in the grammar of the recipient language. This is taken to be true even for lexical items that occur only once in a given corpus, known as "nonce-loans" (Sankoff et al. 1990). In the insertional view, by contrast, codeswitches and borrowings differ not on the surface but only in their status in relation to the mental lexicon of the recipient language, because borrowed forms are entered in this mental lexicon but codeswitched forms are not (Myers-Scotton 1993:207). The argument over codeswitching and borrowing is thus very much a question of the attribution of a given lexical item to a particular language and of defining boundaries between the languages in contact. What both models have in common is that they treat lexical items as inherently "belonging" to a given language. They do not see the attribution of lexical items to languages as an active process that speakers may be conscious of and have control over.

Investigations into these and other language contact phenomena have relied almost exclusively on spoken data, as revealed by the title of Muysken's (2000) book *Bilingual speech*. One reason for this may be that contact phenomena such as codeswitching are generally regarded as characteristic of informal speech styles, which can be observed only in informal settings. From this perspective, written language does not appear to be a suitable site for codeswitching research because it is generally more formal and more standardized than spoken language. For example, Milroy & Milroy (1991:66–67) emphasize that the ideology of "correct" orthography strongly inhibits variation in written language. This also relates to language choice: Sebba 2002 identifies what he calls "the tyranny of written monolingualism," a norm which dictates that printed texts have to be monolingual. He argues that written language alternation can be found only in unregulated, "peripheral" genres of writing such as graffiti, advertisements, or computer-mediated communication, where the ideology of language standardization is less powerful or is openly opposed. Poetry and fiction represent another exception, since codeswitching may be used intentionally as a literary device.[1]

In largely ignoring written data, the field of language contact studies is following a tradition in sociolinguistics that views informal speech as the ideal source for sociolinguistic data. Linguists rarely pay attention to written language use, and if they do, they often don't study it for its own sake, but rather as a substitute

for spoken language when speech data are not available. However, in recent years sociolinguists have begun to pay more attention to writing as they have come to recognize variability in orthography as a socially conditioned phenomenon (e.g., Jaffe 2000). One form of variability in writing exists in the possibility of variation between different writing systems. The use of multiple writing systems within the same speech community has been termed DIGRAPHIA (Zima 1974, Dale 1980, De Francis 1984, Grivelet 2001), by analogy with "diglossia" (Ferguson 1959). Script variation in digraphic contexts has been the subject of some sociolinguistic research, for instance on variability in written Japanese (Smith & Schmidt 1996). However, in allusion to Fishman's (1967) extension of the notion of diglossia, it can be noted that there is also "digraphia with bilingualism," where language contact involves two languages that are commonly written in different scripts. In fact, this is the case with many instances of language contact throughout the world, but relatively few studies of script variation have been conducted in such contexts – for example, involving English in contact with Hebrew (Lubell 1993), with Chinese (Cheung 1992, Li 2000), or with languages of the Indian subcontinent (Banu & Sussex 2001, LaDousa 2002). In such language contact situations, bilingual texts are produced which include elements from both languages and use both writing systems. Such alternation between writing systems within the same text has so far received little attention from linguists. Gazda (1998:163) uses the term GRAPHICAL TRANSPLANTATION (*grafická transplantace*) to describe the occasional occurrence of roman script for foreign words in Russian texts. Examining written data from Bangladesh, Banu & Sussex (2001:54) speak of GRAPHOLOGICAL CODE-SWITCHING.

In such bilingual writing, authors need to negotiate two different standardized orthographies, each one tied to a different writing system, a different visual form in which words may be represented. When authors attempt to integrate elements from both languages within one text, they may either alternate between writing systems, or they may transliterate words from one language into the writing system associated with the other language. In this, script choice parallels the distinction between codeswitching and borrowing. Codeswitching, as defined for example by Poplack & Meechan (1998:129), "should show little or no integration into another language." In other words, elements from different languages can be combined while maintaining their original form. In writing, this original form is dependent on a particular script that is typically used for a given language. Borrowed forms, in contrast, are nativized, made to fit the form of the recipient language ("'Borrowing' is the ADAPTATION of lexical material to the . . . patterns of the recipient language"; Poplack & Meechan 1995:200). In writing, this adaptation can be achieved by transliterating a word into the writing system of the recipient language, giving it a new form that fits the recipient's orthographic norms. The need to choose a particular writing system thus generally forces bilinguals to attribute a word to a given language, because each language is tied to a particular script and a particular orthography. When there is

uncertainty as to which language a lexical item "belongs" to, this orthographic choice may be taken as evidence that the writer attributes a form to a particular language, whether this reflects an unconscious categorization or an intentional choice.

In this article, I investigate the relationship between orthographic choice and language attribution by analyzing script choice in texts produced by and for Russian speakers living in the United States. Russian-English bilingualism provides an example of language contact involving different writing systems, as Russian is written in Cyrillic and English is written in roman script. Where English-origin items are included in texts that are otherwise in Russian, they may be represented in roman script, corresponding to standard English orthography, or they may be represented in Cyrillic script. In order to conduct a quantitative analysis of this script choice, I compiled a data set of 1,263 tokens, mostly from classified advertisements in Russian-American newspapers but also from news articles and advertising brochures. This data set was further supplemented with photographs of Russian-language signage. As I will discuss below, the data prove to be almost evenly divided between the two writing systems but also include a number of mixed forms in which elements of both scripts are combined. To determine the constraints that govern the variation in script choice, I conducted a multivariate analysis in which I identified both variable and categorical constraints. The results suggest that linguistic factors such as morphological integration play an important role in conditioning script choice. At the same time, the variation is also constrained by social factors, indicating that the written representation of linguistic items may be both variable and ideologically contested among members of a linguistic community. This is emphasized in particular by some exceptional cases in which authors avoid a clear choice by using elements from both writing systems within a single word, whether intentionally or unintentionally.

Demonstrating the variable nature of such categorizations in writing, the analysis of script choice thus has implications for bilingual speech as well. The findings of this study pertain to the role of language boundaries in bilingual language use, especially in that they illustrate the theoretical possibility that these boundaries remain variable – that is, that a particular lexical item may be attributed to different languages by different speakers, or even by the same speaker in different contexts. At the same time, the study addresses the differences between bilingual writing and bilingual speech by pointing to the role of standardization in written language in determining the degree to which ambiguity and hybridity are possible and metalinguistic categorization may be avoided.

I will begin with a general discussion of language use and script choice among Russian-speaking New Yorkers, focusing on variation in the written representation of English-origin lexical items. Drawing comparisons between the phenomena of lexical borrowing and transliteration, I will investigate the degree to which script choice can be taken as an indication of language membership. To do so, I will discuss the factors that constrain this variation, beginning

with linguistic factors, which may in part be taken to reflect unconscious metalinguistic categorizations, before turning to social factors, which point toward conscious sociolinguistic choice. In the final section, I will discuss hybrid forms which challenge both the cognitive and the ideological bases for the separation of the two languages.

LANGUAGE AND WRITING SYSTEM: ASSOCIATION AND DISJUNCTION

Over the past three decades, the number of Russian speakers in the United States has greatly increased, a result of continuous immigration from various parts of the former Soviet Union (see Gold 1995, Hinkel 2000, Orleck 2001). In the 2000 U.S. Census, approximately 700,000 respondents claimed to speak Russian at home. About one-third of them live in the Greater New York area, making Russian the fourth most commonly spoken language in the city, after English, Spanish, and Chinese.[2] The majority of Russian-speaking immigrants in the city live in southern parts of Brooklyn. These neighborhoods include not only the well-known "enclave" of Brighton Beach (Andrews 1998:6), but also other areas such as Midwood (Kings Highway), Sheepshead Bay, and Ditmas Park, where Russian speakers represent approximately one-fourth of the total population. In these neighborhoods, shop signs and other advertisements are mostly written in both Russian and English, as illustrated in Figures 1 and 2. Some stores, such as bookstores or music and video stores, advertise exclusively or predominantly in Russian, particularly if they specialize in Russian media products. Newsstands sell a wide variety of Russian-language publications, including numerous newspapers published in the United States, alongside newspapers and magazines imported from Russia.

Figures 1 and 2 show bilingual signage for various businesses in Brooklyn. Each sign consists of two parallel monolingual texts, one in English and one in Russian, which are translation equivalents. In each case, the English text is placed above the Russian text, but fonts or script size may make the Russian text more prominent (e.g., in the case of *flowers* and цветы/ *cvety*).[3] The images are emblematic of the fact that Russian-speaking immigrants in New York are confronted with two languages and two writing systems in their daily lives. Generally, as shown in Figures 1 and 2, the roman alphabet is used for English and the Cyrillic alphabet for Russian. However, as will be shown below, Russian-speaking immigrants alternate between the two alphabets in ways that transcend this distribution. In particular contexts, Russian words may be represented in the roman alphabet, or English words in Cyrillic. Figure 3 shows the possible distribution of languages and scripts, using the words *advokat* 'lawyer' and *lawyer*.

To examine the distribution of script choice in Russian-American writing, I collected data from Russian-language newspapers published in New York City, as well as from brochures, advertising, and signage written by and for Russian

FIGURE 1: Advertisement on Kings Highway, Brooklyn.

FIGURE 2: Advertisement in Brighton Beach, Brooklyn

speakers in the United States.[4] In newspapers, I considered primarily classified advertisements, but also articles about local events in New York. This was done to make sure that the texts were in fact written by authors living in the United States. These data were supplemented by photographs of advertising and other

https://doi.org/10.1017/S0047404505050190 Published online by Cambridge University Press

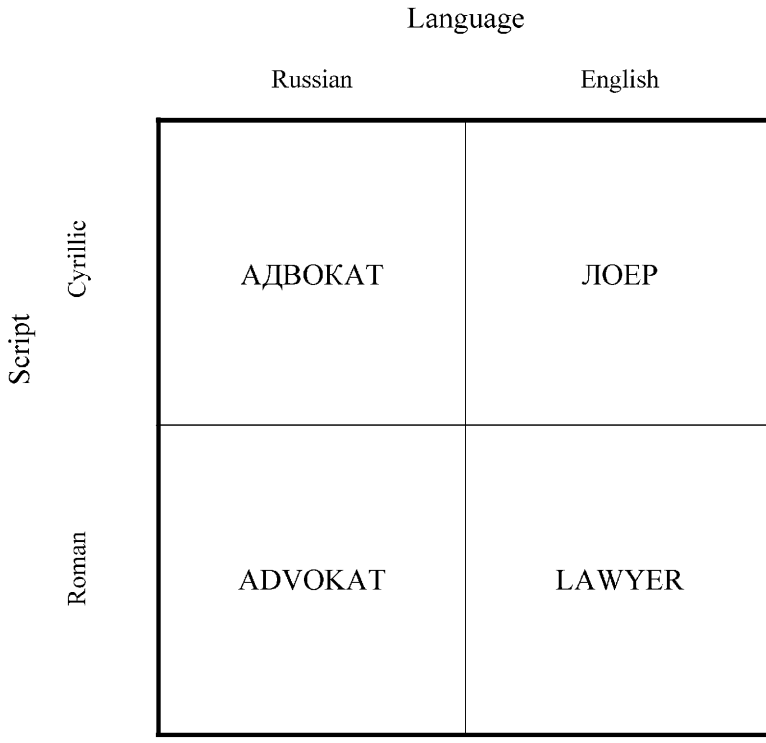

FIGURE 3: Combinations of languages and scripts (cf. LaDousa 2002: 224)

signage taken in various neighborhoods of southern Brooklyn. These photographs are included primarily for illustrative purposes; they are not part of the data set that was analyzed quantitatively. Advertisement has increasingly been recognized as a site of language contact (Banu & Sussex 2001, LaDousa 2002, Sebba 2002, Piller 2003) and as a rich source for sociolinguistic data, particularly with regard to the stereotyping of linguistic groups. At the same time, advertisement is arguably less constrained by norms of standardization than are other forms of printed written language. Sebba 2002 theorizes "the world of written texts … as a set of spaces in which the ideology of standardization is imposed to varying degrees," and describes advertisement as a less regulated space, where orthographic conventions can be broken and languages mixed. This would appear to be especially true of classified advertising, which differs from other genres of texts in that it is short-lived, tends to receive little or no editing, and is written for a small, highly specific target audience.

