

# The phonetics of EAP latinisms in EFL e-dictionaries

Włodzimierz Sobkowiak

"When someone presumes to correct your [Latin – WS] pronunciation, a knowing smile is an appropriate response." (Hickman 1996)

## Abstract

Latinisms are commonly used in scientific registers of English. Foreign learners of English for Academic Purposes (EAP) must be able not only to understand their meaning for passive processing while reading (silently) scientific literature, but also to pronounce them according to English standards should they want or need to read such texts aloud or present their own homework or research publicly in speech. EFL dictionaries are the most natural resource to turn to in this situation.

A sample of forty EAP latinisms have been culled from various sources and looked up in the 'big-five' EFL monolingual dictionaries (MEDAL, LDOCE, OALD, CALD and Cobuild) and in one pronouncing dictionary (CEPD). Only the last dictionary offers an almost complete coverage of the forty latinisms. The other resources are found wanting on a number of counts, mostly in their inconsistent coverage (both in terms of specific latinisms and different modes of phonetic presentation) and problems in aligning phonetic transcription with audio recordings.

## Introduction

According to OED, a *latinism* is "an idiom or form of expression characteristic of the Latin language, esp. one used by a writer in another language" (OED online). In this contribution I will look into latinisms commonly encountered by non-native writers of English in the context of the academic uses of English as a Foreign Language (EFL). Acronymically, then, this is the field of Latin in EAP, or LEAP (Latin in English for Academic Purposes). EFL learners, in their many passive/decoding and active/encoding encounters with LEAP take recourse to monolingual EFL pedagogical dictionaries, currently often in their electronic format. It is the phonetic treatment of LEAP in EFL e-dictionaries which is in the focus of this paper.

The present topic is not one which would attract a lot of scholarly or pedagogical attention. Writing EFL at an academic level (so-called WEAP) is not a skill or activity which is popular enough to warrant serious commercial interest going as far as detailing academic latinisms as a topic worthy of study. Standard EFL textbooks at an advanced level have little or nothing to say about latinisms and how they function in the academic register of English. EFL teacher training manuals and resources are similarly tacit on this and other aspects of WEAP. Even specifically WEAP-oriented resources seldom touch upon latinisms (MacPherson 1994 appears to be one such exception among Schenck 1988, Leki 1989, Hult 1996, Jordan 1997, Oliver 1999, Oshima & Hogue 1999, Młyniec & Ufnalska 2003, Necka & Stocki 2003). It goes without saying, then, that the phonetic aspect of latinisms in English is an issue completely ignored, both in general EFL and in (W)EAP alike.

The web does offer a few hints, mostly in the contexts of (a) classical studies done by native English students, and (b) biological terminology as used in speech at respective congresses and conferences. In the former context, the issues are rather esoteric from our current point of view, for example the differences in pronouncing the so-called "classical", as opposed to the so-called "ecclesiastical" Latin. In the latter context, the questions similarly concern the many different phonetic renditions of biological terminology, most of which is only ever known and used by the respective specialists, rather than for general academic writing purposes (so-called GEAP).

Consider an example. David Cramer claims that "it is actually much easier to learn to pronounce Latin than English", and urges (native English) learners of Latin to pronounce "c always hard: civis" (<http://www.utexas.edu/courses/cc303/sounds/sounds.html>), which in the English orthoepic tradition means a velar stop /k/. And yet, in the Polish (and German) tradition of pronouncing Latin this same letter would be pronounced as an affricate /ts/, which is its normal graphophonemic rendition in Polish, thus: /'tsivis/. Should that not be confusing enough, latinate cognates in English dictionaries are phonetically transcribed with a 'soft' /s/: *civitas* 'siv.i.tæs, 'kiv- (CEPD), at least as the first (preferred) variant. John Wells conducted a pronunciation poll for some words in his 1990 *Longman pronunciation dictionary*; for *schism* he notes that "the traditional 'siz- is being displaced, except perhaps among the clergy, by 'skiz- BrE poll panel preference: 'skiz- 71%, 'siz- 29%" (Wells 1990:625). How should the EFL learner, then, interpret the following quote: "the failure to pronounce the first syllable of *schism* like the first syllable of its congener *scissors* [is a] pronunciation error", found on the website of [www.orlapubs.com](http://www.orlapubs.com), which, surprisingly, does turn out to belong to Eastern Orthodox Christianity? After all, EFL learners could not be bothered disentangling "classical" from "ecclesiastical"<sup>1</sup>, "continental"<sup>2</sup>, "scientific"<sup>3</sup> or "anglicised"<sup>4</sup> pronunciations of Latin.

