Language Use and Linguistic Structure
Proceedings of the Olomouc Linguistics Colloquium 2013
OLOMOUC MODERN LANGUAGE SERIES (OMLS) publishes peer-reviewed proceedings from selected conferences on linguistics, literature, and translation studies held at Palacký University, Olomouc, Czech Republic.

Also in this series:
OMLS, Vol. 1: Teaching Translation and Interpreting in the 21st Century
OLOMOUC MODERN LANGUAGE SERIES
Vol. 3

Language Use and Linguistic Structure

Proceedings of the Olomouc Linguistics Colloquium 2013

organized by

Department of English and American Studies
Philosophical Faculty, Palacký University, Olomouc, Czech Republic
June 6–8, 2013

olincol

Olomouc Linguistics Colloquium

Edited by
Joseph Emonds and Markéta Janebová

Palacký University
Olomouc
2014
Loanwords and Foreign Proper Names in Czech: A Phonologist’s View

Tomáš Duběda\textsuperscript{1}, Martin Havlíček\textsuperscript{2}, Lucie Jilková\textsuperscript{3}, and Veronika Štěpánová\textsuperscript{4}

\textsuperscript{1}Charles University in Prague, Czech Republic; \textsuperscript{2}Institute of the Czech Language of the Academy of Sciences of the Czech Republic
\textsuperscript{3}dubeda@ff.cuni.cz; \textsuperscript{4}havlik@ujc.cas.cz; \textsuperscript{5}jilkova@ujc.cas.cz; \textsuperscript{6}stepanova@ujc.cas.cz

Abstract: The objective of the present paper is to analyze phonological aspects of orthographically non-adapted loanwords and foreign proper names on a non-normative basis. A system of eight adaptation principles is put forward (1. phonological approximation; 2. spelling pronunciation; 3. original pronunciation; 4. analogy with the donor language; 5. analogy with the recipient language; 6. the influence of a third language; 7. the influence of universals; 8. unclearly motivated pronunciation). This system is then applied to a sample of Anglicisms taken from a recently published dictionary. We show that the most important principles are phonological approximation and, to a lesser degree, spelling pronunciation. The “secondary” principles (4–8) affect only a small number of items. Differences between British and American pronunciation are unproblematic for the system.

Keywords: phonology; pronunciation; loanwords; proper names; Czech

1. Introduction

Orthographically non-adapted loanwords and foreign proper names constitute a peripheral yet dynamic and fairly conspicuous area of the Czech lexicon. They are characterized by a number of specific features, including:

a) formal markedness (e.g., the presence of the peripheral phonemes /l/, /q/, /ʢ/, /ʁ/, /ː/ or /ʁː/, unusual phonotactic patterns such as word-initial /s/ or specific morphophonological patterns);\textsuperscript{1}

b) a less transparent relationship between pronunciation and spelling, which contrasts with the phonological character of Czech spelling. It is, for instance, probable that some Czech speakers who watch the TV series The Simpsons are not aware that the name of the main character, Homer, pronounced [ˈhɒmə(r)] in the Czech version of the show, is a reference to the Greek author, whose normal Czech pronunciation is [ˈhɔmər];

c) intrinsic instability of the phonological form as a result of the lack of fixation by orthography or by other words from the same derivational family (Mathesius 1947, 99). In the case of foreign words, the number of attested pronunciations is usually higher than for Czech words; cf. the many attested pronunciation forms of the Gallicism croissant, as described by Řihová (2004);

\textsuperscript{1} The IPA transcription of Czech is based on Dankovičová 1997. English is transcribed according to Roach 2000.

313
d) extrinsic instability of the phonological form, which may be subject to influences from socio-professional groups (academic bodies, media, business) with a varying degree of erudition or pragmatism (cf. the recent decision of the Hyundai Group to present their car brand as [ˈfjonːdː] in the Czech media, despite the well-established form [ˈfjunda]; see Žemlička 2012);
e) sociolinguistic implications (e.g., prestige/stigmatization or socioprofessional stratification).