Of the four possible permutations shown in Figure 3, the use of Russian in Cyrillic script predominates in the data, which is not surprising since all news

FIGURE 4: *The Zamore*: Sign of co-op apartment building in Brighton Beach, Brooklyn (*zamor'e* 'country overseas, abroad').

FIGURE 5: *Primorski Restaurant* in Brighton Beach, Brooklyn (*primorskij* 'by the sea').

articles and nearly all advertisements found in Russian-American newspapers are written in Russian. By contrast, transliteration of Russian words into roman script is very rare. It is found with Russian words that are part of proper names, particularly when the intended audience includes non-Russian speakers, for example in signage of Russian-owned businesses or institutions, or in the publishing and pricing information underneath the title of a newspaper. Figures 4 and 5 show uses in signage. In addition, it is found when the use of the Cyrillic alphabet is not available for some pragmatic reason, for example in 1-800 telephone numbers or web addresses (see also Fig. 18).

The representation of English words in Russian texts is less predictable. As shown in Figure 3, English words or proper names may either be transliterated into Cyrillic script, or they may be maintained in their original spelling

FIGURE 6: Sign in a store in Brighton Beach, Brooklyn.

FIGURE 7: Sign in a store on Ditmas Avenue, Brooklyn (фудстемпы *fudstempy*).

in roman script. In my data, there is considerable variation between the two strategies. This is illustrated in Figures 6–9, which show instances of the English compound noun *food stamp* used in Russian sentences, in either Cyrillic or roman script.

https://doi.org/10.1017/S0047404505050190 Published online by Cambridge University Press

FIGURE 8: Sign in a store in Brighton Beach, Brooklyn.

FIGURE 9: Sign in a store in Brighton Beach, Brooklyn (фудстемпы и кредит карты *fudstempy i kredit karty*).

All the images include signs that inform customers about a store's policies regarding food stamps.[5] The Russian text in Figure 6 can be transliterated and translated as follows: *kulinarija na* FOOD STAMP *ne prodaetsja* 'prepared food is not sold for food stamp'.[6] The other signs all include the words *my prinimaem* 'we accept,' followed either by *food stamps* in roman script (Fig. 8) or by *fudstempy* 'food stamps' in Cyrillic (Figs. 7 and 9), with Russian case marking for the accusative plural. The signs in Figures 6 and 7 are both bilingual signs in English and Russian, of the type discussed above (see Figs. 1 and 2), which consist of two separate parallel texts. The sign in Figure 8 is also bilingual (*we accept* is written in small letters at the top), but *food stamps* is written only once, being simultaneously part of two different utterances, one in Russian and one in English.[7] The sign in Figure 9 is entirely in Russian. The texts differ with respect

https://doi.org/10.1017/S0047404505050190 Published online by Cambridge University Press

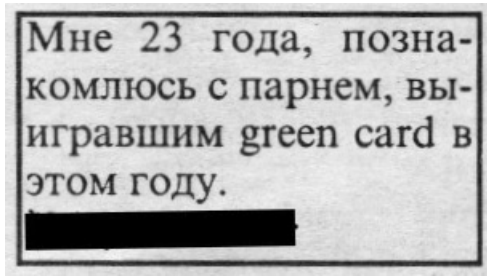

FIGURE 10: Personal ad in *Vecherniy New York*, January 24–30, 2003.

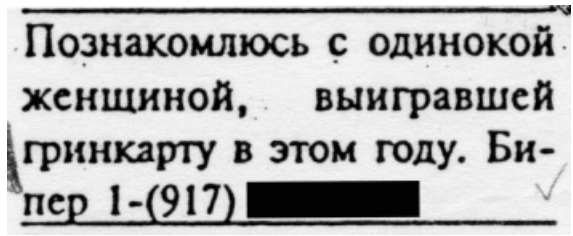

FIGURE 11: Personal ad in *Russian Bazaar*, January 18, 2002 (гринкарту *grinkartu*).

to the representation of *food stamp*: While it is transliterated into Cyrillic in Figures 7 and 9, the texts in Figures 6 and 8 maintain the English orthography in roman script, thus alternating between the roman and Cyrillic alphabets within the same sentence. Figures 6 and 8 thus represent examples of "graphological code-switching" (Banu & Sussex 2001: 54).

Figures 10 and 11 show similar examples from personal ads placed in Russian-American weekly newspapers. The text in Figure 10 has the English lexical item *green card* in roman script, and Figure 11 has it in Cyrillic. The two forms occur in virtually identical syntactical environments: In both cases the noun is the direct object of the same participle (*vyigravšim/vyigravšej* 'having won+INSTRUMENTAL.m./f.') and is followed by a prepositional phrase. The ad in Figure 10 can be transliterated as *Mne 23 goda, poznakomljus' s parnem, vyigravšim GREEN CARD v ètom godu* 'I'm 23 years old, looking to meet a guy, who has acquired the green card this year'. The text in Figure 11 reads *Poznakomljus' s odinokoj ženščinoj, vyigravšej grinkartu v ètom godu. Biper* 'Looking to meet with a lonely woman, who has acquired the green card this year. Beeper'.

As mentioned above, the categorization of lexical items like *food stamp* or *green card* in the above examples has been the subject of much controversy among

researchers studying bilingual speech. While it is a fairly trivial task to identify such items as having an "alien" origin that distinguishes them from the structures in which they appear, linguists differ in their interpretation of the processes that lead to the usage of these lexical items in their given context. Muysken (2000:69), for example, distinguishes between the processes of "inserting alien words or constituents into a clause" and "entering alien elements into a lexicon."

The question of whether or not a form has been entered into a lexicon has traditionally, though not exclusively, been approached from the perspective of the linguistic community, with forms being uncontroversially classified as "established" borrowings (or loanwords) if they are habitually used by speakers of a given linguistic variety (Myers-Scotton 1993:16, Poplack & Meechan 1995:200). Codeswitching, in contrast, is generally illustrated by examples used by individual speakers in particular contexts, especially when the argument is made that an "alien" word is inserted in order to invoke specific social connotations that are associated with the word's language of origin (see, e.g., Auer 1998:7). The fact that these related phenomena are most easily approached from contrasting perspectives points to the interplay between the community norms concerning the metalinguistic categorization of lexical items on one hand, and the attribution of social values to particular usages on the other. It suggests that the social significance of individual usage is more salient if it contrasts with community expectations – that is, if a lexical item is used in a way that differs from the way in which it is used most commonly by other bilingual speakers of the same varieties. However, this should not preclude the theoretical possibility of widespread, repeated codeswitching, or isolated instances of ("nonce") borrowing.

As hypothesized above, script choice can provide a new perspective on such questions owing to the association between languages and writing systems. Community norms are particularly relevant to written language use, yet script choice also provides individual authors with a highly salient tool for identifying the language to which they choose to attribute a given lexical item, regardless of whether this choice contrasts with community norms. In order to investigate the relationship between script choice and language attribution, I set out to conduct a statistical analysis of script choice for English-origin items in Russian texts. I collected examples of English lexical items which appeared in Russian sentences, mostly from classified advertisements in Russian-American newspapers. Altogether, I compiled a data set of 1,263 tokens representing 282 different lexical items. These items were almost exclusively nouns or noun phrases.[8] As shown in Table 1, the tokens are quite evenly divided between the two alphabets. In addition, a small number of examples are orthographically mixed. I will return to them later, after discussing the factors that condition the variation between the two major, unmixed categories.

In the following section, I will discuss a number of different criteria that have been described in the codeswitching literature as characteristic of borrowing (see Muysken 2000:60–85). I will show how they relate to the data with respect to

TABLE 1. *Script choice for English-origin items in Russian texts.*

Script	Tokens	Example
Roman	646	*green card* in Figure 10
Cyrillic	605	гринкарту *grinkartu* in Figure 11
Mixed	12	барбер *shop*, *barber* SHOP in Figure 17

script choice, beginning with a characteristic that is at the heart of the codeswitching debate: the question of morphological adaptation of foreign-origin items.

MORPHOLOGICAL ADAPTATION AND SCRIPT CHOICE

As Table 1 shows, English-origin items in Russian texts occur frequently in either script. However, when we classify the tokens according to certain linguistic characteristics, clear preferences for one or the other script emerge. In particular, case marking constrains alphabet choice in a powerful way. Like all other Slavic languages, Russian has a rich nominal inflectional system in which noun phrases are marked for case, gender, and number, and there is agreement between nouns, adjectives, and demonstratives, but also between the NP and its associated participles or relative pronouns. In a study of Ukrainian-English bilingual speech, Budzhak-Jones 1998 notes that English-origin nouns frequently, but not always, receive overt Ukrainian morphology. Gregor (2002:72) makes a similar observation for Russian-English bilingual speech. The same is true for the written data under consideration here: When an English noun is used in a Russian sentence, it may or may not receive a Russian case ending. However, variation between these two alternatives strongly correlates with script choice. In those environments where Russian syntax requires overt case marking, it never shows up on English nouns that are written in roman script (0/225), but it is almost always present on those written in Cyrillic script (367/379), as shown in Table 2.

This finding can be illustrated by comparing Figures 12 and 13, both classified ads containing the English word *housekeeper*. Figure 12 reads *Opyt, rekomendazii, rabotu* HOUSEKEEPER 'experience, recommendations, work as housekeeper'. The text in Figure 13 can be transcribed *Ol'ga iz Sankt-Peterburga iščet rabotu xauskipera ili pomošč' na domu* 'Olga from Saint Petersburg seeks work as a housekeeper or help in the house' Both instances of the English-origin noun *housekeeper* appear in the same syntactic construction, as complements of the noun работу *rabotu* 'work+ACC.,' for which the genitive case is required. However, while the Cyrillic form хаускипера *xauskipera* (Fig. 13) receives the genitive case ending *-a*, the roman form in Figure 12 does not. The same can be observed with *green card* and гринкарту *grinkartu* in Figures 10 and 11, or with

TABLE 2. *Case marking and script choice of English-origin NPs in Russian texts.*

	Alphabet			
Morphology	Cyrillic		Roman	
Overt case marking present	367	100%	0	
Overt case marking required but absent	12	5%	225	95%
No overt case marking required	226	34%	421	64%
TOTAL	605	48%	646	51%

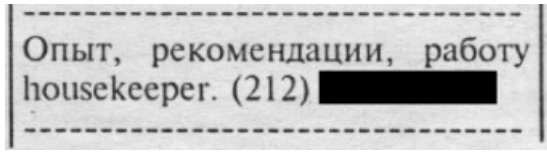

FIGURE 12: Classified ad in *Russkaya Reklama*, January 24–30, 2003.

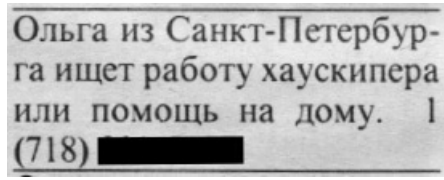

FIGURE 13: Classified ad in *Russian Bazaar*, April 10-16, 2003 (хаускипера *khauskipera*).

food stamps and фудстемпы *fudstempy* in Figures 8 and 9. Again, the contrasting forms occur in the same syntactic environment, but only the transliterated forms receive Russian case marking. This finding is clearly reminiscent of the so-called Free-Morpheme Constraint proposed by Poplack 1980, in which she posits that codeswitching cannot occur between a root and a bound morpheme, as illustrated in ex. (1):

(1) *EAT – iendo (Poplack 1980:586)

In a similar way, the findings of this study suggest that alternation between scripts cannot occur between a root and a bound morpheme, as illustrated by the

contrast between (2a) and (2b). Of the 604 English-origin NPs for which Russian requires overt case marking, 367 (61%) follow the pattern in (2a): the case-marking suffix is present and both the root and the suffix are in Cyrillic. The data set includes no instances of the pattern in (2b) – a root in roman script and a case-marking suffix in Cyrillic.

