

Slavic Lexical Borrowings in English: Patterns of Lexical and Cultural Transfer

DANKO ŠIPKA

5052 W. Kesler Ln, Chandler, AZ 85226, USA

E-mail: danko.sipka@asu.edu

Abstract: Cultural borrowing between the English and Slavic languages in the last one hundred years, and especially in the latter half of the twentieth century was mostly a unidirectional process with the English as the source and Slavic languages as the target. This paper is an attempt to fill this void in examining the other direction of lexical transfer between English and Slavic languages. The following general conclusions can be drawn from the analysis: a) Lexical and cultural influence from subordinate to dominant language is by and large limited to the culture-bound items. Borrowed vocabulary items remain marginal in the overall English vocabulary. Several exceptions from this trend, i.e., the words which have made it to the core of the English vocabulary, are result of the butterfly effect and cannot be accounted for by some general trend; b) Lexical influence of each particular language is directly proportionate to language size. Exceptions from this general trend occur, as demonstrated by East Slavic languages, when one Slavic language clearly dominates others; c) The timeline of borrowing is directly proportionate with the growth and deepening of international communication networks in the nineteenth and in particular twentieth centuries.

Keywords: Slavic lexical borrowings, cultural transfer, external language history, English lexicology

1. Introduction

The process of lexical borrowing has been commanding considerable and constant interest in linguistics in the last one hundred years or so. It suffices to say that the MLA (2004), which covers the publications from 1963 onwards, lists over 2,400 bibliographic units with “lexical borrowing” in the subject field. The quantitative increase was accompanied by significant methodological improvement in the latter half of the twentieth century, primarily in the monumental work of Rudolf Filipović (see Filipović 1997, 1986) and his theory of languages in contact as well as his continuing work in that field under the umbrella of “contact linguistics” (see, for example, Winford 2002).

Lexical borrowing has been approached from different angles ranging from the adaptation of the form to semantic modifications and expansions. In this paper lexical borrowing will be addressed within the framework of cross-cultural linguistics, the approach stemming from Wundt (1911) and Humboldt (1959), established by Sapir (1921), and presently unfolding in the voluminous opus of Anna Wierzbicka (see in particular Wierzbicka 1992).

Sapir (1921: ch. 9, s.v.) makes the following important observations: “When there is cultural borrowing there is always the likelihood that the associated words may be borrowed too. (...) It is generally assumed that the nature and

extent of borrowing depend entirely on the historical facts of culture relation (...) It seems very probable that the psychological attitude of the borrowing language itself towards linguistic material has much to do with its receptivity to foreign words.”

Cultural borrowing between the English and Slavic languages in the last one hundred years, and especially in the latter half of the twentieth century was mostly a unidirectional process with the English as the source and Slavic languages as the target. Both historical events and the attitudes have militated toward strong cultural and ipso facto lexical influence of the English language to all Slavic languages with only insignificant reciprocation. It is then quite understandable that the volume of literature about lexical borrowing comprises almost exclusively studies about English borrowings in Slavic languages. Similar is the relation of Slavic languages to other sources of cultural influence and their respective languages, i.e., German, French, and Italian. While there exists a myriad of monographs devoted to Western European lexical influences (their multitude makes the enumeration here impossible and one should consult MLA 2004 for more information), studies about Slavic elements in other languages are considerably less common and mostly limited to such contact languages and dialects as Rumanian, Hungarian, Baltic languages, German dialects, etc. (see, for example Dini 1993; Lallukka 1985; Papp 1973; Pohl 1990; Vrabie 1992). The most notable exception in this regard are studies about Spanish Slavicisms (Alvarado 1989 and 1990).

English Slavicisms have not been adequately addressed yet. Their treatment is limited to marginal observations within philologico-political works such as Mencken (1921), where we find statements like “The Italians, the Slavs, and above all, the Russian Jews, make steady contributions to the American vocabulary and idiom” (Mencken 1921: ch. 4, s.v.).

This paper is an attempt to fill this void in examining the other direction of lexical transfer between English and Slavic languages. The discussion revolves around the following questions which bear broader theoretical importance:

- a) What is the extent of the lexical influence of subordinated languages to their dominant counterpart?
- b) What is the structure of such influence, what are the subject matter fields involved in this lexical transfer?
- c) What are the temporal and historical parameters of this influence?
- d) What is the proportion of the influence between different Slavic languages and is that proportion related to the number of speakers of each language?