In this situation an advanced learner of EFL at the academic level may be excused to go along with the following old quote: "One may hear educated Englishmen say that every nation has its own way of pronouncing Latin" (Crawford 1885:54). This will not do, however, because, as Crawford is quick to remark: "This, as an excuse for English pronunciation of that language, is nonsense. Each foreign nation makes some slight variation, but each has fixity within its own lines", and later: "How an educated Englishman can venture to quote Latin in the presence of a foreigner baffles comprehension. The effect must be inexpressibly ludicrous, on the supposition that the foreigner can comprehend the utterance. Possibly he takes it for some unintelligible gibberish. [...] The Latin of any Continental nation is understood by the scholars of any other: that of the English is intelligible to none" (*ibidem*; see line 14 in Table1). But, quite clearly, Crawford's words only make sense in the native English context,

---

<sup>1</sup> "In most countries, those who speak Latin for liturgical or other ecclesiastical purposes use the pronunciation that has become traditional in Rome, giving the letters the value they have in modern Italian" (Wikipedia).

<sup>2</sup> "*Continental pronunciation* - A method of pronouncing Latin and Greek in which the vowels have their more familiar Continental values, as in German and Italian, the consonants being pronounced mostly as in English" (Merriam-Webster Online Dictionary at [www.Merriam-Webster.com](http://www.Merriam-Webster.com)).

<sup>3</sup> "The 'scientific' pronunciation: this pronunciation is most often used by those who only use Latin in names (of flowers, animals, materials) or expressions (in law or just to appear more educated). Latin is pronounced like English or whatever language you're speaking" (<http://snow.prohosting.com/sprach/latin/pronunciation.htm>).

<sup>4</sup> "This guide attempts to provide a "preferred" pronunciation that follows a consistent set of rules. As much as possible, this guide follows the so-called Traditional English Method for pronouncing Latin words in English. (<http://www.dinosauria.com/dml/names/plesi.html>)

not when Latin is used through English by non-native speakers/writers of the latter. EFL learners, when they speak English, must not mispronounce Latin according to their L1 model, just like they should not mispronounce English borrowings from French in such a way. If anything, they are encouraged to mispronounce both as if English was their native tongue. This is a whole new meta-level of competence, where one foreign (or dead) language, which may or may not be completely unknown to the learner, is observed from the platform of another foreign language, currently under study. It is no wonder that in this complex situation all of learners, teachers and resource makers feel helpless. Dictionaries, as always, are the last resort.

## 2. Phonetics of LEAP in EFL e-dictionaries

The forty latinisms appearing in Table 1 were collected from a number of internet sources, such as [http://www.appliedlanguage.com/languages/latin/latin\\_phrases.shtml](http://www.appliedlanguage.com/languages/latin/latin_phrases.shtml), <http://www.slu.edu/colleges/AS/languages/classical/latin/tchmat/grammar/vocabulary/evlatin.pdf> or MacPherson 1994, with a bias on WEAP terms, or – more specifically – on WEGAP, i.e. writing English for general academic purposes, with no focus on a particular science, such as entomology, for example. It would, of course, be easy enough to collect a much larger sample of WEGAP latinisms, including proverbs, sayings and clichés of more than two or three words each. My modest sample of forty items, however, is small enough to be able to scan every single latinism in its own right, but numerous and representative enough to draw conclusions about the treatment of LEAP in EFL dictionaries at large.

All forty latinisms were looked up in the electronic versions of the five leading EFL dictionaries, as well as in one e-dictionary of pronunciation, the CEPD, which is also commonly used by advanced EFL learners in the academic setting in Poland. As there are two modes of phonetic representation in electronic dictionaries, the phonetic transcription and the audio recording, information about which mode is used for each latinism in each dictionary is provided in the respective cells. Because my edition of Cobuild (2001) does not provide phonetic transcription as a matter of policy, I have counted its only phonetic representation mode as the maximum possible, i.e. what in other dictionaries is tagged as "both". While there are newer editions of some dictionaries listed in Table 1 (and new ones are in preparation at all times), it is doubtful that the phonetic treatment of latinisms therein has changed significantly.