(2a) с гринкарт – ой
s grinkart – oj
with greencard – INSTR.
'with a greencard'

(2b) **c greencard*-ой

(2c) **c greencard-oj*

However, this constraint is not just directed at alternating scripts, since forms like (2c), with the case morpheme transliterated, are not found either. Clearly, the presence of Russian morphology in a Russian text requires the use of the Cyrillic script. If an English NP occurs in the roman alphabet, it is never case-marked. An example is given in (3a), which is representative of 37% of those tokens for which case marking is required by Russian syntax. By contrast, forms in Cyrillic where overt case marking is absent – as in (3b) – are rare. Only 12 of the 604 NPs in the data set (2%) fit this pattern.

(3a) *c greencard*

(3b) с гринкарт

Although numerous authors have reported counterexamples to the Free-Morpheme Constraint in bilingual speech (e.g. Eliasson 1990; Bentahila & Davies 1991; Myers-Scotton 1993:30–34), the categorical absence of overt case marking on roman script forms points toward an important difference between writing and speech. This suggests that bilingual writing is more constrained than bilingual speech when it comes to combining elements from two languages within one word, a finding that clearly corresponds to observations on the limits of variability in writing (Milroy & Milroy 1991, Sebba 2002). At the same time, the discreteness of the two languages is arguably more salient in bilingual writing than it is in speech because standard orthography (and in the biscriptal case, the association with a particular writing system) marks words as belonging to a given language. This in turn suggests that the proposed Free-Morpheme Constraint has explanatory relevance after all, despite being "violable" in spoken language use. Instead, its validity may be taken to depend on the degree to which the discreteness of the languages is contextually salient.

In Poplack's analysis, the morphological adaptation of foreign-origin lexical items is taken as evidence that such items are borrowed. Likewise, Budzhak-Jones 1998 concludes for her data that English nouns with Ukrainian case marking must be examples of lexical borrowing because they display all properties of Ukrainian-origin nouns. In my data set, inflected English nouns like *green card*

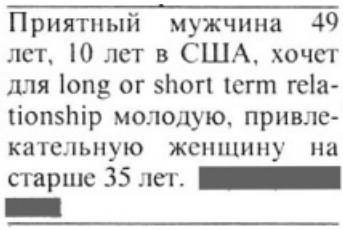
Приятный мужчина 49 лет, 10 лет в США, хочет для long or short term relationship молодую, привлекательную женщину на старше 35 лет. [redacted]

FIGURE 14: Personal ad in *Russian Bazaar*, April 10–16, 2003.

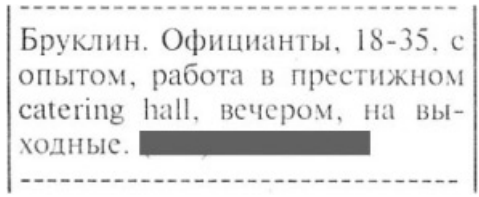
Бруклин. Официанты, 18-35, с опытом, работа в престижном catering hall, вечером, на выходные. [redacted]

FIGURE 15: Help wanted ad in *Russkaya Reklama*, January 24–30, 2003.

in (2a) share an additional property with Russian-origin nouns: script choice. Following this line of analysis, the data thus suggest that writers use script choice to mark a lexical item as belonging to a given language; that is, they use the Cyrillic alphabet to mark a word as borrowed.

Figures 14 and 15 are included to illustrate instances where case marking is absent after prepositions that require a specific case. Figure 14, a personal ad, contains the following text: *Prijatnyj mužčina 49 let, 10 let v SŠA, xočet dlja* LONG OR SHORT TERM RELATIONSHIP *moloduju, privlekatel'nuju ženščinu na* [sic] *starše 35 let* 'pleasant man, 49 years old, 10 years in the US, wants a young, attractive woman, not older than 35, for a long or short term relationship'.[9] The English NP *long or short term relationship* is not case-marked despite being the object of the preposition *dlja*, which requires the genitive case.

Figure 15 contains a help-wanted ad which can be transliterated and translated as follows: *Bruklin. Oficianty, 18–35, s opytom, rabota v prestižnom* CATERING HALL, *večerom, na vyxodnye* 'Brooklyn. Waiters, 18–35, with experience, work in prestigious catering hall, evenings, on weekends'. Here the English compound *catering hall* is not overtly case-marked for the prepositional case, but the adjective *prestižnom* is. However, the suffix *-om* expresses not only prepositional case but also masculine or neuter gender, which it receives from the noun. Thus, even though *catering hall* is not morphologically adapted itself, it appears to have an inherent Russian gender that the adjective *prestižnyj* can agree with.[10]

Forms such as *catering hall* (Figure 15), which lack system morphemes required by the recipient language, have been termed BARE FORMS in the codeswitching literature, but their classification as borrowing or codeswitching has been subject to debate. Myers-Scotton (1993:192), for example, states that "it may be that, while both B [borrowing] and CS [codeswitching] forms may be bare forms, there are significantly more such forms under CS [codeswitching]." Poplack & Meechan (1995:222) suggest that bare nouns are best classified as belonging to that language which shows a higher frequency of bare nouns in unmixed speech (which would surely be English here, given Russian case-marking). Again, if we take script choice as a marker of language attribution, both suggestions are supported here, since bare forms occur overwhelmingly in roman script (95%), as shown in Table 2. At the same time, the fact that such bare forms can participate in agreement phenomena shows that the distinction between codeswitching and borrowing has to be viewed as gradual rather than categorical.

Exx. (4) and (5) show apparent counterexamples to the generalization made in (2b), where Russian case marking was found only on roots written in Cyrillic. In these cases, quoted from Gazda (1998:164), a foreign-origin lexical item is rendered in roman script but receives a Russian case-marking suffix in Cyrillic. In these and all other examples given by him, the suffix is separated from the root by an apostrophe, orthographically marking the transition between the two scripts.

(4) CD-ROM'ных
CD-ROM'nyx
CD-ROM+adjectival-gen.pl.

(5) с вставными зубами и неизлечимым
s vstavnymi zubami i neizlečimym
with false+INSTR-PL teeth+INSTR-PL and incurable+INSTR-SG

tic'ом
tic'om
tic+INSTR-SG
'with false teeth and an incurable tic'
(A. P.Chekhov, *Skuchnaya Istoriya* 'A Dreary Story', 1889).

Gazda's examples come from newspapers published in Russia, as well as from Russian literature, in this case from Chekhov. These are texts produced for a monolingual Russian audience, albeit one with some knowledge of a socially prestigious second language. The different treatment of such forms in the bilingual Russian-American media as opposed to examples (4) and (5) thus hints at a difference between bilingual and monolingual texts. In particular, it suggests that the absence of case marking is less acceptable in monolingual language use than it is in bilingual language use, where the availability of a second grammar (here English) facilitates the temporary suspension of the structural constraints of Russian morphology.

However, both examples also point to the social meaning of script choice. The choice of roman script allows the author to invoke the prestige of English in the computer age (4) or the prestige of French in tsarist Russia (5), while defining the reader as one who shares specialized knowledge of prestige culture. Representing these words in Cyrillic arguably would not have the same effect. This raises the question: To what degree do authors make conscious use of script choice to mark words as Russian or as foreign? On the one hand, the categorical absence of case marking on roman forms in the data set suggests that script choice is to some degree a function of unconscious categorization. On the other hand, the attested variation between uninflected roman forms and inflected Cyrillic forms (e.g., *food stamp, green card*, or *housekeeper* in the examples above) shows that the metalinguistic categorization of these lexical items is nevertheless variable on the community level, and it suggests that authors may consciously choose a script to achieve a particular effect. In the following section, I will investigate variable constraints on script choice in order to assess the factors that lead authors to choose a particular written form.[11]

VARIABLE CONSTRAINTS ON SCRIPT CHOICE

Case marking thus constrains alphabet choice, but as indicated in Table 2, overt case marking is not always required by Russian syntax. This is often true when English nouns occur as subject or direct object, or in an apposition. In these instances, the same lexical form could be represented in either alphabet. In order to examine script choice in these environments, I removed all examples where case marking was required, as well as invariant frequent types. A data set of 514 examples remained.[12] To assess the factors that condition variation in script choice, I conducted a multivariate analysis using Goldvarb (Rand & Sankoff 1990). The results are shown in Table 3.

Two linguistic factor groups were identified by Goldvarb as statistically significant: the syntactic constituency of a lexical item, and the frequency with which it appears in the data. Both factors have also been proposed as diagnostic criteria for borrowing (cf. Muysken 2000:73). As can be seen in Table 3, the use of the Cyrillic alphabet is clearly favored for single nouns compared to compound nouns or multiword expressions.[14] In the debate over codeswitching and borrowing, it has often been argued that borrowing typically involves single lexical items but that compound nouns or multiword fragments tend to be instances of codeswitching (cf. Poplack et al. 1988:52; Poplack et al. 1990:71–72; compare also Myers-Scotton's [1993] notion of "EL-island"). Another linguistic factor identified by Goldvarb as statistically significant is the frequency with which a given lexical item is used. As Table 3 shows, the more frequently a lexical item occurs in the data, the more likely it is to be transliterated into the Cyrillic alphabet. Again, it is a commonplace assumption in language contact studies that frequent forms are borrowings (cf. Myers-Scotton 1993:16; Thomason 2001:134). The linguis-

TABLE 3. *Goldvarb results, use of the Cyrillic script for English lexical items in Russian texts in environments where variation was unconstrained by morphology.*[13]

	Factor Weight	Percent	n
Constituency			
Single noun	.704	40%	108/270
Compound noun/multiword expression	.277	9%	23/244
Frequency			
more than 20 occurrences	.726	42%	56/132
7 to 18 occurrences	.570	29%	38/133
1 to 6 occurrences	.339	15%	37/249
Source			
Novoye Russkoye Slovo/Russian Forward	.807	32%	8/25
Russian Bazaar	.756	41%	64/157
Russkaya Reklama	.466	24%	50/208
Vecherniy New York	.443	14%	6/43
Kurier	.047	2%	1/58
Type of Advertisement/Article			
Personal ad	.896	75%	12/16
Seeking job	.483	28%	30/109
Help wanted	.483	23%	89/389
Total: Input	.157	25%	131/514
Log likelihood = −212.176 Significance = 0.005			

tic factors found to condition script choice to a statistically significant degree thus suggest further parallels between transliteration and borrowing, since those types of forms that have been argued to be more likely to be borrowings favor transliteration, while the forms that have been argued to be likely codeswitches favor the use of the roman alphabet.