The answers to the aforementioned questions may shed some light on the general mechanisms of lexical influences which non-contact subordinated languages exert on their dominant counterparts.

2. Procedure

The analysis of Slavic lexical influence on the English language has been performed using the material from the OED (1989). There are two principle reasons supporting the choice of this dictionary. First, the sheer size of 231,100 main entries offers the most comprehensive insight into the English vocabulary. Second, the dictionary registers the first appearance of each entry word, which enables the data about temporal parameters of the influence to be collected.

A two-step procedure was employed to extract words of Slavic origin. First, the keywords referring to Slavic languages in general and Old Church Slavonic (Slavonic, Slav., OSlav.) and the present-day Slavic languages (e.g., Russian, Russ., Polish, Pol., etc.) have been used, which resulted in 1,211 extracted entries. These entries were consequently manually filtered to exclude the instances where the etymological labels relating to Slavic languages are used only to corroborate the Indo-European origin of Old English words and to eliminate repeated entries.

A rather liberal criterion was used, whereby the word is considered Slavicism if one of Slavic languages was a member of the borrowing chain. This encompasses the words borrowed into the English directly, the non-Slavic lexemes borrowed through a Slavic language, and Slavic words borrowed through a non-Slavic language.

The ensuing part of the analysis entailed determining the source language, period of borrowing, and its subject matter field. These data were consequently processed statistically.

In determining the language of origin, the category of “general Slavic” has been established. This category comprises both references to Slavic languages in general, several Slavic languages, and Old Church Slavonic. This category is meant to separate the common Slavic lexical heritage from those belonging to the present-day Slavic languages and their dialects. One should also note that Serbo-Croatian was treated as one unit. The OED labels Serbo-Croatian as either Croatian or Serbian or Serbo-Croatian without any clear criterion. In all instances of Serbo-Croatian borrowings the word exists in all ethnic standards of the language (Bosniac, Croatian, and Serbian) and it is impossible to determine which variant was the actual source of borrowing. It is then only logical to treat the language in its entirety rather than breaking the analysis down to its ethnic variants.

3. Results

3.1 General Parameters

There were 642 entries attested in the OED as Slavicisms. This constitutes 0.28% of all entries (i.e., 642 out of 231,100 entries). For comparison, even Slavic languages with strong purist tradition contain considerably more Germanic lexemes. For example, Anić (1994), a Croatian monolingual dictionary,

lists 1197 entries with Germanic etymologies out of 51,263 entries in the dictionary, which constitutes 2.34%. This disproportion is considerably higher in the textual frequency of the entries. Numerous Germanic borrowings in Slavic languages are highly frequent lexical items, which is not the case with the English Slavicisms as it will be demonstrated further in this text. Most English Slavicisms remain peripheral, marginal in the English vocabulary not only by the virtue of their low frequency, but also owing to their monosemantic character and lack of word-formation nests.

3.2 Source Languages

The distribution of languages is presented in the *Table 1* and *Chart 1*.

Table 1: Distribution of source languages

| | Example | Count | Percent |
|----------------|--------------------------------|-------|---------|
| Russian | <i>borsch, pirog, samovar</i> | 481 | 74.92 |
| Polish | <i>Mazurka, Sejm, szlachta</i> | 56 | 8.72 |
| S-Cr | <i>dinar, polje, ponor</i> | 33 | 5.14 |
| Czech | <i>dobro, koruna, robot</i> | 29 | 4.52 |
| General Slavic | <i>knez, tsar, vaivode</i> | 24 | 3.74 |
| Bulgarian | <i>lev, Gamza, Pomak</i> | 13 | 2.02 |
| Slovenian | <i>Slovene</i> | 3 | 0.47 |
| Slovak | <i>Tauberian</i> | 2 | 0.31 |
| Ukrainian | <i>gley</i> | 1 | 0.16 |
| Totals | | 642 | 100,00 |

