What can Lexicography Gain from Studies of Loanword Perception and Adaptation?

Mirosław Bańko, Milena Hebal-Jezierska
Institute of Polish Language, Institute of West and South Slavic Studies, University of Warsaw
m.banko@uw.edu.pl, milena.hebal-jezierska@uw.edu.pl

Abstract

In normative dictionaries and usage guides, many loanwords are dismissed as ‘unnecessary’ on the grounds that they have a native synonym, which should be favourred instead. If there is no native equivalent of a loanword in the recipient language, one is often invented in order to eradicate the unwanted loan. However, more detailed studies, in particular, those referred to in this paper, suggest that there is no such thing as fully equivalent words: even if two words have the same designative meaning, they differ in other respects, which makes them mutually unexchangeable except for contexts where the difference between them is inessential.

The goal of this paper is to demonstrate that the meaning potential of lexical loans is different from that of their native synonyms, just because their form is different and differently perceived by language users. The different perception of loanwords can in turn affect their semantic development, thus causing a loanword and its native synonym to diverge. The authors of normative dictionaries and language guides should, therefore, give more consideration to lexical borrowings before they condemn them as ‘unnecessary’ or ‘snobbish’.

Keywords: loanwords; purism; synonymy; variance

1 Introduction

Lexical loans can be divided into two groups: those which are motivated by nominative needs, i.e. the necessity to name a new object or a new phenomenon of foreign origin, and those which appear for expressive needs, because they seem to introduce certain stylistic values, emotional overtones, etc. Polish smartphone and Russian смартфон, both coming from English smartphone, are examples of the former group, while the English interjection wow, now used in many other languages, can exemplify the latter group. There are obviously more ways to fill lexical gaps, as the French alternative names for smartphone – ordiphone and téléphone intelligent – show, but we will not be dealing with them here. Instead, we will restrict ourselves to borrowings proper, i.e. words taken from a donor language with possible alterations in their pronunciation, spelling, morphological and syntactic features, sometimes also in their meaning.
By definition, loanwords borrowed for nominative needs have no commonly known synonyms in the recipient language, not at least at the moment they enter it. Loanwords adopted for expressive needs, again by definition, do have some synonymous words and that puts them in a disadvantageous position whenever purist attitudes or specific concern about the ‘economy’ of language come into play. Many such loanwords are dismissed as ‘unnecessary’, criticized as ‘overused’, pointed out as examples of bad taste and snobbery. Negative assessments of them are expressed in the popular press, scholarly literature, academic textbooks, language guides and dictionaries alike.

It is not our intention to question the efforts of normatively oriented linguists and lexicographers who aim at reducing the number of loanwords in a language. Critical assessment of word borrowing – and of particular borrowings – is something needed, if only because there are readers waiting for such criticism. Dictionaries have to account not only for how words are really used, but also for what the language users think about their correct use. On the other hand, lexicographers should be aware that cases of full equivalence between a loanword and its native synonym are practically non-existent. Even if two words have the same designative meaning, they differ in other respects and evoke different associations in the minds of the language users. Over the years, such associations may stabilize and become part of the designative meaning, thus making the originally synonymous words diverge.

The goal of this paper is not so much to remind us of these relatively simple truths, as to demonstrate that the adaptation of loanwords in the recipient language is guided, at least to a certain extent, by the interplay between their form and meaning. A tendency can be observed to maintain harmony between the form and meaning of loans, which can manifest itself, inter alia, in how the original spelling of a loan is assimilated in the recipient language and how its meaning is shaped in relation to its native synonyms. The tendency will be illustrated with a number of examples later on in the paper. Let us begin, however, with examples of purist attitudes from Polish, German and Czech lexicography and linguistics.