In addition to the criteria discussed so far, a number of further diagnostic criteria for borrowing have been proposed in the literature, including semantic change, synonym replacement, culture specificity, and phonological adaptation. They could not be tested on a quantitative basis, but I will briefly address them here. Weinreich 1968 noted early on that a borrowed word may undergo semantic change compared to its meaning in the donor language, typically involving a specialization of meaning. The data contain several examples where this may be the case. For example, *status* is used to refer to immigration status and occurs exclusively in Cyrillic in the data set (20/20), again suggesting a patterning of transliteration with borrowing.[15] Another example is the word *cash*, which is identified by Andrews (1998:10) as a borrowing whose meaning has been narrowed in the speech of Russian immigrants to describe a form of payment, but

https://doi.org/10.1017/S0047404505050190 Published online by Cambridge University Press

which cannot be used as a synonym for money more generally. In the data set it occurs in both scripts, but more frequently in Cyrillic (23/33).[16]

Furthermore, it is often argued that borrowing has occurred if native synonyms have been replaced (cf. Poplack & Sankoff 1984:129). I did not code tokens for this category because I did not investigate systematically the native Russian vocabulary used in the newspapers. However, the advertisements provided some evidence for cases where synonym replacement appears not to have occurred – for example, *housekeeper*, which occurs in both scripts but more often in roman (34/46).[17] Other such forms, for which a Russian equivalent was found in the data set, include *experience, waitress*, and *salesperson*, all of which occur only in roman script (5, 1, and 12 tokens respectively).[18] Judging from spoken usage, some lexical items could be identified where synonym replacement does appear to have occurred. These include some of the most frequent forms – *office* and *van*, which are invariably written in Cyrillic. These examples thus appear to confirm the parallel between borrowing and transliteration, in that English lexical items that occur alongside Russian equivalents tend to be spelled in roman script, but those that do not so occur tend to be spelled in Cyrillic.

A related phenomenon is the borrowing of culture-specific terms for which the recipient language does not have an equivalent, or so-called cultural borrowing (Myers-Scotton 1993:169). I did not code tokens for this category either, because the question of whether a Russian equivalent exists appeared often less than clearcut. For example, Andrews (1998:71–72) claims that Russian *njanja* is not "an appropriate equivalent" of *bebisiter* 'babysitter' because of the different connotations that it evokes. However, most English nouns are likely to have different connotations from their Russian counterparts, making the question of their equivalence one of degree, not one that can always be answered categorically.[19] Nevertheless, there are some terms that can be clearly identified as culture-specific because they refer to American institutions, customs, or policies. These include terms like *green card* or *food stamp*, which occur in either script, but also a number of items that occur only in roman script, such as *paralegal* (4), *tax season* (3), or *metro card* and *Social Security*, which occur only once.[20] There are too few clear examples to draw a conclusion, but they nevertheless suggest the possibility that the culture-specific nature of a referent, its obvious Americanness, may serve to preserve the roman spelling despite the fact that it is habitually used and is without "competition" from a Russian synonym.

Finally, another criterion that is regularly mentioned as characteristic of borrowed forms is phonological adaptation, though it may be found in codeswitching as well (cf. Poplack & Sankoff 1984; Myers-Scotton 1993:176). In writing, phonological adaptation appears to be an almost inevitable consequence of transliteration, since Cyrillic forms are pronounced according to the Russian pronunciation of the Cyrillic alphabet.[21] In particular, evidence of adaptation is found in the representation of English speech sounds that have no equivalent in Russian, such as [h].[22] On the other hand, occasional nonstandard orthography in

TABLE 4. *Characteristics of borrowing and codeswitching and their patterning with alphabet choice.*

Cyrillic Script (Transliteration)	Roman Script (Script Alternation)
recipient-language morphology	no recipient-language morphology
frequent use	infrequent use
single word	multiword expression
semantic change	no semantic change
displacement of recipient-lg. synonym	use alongside recipient-lg. synonym

roman forms may also suggest phonological adaptation,[23] and even the use of standard English orthography cannot be taken as evidence for the absence of phonological adaptation in spoken language use. In any case, the relationship between spelling and pronunciation is complex and indirect, making it difficult to draw conclusions from written data alone. As pointed out by Kress (2000:166), spelling (especially in English) relies not primarily on sound-letter correspondences but rather on "an understanding of words as visual units."

In summary, while the findings are not equally conclusive for all of the factor groups discussed in the previous section, a number of linguistic criteria have been identified that favor or appear to favor the use of the Cyrillic alphabet for English-origin lexical items. Table 4 gives a summary and schematic overview of the factors discussed in this section.

With the possible exception of cultural borrowing, it can be observed that the criteria found to be characteristic of transliteration are precisely those that have been described in the codeswitching literature as diagnostic for borrowing. Thus, it seems reasonable to conclude that the association between languages and writing systems extends to the distinction between borrowing and codeswitching. The findings suggest that a word is treated as Russian (i.e., as borrowed) if it is written in Cyrillic, and as English (i.e., codeswitched) if it is written in roman characters. In bilingual writing, alphabet choice may thus function as an indicator of metalinguistic categorization for a given lexical item.

This generalization has important implications for the interpretation of codeswitching in the data, as well as for language contact studies in general. First of all, we can note that the data set contains no instances of word-internal codeswitching, and that English-origin nouns with Russian case marking represent cases of borrowing. As mentioned above, this can be taken as evidence in support of Poplack's (1980) Free-Morpheme Constraint and of subsequent analyses by her and her associates. However, contrary to claims made, for example, by Poplack & Meechan (1998:135) that "most lone other-language items are borrowings," there are in fact many instances of single-item codeswitching. In the restricted

data set used for the Goldvarb analysis, 60% of single nouns are written in roman script; that is, they can be considered to be codeswitched, not borrowed.

Most important, though, the generalization forces us to explain why some lexical items occur in both scripts. As mentioned above, the topic of lexical borrowing has traditionally been approached from the perspective of the speech community, and linguists have generally categorized lexical entries across the board.[24] If a word occurs several times in the same corpus, it is either always classified as a borrowing, or never, though there is a strong tendency to treat multiply occurring forms as borrowings (e.g., Myers-Scotton 1993:204). However, if we accept a connection between borrowing and script choice, we have to conclude that forms such as *food stamp, green card*, and *housekeeper*, which may occur in either script, may be either borrowed or codeswitched. For example, the English-origin item *housekeeper* can appear either as a borrowed (Russian) form хаускипер *xauskiper* (Fig. 13) or as a code-switched (English) form *housekeeper* (Fig. 12). All in all, of the 282 different lexical items in the data set, 34 (12%) appear in both scripts. Also, several of the lexical items identified by Andrews 1998 as established loanwords in Russian émigré speech appear in both scripts in the data.[25] This suggests that the distinction between borrowing and codeswitching has to be made on a token-by-token basis and cannot be made across the board for each lexical entry, as is common practice.

This finding should not be surprising. If Russian-English bilinguals have borrowed the word *housekeeper* into Russian, this does not mean that the word ceases to be part of their English lexicon. Instead, it is now part of both mental lexicons, as noted by Muysken (2000:69): "Bilinguals dispose of two grammars and lexicons, and the lexicons can be viewed as one large collection that consists of several subsets. Thus lexical borrowing could be termed lexical sharing." Borrowed items thus remain available for codeswitching, and bilingual authors and speakers are able to choose between borrowing and codeswitching. As discussed above, this choice is constrained by a variety of factors and may be quite predictable. In fact, more often than not it may not be a conscious choice at all, since most factors described in the previous section can be assumed to constrain alphabet choice in a way that operates below the level of metalinguistic awareness. But no matter how well established a borrowed form is, the possibility remains for the bilingual speaker to treat it as foreign, and the bilingual author can do so through script choice.[26] However, if a speaker or author can consciously opt to mark a lexical item as belonging to a particular language, it follows that this categorization (and by extension script choice itself) must be available for social evaluation. This point will be explored in the following section.

SOCIAL EVALUATION OF SCRIPT CHOICE

In recent years, sociolinguists have increasingly come to view variation in writing as socially meaningful. Jaffe (2000:499), for example, writes that "ortho-

graphic choices and their interpretation are read as meta-linguistic, socially conditioned phenomena which shed light on people's attitudes towards both specific language varieties and social identities and on the relationship between linguistic form and the social world in general." With regard to multilingualism, LaDousa 2002 identifies the indexical nature of the language/script combinations Hindi/Devanagari and English/Roman in India. In keeping with such observations, this study finds variation in alphabet choice to be constrained by social factors as well as by linguistic ones. In the Goldvarb analysis shown in Table 3, both the source (the type of newspaper) and the genre of the text were found to be statistically significant factors. Transliteration into Cyrillic was favored by "quality" newspapers with low advertising content, while alphabet-switching was most frequent in newspapers with high advertising content and little original editorial content. Transliteration was also favored by personal ads as opposed to work-related ads, and it was categorical in local news articles, though too few tokens remain in the reduced data set.[27]

As mentioned above, Sebba 2002 notes that certain genres of writing are less regulated than others and therefore allow deviance from standard conventions, including language alternation, in contrast to regulated genres where such deviance is not permitted. He includes journalistic texts under the most highly regulated genres and advertisements under the less highly regulated genres of writing. His predictions are borne out in the data. Moreover, advertising in general may be seen as seeking to appeal to a specific group of potential addressees by invoking a common social, cultural, and linguistic identity. As such, it is perhaps more capable of reflecting nuanced identities than other forms of written language are.

It remains to ask just how unregulated advertisement writing is. The striking differences among some of the newspapers suggest that script choice may sometimes be a newspaper editor's decision rather than that of the individual placing the ad. However, there is also variation within newspapers: The same English-origin word may be found in both scripts on the pages of the same newspaper. But what might be the social motivation for writers or editorial boards to choose one alphabet over the other? Research on codeswitching has often shown that in multilingual societies languages are tied to social categories, and that the choices that individuals make between languages can be understood only in the context of these social categories. It can be assumed that the same holds true for lexical choices between borrowing and codeswitching, as well as for choices between scripts.

A recurrent theme in the ethnographic literature about Russian-speaking immigrants to the United States is the observation that they are on average highly educated, many of them being physicians, musicians, lawyers, or Ph.D.s (Levkov 1984:110; Gold 1995:48; Andrews 1998:54; Hinkel 2000:358; Orleck 2001:120). Having received their higher education in Russian, these immigrants can be expected to have prescriptivist attitudes toward language contact phenomena, and in fact, Andrews (1998:56) observes that "there are some very vocal purists …

who decry the intermixture of English into their native Russian."[28] However, my method of collecting data did not permit me to collect adequate information about authors' social and linguistic backgrounds, so any social interpretation of the data must remain speculative.[29] However, assuming the prevalence of prescriptivist attitudes, we can surmise that some bilingual authors find themselves in a bind: Alphabet-switching is "bad Russian" because it introduces a foreign element into a Russian text, but transliteration is writing "English with a Russian accent." One might expect that neither of these two strategies appeals to language purists. Cross-linguistically, prescriptivists have argued both against the introduction of foreign words and against their nativization into the conventions of the recipient language. The most striking example of the tendency to mark loanwords orthographically as foreign may be the use of *katakana* as a separate set of characters for loanwords in Japanese (see Smith 1996). Other examples involve the maintenance of the spelling that the borrowed word has in its language of origin. Johnson (2000:113), for example, notes that opponents of orthography reform in Germany have advocated the maintenance of the source-language orthography for loanwords. A similar argument was made in language planning in the Soviet Union, where the introduction of the Cyrillic alphabet for the Turkic languages of Central Asia was accompanied by a rule requiring Russian loanwords to be spelled according to Russian orthography (Sebba 2003:5). In the data examined here, script alternation enables authors to maintain standard English orthography for English words in Russian texts. However, text-internal script alternation is a highly unusual phenomenon which conflicts both with pragmatic considerations and with standard language ideologies. Not only does it require readers to be literate in two scripts (thus reducing the size of the potential audience), it also represents a technological challenge for printers, typesetters, and many users of word-processing software (albeit a decreasing challenge).[30] As a result, it is explicitly discouraged in many contexts – for example, style guidelines for academic writing, including the instructions for contributors to *Language in Society*.[31] It therefore seems warranted to extend Sebba's (2002) notion of the "tyranny of written monolingualism" to include a "tyranny of monoscriptalism."