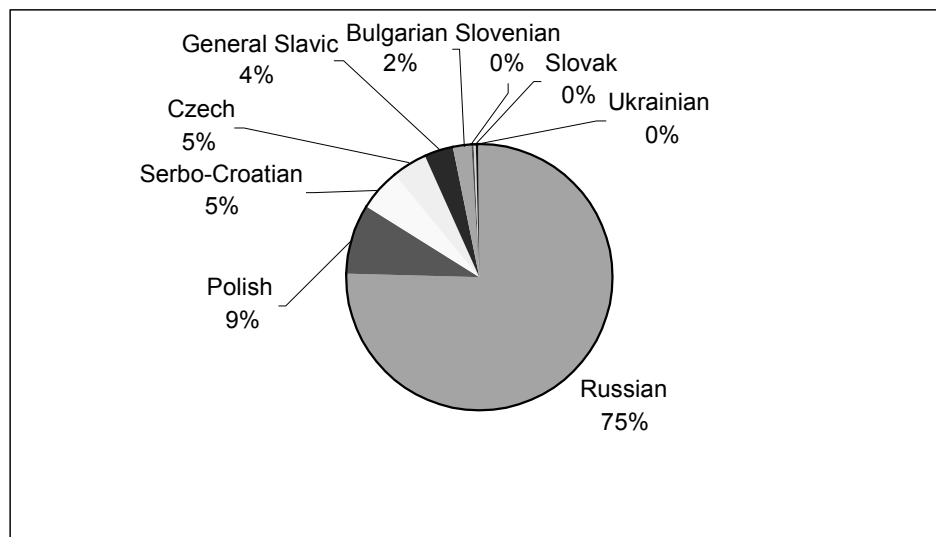


Chart 1: Distribution of source languages

Several observations should be made with regards to the source languages. First, there is a strong statistically significant correlation between the number of speakers in each language and its proportion in the English Slavisms. Having excluded general Slavic sources and included the languages without any borrowings in English, the data presented in *Table 2* was used to compute the correlation coefficient between the number of borrowings in English and number of speakers in each language. The Pearson R value of 0.96 has been obtained, statistically significant at $p < 0.01$. This general pattern is followed by all languages except Ukrainian and Belorussian, which remain in the shadow of the Russian cultural influence.

Table 2: Number of English Slavisms versus number of speakers

| Language | Number of borrowings | Number of speakers in millions |
|----------------|----------------------|--------------------------------|
| Russian | 481 | 167.0 |
| Polish | 56 | 44.0 |
| Serbo-Croatian | 33 | 21.0 |
| Czech | 29 | 12.0 |
| Bulgarian | 13 | 9.0 |
| Slovenian | 3 | 2.2 |
| Slovak | 2 | 5.5 |
| Ukrainian | 1 | 47.0 |
| Belorussian | 0 | 10.0 |
| Macedonian | 0 | 2.0 |
| Lower Sorbian | 0 | 0.01 |
| Upper Sorbian | 0 | 0.1 |

The data suggest that the number of speakers, the key determinant of the borrowing magnitude, can be mediated by other factors, e.g., the dominance of a neighboring language.

3.3 Subject-matter Areas

The data concerning the subject matter fields of the borrowings are presented in *Tables 3* and *4* and in *Chart 2*. Each of these fields will be discussed in turn. The structure of the lexemes in each of these fields is the most reliable indicator of intercultural transfer as the nature of the vocabulary mirrors the cultural entities being transferred.

Table 3: Subject-matter fields of the borrowings

| Subject-matter field | Count | Percent |
|--------------------------------|-------|---------|
| History | 141 | 21.96 |
| Ethnic designations | 92 | 14.33 |
| Geography, Geology, Mineralogy | 75 | 11.68 |
| Science and technology | 61 | 9.50 |
| Food | 46 | 7.16 |
| Zoology | 34 | 5.30 |
| Music | 28 | 4.36 |
| Literature and Art Theory | 27 | 4.21 |
| Politics | 27 | 4.21 |
| Social relations | 27 | 4.21 |
| Objects | 23 | 3.58 |
| Garments | 20 | 3.11 |
| Religion | 17 | 2.65 |
| Measures | 15 | 2.34 |
| Botany | 9 | 1.40 |
| Totals | 642 | 100.00 |

