Loanword adaptation patterns: the case of English loans in Polish and Czech

Mirosław Bańko
Faculty of Polish Studies
University of Warsaw

Diana Svobodová
Pedagogical Faculty
University of Ostrava

Introduction

This paper addresses the question of whether there exists some typical order by which different characteristics of loanwords are adapted to the recipient language. Does, for example, the adaptation of spelling typically precede the adaptation of morphological structure of loans or vice versa? Generally speaking, adaptation processes may affect the meaning of loans, their syntactic behavior and their formal characteristics (to mention just the main possibilities). In this paper, only certain formal characteristics of loans are examined – their pronunciation, inflections, derivational potential, and spelling. The question is therefore: is there some typical order by which the adaptation of a loanword's pronunciation, morphology and spelling proceeds? The answer depends on the typological features of the donor and recipient languages, their writing traditions, the script they use, etc. The question is a general one, but the present study is limited to English loans in Polish and Czech.

The approach we use can be classified as empirical and quantitative. We divide a sample of loanwords into groups corresponding to different adaptation patterns. We then analyze the patterns – in particular, the number of loans assigned to them – in order to reveal a typical adaptation order these patterns might be a result of.

Examples of loanword adaptation processes

Words migrate between languages, but are usually not adopted unchanged. They change to adjust to the recipient language: its vocabulary, grammar, pronunciation and writing. They therefore may change their meaning, their syntactic dispositions and their form, to name just the main possibilities.

For example, the word *puzzle* can be used as a noun or a verb in English, but was borrowed into Polish as a noun only, taking the form *puzzel*. Moreover, the only meanings of *puzzle* transferred to Polish are 'jigsaw puzzle' and 'something difficult to understand' (the latter could also have been formed in Polish by metaphorical extension). On the other hand, a piece of a jigsaw puzzle can be named *puzzel* in Polish, which is not true for *puzzle* in English. As can be seen, the Polish word is more restricted in its syntactic functions and meaning than its English source, but has one extra meaning which is not attested in English. It can also be noted that its Polish spelling is modeled partly on the English spelling, partly on the English pronunciation. The word inflects

and is active in word formation processes: the adjective *puzzlowy* is based on it and so may be the noun *puzzlomania* (unless borrowed directly from English *puzzlemania*).

The difference in meaning and syntactic behavior is greater in the case of *skate*, another English loan in Polish. Again, the word can be a noun or a verb in English and only a noun in Polish, but its only meaning in Polish does not correspond to any of its meanings in English: according to dictionaries of Polish, *skate* means 'skater'.

The Czech borrowings from English *puzzle* and *skate* are similar to those in Polish, but not identical. *Puzzle* became *puzzle* or *pucle* in Czech, both words able to refer to a whole jigsaw puzzle, as well as a piece of it. The variant *pucle* (which is occasionally encountered in Polish, too) is used rarely, in substandard Czech, and seems to be based either on the Italian pronunciation or on a certain analogy to German pronunciation, even though in German the word is not pronounced this way. Both *puzzle* and *pucle* inflect and are active in word formation, cf. the adjective *puzzlový*, the verb *puzzlovat* and the deminutive *puzzlík*/*puclík*, which refers to a piece of a jigsaw puzzle. As for *skate*, it has a different meaning in Czech than in English, yet not the same as in Polish: in Czech, *skate* means 'skateboard'.

More could be said about how *puzzle* and *skate* were adapted in Polish and Czech, but the above is enough to illustrate the changes that may occur in loanword adaptation.

Assumptions

Setting meaning and syntactic functions aside, this paper focuses on certain formal aspects of loanword adaptation, i.e. their pronunciation (P), inflection (I), derivational potential (D) and spelling (S). These four dimensions of loanword adaptation are logically independent, so in principle a word can be adapted along any of them and not adapted on others. Indeed, it is their logical independence that allows for the use of the term *dimensions* here.

A loanword may be adapted to a greater of lesser degree, rather than simply adapted or not. However, to simplify the issue a bit, we will assume a binary assessment on each particular dimension of loanword adaptation. Each given loan can therefore be represented as a four-digit binary number in which the digit 0 on a particular position means that the loan is not adopted on the corresponding dimension while the digit 1 has the opposite meaning. For example, if the dimensions are considered in the order PIDS, the number 1100 will represent a loanword adapted with respect to its pronunciation and inflection, but not with respect its derivational potential or spelling.