In an ideology that rejects both transliteration and alphabet-switching, avoidance of foreign elements thus seems to be the only viable alternative. In fact, those "quality" newspapers that show a lower rate of alphabet-switching (i.e., use of roman script) also appear to use fewer English words in general. Furthermore, in the daily *Novoye Russkoye Slovo*, transliterated English nouns are sometimes placed in quotation marks, as if the author were apologizing for their use.[32] However, some amount of alphabet-switching is unavoidable. In the New York City subway system, some lines are identified by letters of the roman alphabet, and so are some of the avenues in the Brooklyn neighborhoods where many Russian speakers live. When a text refers to the F train or to Avenue U, alphabet-switching is unavoidable. Transliteration into Cyrillic is not an option, because

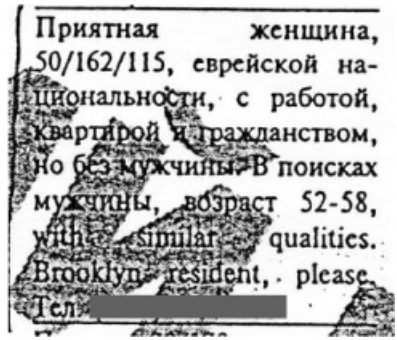
Приятная женщина, 50/162/115, еврейской национальности, с работой, квартирой и гражданством, но без мужчины. В поисках мужчины, возраст 52-58, with similar qualities. Brooklyn resident, please. Тел:

FIGURE 16: Personal ad in *Russian Bazaar*, January 18–24, 2002.

the characters no longer represent sounds but are iconic symbols in their own right, independent of their phonetic value in English.[33] Such forms point to the structural impossibility for Russian-English bilinguals to fully separate their languages and scripts. Moreover, they illustrate that there may at times be a need to integrate and combine the two writing systems.

HYBRIDITY IN LANGUAGE AND SCRIPT CHOICE

In the previous sections, I have largely presented the choice between the two scripts, and by extension the choice between codeswitching and borrowing, as a binary one. However, as shown in Table 1, a few tokens were found to be orthographically mixed, and in the discussion of bare forms I pointed out that even an uninflected roman form like *catering hall* in Figure 15 can take on some characteristics of nativized (i.e., borrowed) forms, such as inherent grammatical gender. In the following section, I will discuss the importance of such ambiguous categorizations in the data set and its implication for the analysis of language contact phenomena in general.

Woolard 1999 has noted that bilingualism research has too often treated two languages as in opposition to each other, as discrete entities which are juxtaposed. Instead, as she demonstrates with examples from Catalan-Spanish bilingualism, individual speakers may use language in a way that is not easily attributable to a specific code. Woolard (1999:17) identifies codeswitching as one such strategy. For her, codeswitching can have a sequential or a simultaneous interpretation, with speakers either juxtaposing two languages and two social identities as different from each other, or simultaneously invoking both languages and both social identities. An example of such "simultaneity-simulating codeswitching" can be found in Figure 16, a personal ad placed in the newspaper *Russian Bazaar*.

The text reads *prijatnaja ženščina, 50/162/115, evrejskoj nacional'nosti s rabotoj, kvartiroj i graždanstvom, no bèz mužčiny, v poiskax mužčiny, vozrast*

https://doi.org/10.1017/S0047404505050190 Published online by Cambridge University Press

52–58, WITH SIMILAR QUALITIES. BROOKLYN RESIDENT, PLEASE 'nice woman, 50/162/115 [age/height/weight], of Jewish ethnicity, with employment, apartment, and citizenship, but without a man, in search of a man, age 52 to 58, WITH SIMILAR QUALITIES. BROOKLYN RESIDENT, PLEASE'. Along the lines of Woolard's (1999:3) analysis, it can be argued that the author of this personal ad makes a "simultaneous claim to more than one social identity." Not only does she characterize herself as a bilingual, but she is also looking for a partner "with similar qualities" – that is, someone who reads both Russian and English and is comfortable with their alternating usage.

In her discussion of language contact phenomena, Woolard draws on Bakhtin's (1981:358) notion of hybridization, by which he describes the "mixing of two languages within the boundaries of a single utterance," a mixture "between two different linguistic consciousnesses." While Woolard takes codeswitching to exemplify hybridity on the discourse level, it can also be identified on the level of the word. In his book *Discourse in the novel*, Bakhtin (1981:305) writes:

> It frequently happens that even one and the same word will belong simultaneously to two languages, two belief systems that intersect in a hybrid construction – and, consequently, the word has two contradictory meanings, two accents.

Woolard (1999:7) uses the term "bivalency" to describe "the use by a bilingual of words or segments that could 'belong' equally, descriptively or even prescriptively, to both codes." Similarly, Muysken (2000:3–8) speaks of "congruent lexicalization" to describe cases of bilingual language use in which a given structure is shared by both languages. Although such bivalent or shared forms may be particularly frequent in the usage of bilinguals who speak closely related languages such as Castilian Spanish and Catalan, it is nevertheless relevant here as well, particularly in the treatment of loanwords. Arguably, hybridity manifests itself also in forms that do not fully belong to either language, but partially belong to both.

This can also be observed in a small number of compound nouns in the data set that are alphabetically mixed. Figure 17 includes the compound noun *barber shop*, with *barber* in Cyrillic and *shop* in roman script. This compound noun is a hybrid, with its hybridity manifested in orthography. This form is half Cyrillic and half roman; it doesn't wholly "belong" to either writing system.[34] The text reads *V barber SHOP trebuetsja ženščina dlja raboty mužskim masterom. Neobxodimy lajsens i anglijskij. Miša* 'Barber shop is looking to hire a woman for work as coiffeur for men. License and English required. Misha'.

ROMAN-CYRILLIC GRAPHEMIC BIVALENCY

The previous examples have illustrated ways in which hybridity prevents the unambiguous attribution of a text or word to one particular language or writing

В Барбер shop требуется женщина для работы мужским мастером. Необходимы лайсенс и английский. [redacted]
Миша

FIGURE 17: Help wanted ad in *Russian Bazaar*, April 10–16, 2003. (барбер "barber")

FIGURE 18: Advertisement for a law firm, at a bus stop in Brighton Beach, Brooklyn.

system. But hybridity in writing can be even more detailed, as shown in Figure 18. Here we find the two alphabets mixed within a single lexical item. This advertisement includes a phone number that is spelled out as 1-800-A-D-B-O-K-A-T.[35] However, a Russian-speaking reader would certainly pronounce this word [advokat] 'lawyer', with [v] instead of [b]. The Cyrillic spelling of the Russian word is shown in (6a), its expected transliteration into roman script in (6b). The word on the advertisement shares all characters with the Cyrillic form except for the Cyrillic letter Д, and it shares all characters with the roman form except for

the roman letter *V*. The third letter of the 1-800 number thus represents a bivalent element that has two alphabetic interpretations at once: In Cyrillic script it refers to the speech sound [v] as part of the word *advokat*, and in roman script it refers to the letter *B* and its correspondence to the number 2 on the keypad of American touch-tone phones.

(6a) Standard Russian: АДВОКАТ

(6b) Transliterated Russian: ADVOKAT

(6c) Orthographically mixed: ADBOKAT

I set out to identify further examples of Russian 1-800-numbers, which are given in (7). Most were transliterated into roman script, but one other, 1-800-DOKTOP-4, was orthographically mixed, and one number managed to recreate a standard Cyrillic spelling using only characters that also exist in the roman alphabet, 1-877-KPACOTA.

(7) Roman:	1-888-AVARIYA	(lawyer; Авария 'accident')
	1-866-ZVONITE	(telecommunications; звоните 'call!')
	1-866-PODAROK	(gifts, flowers; подарок 'gift')
	1-800-25-SLOVO	(newspaper; слово 'word')
	1-888-KONTAKT	(law firm; контакт 'contact')
	1-86N-ORBEKOV	(motivational speaker; personal name Норбеков)
	1-888-3-DNIPRO	(travel agent; named for the river Днипро Dnieper)
Cyrillic:	1-877 KPACOTA	(plastic surgeon; красота krasota 'beauty')
Mixed:	1-800 DOKTOP-4	(physician; доктор doktor 'doctor')

Although these forms are clearly produced intentionally, it is important to remember that they depend on the availability of a given "underlying" phone number. Perhaps the number corresponding to 1-800-ADVOKAT was already in use, forcing the advertising law firm to be more creative in its search for a memorable number.[36] In general, lettered phone numbers are an almost uniquely American phenomenon, closely tied to American business and advertisement practices. As such, they allow Russian-owned business to display their "Americanness" even without resorting to the use of English. Lettered phone numbers that refer to the Cyrillic alphabet can thus be interpreted as bivalent elements which express a business-owner's claim to a dual social identity as Russian-American.

The mixed roman and Cyrillic forms are possible because of the overlap between the graphic inventories of the two alphabets, a consequence of their common origin in the Greek alphabet.[37] Figure 19 shows the inventories of both alphabets and the areas where they overlap, including characters like A, O, K, and T which have the same phonetic value in the two alphabets, and characters like B and P which have different phonetic values.

The phenomenon of shared characters has received some attention, both from linguists and non-linguists, and it has at times been the object of social and political contention. At the beginning of the 18th century, Tsar Peter the Great introduced an orthography reform as part of his drive to westernize Russia. One

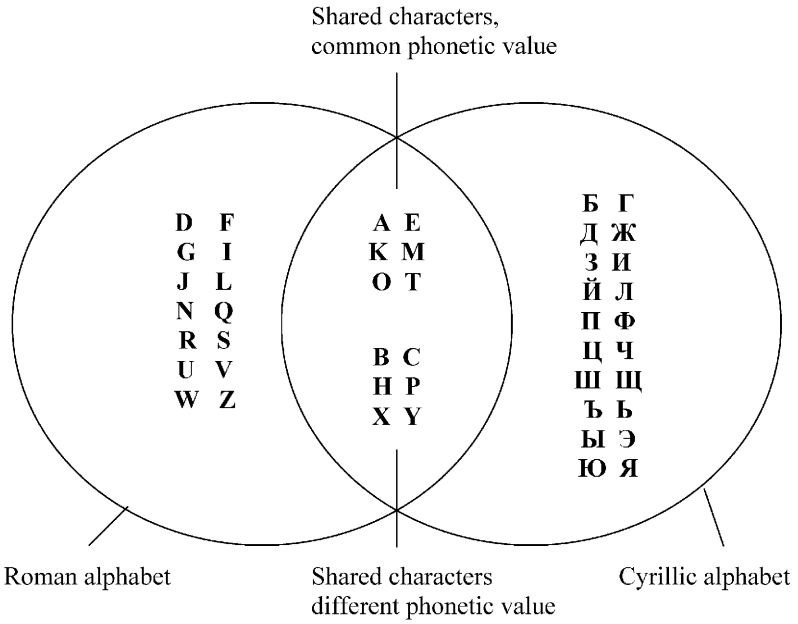

FIGURE 19: The letters of the roman alphabet, as used for Modern English, and the Cyrillic alphabet, as used for Modern Russian. (The diagram includes only capitals. The distribution of shared letters is different with lower case letters as well as in cursive writing) (cf. Feldman & Barac-Cikoja, 1996).

aim of the reform was to eliminate homographs from the Old Russian Cyrillic alphabet. As Spraul (1999:81) notes, Tsar Peter preferred the use of those letters that resemble characters of the roman alphabet and advocated the use of *I* instead of *И* (both pronounced [i]), as well as of *S* instead of *C* (both pronounced [s]). However, these westernizing efforts met resistance, and not all parts of the orthography reform were accepted in the long run. The characters resembling roman *I* and *S* are no longer used in Russian, though they are still used in other Cyrillic alphabets.[38] The overlap between the two alphabets has also been of concern in Yugoslavia and its successor states, where the Cyrillic alphabet is used by Serbs and the roman alphabet by Croats. According to Magner (2001:22), the Office of the High Representative in Bosnia and Herzegovina ruled that automobile license plates could contain only the letters *A*, *E*, *J*, *K*, *M*, and *T*, as they are common to both alphabets. This was done to avoid a conflict over which script to choose and to protect individuals from being identified as belonging to a particular ethnic group through the license plates on their cars.