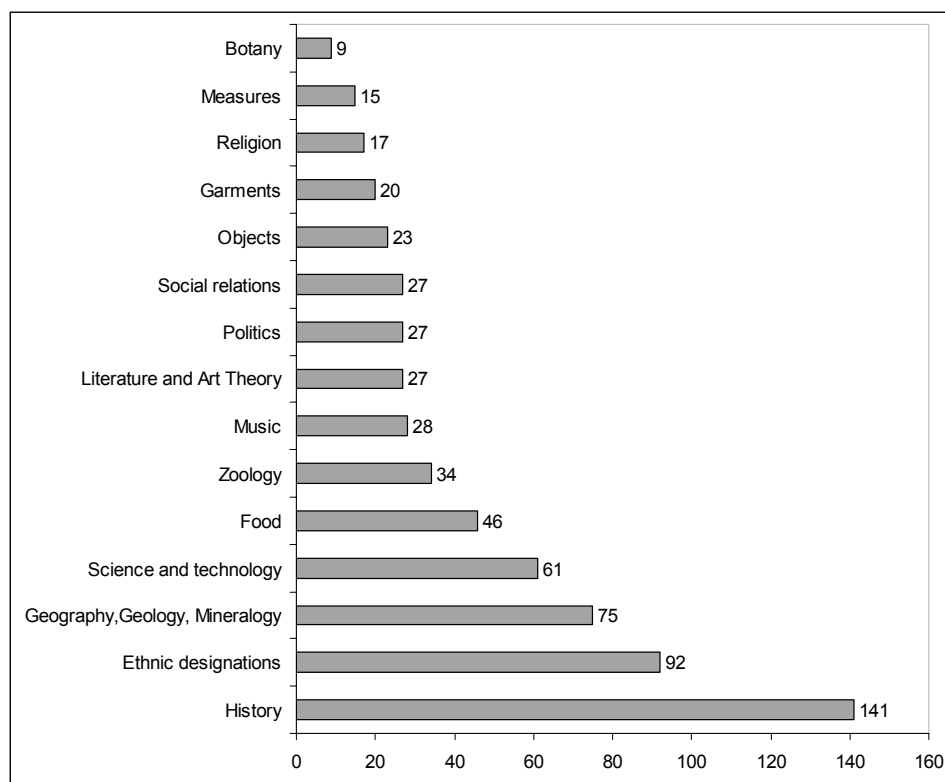


Chart 2: Subject-matter fields of the borrowings

Table 4: Subject-matter fields in each language

| | Rus- sian | Slavic | Polish | S-Cr | Bul- garian | Czech | Ukrai- nian | Slovak | Slove- nian | Totals |
|-------------|--------------|--------|--------|------|----------------|-------|----------------|--------|----------------|--------|
| History | 113 | 8 | 11 | 5 | 0 | 4 | 0 | 0 | 0 | 141 |
| Ethnic | 62 | 4 | 9 | 7 | 3 | 3 | 0 | 1 | 3 | 92 |
| Geography | 62 | 0 | 1 | 5 | 1 | 5 | 1 | 0 | 0 | 75 |
| Science | 43 | 2 | 9 | 1 | 0 | 5 | 0 | 1 | 0 | 61 |
| Food | 36 | 1 | 2 | 3 | 3 | 1 | 0 | 0 | 0 | 46 |
| Zoology | 25 | 4 | 4 | 0 | 0 | 1 | 0 | 0 | 0 | 34 |
| Music | 8 | 1 | 9 | 4 | 0 | 6 | 0 | 0 | 0 | 28 |
| Literature | 22 | 0 | 0 | 4 | 0 | 1 | 0 | 0 | 0 | 27 |
| Politics | 23 | 0 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 27 |
| Social life | 21 | 1 | 3 | 0 | 1 | 1 | 0 | 0 | 0 | 27 |
| Objects | 21 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 23 |
| Garments | 15 | 1 | 3 | 1 | 0 | 0 | 0 | 0 | 0 | 20 |
| Religion | 13 | 1 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 17 |
| Measures | 9 | 0 | 1 | 1 | 2 | 2 | 0 | 0 | 0 | 15 |
| Botany | 8 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 9 |
| Totals | 481 | 24 | 56 | 33 | 13 | 29 | 1 | 2 | 3 | 642 |

History

There are two major groups of historisms. First, there are the terms such as *hussar*, *szlachta*, *boyar*, *heyduk*, *tsarevich*, etc. related to the pre-twentieth-century, often medieval past of various Slavic peoples. Second are items such as *Bolshevik*, *commissar*, *Gosplan*, *kolkhoz*, *sovkhoz*, *subbotnik*, *udarnik*, etc. related to the Soviet communist experiment. The most recent loanwords in this group, e.g., *perestroika* come from the practice of abandoning the system. Most lexemes in this group are so-called culture-bound words, i.e., the definition will contain a reference “in Russia”, etc. In addition, being historisms these lexemes are also period-bound, which is reflected in the usage label “Hist.”. A natural result of this limited reference is low frequency of practically all items in this group. It suffices to say that out of 141 lexemes in this group only 27 (or 19.15%) was recognized by the English spell checker in Microsoft Word 2003.