Theoretically, there are 16 four-digit binary numbers, from 0000 (meaning no adaptation on any dimension) to 1111 (denoting adaptation on all dimensions). In practice, however, the four dimensions are not independent, which makes some numbers more probable than others. For example, in the case of English loans in Polish or Czech, one can expect loanwords to be adapted in the PIDS order: from pronunciation through inflection and derivational potential to spelling. There is no need, of course, for a loanword to go through all these four adaptation phases, but if the above order prevails, then, e.g., the number 1110 (derivation prior to spelling) should be encountered more often in a sample of loans under investigation than 1101 (spelling prior to derivation), and the number 1101 (inflection prior to derivation) should be more frequent than 1011 (derivation prior to inflection). This is, basically, the idea underlying the present study: by counting the instances of each adaptation pattern, coded as four-digit binary numbers, we expect to identify the most typical sequence of loanword adaptation phases in Polish and Czech.

Wanting to explain why in Polish and Czech the adaptation of loanwords should proceed in the PIDS order, one could use the following reasoning. First of all, pronunciation has to be adjusted to the recipient language, because the phonology of English is so different from that of Polish or Czech that pronouncing a word the English way would be too difficult for most speakers and would sound snobbish anyway. Next, inflections are very likely to appear, since in a highly inflected language it is awkward not to inflect words, in particular nouns, and it is usually nouns that are borrowed. In the third place, derivational potential is unlocked, because some kinds of derivations are necessary for communicative purposes, e.g. denominal adjectives or feminine counterparts of masculine personal nouns. Spelling adaptation comes at the very end, because English loans can be written the English way in a language using a Latin-based script, and often the original spelling is preserved long after a word is borrowed, e.g. to make the word look foreign and better harmonized with its meaning.

There may be deviations from this general scheme, for example, spelling adaptation may be necessary for a word to be able to inflect. Another reason for deviations is that original spelling makes a word look foreign, which may be desired in the case of some loans, but not all. In addition, the order of adaptation phases may be different for common nouns than for proper names, the latter being more likely to retain their original spelling and nearly-original pronunciation. Nevertheless, our initial stance was that the PIDS order can be reasonably assumed for English loans in a highly inflected language like Polish or Czech, unless analysis concludes otherwise.

The assumption above is incongruent with some observations in the literature on the subject. For example, Bartmińska and Bartmiński (1997: 40) give a different adaptation order in Polish, namely IDPS. The difference probably results from the fact that the core of their material was foreign family names. Such names are usually assimilated more slowly than proper names (mainly to facilitate the recognition of referents, see Bartmiński 1992), and their pronunciation is often very close to the original. This may explain why pronunciation was placed as late as in the third place in the adaptation pattern postulated by these authors.

In Czech, spelling adaptation is traditionally considered to be interrelated with other forms of loanword adaptation, with some authors discussing “orthographical-orthoepical” adaptation (Mravinacová 2005: 190). However, in many English words the spelling is either not adapted at all, or adapted in colloquial (substandard) variants that are not acceptable in formal texts (Bozděchová 1997: 278). The adaptation order is usually given in the PIDS pattern, as uninflected forms are not organic parts of the synthetic language sentence structures (Daneš 2002: 32). Nevertheless, the adapted forms of adjectives and verbs have usually denominative derivational character (Světlá 2005: 99, Martincová 2005: 119), which would in fact lead to the PDIS order.

To find out whether our assumption about the typical adaptation order in Polish and Czech is correct, or perhaps the other authors are right when taking a different position, we decided to analyze data from our previous study concerned with English loans in Polish and Czech.

Evidence

The previous study aimed to test the hypothesis that in Czech, lexical loans are more strongly integrated with the spelling, pronunciation and grammar of the language than they are in Polish (Bańko, Svobodová 2015). We investigated a hundred English words which had been borrowed into both Polish and Czech and assigned points to them, according to how well they were adapted to the pronunciation, grammar and spelling of the recipient language. The total score of

a word was used as a measure of its adaptation and the overall score of all words selected for the study was an index of the recipient language's ability to adapt lexical loans.