**Table 1. Selected LAEP words and phrases in the leading EFL e-dictionaries**

	Latin entry	MEDAL	LDOCE3	CALD	OALD7	COBUILD	CEPD	BNC
1.	<i>a fortiori</i>	both	---	---	both	---	both	1
2.	<i>a posteriori</i>	both	both	---	both	---	both	3
3.	<i>a priori</i>	both	both	both	both	audio	both	143
4.	<i>ad hoc</i>	both	both	both	both	audio	both	406
5.	<i>ad hominem</i>	---	---	---	both	---	both	14
6.	<i>caveat</i>	both	both	both	both	audio	both	121
7.	<i>ceteris paribus</i>	---	---	---	---	---	---	46
8.	<i>de facto</i>	both	both	both	both	audio	both	267
9.	<i>e.g.</i>	both	both	both	both	audio	both	7034
10.	<i>ergo</i>	both	both	both	both	audio	both	45
11.	<i>et al</i>	both	both	both	both	audio	both	4169
12.	<i>etc.</i>	both	both	neither	both	audio <sup>9</sup>	both	7366
13.	<i>i.e.</i>	audio	both	both	both	neither	both	6295

14.	<i>ibid.</i>	audio	both	both	neither	audio	both	248
15.	<i>infra</i>	---	trans	---	neither	---	both	18
16.	<i>inter alia</i>	both <sup>1</sup>	both	both	both	audio	both	310
17.	<i>ipso facto</i>	both	both	both	both	audio	both	45
18.	<i>magnum opus</i>	audio	audio	neither	neither	audio	both	12
19.	<i>ms. / mss.</i>	---	neither	neither	both	---	both	318
20.	<i>mutatis mutandis</i>	---	---	---	both	---	both	0
21.	<i>n.b.</i>	audio	audio	both	both	audio	both	5
22.	<i>non sequitur</i>	both	both	both	both	audio	both	9
23.	<i>op. cit.</i>	both	both	both	neither	neither	both	10
24.	<i>pace</i>	both <sup>2</sup>	---	---	both <sup>7</sup>	---	both	0
25.	<i>passim</i>	---	both	---	both	audio	both	55
26.	<i>per se</i>	both	both	neither	both	---	both	345
27.	<i>post hoc</i>	both	---	---	both	---	both	43
28.	<i>prima facie</i>	both	both	both	both	audio	both	320
29.	<i>pro forma</i>	both	both	both	both	audio	both	48
30.	<i>qed</i>	both	both	both	both	---	both	36
31.	<i>sc.</i>	---	---	---	---	---	---	10
32.	<i>sic</i>	both	both	both	both	audio	both	396
33.	<i>sine qua non</i>	both <sup>3</sup>	both <sup>5</sup>	both <sup>6</sup>	both	audio	both	38
34.	<i>sui generis</i>	---	---	---	both	audio	both	32
35.	<i>supra</i>	---	---	---	---	---	both	35
36.	<i>v. (vide)</i>	---	---	---	neither	---	both	? <sup>10</sup>
37.	<i>verbatim</i>	both	both	both	both	audio	both	101
38.	<i>versus</i>	both	both	both	both	audio	both	928
39.	<i>vice versa</i>	both	both	both	both	audio	both	652
40.	<i>viz.</i>	both	both	both	both <sup>8</sup>	audio	both	217
	neither	0	1	4	5	2	0	
	both	26	26	23	32		38	
	audio only	4	2	0	0		0	
	trans only	0	1	0	0	24	0	
	not in dictionary	10	10	13	3	14	2	

<sup>1</sup> Linking /r/ in trans, but not in audio. /æliə/ in trans, /eɪliə/ in audio.

<sup>2</sup> AmE audio: /peɪs/.

<sup>3</sup> Trans: /'bɒn/.

<sup>4</sup> *Lonely hearts* (!) appears as an entry presumably *containing* the Latin phrase, but it does not.

<sup>5</sup> Trans: /nəʊn/, audio: /nɒn/.

<sup>6</sup> Trans: /,smeɪkwɑ:'nəʊn/, audio: /,smɪkwɑ:'nɒn/.

<sup>7</sup> Different transcription variants; one erroneous audio: /peɪs/.

<sup>8</sup> AmE audio: /vi:/.

<sup>9</sup> Audio: /i:ti:si:/.

<sup>10</sup> *V.* is multiply homonymous, so it was hard to get the correct number.