Only in rare instances will a lexeme assume a general meaning. The entry *tsar* with the following two meanings exemplifies this point.

- a) Hist. The title of the autocrat or emperor of Russia; historically, borne also by Serbian rulers of the 14th c., as the Tsar Stephen Dušan.
- b) transf. A person having great authority or absolute power; a tyrant, ‘boss’, orig. U.S.

Ethnic designations

There are two major groups of designations in this category: a) the names of the Slavic peoples, such as *Bosniac*, *Czech*, *Slovak*, etc., and b) the names of minority groups

inhabiting Slavic lands, e.g., *Koryak*, *Mordvin*, *Nenets*. The first group of ethnic names is considerably more frequent than the second one. The high frequency is particularly observable with the Russian ethnic name, which has even developed its derivational nest, e.g., *Russify*, *Russianize*, *Russianist*, *Russianism*, etc.

Geography, Geology, Mineralogy

One part of the lexemes in this category refers to the morphological features of terrain and types of soil, while the other designates various minerals. The terms from the first group refer to specific morphological configurations and soil types of Russia (e.g., *chernozem*, *sierozem*, *step*, *taiga*, etc.) and the Balkans (e.g., *ponor*, *polje*, *uvala*). Terms for minerals (e.g., *nenadkevite*, *nifontovite*, *nordite*, *obruchevite*, *samarskite*, etc.) refer to the minerals and their components found in the Slavic countries, mostly Russia and/or discovered by the Slavic scholars. This group refers to highly specialized and region-bound entities, which makes them marginal in the English vocabulary.

Science and technology

Most lexemes in this category belong to either Soviet space-era vocabulary (e.g., *lunokhod*, *sputnik*, etc.), or various discoveries of scientists from the Slavic countries, e.g., *kurchatovium*, *mendelevium*, *Tesla*, *Markov*. Most of these lexemes are highly specialized and thus marginal in the English vocabulary. The most central vocabulary item in this category, the word *robot* demonstrates the role of the so-called butterfly effect in the process of borrowing. As stated in the OED (2004: s.v.), the word comes from “Czech, f. *robota* forced labour; used by Karel Čapek (1890–1938) in his play *R.U.R.* (*‘Rossum’s Universal Robots’*) (1920).” This particular item defies the general pattern of borrowing both in this thematic category and in general. An unpredictable act of creative use of language has yielded a highly frequent borrowing.

Food

Vocabulary items from this category fall under the general pattern of the culture-bound nature of the borrowings. The lexemes are either food products (e.g., *borsch*, *kielbasa*, *pierog*, *shaslik*, *stroganoff*) or spirits (e.g., *rakia*, *vodka*, *zubrowka*) characteristic for Slavic cultures. Some of these items (e.g., *vodka* and *paprika*) have gained general popularity in the English speaking countries, yet most of the lexemes remain marginal vocabulary items.

Zoology

This category comprises faunae, such as *beluga*, *borzoi*, *zubr*, etc. endemic to or characteristic of biomes of the Slavic countries. They remain marginal in the English vocabulary. The structure of this subject-matter category is akin to the botanical field.

Music and Dance

There are two groups of lexemes in this subject-matter area. First, there are traditional music forms and instruments, and dance formations such as *gusle*, *balalaika*, *tamburitza*, *kolo*, *kazachoc*, etc. Second, there are the terms related to the most prominent Slavic classical composers, e.g., *Chopinesque*, *Dvořakian*, *Tchaikovskian*. Both groups remain peripheral in the English vocabulary.

Literature and Art Theory

This group contains the lexemes relating to the traditional Slavic literary traditions, such as *bylina*, *skaz*, *Glagolitic*, etc., those belonging to the arsenal of the twentieth-century avant-garde artistic movements, e.g., *constructivism*, *Cubo-Futurism*, *Suprematism*, etc., and items relating to philological theories (*defamiliarization*, *Dostoyevskian*, *Trubeckoyan*, and others). All three groups remain marginal in the English vocabulary.