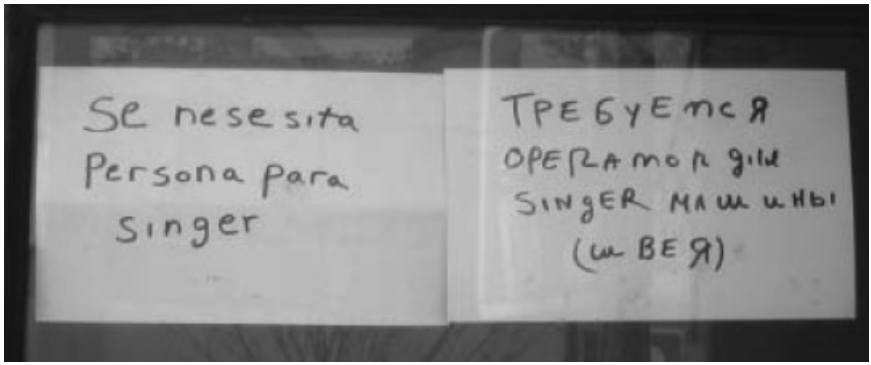

FIGURE 20: Handwritten sign in a store window on Kings Highway, Brooklyn.

The shared characters have also been identified as a source of psycholinguistic confusion. Experimental research by Feldman and associates has shown that biscriptal Serbo-Croatian speakers have delayed word recognition for words consisting only of shared letters when they are presented with such words in isolation (cf. Feldman & Barac-Cikoja 1996). Marian & Kaushanskaya 2004 achieved similar results in experiments with Russian-English bilinguals.[39] These findings suggest that language attribution plays a role in word recognition and processing, but also that shared characters are recognized visually before they are attributed to a given script. This creates the possibility for confusion which would appear to be the source of the bivalent sign shown in Figure 20. The Russian text on the right reads *trebuetsja* оператор *dlja* SINGER *mašiny (šveja)* 'Operator for Singer machine needed (seamstress)'. The Spanish text on the left translates as 'Person for Singer needed'.[40] The author of this sign is looking for individuals who can operate a Singer sewing machine. The bivalent form here is the word *operator* in the Russian text. It is written partially in Cyrillic and partially in roman, but the author also alternates between Cyrillic cursive and print, and between uppercase and lowercase letters, as illustrated in (8).

(8a) Cyrillic lower case, print: оператор 'operator'

(8b) Cyrillic lower case, cursive: *onepamop*

(8c) Roman capitals, print: OPERATOR

(8d) Orthographically mixed: OPERA*m*OR

In Figure 20, the letter that looks like a lower-case roman *m* is actually a Cyrillic lower-case т, but in cursive rather than in type, the lower-case т being one of several characters for which Russian cursive script diverges from type (see 8b).[41]

Of the remaining letters, the vowels are all shared characters with comparable pronunciation. The ambiguous shared letter *P* must be pronounced as a roman character here, and only the upper-case roman *R* is unambiguous, except for the fact that the second one looks almost like a *p* whose bow was left open (note the difference from the *R* in *SiNGER*, but also from the Cyrillic *p* in *TPEБуEmcЯ*). Nevertheless, despite the graphemic confusion, the mixing remained initially undetected by me, and presumably also by many other readers of the sign. In contrast to ambiguous spellings used in psycholinguistic experiments, the forms in Figures 18 and 20 occur not in isolation but in a context that prevents them from being misinterpreted.

In earlier models of bilingual language use, the mixing of scripts found in Figure 20 would have been characterized as interference (cf. Weinreich 1968), or as triggered by common elements such as the shared letters *O*, *E*, and *A* (Clyne 1967).[42] Yet the difference between Figures 18 and 20 lies mainly in the attribution of intentionality. Where the text is printed and the letters correspond to the digits of a phone number, the word-internal mixing of scripts appears to be intentional. In a handwritten, presumably unedited ad posted in a store window, it is likely not done on purpose. Yet both forms attest to the fact that the distinction between the two scripts and the two languages can cease to be relevant in the experience of biscriptal bilinguals, at least momentarily and at points where the respective inventories overlap. It is certainly not a coincidence that both cases involve cognates – the Latin-origin loanwords *advocate/advokat* and *operator* – and both spellings contain a majority of shared letters with a common phonetic value, including all the vowels. For these forms, word recognition is arguably not script-dependent because the words "belong" to both languages. The ambiguous, mixed spelling thus reflects and preserves an ambiguity in language attribution.

These examples show that there are aspects of bilingual language use that go beyond a mere combination of two monolingual patterns.[43] Moreover, the data demonstrate that such simultaneity in bilingualism can extend to writing, as bilinguals blur the lines between languages and writing systems. Just as languages are dynamic systems without clear boundaries, writing systems are not discrete entities either; their character inventories can be recombined and reinterpreted. As a consequence, it seems warranted to revise the distribution given in Figure 3 in order to make room for a bivalent, hybrid space of Russian-American writing that belongs to neither language or writing system, or to both at the same time. Figure 3′ illustrates the location of this hybrid space in the middle of the bilingual experience, a space that can be inhabited by biscriptal bilinguals who lay a claim to multiple social identities. This hybrid space may be more easily recognized and inhabited in language contact situations that involve languages and orthographies that are similar to each other, but it can arguably be found and explored in any contact situation where bilingual writers seek a bilingual spelling.

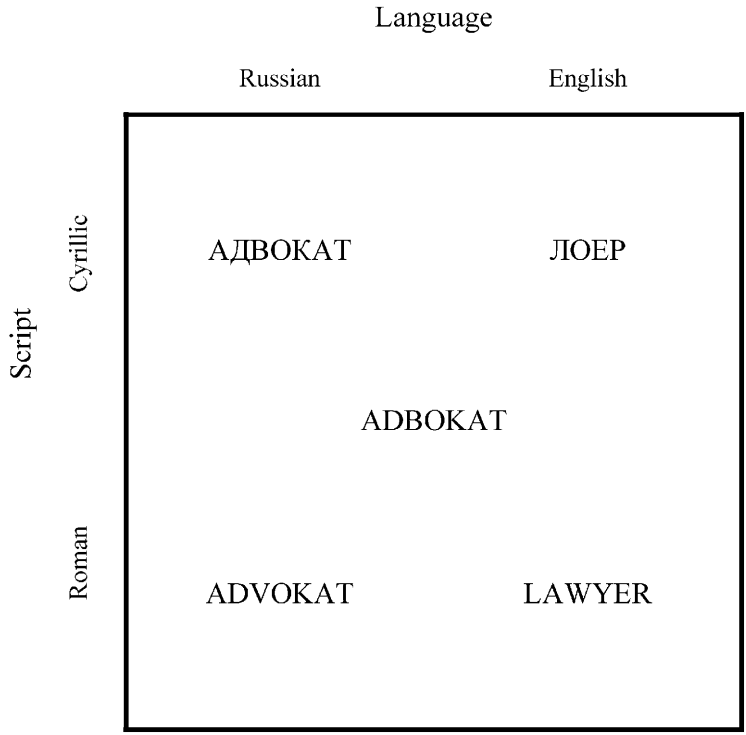

FIGURE 3′: Distribution of languages and scripts.

CONCLUSION

In conclusion, this article has documented the existence of hybridity in Russian-American writing on several levels: on the level of the text, in the codeswitching of Figure 16; on the level of the word, in the mixing of alphabets in compound nouns (Fig. 17); and on the level of the grapheme, in the intentional or unintentional use of characters that are ambiguous in the two alphabets (Figs. 18 and 20). However, these hybrid aspects are clearly exceptional in the written data examined here. For the data set as a whole, my analysis has demonstrated that Russian-American bilingual authors show a strong tendency to categorize lexical items as either Russian or English, even as they mix elements from both languages in the same text. This study thus points to the validity of the distinction between borrowing and codeswitching by showing that lexical items behave differently depending on their attribution to a particular language, and by suggesting that in biscriptal, bilingual writing, authors use script choice to attribute a lexical item to a particular language. Through script choice, authors can – in-

tentionally or unintentionally – mark lexical items in one of three ways: (i) through transliteration into Cyrillic, a word may be marked as Russian, i.e. borrowed; (ii) through the use of the roman script, a word may be marked as English, i.e. codeswitched; and (iii) through the mixing of the Cyrillic and roman alphabets, a word may be marked as ambiguous with regard to language membership.

The findings of this study thus pertain to the role that written data can play in studies of language contact, particularly when it involves the use of multiple writing systems. Where different writing systems exist in the same community, written data can be drawn upon to illuminate questions of language boundaries from the perspective of the bilingual individual or of the linguistic community as a whole. Furthermore, the findings suggest that speakers have a choice between treating a given lexical item as borrowed or switched. Even after a word has been borrowed, it is still available for codeswitching. This theoretical possibility has too often been ignored because most bilingualism researchers have tended to classify all instantiations of a given lexical item alike.

The attribution of linguistic forms to a particular language has been a central issue in the debate about codeswitching and borrowing, but also in language contact studies in general. Focusing primarily on bilingual speech, Gardner-Chloros 1995, Woolard 1999, and others have challenged the notion of the discreteness of linguistic systems, arguing that it may be introduced by researchers in ways that are not meaningful to community members. Along similar lines, Auer 1999 argues for a typological distinction between types of bilingual speech where codes are meaningfully juxtaposed (codeswitching) and where they are not (mixing). These and other studies have relied primarily if not exclusively on bilingual speech. In written data, the distinction between languages is ideologically emphasized through the standardization of written language (Milroy & Milroy 1991, Sebba 2002). Standard orthography provides a norm that not only defines a "correct" representation of a given lexical item but also serves to attribute that item to a particular standard language. As a consequence, the discreteness of linguistic systems is an inherent aspect of bilingual writing, even if it may not be present in bilingual speech. Arguably, this is what motivates the categorical adherence to the Free-Morpheme Constraint that was found in the present data.

Nevertheless, occasional hybrid forms exist, facilitated here by the overlap between the two alphabets. The three observed patterns of script choice thus mirror the three basic processes identified by Muysken 2000 as active in bilingual language use – INSERTION, ALTERNATION, and CONGRUENT LEXICALIZATION. Bilingual authors, like bilingual speakers, may integrate elements from one language into structures of the other, they may alternate between languages, or they may create hybrid forms that are wholly or partially shared by both languages. Thus, while bilingual writing is clearly different from bilingual speech in important ways, the two share underlying principles, and both are available to linguists who seek to identify what these principles are.

NOTES

* Versions of this article were presented at two conferences: "*Alphabetics: Interpreting letters*" at Harvard University, 26–27 April 2003, and *NWAVE 32*, Philadelphia, 9–12 October 2003. I thank the audiences for their valuable insights and observations, especially Erika Boeckeler and Daniel Kokin. I also owe thanks to Katya Korsunskaya, Vladislav Rapoport, Doris Stolberg, Mario Geiger, and Tobias Kuhn. I am grateful to John Victor Singler, Mark Sebba, and Jannis K. Androutsopoulos, as well as to Jane H. Hill and to two anonymous reviewers for their valuable comments and suggestions. All errors and omissions remain my own.

[1] Codeswitching in literature has occasionally been the subject of linguistic research. A well-known example is Timm's (1978) study of switching between Russian and French in Tolstoy's *War and Peace*. However, her study contains no reference to the alternation between the Cyrillic and roman alphabets.

[2] Source: U.S. Census Bureau, Census 2000 Summary File 3, Matrix PCT10. (See http://factfinder.census.gov/).

[3] In transliterating Cyrillic forms into roman script, I follow the ISO system; see Cubberley (1996:351).