2 Examples of purist attitudes

In a textbook for students of Polish language and literature, Andrzej Markowski (2005), a prominent linguist, chairman of the Council of the Polish Language, gives long tables of ‘overused words’ and ‘vogue words’, usually of foreign origin, and demonstrates, by means of invented examples, how they can be replaced by other words, most of them native or borrowed so long ago that their foreign origin is no longer recognizable. The same or similar loanwords were reviled earlier in a standard dictionary of Polish usage, edited by the same author (Markowski 1999), in some of his other books, and in many popular dictionaries and usage guides, compiled by others. It is worth stressing that Markowski’s position is far from extreme purism. His judgments are of a ‘better/worse’, not ‘yes/no’ type, yet his decisions are clearly not based on detailed analyses of the meaning and use of the particular words he
paired. Had he looked at them more carefully, he would have found distinctions which make the words mutually unexchangeable, except for contexts where the difference between them is inessential.

The attitude to loanwords varies depending on the political and sociolinguistic situation, as well as the normative tradition in particular countries. In Germany, around 300 dictionaries of loanwords were published between 1801 and 1945, around half of them belonging to the class known as *Verdeutschungswörterbücher*, literally ‘Germanizing dictionaries’. They were not intended to explain the meaning and illustrate the use of borrowed words, but rather to demonstrate how these could be replaced by native words, some of them specifically invented for this purpose (Lipczuk 2007, 2011). No doubt the regional disintegration of Germany before 1871 favoured the development of national purism, but purist attitudes developed within German society even after the unification of the country, because the rising power of the state created favourable conditions for German nationalism. Also the romantic tradition of treating the language as the embodiment of the spirit of a nation caused many Germans to believe that loanwords posed a threat not only to their language, but also to their national identity.

The example of Germany, where even international words became the object of purifying actions (cf. *Rundfunk* and *Fernsprecher*, coined to replace *Radio* and *Telephon*, respectively), is an extreme one, but similar ‘nativizing’ dictionaries are known from the lexicographic tradition of many countries. In Poland, which from 1795 to 1918 was partitioned among Russia, Prussia and Austria, the concern about the language was steadily expressed at the time and took on different forms. At one end of the scale was Linde’s (1807-1814) six-volume dictionary of the Polish language, based on citations from about 800 sources, an attempt to save the treasures of the language and help the nation to survive the difficult time (Adamska-Sałaciak 2001). At the other end there were a number of much smaller dictionaries and usage guides whose aim and content made them similar to German *Verdeutschungswörterbücher*. Among them, Kortowicz (1891) is a good example, see Leszczyński (2000) and Czesak (2007) for information about his dictionary.

In the history of the Czech language, purist attitudes appeared in the times of Jan Hus and have been present continuously thereafter, up to the present day (Engelhardt 2001: 235). The purist trends were particularly strong at the end of the 19th century and between 1920s and 1940s. At the end of the 19th century purists tried to eliminate words of foreign origin, particularly Latin and Greek internationalisms, as well as German and French loans. A number of neologisms were formed on the basis of native words, but the new coinages often replicated the structure of the words they were supposed to replace (Engelhardt 2001:237). For instance, the Czech lexical innovation *pololetí*, patterned on the German word *Halbjahr* (literally ‘half year’), was invented to substitute for the Czech internationalism *semestr* (cf. Latin *semestris* ‘six-monthly’). As was often the case, the substitution failed, with both *semestr* and *pololetí* being used in present-day Czech.

At the beginning of the 20th century the purist tendencies in Czech linguistics became even stronger. Linguists aimed to eliminate not only proper loans, but also lexical and syntactic calques. Many usage
guides warning language users against loans of different kinds were published. Purists were particularly eager to identify German loans everywhere, even in native constructions (Král 1917), and they formed bizarre neologisms to replace them. Many of the new coinages had a short life, but some have survived to the present. A good example is the word *rozhlas*, which was introduced with the intention to replace *broadcasting* and *radio* (from the verb *hlásit* ‘report’ and the suffix *roz-,* denoting the spread of something from one place).

After the communists came to power in the Czech state, especially after 1948, English loans were fought most vigorously and replaced with native words or their spelling was changed to conceal their western origin. Many native neologisms were created at that time, e.g. *silostroj* (a compound word of the meaning ‘power and machine’) was introduced to replace *motor*, and *samohyb* (another compound combining the meanings of ‘itself’ and ‘move’) was intended to take the place of *auto*, cf. Svobodová (2009: 33). Nowadays the tendencies to purify the Czech language are not so strong, but protective attitudes can still be observed, because some linguists are afraid that the increasing presence of foreign words poses a threat to the Czech language.