The present paper draws on the same data and can be thought of as a by-product of the earlier research. As a by-product, it inherits some features of the original work which should be briefly explained now.

First, as in the previous study, we decided to ignore cases of loans that are pronounced exactly the English way (the rationale is that the original pronunciation of English loans in Polish or Czech is rare and sounds unnatural). As a consequence, of the 16 theoretically possible assignments of values 0 and 1 to P, I, D, and S dimensions, half are not considered here, namely those which have value 0 in the position P (indicating they are un-adapted phonologically).

Second, though in the previous study loans scored one point even when they inflected only the English way (e.g., *talkshow*, *talkshows*) and two points when they had native inflections (exclusively or alternatively to the English ones), we decided to ignore English inflections here and treat a loan as adapted to the inflection of the recipient language only when it inflected by adding native morphemes. Moreover, unlike in the previous study, we decided to pay no heed to suffixation, i.e. a process in which a native suffix is added to the base form of a loan, as in Polish *celebryta* and Czech *celebrita*, both coming from English *celebrity* and both formed by means of the inflectional ending *-a*. In sum, we assigned value 1 to dimension I if and only if a loan inflected in a native way, ignoring any purely English inflection and suffixation.

Third, though the score for pronunciation and spelling in the previous study ranged from 0 to 2, and the score for derivational potential ranged from 0 to 3, we decided that loans having positive scores on these dimensions would not be differentiated here. In effect, a value of 1 was assigned on the corresponding dimension if the loan's score in the previous study was greater than 0.

Results

The results are given in tables 1 and 2. In each table, the first column contains the loanword adaptation patterns, based on the PIDS order of adaptation phases that was assumed in this study. The patterns are coded as four-digit binary numbers, according to the principles presented above. For example, the number 1000 corresponds to loans which are adapted only with respect to their pronunciation, while the number 1010 corresponds to loans which are adopted on the pronunciation and derivation dimensions. The second column indicates how many loans from our one-hundred sample fit to a given pattern, while the third column gives a list of these loans (with variant forms in parentheses, and forms unattested in dictionaries, but common enough to be included, preceded with an asterisk). The lists are exhaustive (except for the last one) and arranged in the same order as in our previous study (for ease of reference). In each table, the first line represents patterns beginning with 0 which – for reasons explained above – were excluded from the investigation.

Table 1. Loanword adaptation patterns in Polish, based on PIDS adaptation order

Adaptation pattern	Number of instances	Examples
0–––	0	—
1000	6	Barbie, CD, DVD, party, smiley, talk-show (talk show, *talkshow)
1001	2	interview (interwiew), show (*szoł)
1010	4	cool, CV, outdoor, PC
1011	2	OK (O.K., okej, okay), online (on-line, on line, *onlajn)
1100	3	coach, deadline, remake
1101	12	bodyguard (*bodygard), briefing (*brifing), cornflakes (kornfleksy), disc jockey (dyskdżokej, *discjockey), display (displej), drive (drajw), jeep (dżip), *loser (*luzer), make-up (*mejkap), piercing (*pirsing), rockers (*rokers), Skype (*skajp)
1110	27	basketball, break, broker, dubbing, frontman, golf, hamburger, hardware, hit, Internet (internet), logować się, marketing, net, notebook, *off-road (*offroad, *off road), paperback, poster, reset, roaming, spam, stretching, surf, tuning, underground, VIP (*vip), WAP, Web
1111	44	*others, e.g.* backhand (bekhend), billboard (bilbord, *billbord, *bilboard), celebryta (*celebryt), klik, komiks, komputer, kontener, kowboj (cowboy), dealer (diler), deweloper (developer), gej (gay), haker (hacker), playboy (*plejboj), puzzel, skan, skaner, skuter, skrecz (scratch), singiel (singel)
Total	100	