The phonetic and phonological characteristics of these lexical items are currently understudied in Czech linguistics. The last extensive survey of the pronunciation of foreign words was carried out in the 1960s and 1970s; its results were used as a basis for the 1978 *Výslovnost spisovné češtiny – Výslovnost slov přejatých* (Pronunciation of Standard Czech – Foreign Words). In recent lexicographic works, the pronunciation of new words is mostly based on the intuitions of the authors, almost all of whom are non-phonicians. This does not mean, however, that these intuitions do not lead to a globally satisfactory result. Foreign proper names are also covered by some more recent sources (e.g., Kučera and Zeman 1998).

The objective of this paper is to analyze phonological aspects of orthographically non-adapted loanwords and foreign proper names on a non-normative basis, and to examine the adaptation processes in a sample of Anglicisms taken from a recently published dictionary.

Our analysis is partly inspired by the concepts proposed by Loanword Phonology (e.g., Calabrese and Wetzel 2009), which considers the adaptation process as a phonological repair of an illegal input. The purpose of the repair is to make the word more native-like (i.e., to bring its phonological properties in line with native phonology); however, cases of divergent repair (i.e., adaptation which is not explicable by Czech phonology) and unnecessary repair may be observed as well. In Calabrese and Wetzel (2009, 1), two scenarios of nativization are outlined: the nativization-through-production principle supposes that a speaker who knows the donor language will pronounce the new word in the recipient language by applying native phonological rules to it. According to the nativization-through-perception principle, on the other hand, a speaker who has no knowledge of the donor language utters the new word by imitating the original phonetic form. Surprisingly, the authors do not mention a third logical possibility: that speakers who neither know the donor language nor have overheard the phonetic form of the word base their pronunciation solely on the spelling. Furthermore, it should be noted that the nativization of a loanword is not an instantaneous and individual act, but a process which is socially anchored and in which other factors such as tradition or analogy also play a role.

2. Adaptation Principles
The scale of integration of lexical items can be viewed in the following terms (the object of our analysis being points d, e, and f):

a) Czech words (*město* “city,” *Vladislav* “Czech male first name”)
b) “unrecognizable” (i.e., highly integrated) loanwords (*must* “must,” *Petr* “Peter”)
c) loanwords with Czech spelling (*tramvaj* “tram,” *Žaneta* “Jeanette”)
d) loanwords with double spelling (*jazz*/*žez*, *Kristina/Kristýna* “Christine”)
e) loanwords that have retained their original spelling (*croissant, Edward*)
f) foreignisms (*cherchez la femme*)
g) code mixing
Eight different (but combinable) principles are observed in the phonological adaptation of loanwords.

1. **Phonological approximation.** This process, which is the most frequent and is presented as the default method in pronunciation manuals, denotes the substitution of non-native sounds with their nearest counterparts in Czech, together with the application of Czech prosodic, phonotactic, and morphological rules, e.g., *Windows* ['vindōs] → ['vindoš]. The basic rules of phonological approximation for British English phonemes (not mentioning most consonants for which the conversion is obvious, e.g., /m/ > /m/) can be summarized in the following way: /t o ð s/ > /t o o o s/; /θ: u: x: a:/ > /θ: u: o a:/; /e æ e z/ > /e/ (three phonemes merging into one); /æt ɔt ou/ > /aj ɔt ou/; /wa wa ou/ > /wa wa ou/; /ɔ ð æ s/ > /θ æ s/; /θæt ðæt ou/ (r grapheme is always pronounced as /r/, even in positions where it is elided in non-rhotic accents); /θ ok/ > /θ k/; /θ s/ > /θ s/; /θæt ðæt ou/ > /θæt ðæt ou/; /θæt ðæt ou/ > /θæt ðæt ou/. Cases in which phonological approximation is not “automatic” include vowel length before /i/ (software can be pronounced both [′softver] and [′softvr:] or [′softvr:], other cases of vowel length (bypass can be pronounced both [′baipas] and [′bajpa:s]), potentially syllabic sonorants (pixel can be pronounced both [′piksə] and [′piksrl]), and /θ ɔi/ (Smith can be pronounced both [′smi] and [′sma:j]).