[4] The newspapers included were the daily Новое Русское Слово (*Novoye Russkoye Slovo*) and the weeklies Форвертс (*Forverts*, *Russian Forward*), Русская Реклама (*Russkaya Reklama*), Русский Базар (*Russkij Bazar*, *Russian Bazaar*), Вечерний Нью-Йорк (*Vecherniy New York*), Курьер (*Kurier*), Полезная Газета (*Poleznaya Gazeta*) and В Новом Свете (*V novom svete*).

[5] Food stamps are coupons which can be used to purchase groceries as part of a federal welfare program administrated by the U.S. Department of Agriculture.

[6] The sign refers to food items prepared in the store, such as roasted chicken or soup.

[7] In his study of bilingual signage in Quebec, Shell (1993:50–51) uses the term "semiotic mediator" to describe elements such as *food stamps* in Figure 8, which participate in two texts at once and act as a connection between them.

[8] The data contain 26 tokens that are not nouns or noun phrases. All are content morphemes with an adjectival character, e.g. forms qualifying types of employment such as *part-time*, *full-time*, or *long-distance*. All tokens are written in roman script except for one: парт-тайм работу *part-tajm rabotu*, 'part time work+ACC.' in an ad in *Vecherniy New York*.

[9] The text contains an apparent typographical error, with *na* 'on' instead of *ne* 'not.'

[10] This may well be a case of gender marking by default, since the adjective cannot be case-marked without also being gender-marked, and no bare (i.e., suffixless) form of the adjective is available.

[11] Script choice for borrowings or single-item codeswitches can also be a concern for linguists studying bilingual speech. While Andrews 1998 and Budzhak-Jones 1998 avoid the issue by transliterating all Russian or Ukrainian forms into roman script, Gregor (2002:20–22) specifically discusses the merits of script alternation and transliteration, opting to write in Cyrillic those English-origin words that she classifies as borrowed, and in roman script those that she classifies as codeswitched. She bases the distinction between the two on criteria of morphological and phonological adaptation, leading her to use the Cyrillic script for all English-origin items that are marked with Russian inflectional affixes. Her usage thus mirrors that of the Russian-American press.

[12] In addition to forms that require case endings, I also removed all tokens of lexical items that occur frequently but always in the same alphabet. For example, as a lone lexical item, *office* is one of the most frequent words in the data set, with 127 occurrences, but it is always written in Cyrillic (офис). A majority of the other frequent invariant forms were also always in Cyrillic, namely вэн *vèn* 'van' (30 occurrences), лайсенс *lajsens* '(driver's) license' (20), статус *status* 'immigration status' (20), мини-вэн *mini-vèn* 'minivan' (11), менеджер *menedžer* 'manager' (13), and шитрок *šitrok* 'sheetrock' (13). The following forms were always written in roman script: *salesperson/salesman/sales people* (20), *medical assistant* (10), *dental assistant* (10), and *receptionist* (13).

[13] The factor groups were identified as significant in the following order: (1) constituency, (2) source, (3) frequency, (4) type of advertisement.

[14] These categories were defined orthographically. If a given English noun was spelled as one word at least part of the time, it counted as a single noun.

[15] Another potential example is *sponsor*, which is used to refer to an "immigration sponsor" but which occurs only once in the data set, albeit in Cyrillic.

[16] This usage is illustrated by two ads, both from *Russian Bazaar* (April 10–16, 2003, p. 71A). Надежен. Хорошее знание ангийского и компьютера. Ищу соответствующую работу на кеш или чек, *Nadežen. Xorošee znanie anglijskogo i komp'yutera. Išču sootvetstvuyuščuyu rabotu na keš iliček* 'Reliable. Good knowledge of English and computers. I seek suitable work for cash or check'. Программист-электронщик, работу, можно на чек или cash, p/t, *Programmist-èlektronščik, rabotu možno na ček ili* CASH, P/T, 'Programmer-electrician, work, may be for check or cash, part time'.

[17] See in Figure 13 the form помощь на дому *pomošč' na domu* 'help in the house' which is used alongside *housekeeper*, arguably its equivalent.

[18] Compare the uses of опыт *opyt* 'experience' in Figure 12, and официанты *oficianty* 'waiters' in Figure 15. Instead of *salesman*, *saleswoman*, or *salesperson*, продавец *prodavec* 'salesman' and продавщица *prodavščica* 'saleswoman' are used frequently.

[19] In any case, the majority of the English lexical items that occur in the data do not appear to be culture-specific, including such frequent forms as офис *ofis* 'office,' трак *trak* 'truck,' вэн *van* 'van,' or лайсенс *lajsens* 'license.'

[20] A *metro card* is a ticket for New York City's public transport system.

[21] Transliteration is mostly phonetic, intended to approximate the pronunciation of a given word in American English. In some instances, transliteration may also be based on orthography, for example in the spelling колледжа *kolledža* 'college+GEN,' with Cyrillic double л corresponding to roman double *L*. Describing cases like these, Banu & Sussex (2001:53) speak of letter-by-letter transliteration, i.e. transliteration based on equivalence between characters. Transliteration may also be based purely on graphic similarities between letters. Androutsopoulos 2002 cites examples from Greek computer-mediated communication in which Greek letters are transliterated by graphically similar numbers (ξ as 3, θ as 8).

[22] English *h* [h] is typically transliterated as *x* [x] in Russian (see хаускипера *khauskipera* 'housekeeper+GEN' in Figure 11). Transliteration is also not uniform. For example, the data contain four alternate spellings of *babysitter*: бебиситер *bebisiter*, бэбиситер *bèbisiter*, бебиситор *bebisitor*, and бэбиситор *bèbisitor*.

[23] For example, a classified ad in *Russkaya Reklama* (January 24–30, 2003) contains the spelling *countertabs* for *countertops*. This spelling may be taken to reflect Russian phonology both in the misinterpretation of the final, unstressed vowel and in the choice of *b* instead of *p*, possibly resulting from hypercorrection that assumes an underlying voiced labial stop and attributes the surface voicelessness to Russian final devoicing.

[24] For example, in their study of English-origin borrowing in the French of Ottawa and Hull, Poplack et al. (1988:54) state that "all occurrences of a given English-origin word were considered tokens of the same lexical type."

[25] They include кеш *keš* 'cash' (Andrews 1998:10) with 23 occurrences in Cyrillic and 10 in roman script, бебиситер/бэбиситер/бебиситор/бэбиситор *bebisiter/bèbisiter/bebisitor/bèbisitor* 'babysitter' (p. 71; 59 Cyrillic, 126 roman), трак *trak* 'truck' (p. 77, 15 Cyrillic, 3 roman), and уик-энд *uik-ènd* 'weekend' (p. 28), with one occurrence in each script, as well as фудстемп *fudstemp* 'food stamp,' as evidenced by signage shown in Figures 6–9.

[26] The possibility for bilingual speakers to treat borrowed forms as codeswitches has been pointed out previously by Hill & Hill (1986:356) and Heath (1989:24). The choice between borrowing and codeswitching may become apparent to any bilingual who speaks two languages that have exchanged loanwords, for example a French-English bilingual using the expression *déjà vu* in an English sentence, or a word like *weekend* in a French sentence. In either case, the borrowed form and the codeswitched form are marked by clear differences in pronunciation. Compare also Thomason's (2001:134) discussion of the pronunciation of the name *Bach* in English.

[27] In the total data set, all 17 English-origin lexical items in local news articles occur in Cyrillic. The tokens are primarily from two newspapers, the weekly *Russkaja Reklama* and the daily *Novoe Russkoe Slovo*. In its articles, the latter has the tendency to place transliterated English-origin lexical items in quotation marks, a phenomenon not found in any of the other publications (see note 32).

[28] While most of the Russian-speaking immigrants are Jewish, according to Birman (1979:49) they "had been fully acculturated into the Russian tradition … hav[ing] embraced Russian culture, literature and even history as their own" (quoted by Andrews 1998:44).

[29] After immigration to the United States, many Soviet immigrants were unable to maintain the socioeconomic status that they had had in their homeland (Levkov 1984:142; Orleck 2001:120).

According to Levkov 1984, many immigrants who arrived in the 1970s and 1980s considered their social and cultural status in the United States lower than it had been in the Soviet Union. Thus it can be presumed that for people who are "highly literate in Russian" (Hinkel 2000:356), the Russian language is tied to the social status that they had achieved in the Soviet Union by means of higher education. Maintaining Russian, and maintaining unmixed Russian in particular, may thus be a way for immigrants to hold on to the social status that they have lost. The relevance of pre-emigration professions and socioeconomic standing is also illustrated by a personal ad found in the newspaper *Kurier*, January 24, 2003, page 67A, in which a woman is identified as a doctor, followed by the word здесь *zdes'* 'here' in parentheses: Она просто суперженщина! 32/166, доктор (здесь), красотка на загляденье, с густыми золотистыми волосами… *Ona prosto superženščina*! 32/166, *doktor* (*zdes'*), *krasotka na zaglyaden'e*, *s gustymi zolotistymi volosami*, 'She's simply a superwoman! 32/166, doctor (here), a beauty to look at, with thick golden hair'.

[30] In addition to these considerations, Gazda 1998 contends that alphabet-switching impedes readability.

[31] Under the heading for "Citations and forms of emphasis," it is stated that "[n]ormally, the Latin alphabet is to be used." Cf. http://assets.cambridge.org/LSY/lsy_ifc.pdf

[32] For example, an article in *Novoye Russkoye Slovo* of January 27, 2003, p. 8, includes in quotation marks the words толлы *tolly* 'tolls', толл-плазы *toll-plazy* 'toll-plazas', and и-зи пассов *i-zi passov* 'E-Z Passes+GEN.' All terms refer to a computerized form of toll payment used on some U.S. highways. Gazda (1998:165) reports the use of quotation marks for foreign names rendered in roman script in monolingual Russian texts, often followed by a translation and "explanation" in Russian. In both cases, the quotation marks can be seen as a flagging, as a marker of an unexpected element which is quoted, i.e. attributed to the voice of another.

[33] Acronyms represent a similar case (cf. Gazda 1998:166). The letters of acronyms are abbreviations for words, and as such they no longer directly represent speech sounds but are in a sense logographic symbols. Because transliteration between roman and Cyrillic does not generally establish an equivalence between letters but rather between speech sounds, it is not an option for acronyms here. They must thus either be translated, or they are maintained in the original language and alphabet, which is in fact what we find in many cases in the Russian-American print media. Conversely, many roman-script readers are no doubt familiar with both the Cyrillic original and the roman-script translations of Soviet-era acronyms such as CCCP and USSR. A third option employed occasionally in the Russian-American print media is to spell out the pronunciation of a roman acronym in an effort to familiarize readers with it. For example, in a front-page article of March 24, 2003, the daily *Novoye Russkoye Slovo* introduces the name of a British television station in quotation marks as Ай-ти-эн *Aj-ti-èn*. From then on, the roman letters *ITN* are used exclusively throughout the article. The practice of spelling out acronyms is also reported by LaDousa (2002:224) for Hindi/English texts.

[34] Gazda (1998:164) reports similar examples from monolingual Russian texts. However, all the examples quoted by him use roman script for the first element and Cyrillic script for the second element, which is often a more general term, e.g. WEB-сервер *WEB-server*, health-клуб *HEALTH-klub*.

[35] On American telephones, each number is associated with a group of letters of the roman alphabet, allowing words to represent a telephone number. For example, *A*, *B*, and *C* correspond to the number 2, *D*, *E*, and *F* to 3, and so on. The word ADBOKAT thus represents the phone number 232-6528. As can be seen in Figure 18, the numbers are indicated in the ad in smaller print, presumably as a "translation" for immigrants who are not yet familiar with the letter format. So-called 1-800 numbers are phone numbers that can be called free of charge (as are numbers beginning with 1-888, 1-866 or 1-877).