The two e-dictionaries boasting the best coverage of latinisms are CEPD and OALD7. Of the 'big-five' pedagogical dictionaries of EFL Oxford has no competition, with only three latinisms missing: *ceteris paribus* (which fails to appear in all of the six dictionaries studied), *sc.* and *supra*. At the same time, however, OALD7 has the highest number of latinisms without any indication of pronunciation: five. In all dictionaries there are 12 such cases – this is clearly unsatisfactory from the point of the advanced EAP learner, who, while possibly having good intuitions as to the expected orthoepy of these strings, deserves to be offered full guidance. Similarly, the six cases of audio-only entries (not counting Cobuild), while better than nothing, may certainly be less than enough in fixing the categorical phonemic values of

the respective phonetic representation in the minds of the learners. For example: are the unstressed vowel values in both words of the recorded *magnum opus* 'really' schwas or /ʊ/'s? As a matter of fact, they sound much more peripheral in the BrE recording than in AmE in both MEDAL and LDOCE3. The inquisitive learner may well ask if this is a categorical accentual difference or simply free speaker variation? Transcription would solve this problem.

18 latinisms appear in all dictionaries, in both presentation modes (with the proviso concerning Cobuild, as above). They are shaded in Table 1. This is less than half of all investigated latinisms. It is not clear why two of the forty latinisms fail to appear in any presented dictionary. CEPD features *scilicet*, which is presumably less common than its abbreviation in English writing (the former is not attested in the written component of the BNC corpus; taken from <http://www.kilgarriff.co.uk/bnc-readme.html>). *Ceteris paribus* simply does not exist, as far as EFL dictionaries are concerned. And yet it occurs 46 times in 90 million running words of the written-BNC (see last column in Table 1), i.e. it shows higher frequency than some other latinisms on the list, which are much better represented in the EFL dictionaries, such as *sine qua non*, which appears in all dictionaries here under scrutiny. Even *mutatis mutandis* and *pace*, with their zero frequency in BNC, do appear in OALD7.

To see if there is any correlation between the number of dictionaries containing the given latinism and its corpus frequency, I ran the Pearson test over all 39 items (# 36 excluded; BNC frequencies converted to log values). The result stands at  $r=0.5$ , which is significant at  $p=0.001$  for a one-tailed test, at  $df=37$ . This is a respectably high correlation, showing that the more frequent latinisms do indeed have a better chance of being included in an EFL dictionary<sup>5</sup>. All in all, however, a working hypothesis to explain the observed inconsistencies, discrepancies and omissions might be that in this, rather exclusive, lexical area the control of EFL dictionary makers over the frequency aspects of covered vocabulary is looser than in the central areas.

As seen in the notes to Table 1, there are some problems with phonetic representation of the entries under study here. Characteristically, in some cases, audio and transcription are not matched. *Sine qua non* suffered most in this respect: (a) in MEDAL an obvious typo occurred in transcription, (b) in LDOCE3 and CALD the two pronouncing variants of *non* are clearly misaligned across the two modes of presentation, and (c) in CALD also the two variants of the final vowel in *sine* are similarly mismatched. Such errors have more serious consequences than could be envisaged prima facie. A phonetic search for /\*nɒn/ in MEDAL yields "No Results"; in LDOCE – three entries are retrieved: *anon*, *non-*, *xenon*. Only CEPD correctly retrieves *sine qua non*, together with a handful of other, mostly Greek, loanwords, such as, for example: *Agamemnon*, *Lebanon*, *noumenon*, *organon*, *Parthenon*, *prolegomenon*, *turn-on*. Thus, an error or inconsistency in an electronic dictionary will usually make the given entry completely invisible to the search mechanism, with all the ramifications of this fact. This is unlike in the traditional printed dictionaries, where errors and inconsistencies are easily normalized by the reader in most cases in the process of lookup (but not global search).

Such misalignments of the two phonetic presentation modes are quite common in electronic dictionaries, as I have amply demonstrated in a number of contributions (see <http://ifa.amu.edu.pl/~swlodek/public.htm> for a full list). The main culprit appears to be, as above, rather sloppy editing, whereby phonetic transcription editors apparently have little contact with people overseeing the audio recording of the entire contents of the dictionary

---

<sup>5</sup> My gratitude goes to Robert Lew for help with statistics.

macrostructure (not a trivial task!), and vice versa. On a somewhat higher level, the tension between the nomothetic nature of phonetics and the idiographic nature of lexicography may also be coming to the fore, as I speculated in my 2002 contribution to the *Symposium on lexicography XI* (Sobkowiak 2005).

Assignment of phonetic transcription to dictionary entries is nowadays mostly done (semi)algorithmically, while recordings continue to be made on a one-by-one basis. There are three potential methods to solve this problem: (1) transcribe entries from recorded audio, (2) record entries by reading their phonetic transcription, (3) turn to text-to-speech (or transcription-to-speech) synthesis. This last solution is now becoming feasible, with TTS reaching a level of authenticity where it is no longer distinguishable from natural speech.