Politics

This category comprises names of the parliaments and territorial units in Slavic countries, e.g. *duma*, *Sejm*, *Sobranie*, *Skupština*, *oblast*, etc. as well as Soviet-era general political terms which are still being used (unlike previously mentioned historicisms), e.g. *apparatchik*, *disinformation*, *nomenklatura*, *residentsia*, etc. This group, as a whole, is somewhat less marginal than most other groups of English Slavicisms.

Social relations

This category comprises of mostly marginal vocabulary used to describe features and social roles, e.g. *babushka*, *kulturny*, *niekulturny*, *specialist*, etc. The category falls out of the general pattern of borrowing in which the Slavic words refer to culture-bound concepts. The only non-marginal lexeme in this group is *vampire*, which is far more frequent than other lexemes and which has developed a rich polysemic structure.

Objects

This category follows the general pattern of the culture-bound character of the borrowings and contains names of edifices and other objects related to Slavic and their neighboring cultures, e.g., *dacha*, *isba*, *terem*, *troika*. The group in general remains marginal, with the only exception of *troika*, which has developed a new sense 'a group or set of three persons (rarely things) or categories of people associated in power; a three-person commission or administrative council. Also attrib.' (OED 2004: s.v.)

Garments

Similar to the previous category, this group of words comprises garments specific to Slavic and their neighboring cultures, e.g., *rubashka*, *shapka*, *shuba*, *parka*, etc. The latter item is the only one which gained popularity beyond the borders of the Slavic cultures.

Religion

This category contains items relating mostly to Orthodox Christianity, different schools of thought, customs, and sacral object, e.g., *starover*, *slava*, *iconostas*, *riza*. These lexemes follow the general pattern of borrowing. They are culture bound and marginal in the English vocabulary.

Measures

There are two kinds of lexemes in this category. First, parochial measures, such as *arsheen*, *verst*, *tchetvert*, and second local currencies, e.g. *dinar*, *koruna*, *rouble*, *zloty*. Practically this entire category is destined to become obsolete. The measures are replaced with the metric systems and the currencies will be replaced by the euro.

Botany

This category comprises faunae, such as *badiaga*, *kamish*, *zelkova*, etc. endemic to or characteristic of biomes of the Slavic countries. They remain marginal in the English vocabulary. The structure of this subject-matter category is akin to the zoological field.

3.4 Timeline

The temporal distribution of the English Slavisms presented in the *Table 5* and *Chart 3* mirrors enormous communication improvements of the nineteenth and in particular twentieth centuries. Almost one half of all Slavisms falls in the twentieth century, the last two centuries comprise over 87% of all lexical borrowings.

Table 5: English Slavisms through centuries

| Century | Number | Percent |
|---------|--------|---------|
| 1200s | 1 | 0.16 |
| 1300s | 5 | 0.78 |
| 1400s | 1 | 0.16 |
| 1500s | 49 | 7.63 |
| 1600s | 32 | 4.98 |
| 1700s | 57 | 8.88 |
| 1800s | 191 | 29.75 |
| 1900s | 306 | 47.66 |
| Totals | 642 | 100.00 |