2. **Spelling pronunciation.** According to this principle, Czech pronunciation rules are applied to the foreign spelling form (e.g., *Superman* [′supərman], but *Batman* [′batman] and *Spiderman* [′spajdman], pronounced according to Principle 1, probably because they are more recent).

3. **Original pronunciation.** This kind of pronunciation, according to which the phonological and phonetic rules of the donor language are maintained, is sometimes used in citations (*Výslovnost spisovné češtiny* 1978, 30), in scientific communication (Hůrková 1995, 69), and informally: by youngsters talking about pop music, for instance. Technically, this option leads to code mixing, and, in inflected forms, to phonetic hybridization, as Czech phonemes must appear in the endings of inflected forms.

4. **Analogy with the donor language.** In this case, the adapted form is the result of the (often incorrect) application of a phonetic analogy from the source language (e.g., *Robert* [′roebt], a widespread pronunciation variant, commonly heard in the media, may be considered a hypercorrect form of [′robût]).

5. **Analogy with the recipient language.** According to this principle, the phonological changes made to the word that has been adopted are motivated by analogy with Czech words, or, more generally, by analogy with sufficiently integrated words of any origin. This principle accounts for what is usually called folk etymology; for example, the word *protézovat* ("to favor" < French *protéger*) is often pronounced (and even spelled) as [porçovat], under the influence of Czech words such as *vytěžovat* and *zatěžovat*, which share a number of semantic features. Likewise, the French specialty *salade niçoise* (named after the city of Nice), is often interpreted as “Nicosia” salad because of its complicated spelling; the situation regarding this form has become even more confusing since a supermarket chain in the Czech Republic started selling this very salad under the name *Nicosia s tuňákem* (“Nicosia with tuna”).

6. **Influence of a third language.** Words may be affected by the phonology of a third language, either because they were adopted via this language (e.g., *lajtnant*, adopted through German *Leutnant* from the French *lieutenant*) or by analogy (e.g., *puzzle*, often pronounced as [′putiz] in Czech). This last form may have come about through analogy either with
German pronunciation rules or with the similar-sounding Czech word pučlik ("chubby child"; Štěpánová 2013).

7. **Influence of universals.** An example is the word peloton, which is often pronounced as ['peleton] and sometimes spelled peleton. The presence of an [r] in the second syllable can be explained by vowel harmony.

8. **Unclearly motivated pronunciation.** This last category, which is technically not a principle, includes cases for which there is no obvious explanation (e.g., country pronounced as ['kɑntrɪ]).

Principles 1–2 are of central importance in the system (Výslovnost spisovné češtiny 1978, 27), while Principle 3 is peculiar in that it is socially or individually conditioned, potentially gradual, and exists as an alternative to all the other principles listed above. Principles 4–8 may be considered as secondary, as their effects are usually local and they are problematic with respect to the norm, at least for recent words. Elements of the system that we have presented can be found in several sources (Výslovnost spisovné češtiny 1978; Kučera and Zeman 1998; Ološtiak et al. 2006); however, the advantage of our list is that it presents the principles in a structured, exhaustive, and non-normative way.

In many cases, foreign words exhibit more than one of the principles: 1 and 2 in Charleston [ˈʧərəstən], 1 and 4 in Robert pronounced as [ˈrōbərt], and 1, 2, and 5 in heavy metal [ˈhɛvjmɛlt]. In this last case, heavy is adapted according to Principle 1, while the pronunciation of metal as [mɛlt], rather than [mɛlt], is motivated both by orthography and by other Czech words containing metal like metalurgie (“metallurgy”) and metaliza (“metallic paint”).