3 Why are native equivalents never fully equivalent to lexical loans?

Many linguists, philosophers and literary historians have claimed that no two words can be fully equivalent with respect to their linguistic function. One can find the same opinion among lexicographers, cf. the often-quoted passage from Urdang’s introduction to *The Synonym Finder*:

> Those who work with language know that there is no such thing as a true ‘synonym’. (...) Even though the meanings of words may be the same – or nearly the same – there are three characteristics of words that almost never coincide: frequency, distribution, and connotation. (Urdang 1978)

Ullmann takes a less extreme position on this point and explains why cases of absolute synonymy are very rare:

> (...) it is perfectly true that absolute synonymy runs counter to our whole way of looking at language. When we see different words we instinctively assume that there must also be some difference in meaning, and in the vast majority of cases there is in fact a distinction even though it may be difficult to formulate. Very few words are completely synonymous in the sense of being interchangeable in any context without the slightest alteration in objective meaning, feeling-tone or evocative value. (Ullmann 1964: 142)

As ‘completely synonymous’ he mentions technical terms, e.g. *caecitis* and *typhlitis* can both be used with reference to the inflammation of the blind gut. However, even such names differ with respect to non-designative features, e.g. they evoke different associations in the minds of the language users.
One often hears that language does not tolerate fully equivalent words and differentiates them, thus eliminating cases of full equivalence. The tendency to avoid redundant means of expression is said to be evidence that language is governed by laws of economy (Nagórko 2004: vii). However, such propositions do not explain the inner mechanism of linguistic economy and, in particular, they do not explain why there should be a difference in meaning, broadly understood, between loanwords and their native synonyms. Our position is that this has something to do with the word form itself. The unfamiliar forms of lexical loans are perceived differently from the familiar forms of their native synonyms and the difference in perception may result in different semantic development of such words.

For example, shortly after kurzort, a 19th-century borrowing from German, appeared in Polish, a native term uzdrowisko (literally ‘health resort’) was coined with the intention to relegate the unwanted loan from the language. However, this effort failed to have the desired effect: instead of disappearing, kurzort changed its meaning to ‘popular and snobbish holiday place’. The change was very likely directed by the connotations of the word: some pre-war dictionaries (e.g. Słownik wyrazów obcych of 1921) informed that kurzort was most often used with reference to health resorts in Germany and the preference for foreign places in its use is still visible in modern texts. Furthermore, the collocation image of kurzort includes such features as exclusiveness (strangely enough, not in conflict with popularity), modernity, reputation and elegance, whereas in the collocation image of uzdrowisko it is tradition and aesthetic values that are best seen (see Bańko 2013a for a more detailed analysis of both words).

The influence of word form on word meaning can be studied in texts and other cultural artifacts (e.g. by collocation analysis or Google image inspection), but it can also be brought to light by means of experiments. The results published by Song and Schwarz (2010) are worth quoting here. They experimented with nonce words, some of them familiar in shape, some strange, and observed a correlation between the familiarity of a word and its perception. For instance, fictitious food additives with names difficult to pronounce were evaluated as more harmful than food additives with easy names. Similarly, roller-coaster rides in a fictitious amusement park were judged as more risky and more exciting when their names were strange and difficult. Song and Schwarz explain the effect with a mistaken projection of the difficulties the subjects experienced in processing the unfamiliar words onto the referents of the words: unaware of the source of difficulty, the subjects attributed it to the referents, judging them as more risky, more dangerous, more harmful, etc. (for a critical review of these studies see Rączaszek 2013).