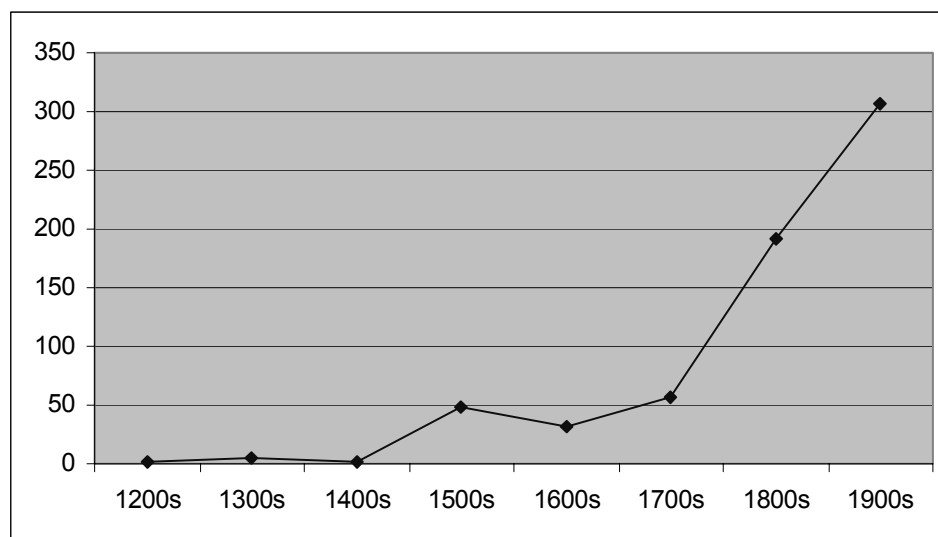


Chart 3: Slavisms through centuries

4. Conclusions

The following general conclusions can be drawn from the analysis presented in section 3.

- a) Lexical and cultural influence from subordinate to dominant language is by and large limited to the culture-bound items. Borrowed vocabulary items remain marginal in the overall English vocabulary. Several exceptions from this trend, i.e., the words which have made it to the core of the English vocabulary, are result of the butterfly effect and cannot be accounted for by some general trend,
- b) Lexical influence of each particular language is directly proportionate to language size. Exceptions from this general trend occur, as demonstrated by East Slavic languages, when one Slavic language clearly dominates others,
- c) The timeline of borrowing is directly proportionate with the growth and deepening of international communication networks in the nineteenth and in particular twentieth century.

References

- ALVARADO, S. (1989) Eslavismos en el lexico espanol. *Boletin de la Real Academia Espanola* 1989 Sept–Dec; 69 (248): 403–416.
- ALVARADO, S. (1990) Mas eslavismos para el lexico espanol. *Boletin de la Real Academia Espanola* 1990 May–Aug, 70 (250): 421–427.
- ANIĆ, Vladimir (1994) *Rječnik hrvatskoga jezika*. Zagreb: Novi Liber, 2nd revised and enlarged ed.
- DINI, P.-U. (1993) Slavizmai Lizijaus ir Engelio katekizmuose: Gretinamoji analize. *Lietuvių Kalbotyros Klausimai* 30 (Vilnius 1993) 33–36.
- FILIPOVIĆ, R. (1997) The Theoretical Background of the Project 'The English Element in European Languages'. *Studia Romanica et Anglica Zagabiensia* 42 (1997) 105–111.
- FILIPOVIĆ, R. (1986) *Teorija jezika u kontaktu*, JAZU.
- LALLUKKA, G. (1985) Сопоставление славянских заимствований в венгерском и финском языках. *Studia Slavica Finlandensia* 2 (1985) 81–109.
- HUMBOLDT, W. (1959) *Bildung und Sprache; eine Auswahl aus seinen Schriften*. Besorgt von Clemens Menze. Paderborn, F. Schöningh.
- MENCKEN, H. L. (1921) *The American Language*, New York: Alfred Knopf (<http://www.barleby.com/185>)
- MLA (2004) *MLA International Bibliography*, electronic database, accessed on April 16, 2004, via http://info.lib.asu.edu/NetAns2/gate.exe?f=doc&p_d=dbdb&state=8q4bq6.2.1
- OED (1989) *Oxford English Dictionary Online* (<http://www.oed.com>)
- PAPP, F. (1973) A Few General Characteristics of Words of Slavic Origin in the Hungarian Language. *VII Międzynarodowy Kongres Slawistów w Warszawie 1973: Streszczenia referatów i komunikatów*. Warszawa 1973. 407–408.
- POHL, H.-D. (1990) Slovenske (in slovanske) izposojenke v nemškem jeziku Koroške. *Slavistična revija* 38/2 (1990 Apr–June) 101–104.
- SAPIR, E. (1921) *Language. An Introduction to the Study of Speech*, New York: Harcourt, Brace (<http://www.barleby.com/186>)

- VRABIE, E. (1992) Slavic Influence on Romanian: A Case of Exaggeration. *General Linguistics* 32/2–3 (1992) 105–110.
- WIERZBICKA, A. (1992) *Semantics, Culture and Cognition: Universal Human Concepts in Culture-specific Configurations*. The Hague.
- WINFORD, D. (2002) *An Introduction to Contact Linguistics*. London: Blackwell.
- WUNDT, W. (1911) *Völkerpsychologie*. Leipzig: W. Engelmann.

Acknowledgement

The author would like to extend his gratitude to Rick P. ROSENBERG for proof reading this paper.