[36] This phone number is currently in use.

[37] However, the combination of different writing systems is possible in other cases as well; see Lubell 1993 for the use of Hebrew vowel points with roman characters, and Smith & Schmidt 1996 for uses of the roman alphabet with Japanese characters. Banu & Sussex (2001:57) report an example in which the apostrophe separating an English genitive *s* from the noun occurs with an English phrase that has been transliterated into Bengali script.

[38] Other versions of the Cyrillic alphabet contain additional characters which are shared with the roman alphabet, owing either to common origin or to borrowing; see Cubberley 1996 and Comrie (1996:700–1).

(i) *Inherited from Old Church Slavonic:*

I i [i] Ukrainian, Belarusian, Old Russian ("civil script" of Peter the Great)
S s [dz]/[s] Macedonian, Old Russian ("civil script" of Peter the Great)

(ii) *Borrowed from roman alphabet:*

J j [j] Serbian, Macedonian

[39] Marian & Kaushanskaya 2004 conducted experiments in the Picture-Word Interference paradigm in which participants were shown a picture and a word. The task was to identify in English the object on the picture while ignoring the word. Compared to English monolinguals, Russian-English bilinguals were shown to be distracted more by words that consisted of characters shared between both alphabets.

[40] The Spanish part of the sign contains a nonstandard spelling of *necesita* '(he, she, it) needs'.

[41] Other discrepancies between cursive and type are: г = *г* [g], д = *д* [d], и = *и* [i], и = *й* [j], and п = *п* [p]. Note that several of the cursive characters are shared with the roman alphabet.

[42] Clyne (1994:959) defines the term "trigger word" as referring to any element that is "identical or nearly identical in the two languages, (it) brings about a linguistic disorientation in the speaker."

[43] They also show that word-internal alphabet-switch is theoretically possible, but only in a way that dissolves and denies the distinction between the two scripts, along with the distinction between the two languages and the one between codeswitching and borrowing.

REFERENCES

Andrews, David R. (1998). *Sociocultural perspectives on language change in diaspora: Soviet immigrants in the United States*. Amsterdam & Philadelphia: Benjamins.

Androutsopoulos, Jannis (2002). Greek on the Internet: Recent findings and research perspectives. Paper presented at Institut für Deutsche Sprache Mannheim, Research Group "Sprachvariation," November 2002.

Auer, Peter (1998). Introduction: *Bilingual Conversation* revisited. In Peter Auer (ed.), *Code-switching in conversation: Language, interaction and identity*, 1–24. London: Routledge.

_____ (1999). From codeswitching via language mixing to fused lects: Toward a dynamic typology of bilingual speech. *International Journal of Bilingualism* 3:309–32.

Bakhtin, Mikhail M. (1981 [1934-35]). Discourse in the novel. In Michael Holmquist (ed.), *The Dialogic Imagination. Four Essays by M. M. Bakhtin*, 259–422. Austin: University of Texas Press.

Banu, Rahela, & Sussex, Roland (2001). Code-switching in Bangladesh. *English Today* 17:51–61.

Bentahila, Abdelâli, & Davies, Eirlys E. (1991). Constraints on code-switching: A look beyond grammar. In *ESF Network on Codeswitching and Language Contact. Papers for the Symposium on Codeswitching in Bilingual Studies: Theory, Significance and Perspectives, Barcelona, March 1991*, 369–403. Strasbourg: European Science Foundation.

Birman, Igor (1979). Jewish immigration from the U.S.S.R.: Some observations. *Soviet Jewish Affairs* 9(2):46–63.

Budzhak-Jones, Svitlana (1998). Against word-internal codeswitching: Evidence from Ukrainian-English bilingualism. *International Journal of Bilingualism* 2:161–82.

Cheung, Yat-Shing (1992). The form and meaning of digraphia: The case of Chinese. In Kingsley Bolton & Helen Kwok (eds.), *Sociolinguistics today: International perspectives*, 207–17. London: Routledge, Chapman & Hall.

Clyne, Michael (1967). *Transference and triggering: Observations on the language assimilation of postwar German-speaking migrants in Australia*. The Hague: Nijhoff.

_____ (1994). Discourse in cross-linguistic and cross-cultural contexts. In R.E. Asher & J.M.Y. Simpson (eds.), *Encyclopedia of language and linguistics*, 2:958–64. Oxford: Pergamon.

Comrie, Bernard (1996). Adaptations of the Cyrillic alphabet. In Daniels & Bright, 700–26.

Cubberley, Paul (1996). The Slavic alphabets. In Daniels & Bright, 346–55.

Dale, Ian (1980). Digraphia. *International Journal of the Sociology of Language* 26:5–13.

Daniels, Peter, & Bright, William (1996) (eds.). *The world's writing systems*. New York: Oxford University Press.

De Francis, John (1984). Digraphia. *Word* 35:325–40.

Eliasson, Stig (1990). Models and constraints in code-switching theory. In *ESF Network on Code-switching and Language Contact. Papers for the Workshop on Constraints, Conditions and Models, London, September 1990*, 17–50. Strasbourg: European Science Foundation.
Feldman, Laura Beth, & Barac-Cikoja, Dragana (1996). Serbo-Croatian: A biscriptal language. In Daniels & Bright, 769–72.
Ferguson, Charles A. (1959). Diglossia. *Word* 15:325–40.
Fishman, Joshua A. (1967). Bilingualism with and without diglossia: Diglossia with and without bilingualism. *Journal of Social Issues* 23:29–38.
Gardner-Chloros, Penelope (1995). Code-switching in community, regional and national repertoires: The myth of the discreteness of linguistic systems. In Lesley Milroy & Pieter Muysken (eds.), *One speaker, two languages: Cross-disciplinary perspectives on code-switching*, 68–89. Cambridge: Cambridge University Press.
Gazda, Jiří (1998). K otázce konkurence grafických kódů v ruském publicistickém textu. [The issue of script competition in Russian journalistic texts.] *Sborník Prací Filozofické Fakulty Brněnské Univerzity, Řada Jazykovědná* A46, 161–68.
Gold, Steven J. (1995). *From the workers' state to the golden state: Jews from the former Soviet Union in California*. Boston: Allyn & Bacon.
Gregor, Esma (2002). *Russian-English code-switching in New York City*. Dissertation, Humboldt Universität zu Berlin.
Grivelet, Stéphane (2001). Introduction. *International Journal of the Sociology of Language* 150 (Special Issue: *Digraphia: Writing systems and society*):1–10.
Heath, Jeffrey (1989). *From code-switching to borrowing: foreign and diglossic mixing in Moroccan Arabic*. London & York: Kegan Paul.
Hill, Jane H., and Hill, Kenneth C. (1986). *Speaking Mexicano: Dynamics of syncretic language in Central Mexico*. Tucson: University of Arizona Press.
Hinkel, Eli (2000). Soviet immigrants in the United States: Issues in adjustment. In Sandra Lee McKay & Sau-Ling Cynthia Wong (eds.), *New immigrants in the United States: Readings for second language educators*, 352–68. Cambridge: Cambridge University Press.
Jaffe, Alexandra (2000). Introduction: Non-standard orthography and non-standard speech. *Journal of Sociolinguistics* 4:497–513.
Johnson, Sally (2000). The cultural politics of the 1998 reform of German orthography. *German Life and Letters* 53:106–25.
Kress, Gunther (2000). *Early spelling: Between convention and creativity*. London: Routledge.
LaDousa, Chaise (2002). Advertising in the periphery: Languages and schools in a North Indian city. *Language in Society* 31:213–42.
Levkov, Ilya (1984). Adaptation and acculturation of Soviet Jews in the United States: A preliminary analysis. In Robert O. Freedman (ed.), *Soviet Jewry in the decisive decade, 1971–80*, 109–43. Durham: Duke University Press.
Li, David C. S. (2000). Phonetic borrowing: Key to the vitality of written Cantonese in Hong Kong. *Written Language and Literacy* 3:199–233.
Lubell, Stephen (1993). Bilingualism in the Hebrew text. *Visible Language* 27:163–204.
Magner, Thomas F. (2001). Digraphia in the territories of the Croats and Serbs. *International Journal of the Sociology of Language* 150:11–26.
Marian, Viorica, & Kaushanskaya, Margarita (2004). Mapping written input onto orthographic representations: The case of bilinguals with partially overlapping orthographies. Poster presented at the 26th Annual Meeting of the Cognitive Science Society, Chicago, August 5–7, 2004.
Milroy, James, & Milroy, Lesley (1991). *Authority in language: Investigating language prescription and standardisation*. London: Routledge & Kegan Paul.
Muysken, Pieter (2000). *Bilingual speech*. Cambridge: Cambridge University Press.
Myers-Scotton, Carol (1993). *Duelling languages: Grammatical structure in codeswitching*. Oxford & New York: Clarendon Press.
Orleck, Annelise (2001). Soviet Jews: The city's newest immigrants transform New York Jewish life. In Nancy Foner (ed.), *New immigrants in New York*, 111–40. Rev. ed., New York: Columbia University Press.
Piller, Ingrid (2003). Advertising as a site of language contact. *Annual Review of Applied Linguistics* 23:170–83.

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.
_____ (1998). Introduction: How languages fit together in codemixing. *International Journal of Bilingualism* 2:127–38.
_____, & Meechan, Marjory (1995). Patterns of language mixture: Nominal structure in Wolof-French and Fongbe-French bilingual discourse. In Lesley Milroy & Pieter Muysken (eds.), *One speaker, two languages: Cross-disciplinary perspectives on code-switching*, 199–232. Cambridge: Cambridge University Press.
_____; Sankoff, David; & Miller, Christopher (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26:47–104.
Rand, David, & Sankoff, David (1990). *Goldvarb, version 2: A variable rule application for the Macintosh*. Montreal. Université de Montréal, Centre de recherches mathématiques.
Sankoff, David, & Poplack, Shana (1984). Borrowing: the synchrony of integration. *Linguistics* 22:99–136.
_____; Poplack, Shana; & Vanniarajan, Swathi (1990). The case of the nonce loan in Tamil. *Language Variation and Change* 2:71–101.
Sebba, Mark (2002). Regulated spaces: Language alternation in writing. Paper presented at the 2nd International Symposium on Bilingualism, Vigo, October 23–26, 2002.
_____ (2003). *Ideology and alphabets in the former USSR*. Center for Language in Social Life, Working Paper No. 123. Lancaster University, Department of English and Modern English Language.
Shell, Marc (1993). *Children of the earth: Literature, politics and nationhood*. New York: Oxford University Press.
Smith, Janet S. [Shibamoto] (1996). Japanese writing. In Daniels & Bright, 209–17.
_____, & Schmidt, David L. (1996). Variability in written Japanese: Towards a sociolinguistics of script choice. *Visible Language* 30:46–71.
Spraul, Hildegard (1999). Graphemik. In Helmut Jachnow (ed.), *Handbuch der Sprachwissenschaftlichen Russistik und ihrer Grenzdisziplinen*, 66–86. Wiesbaden: Harrassowitz.
Thomason, Sarah Grey (2001). *Language contact*. Washington, DC: Georgetown University Press.
Timm, L.A. (1978). Code switching in *War and Peace*. In M. Paradis (ed.), *Aspects of bilingualism*, 302–15.Columbia: Hornbeam.
Weinreich, Uriel (1968 [1953]). *Languages in contact: Findings and problems*. The Hague: Mouton.
Woolard, Kathryn A. (1999). Simultaneity and bivalency as strategies in bilingualism. *Journal of Linguistic Anthropology* 8:3–29.
Zima, Petr (1974). Digraphia: The case of Hausa. *Linguistics* 124:57–69.

(Received 9 December 2003; accepted 18 September 2004;
final revision received 15 October 2004)