It would be premature to claim that Song and Schwarz’s findings can explain all the distinctions observed between a synchronic loan and its native synonyms. More experiments are needed and they should be done on real language data, not on invented words. We will next briefly describe a project designed to perform more systematic research in this regard, using both linguistic and psycholinguistic methods.
4 About the APPROVAL project

The aim of the APPROVAL project is to search for various factors bearing on the psychological perception, social reception and linguistic adaptation of loanwords. Among the possible factors, the relation between word meaning and word form is of particular importance, because we assume that the form of a word is not irrelevant (contrary to the widely accepted view of the arbitrary nature of linguistic signs, a foundation stone of Saussurian linguistics). The form of a word can influence its meaning (cf. kurort above), but also the meaning of a word can affect its form, e.g., by hindering the process of a loanword’s adaptation (the word jazz can be a case in point: though borrowed to Polish and Czech almost a hundred years ago, it still appears mainly in its original spelling in both languages, very likely because the foreign spelling reflects better the symbolic values associated with jazz music in Poland and the Czech Republic, see Bańko and Hebal-Jezierska 2012 for details).

We also assume that fully equivalent words do not exist, so we focus on comparative analysis of word pairs (sometimes triples, quadruples, etc.) in which one element is of foreign origin, the other native, or in which one element has the original spelling, while the other is graphically adapted to the recipient language. Fifty Polish word pairs have been subjected to in-depth analysis, based on language corpora and other data, not excluding the evidence in the language itself (e.g. we treat the frequency of a word as indicative of its importance and we pay attention to secondary uses of a word, its derivatives and idiomatic expressions, because such data reveal some of the typical associations the word calls up in the minds of its users and help to draw the stereotypical image of its referent). As this paper is being prepared, most of the 50 synonym and variant pairs have been already inspected and the results are available on the project website, see http://www.approval.uw.edu.pl/en_GB/start pl.

In order to make the observations more credible, the same research is being done on corresponding Czech word pairs which serve as a control group, e.g. the Polish pair absurdalny – niedorzeczny ‘absurd, nonsensical’ corresponds to the Czech pair absurdní – nesmyslný, in which the first element comes from the same root as the first element of the Polish pair. By composing the research material this way, it became possible to study the adaptation processes in two cognate languages on the basis of comparable examples, using the same methodology, the same kind of data, and even the same description format.

In total, 100 pairs in two languages will have been analyzed by the end of the project, using corpora, as well as dictionaries, web archives, digital libraries, library catalogues, Google images and other sources of language relevant data (e.g., library catalogues are being used to check frequencies of words in book titles). In addition, psycholinguistic experiments are being carried out on the Polish material to enrich and verify the observations based on language corpora and other textual and non-textual sources by means of linguistic methods (see the project website for details).

1 The name of the project comes from ‘Adaptation, Psychological Perception and Reception of Verbal Loans’. It is also meant as a reminder that in normal circumstances loanwords do not pose a threat to a language; to the contrary, they add to its wealth.
The results gained so far are encouraging and they largely support the hypothesis of there being a relationship between form and meaning in the adaptation of lexical loans. For limits of space, a few examples from Polish will have to suffice here. We will focus on selected details, with no intention to account for a full analysis of any of the words mentioned below.

4.1 strofa and zwrotka

The Polish words strofa and zwrotka both mean ‘stanza’, but in technical literature usually the former word is used, while in the general language the latter one is more common. The likely reason is not only that strofa is of Greek origin (borrowed via Latin), but also that zwrotka contains a familiar suffix -k- which in many other nouns (though not in zwrotka) has a diminutive function.

The difference between these two words was first observed in corpus analysis, especially in their collocation images. For example, zwrotka, but not strofa, is used in reference to popular songs and children’s poems, strofa can be the subject of aesthetic evaluation (cf. piękne strofy ‘beautiful stanzas’) and artistic activity (cf. pisać, układać strofy ‘write, arrange stanzas’), while zwrotka is less frequent in such contexts. In addition, strofa can be recited, but zwrotka is sung. Only zwrotka collocates with the word refren (‘refrain’), which confirms its connection to songs.