### 3. An investigation of Adaptation Processes in a Sample of Anglicisms

To the best of our knowledge, none of the available sources considers the relative importance of the aforementioned principles in the lexicon. We therefore decided to analyze a sample of orthographically non-adapted Anglicisms taken from a modern medium-sized dictionary of Czech (Slovník současné češtiny [Dictionary of Contemporary Czech] 2011). We first selected all the entries for which a phonetic transcription is given (the transcription indicates that the word is of foreign origin, since entries for Czech words are not provided with a phonetic transcription in most dictionaries); we excluded, however, entries where the only issue was the pronunciation of di, ti, and ni (pronounced [ɗi tɨ ni] in Czech words but [ɗɪ tɪ ni] in foreign words of Western origin), or the pronunciation of -isma (pronounced as [izmʊʃ]). Such cases are not informative for our study, as they cause almost no problems for Czech speakers. We also decided not to include orthographically adapted loanwords in our sample. In total, we analyzed 225 Anglicisms (24% of which had an alternative spelling and 19% of which were listed as having more than one pronunciation).

The phonetic transcription of the selected Anglicisms given in Slovník současné češtiny [Dictionary of Contemporary Czech] 2011, was then compared to the transcriptions of these words in five other comprehensive dictionaries (Výslovnost spisovné češtiny – Výslovnost slov přejatých [Pronunciation of Standard Czech – Pronunciation of Loanwords] 1978; Slovník spisovné češtiny pro školu a veřejnost [Dictionary of Standard Czech for Schools and the General Public] 2003; Pravidla českého pravopisu [Czech Spelling Rules] 2004; Nová slova v češtině III.
3.1 Agreement among Sources

Figure 1 shows the agreement between the pronunciations listed in the Dictionary of Contemporary Czech and the five other sources. We can see from the graph that the agreement is relatively high (between 77% and 90%).

For 37 entries (16% of the sample) at least one of the dictionaries gives a pronunciation which is different from the Dictionary of Contemporary Czech. These cases can be structured in the following way (all transcriptions have been converted into the IPA):

a) concurrence of Principle 1 (phonological approximation) and Principle 2 (spelling pronunciation): holding [ˈhoʊldɪŋk/ˈhoʊldɪŋk], spam [ˈspæm/ˈspæm];

b) concurrence of a short and a long vowel: software [ˈsəʊtərɛɪ/ˈsəʊtərɛɪ], bypass [ˈbaɪpɑs/ˈbaɪpɑs];

c) concurrence of epenthetic [ə] and a syllabic consonant: hacker [ˈhekər/ˈhekər], pixel [ˈpɪksəl/ˈpɪksəl];

d) concurrence of a (quasi-)diphthong and a monophthong: cornflakes [ˈkɔrnflæks/ˈkɔrnflæks], catering [ˈkeɪtrɪŋk/ˈkeɪtrɪŋk].

![Figure 1](image)

**Figure 1.** Comparison of the pronunciation of words in the sample that was studied with their pronunciation listed in five other lexicographic sources. Explanation of the legend: “same” — the other source lists the same pronunciation(s) for the word in question; “different” — the other source lists (a) different pronunciation(s); “no pronunciation” — no pronunciation is listed in the other source; “word missing” — the word in question is not listed in the other source.