Table 2. Loanword adaptation patterns in Czech, based on PIDS adaptation order

Adaptation pattern	Number of instances	Examples
0–––	0	—
1000	0	
1001	3	OK (ok, O. K., o. k., oukej, okej, okey, okay), show (*šou), talk-show (talk show, talkshow, *talkšou, *tókšou)
1010	1	*Barbie (barbie)
1011	8	CD (*cédé), cool (*kúl), *CV (*síví), DVD (*dévédé), online (on-line, *on line, *onlajn), party (*párty), PC (*pécé, *písí), SMS (*sms, *esemes)
1100	1	WAP
1101	3	cornflakes (kornfleksy, kornfleky), deadline (*dedlajna), puzzle (*pucle)
1110	28	bodyguard, broker, chill-out (chillout), developer, *Facebook (*facebook), frontman, gay, gender, golf, hamburger, hardware, hit, Internet (internet), logovat, net, notebook, offroad (off-road, *off road), outdoor, poster, remake, reset, roaming, spam, superman, surf, tuning, underground, VIP (V.I.P., *vip)
1111	56	*others, e.g.* *backhand (bekhend), *basketball (basketbal), billboard (*bilboard, *bilbord, *billbord), *break (brejk), celebrita, klik, coach (kouč), comics (komiks), computer (komputer), kontejner, *cowboy (kovboj), dealer (*díler, *dýler), *dubbing (dabing, dabink), hacker (haker), marketing (marketink), playboy (plejboj), scan (sken), scanner (skener), skútr, *scratch (skreč), singl (*single)
Total	100	

Discussion

It is worth noticing that the adaptation of English loans in both Polish and Czech rarely stops at the level of pronunciation. The pattern 1000 has only six occurrences in the Polish data set and no occurrences in the Czech data. Patterns consisting of two 0's and two 1's are likewise rare, which means that it is not common for the adaptation process to include pronunciation and only one additional dimension. More often loanwords are adapted on three dimensions and the most frequent pattern in both languages is 1111, indicating adaptation on all dimensions. However, patterns 1011 in Polish and 1101 in Czech deviate from the general rule, having very few examples compared to other patterns with three 1's. We will come back to this observation soon.

A consequence of the above is that not all patterns we could think of as congruent with the PIDS order, i.e. 1000, 1100, 1110 and 1111, are exemplified with approximately the same number of examples. Although these four patterns together cover 80 percent of the Polish data and 85 percent of the Czech data, most examples belong to the 1111 group, which neither confirms nor refutes the hypothesis that the adaptation of loanwords in Polish and Czech follows the PIDS order. We have to look at the other frequent patterns in more detail.

In Polish, the second most frequent pattern is 1110 (27 instances), the third is 1101 (12 instances) and the others have frequency below 10 percent. This means that loanword spelling is indeed usually adopted at the very end, after adaptation on other dimensions takes place, but a situation where the adaptation in spelling precedes the adaptation in derivational potential is not exceptional and has to be considered as a secondary variant. Another fact worth noticing is that the abstract pattern –10– (with any digit at the beginning and at the end) has 15 examples in the Polish data, whereas the abstract pattern –01– is exemplified only 6 times. This means that a loan more often begins to inflect before it derives new words than vice versa. All this supports our hypothesis that the typical loanword adaptation order in Polish is PIDS, but PISD has to be considered a secondary variant, one less common, but not negligible.

In Czech the situation is partly different. Like in Polish, the second most frequent pattern is 1110 (28 instances), but the third one is 1011 (8 instances), which suggests that inflections often fall behind derivations in the morphological adaptation of loans. This supposition is confirmed by the statistics for –10– and –01– patterns: they have 4 and 9 examples, respectively. The total number of examples for the ––0– pattern is also significant: we have only 7 loans in Czech with value 0 in position D (i.e., 7 loans un-adapted with respect to their derivational potential) compared to as many as 23 such loans in Polish. For the ––1– pattern, the proportion is reversed: 93 loans in Czech are adapted on the derivational dimension compared to 77 loans in Polish. This would seem to indicate that the Czech language has a more powerful derivational system than Polish (an observation made already in the literature, see Damborský 1977).

In both Polish and Czech, spelling is usually adapted at the very end, but the order of adaptation on the inflectional and derivational dimensions is different: in Polish, inflections usually appear prior to derivatives, while in Czech the converse is more common. This explains why patterns 1011 in the Polish data and 1101 in the Czech data have much lower frequency than other patterns with three 1's: the former is at odds with the Polish tendency to inflect a loanword before it is used as a basis for new words, the latter is in contradiction with the Czech principle to use a loanword in word-formation processes even before it starts to inflect, cf. *cédéčko* (from *CD*), *coolový* (from *cool*), etc. In sum, PIDS turns out be the most typical adaptation order in the case of English loans in Polish (with PISD as a secondary variant), while PDIS is the most typical order in Czech.