A study of free associations, based on Osgood’s semantic differential, was next carried out. Twenty-two subjects took part in it, each asked to mention up to three associations for one word, so the maximum number of associations for a word was 66. Among the associations noted more than once, refren ‘refrain’, rymy ‘rhymes’, śpiewanie ‘singing’, muzyka ‘music’ and ognisko ‘camp-fire’ were given only for zwrotka, while szkoła ‘school’, poezja ‘poetry’, literatura ‘literature’ and Mickiewicz (the best known Polish poet) were given only for strofa. In addition, though piosenka ‘song’ and wiersz ‘poem’ were mentioned for both words, piosenka had a frequency of 19 with zwrotka and 3 with strofa, while wiersz appeared 16 times with strofa and only 3 times with zwrotka. As can be seen, the results of corpus analysis are in line with the study of free associations.

4.2 helikopter and śmigłowiec

Though helicopter and śmigłowiec both mean ‘helicopter’ in Polish, their stylistic distribution is different. Helicopter is common in spoken language and in many other language varieties, while śmigłowiec tends to be used in technical literature. This is probably the reason why in the domain lego.com/pl-pl, belonging to the producers of Lego bricks, the Google search engine finds far more occurrences of helikopter than śmigłowiec. As far as book titles are concerned, helikopter can be found on the covers of children’s stories, while śmigłowiec appears in the titles of books on aeronautical technology and military science. Among the Google images indexed with the word helikopter, toys and miniature models are more frequent than among images indexed with the word śmigłowiec.
However, a more interesting and more surprising observation about *helikopter* and *śmigłowiec* can be made when comparing their relative frequencies in certain contexts. Though on Polish-language websites *helikopter* is several times more frequent than *śmigłowiec*, the quantitative advantage of *mały helikopter* ‘small helicopter’ and *szybki helikopter* ‘fast helicopter’ over *mały śmigłowiec* and *szybki śmigłowiec*, respectively, is significantly lower. Moreover, *leki helikopter* ‘light helicopter’ is less frequent than *leki śmigłowiec*. Apparently, small, light and fast machines of this type are more often referred to with the word *śmigłowiec* than its relative frequency to the word *helikopter* would suggest. This may be due to the fact that the name *śmigłowiec* is related to the words *śmigło* ‘propeller’, *śmigły* ‘swift’ and *śmigać* ‘move quickly, zip (around)’.

However, the overall picture is not quite clear yet, partly because the relative frequencies of *helikopter* and *śmigłowiec* in two reference corpora of Polish – Narodowy Korpus Języka Polskiego and Korpus Języka Polskiego PWN – are opposite to those found on the Internet and partly because the study of free associations has yielded different results for these two words than obtained in corpus analysis. Further research into the psychological perception of *helikopter* and *śmigłowiec* is planned within the APPROVAL project with the intention to confirm or refute the conjectures made on the basis of corpus data. Whatever the results of the research, there is no doubt a difference between *helikopter* and *śmigłowiec* in their semantic content, if only non-designative components of the word meaning are allowed.

### 4.3 eksplodzę and wybucha, kuriozalny and osobliwy

*Eksplozja* ‘explosion’ and *wybuch* ‘explosion, outbreak, outburst’ may refer to the same kind of events, but the phenomena referred to by the former word are perceived as stronger and more violent. The difference is so distinct that it has even been noted in the definitions for these two words in some dictionaries. There are more synonym pairs in which the referents of a loanword seem larger and more powerful than the referents of its native synonym, cf. *dewastować* and *niszczyć*, both meaning ‘destroy’. Here the first element, cognate with English *devastate*, denotes a purposeful or mindless activity, especially against public property or natural environment.

However, sometimes the difference between a loanword and its native synonym lies not in the referents themselves, but in the way they are talked about, e.g. in the values conveyed. The adjective *kuriozalny* ‘peculiar, bizarre’, related to Latin *curiosum*, is half as frequent in Polish as its native counterpart *osobliwy* ‘peculiar’, but in parliamentary reports the former word prevails overwhelmingly. A closer inspection shows that Polish MPs need it to criticize their political opponents, e.g. *Pana poglądy są kuriozalne, panie pośle* ‘Your views are bizarre, Mr. X’.