3.2 Distribution of Adaptation Processes

The distribution of the primary adaptation processes (Principle 1 and Principle 2) in the Dictionary of Contemporary Czech (2011) is displayed in Figure 2. Phonological approximation is clearly the predominant principle, and accounts for 73% of the entries studied. In 10% of items, the word root is adapted by phonological approximation, while the prefix or suffix has a spelling-
based pronunciation. All the prefixes observed in this category are of Greek or Latin origin and exist in other Czech words (superstar ['superstar], gigabyte ['gigabát]); for suffixes, all the lexical items but one contained *-ing, whose established phonetic form in Czech is [ŋk] (roaming ['rόʊmŋk], happening ['hæppɪŋk]). The adaptation of such prefixed or suffixed Anglicisms is thus not a single-step event, but a compositional process based on two different paradigms. The third group of words (5%) exhibits mixed treatment: phonological approximation and spelling pronunciation are applied within the same word (rock-and-roll ['rǒkəndroll]). The fourth group (9%) includes words with two parallel pronunciation forms, one based on approximation and the other on spelling (gangster ['gæŋkstr, 'gæŋkstə]). Finally, in 3% of words we find a pronunciation which is fully based on spelling (developer ['dɛvələpər]). As expected, there are no instances of Principle 3 (original pronunciation) in the dictionary.

![Figure 2. Distribution of the primary adaptation processes in the words studied.](image)

For the secondary adaptation processes (Principles 4–8) we found the following 19 cases, which make up 8% of the sample:

a) Principle 4 (Analogy with the donor language: 4 items): catering ['kɛtərɪŋ] (by analogy with other English words in which a is pronounced as [æ], and possibly even with the word eat; a general tendency towards diphthong simplification seems to be a less plausible explanation), cookie ['kuːki] ("full" pronunciation of both vowels by analogy with other English words); cornflakes ['kɔrnflæks, 'kɔrnflɛks] (cf. catering; additionally, we may hypothesize that there is an analogy with the Czech word fleky "a type of flat pasta"), forfeiting ['fɔrfɪtɪŋ] ("full" pronunciation not reflecting vowel reduction in the original form ['fɔːfɪtɪŋ]; NB this word is given in the Dictionary of Contemporary Czech (2011) with the wrong spelling forfating);

b) Principle 5 (Analogy with the recipient language: 4 items): hamburger ['hæmbɜrɡə] (analog with the city of Hamburg, pronounced ['hamburk] in Czech), heavy metal ['heviːmɛltəl], heavymetalový ['heviːmɛltəlovɨ] (cf. above), leasing ['liːznɪŋ] (analogy with other foreign words in which intervocalic -s- is pronounced as [z]);

c) Principle 6 (Influence of a third language: 6 items): démičon ['demɨtʃon, 'demɨtʃon] (an adapted word with variable length in the -on ending, typical of French loanwords such as balkon "balcony" or bonbon "sweet"), manager ['mænɪdʒər], managerka ['mænɪdʒɛrkə], managerský ['mænɪdʒɛrskɨ] (probably influenced by the French pronunciation; see Jílková, forthcoming), management ['mɛnɪdʒmənt, 'mɛnɪdʒmənt] (the first variant is based on
phonological approximation, whereas the second is a combination of English and French influences), *puzzle* ['pazl/ˈpʊzəl] (the first variant is based on phonological approximation, and the second one is inspired by German grapheme-phoneme conversion rules, although the standard German pronunciation of the term is ['pazl] or ['pazl]);

d) Principle 7 (Influence of universals: 4 items): *baseball* ['beɪz bol], *baseballový* ['beɪz bolovɪ] (the shortening may be due to the presence of a double grapheme at the end, which implies a short vowel in many European languages, by a general tendency towards the unmarked term in a phonological contrast — see Maddieson 1984, or by analogy with other foreign words ending in -ol, such as *alkohol*), *grunge* ['ɡrʌŋ] (affricate simplification in a consonant cluster), *paintball* ['peɪnt bol] (cf. *baseball*).

e) Principle 8 (Unclearly motivated pronunciation: 1 item): *country* ['kʌntəri] (the lengthening cannot be explained by any analogy or a general tendency; the hypothesis that it is induced by the presence of a double vocalic grapheme is speculative).