It should be noted in closing that pattern 1111, meaning full adaptation on all dimensions, has more examples in the Czech data than in the Polish (56 compared to 44). This can be considered as a general indication that in Czech, English loans are more strongly integrated with the pronunciation, grammar and spelling of the language than they are in Polish. This final conclusion remains in agreement with our previous research on which the present study is based.

Summary

The aim of this study was to identify the typical order in which English words are adapted to the pronunciation (P), inflections (I), derivational potential (D) and spelling (S) of Polish and Czech as recipient languages. Based on some general considerations, we assumed that the PIDS order is most likely to govern the adaptation of lexical loans in Polish and Czech. Then, drawing on the material from our earlier study, we counted how many loanwords fit particular patterns resulting from assigning values 0 ("un-adapted") and 1 ("adapted") to particular positions in the PIDS sequence. Our intention was to use these counts to identify the typical loanword adaptation order in Polish and Czech. The method was, basically, akin to those used by some other authors, who classified loanwords into groups based on varying degrees of adaptation to the recipient language. The difference is that the earlier studies that we are familiar with were not as explicit and systematic as this one.

The analysis of different adaptation patterns has confirmed our initial assumption only in part: PIDS turned out indeed to be the most typical loanword adaptation order in Polish, but not in Czech, where PDIS seems to be more common. For these conclusions to be supported more generally, further observations are needed.

This study was an attempt to investigate quantitatively, on a well-defined data set, what in some previous studies was the object of intuitive judgments. The method we used has its limitations: the data was limited to a hundred loans due to the amount of work needed to describe them and the details of the description might be questioned (e.g., deciding to include or exclude a particular derivative on the basis of its number of occurrences in a reference corpus is arbitrary to a certain extent). On the other hand, the method adopted in this paper seems to be a worthwhile alternative to the older quantitative approaches.

References

Bańko M., Svobodová D. (2015). *Is Czech more effective in loanword adaptation than Polish? Lexical score as a measure of loanword adaptation* [online at http://portal.uw.edu.pl/en_GB/web/approval/anglicyzmy_wyniki, Report 1 and Appendix].

Bartmiński, J. (1992). *Nazwiska obce na tle kontaktów językowych i kulturowych*, in: J. Maćkiewicz, J. Siatkowski (eds.), *Język a kultura*, t. 7, *Kontakty języka polskiego z innymi językami na tle kontaktów kulturowych*. Wrocław: Wiedza o Kulturze, 9–13.

Bartmińska, I., Bartmiński. J. (1997). *Słownik wymowy i odmiany nazwisk obcych*. Bielsko-Biała: PPU „Park".

Bozděchová, I. (1997). *Vliv angličtiny na češtinu*, in: F. Daneš (ed.), *Český jazyk na přelomu tisíciletí*. Praha: Academia, 271–279.

Damborský, J. (1977). *Zestawienia dwuwyrazowe w języku polskim*, in: J. Damborský, *Studia porównawcze nad słownictwem i frazeologią polską i czeską*. Warszawa: PWN, 21–29.

Daneš, F. (2002). *Xenizmy v dnešní češtině*, in: Z. Hladká, P. Karlík (eds.) *Čeština – univerzália a specifika 4*. Praha: Nakladatelství LN, 29–35.

Martincová, O. (2005). *Nová slovesná pojmenování*, in: O. Martincová (ed.), *Neologizmy v dnešní češtině*. Praha: ÚJČ AV ČR, 119–133.

Mravinacová, J. (2005). *Přejímání cizích lexémů*, in: O. Martincová (ed.), *Neologizmy v dnešní češtině*. Praha: ÚJČ AV ČR, 187–211.

Světlá, J. (2005). *Nová adjektivní pojmenování*, in: O. Martincová (ed.), *Neologizmy v dnešní češtině*. Praha: ÚJČ AV ČR, 99–118.

Published on 5/12/2015

This study was carried out within a Polish-Czech project: *Adaptation, perception and reception of verbal loans in Polish and Czech – linguistic, psychological and historic-cultural factors*, supported with a grant from the National Science Centre, Poland, no DEC-2011/03/B/HS2/02279.