The tendency to use difficult and erudite words for hyperbolic effects, whether to make a phenomenon look more powerful or just to convey negative attitudes, can be well explained in the context of Song and Schwarz’s (2010) experiments discussed above.
4.4 dealer and diler

The last of our examples is different from the previous ones, because it is not concerned with a synonym pair. It deals with a pair of spelling variants of the same word, a recently new Polish borrowing from English. However, the situation is much the same as before, because one element of the pair is foreign while the other one is ‘nativised’ (rather than native), and the two elements therefore exhibit the same ‘unfamiliar – familiar’ opposition as in the case of synchronic loans and their native synonyms. Thus, the necessary conditions are met for different associations to evolve around the different words.

As all variants, dealer and diler have the same designative meaning, but their stylistic distribution and the areas of their application are not identical. In the press, dealer is almost twice as frequent as diler and used mainly with reference to car vendors, whereas diler has equal frequency in automobile and drug-related contexts. In literary texts, on the other hand, diler is twice as frequent as dealer and applied usually to drug sellers, whereas dealer is equally often used in automobile contexts. Apparently, the foreign variant is more prestigious and better suited to name the job of authorized vendors in car showrooms; the domestic variant, on the other hand, is unpretentious and corresponds well with the dubious job of drug peddling. The difference can well be observed in the Google image galleries, too, which once more confirms the usefulness of Google images in linguistic analysis (Bańko 2013b).

5 Conclusions for practical lexicography

Our intention was not to question the need for normative assessments in dictionaries, nor to claim that cumulative synonym dictionaries have no raison d’être. The conclusions from our work are relevant to the theoretical foundations of lexicography, but also to lexicographic practice. It is important for lexicographers to be aware that distinctions often may lie where similarity seemingly prevails (in a way, all of language is based on distinctions, and here we are in complete agreement with de Saus-sure). Furthermore, it is important to realise that in the adaptation of lexical loans, form and meaning are interdependent: the form of a word can affect its meaning, but it can also be can influenced by it. In many cases, a tendency to maintain harmony between a word form and word meaning can be observed in the process of loanword adaptation.

Better recognition of differences between near synonyms is essential for both monolingual and bilingual lexicography. Adequate definitions, especially in production dictionaries, should explain differences between synonymous words. Adequate translation equivalents are dependent, among other things, on how well near synonyms of the source language and the target language are discriminated. This is not to say that each dictionary ought to be equally specific in its treatment of word meanings. For example, decoding dictionaries need not focus so much on distinctions between words as may be expected from encoding dictionaries. Lexicographers should consider for themselves to what extent
the observations made in this paper may be useful in their work. In any event, more caution is advised before assessing a loanword as unnecessary and more thoroughness is needed in how new words borrowed from other languages are treated. It is not enough to blame those who use them of snobbery.

A separate question, not to be dealt with here, is how our findings can be incorporated in what are called distinctive synonym dictionaries, which are intended to account for differences among synonyms rather than to gather as many words close in meaning as possible (cf. Dystynktywny słownik synonymów by Nagórko et al. 2004 as an example of works of this type). Another area where it is more important to show differences between words than to identify similarities is the so called synonym discussions, known from some dictionaries (cf. special paragraphs, headed Synonyms, in The American Heritage Dictionary).

6 References

Kortowicz, E. S. (1891). Oczyszczaciel mowy polskiej, czyli słownik obcosłów, składający się z blisko 10,000 wyrazów i wyrażeń z obcych mów utworzonych a w piśmie i w mowie polskiej niepotrzebnie używanych, oraz z wyrazów gminnych, przestarzałych i ziemszczen w różnych okolicach Polski używanych z wysłowieniem i objaśnieniem polskim. Poznań: czcionkami drukarni „Dziennika Poznańskiego”.


Słownik wyrazów obcych. 25.000 wyrazów, wyrażeń, zwrotów i przysłów cudzoziemskich, używanych w mowie potoczonej i w prasie polskiej, 9th ed. Warszawa: Wydawnictwo M. Arcta, 1921.


Acknowledgements
This project has been supported by a grant from the National Science Centre, Poland, no DEC-2011/03/B/HS2/02279.