3.3 Concurrency between British and American English

One of the questions that naturally arises in the study of Anglicisms in Czech is whether the approximated forms reflect British (RP) or American (General American) pronunciation in the event that they differ for a given lexical item. The following six categories were considered:

a) **Rhoticity:** as we mentioned above, all approximated forms reflect the underlying /r/. This is probably the joint influence of spelling and rhotic accents of English;

b) **alternation between [a:] and [æ]:** we found only one item of this kind in our sample: *bypass* ['bæpjɔs], which is based on the British pronunciation form. In other recent words, pronunciation may vary: *grant* is pronounced uniformly as ['ɡrænt], but *Hugh Grant* (despite his British origin) is often realized as ['ɡrænt]. The usual pronunciation of *breakdance* is ['brɛkˌdɑːns];

c) **alternation between [n] and [ø]:** out of the 34 items found in the sample (e.g., *box, copyright, laptop*), all are given with the [ø] vowel, based on the British form. Only one item (*rock-and-roll*) is imaginable with the [a] vowel, reflecting the American pronunciation. Other cases of variability can be found marginally (e.g., the female name *Dolly* is usually pronounced as ['dɔli], but in the 1997 version of *Hello Dolly*, the form ['dɑli] can be heard);

d) the [so/œo] alternation: all 15 approximated items contain [ou] (e.g., *notebook, show*). Any other alternative is hardly imaginable;

e) **[j] deletion:** out of the two items where [j] may elide in American English (*newton, tuning*), one is given with a pronunciation which is based on the British version (['njuːtɒn]), and the other with two alternatives (['tjuːnɪŋk*/tʌnɪŋk]); it is not easy to say whether the second one is motivated by American pronunciation or by spelling;

f) **intervocalic [t] voicing:** out of the six items where [t] may be voiced in American English (e.g., *heavy metal, party*), all are given with a [t]. Marginally, the voiced variant may appear, e.g., in *shut up!*, often pronounced as [ʃʌt 'æp].

On the whole, approximated forms are based on British pronunciation variants, with the notable exception of rhoticity, which is always maintained. Out of the six categories, it is only the [a:/æ] difference which is likely to introduce instability in the system of phonological approximation.
4. Conclusion
The phonology of orthographically non-adapted Anglicisms in Czech seems to be a rather stable system, with phonological approximation as the leading principle (with 73% of the sample that was studied conforming to this principle). As is evident from the items with double spelling (original and adapted), phonologically approximated forms are also the basis for spelling adaptation. Phoneme mapping is mostly straightforward, with some degree of variability for vowel length, sonorant syllabicity, and /θ/ /ð/ conversion. The second principle according to its frequency of occurrence is spelling pronunciation, although it is rarely used alone: in most cases, it is used in combination with phonological approximation (compositional adaptation, mixed adaptation or parallel adaptation). Secondary adaptation principles (analogies, universals, and unclearly motivated pronunciation) concern only 9% of the sample. Aside from the alternation between [a:] and [ae], the concurrence between British and American forms does not seem to perturb the system.

All the aforementioned results should be interpreted with respect to the method by which they were obtained: we investigated the pronunciations given by a recent general dictionary of Czech rather than real usage. However, the comparison with five other sources gives us at least a rough idea about potential variability in pronunciation in real usage. Sixteen per cent of the items that were studied were treated differently by at least one of these five sources. Most of the discrepancies can be explained by the concurrence of phonological approximation and spelling pronunciation, as well as by the intrinsic instability of some approximation rules (especially in the case of vowel length).

Obtaining a complete picture of the subsystem of orthographically non-adapted Anglicisms in Czech would obviously require an investigation of real usage. A comparison with older Anglicisms, which have mostly adopted Czech spellings, may provide diachronic insights into loan-word natization. As for proper names of English origin, we may expect a higher degree of variability and a greater proportion of secondary adaptation principles.

The present analysis may serve as a basis for phonetic predictions about newly adopted Anglicisms; it may also find its application in lexicographic practice.

Funding Acknowledgement
This article was supported by the Czech Science Foundation grant No. 13-00372S “Pronunciation of non-integrated lexical items in Czech.”

Works Cited