As for the baffling case of *ad hominem* in LDOCE3, this is simply a bug in the search engine: e-LDOCE3 behaves erratically when it cannot find a multi-word expression containing spaces. It sometimes offers a fuzzy-search-like list of guesses, at other times simply says "No matching search results", but occasionally produces a new window with links to entries presumably containing the needed phrase. While the issue is not inherently phonetic in any sense, and simply a quirk in the software, it may be of interest to the reader to illustrate this slightly eerie case of *ad hominem* leading to *lonely hearts* in LDOCE3. This is done in Figure 1.

Figure 1. *Ad hominem* ► *lonely hearts*



### 3. Conclusions

In teaching and learning English for Academic Purposes the issue of the pronunciation of latinisms may not be among the most urgent. Arguably, if latinisms are encountered at all it is mostly in reading academic prose, much less often in writing it (WEAP), least of all in reading aloud and speaking. Occasions where the learner will need to actually say a Latin word or phrase while speaking English will, however, arise from time to time. This may be when presenting homework in the classroom or having a paper at a conference. To some advanced learners of EAP no pragmatic excuse of this type is needed at all; it is enough that the words are there to have the urge to conquer them, also phonetically. The standard orthoepic rules of English, presumably by now well fixed in the mind of the advanced EFL learner, will be of little help in trying to guess the correct English pronunciation of Latin. A dictionary is needed.

In this situation only one of the reviewed dictionaries can be recommended without reservation, namely CEPD. The problem is, of course, that it is not – strictly speaking – a pedagogical EFL dictionary at all, but rather an all-round pronouncing dictionary, which could offer to the EAP learner but one aspect of the sought-after latinism: phonetic; meaning would have to be looked for elsewhere. The 'elsewhere' to most advanced learners would probably be their favourite pick among the most popular monolingual EFL dictionaries (see Lew 2004). While not entirely unhelpful, in turn, these dictionaries, at least in their electronic versions (increasingly more often used by all learners), present some obvious inadequacies in their treatment of some latinisms, as was demonstrated in this paper. The main causes appear to be: (a) the special linguistic status of such borrowings (which they share with those coming

from French, for example), (b) the notorious phonolexicographic problems of electronic dictionaries, (c) inconsistent coverage of this well-defined lexical field. All three deserve much more metalexigraphic and linguistic attention than they have attracted so far.

## References

- Cramer, D. (n.d.). *Latin 506 Correspondence: The Pronunciation of Latin*.  
<http://www.utexas.edu/courses/cc303/sounds/sounds.html> (course in Classical Mythology at the University of Texas)
- Crawford, J.C. 1885. "The Maori language, with remarks on the reform of English spelling". *Transactions and Proceedings of the Royal Society of New Zealand 1868-1961*. Vol. 18. Pp. 46-58.
- Hickman, J. (ed.). 1996. *The Jepson Manual: Higher Plants of California*. Berkeley: University of California Press.
- Hult, C.A. 1996. *Researching and writing: an interdisciplinary approach*. Belmont: Wadsworth Publishing Company.
- Jordan, R. 1997. *English for academic purposes*. Cambridge: Cambridge University Press.
- Leki, I. 1989. *Academic writing : techniques and tasks*. New York: St.Martin's Press.
- Lew, R. 2004. *Which dictionary for whom? Receptive use of bilingual, monolingual and semi-bilingual dictionaries by Polish learners of English*. Poznań: Motivex.
- MacPherson, R. 1994. *University English*. Warszawa: WSiP.
- Młyniec, W. & S.Ufnalska. 2003. *Scientific communication, czyli jak pisać i prezentować prace naukowe*. Poznań: Sorus.
- Nęcka, E. & R.Stocki. 2003. *Jak pisać prace naukowe z psychologii: poradnik dla studentów i badaczy*. Kraków: Universitas.
- Oliver, P. 1999. *Jak pisać prace uniwersyteckie. Poradnik dla studentów*. Kraków: Wydawnictwo Literackie.
- Oshima, A. & Hogue, A. 1999. *Writing academic English*. London: Pearson ESL.
- Schenck, M.J. 1988. *Read, write, revise: a guide to academic writing*. New York: St. Martin's Press.
- Sobkowiak, W. 2005. "Lexicographic phonetics or phonetic lexicography?". In H.Gottlieb, J.E.Mogensen & A.Zettersten (eds). 2002. *Symposium on lexicography XI*. [Lexicographica Series Maior 115]. Tübingen: Max Niemeyer Verlag. 511-519.
- Wells, J.C. 1990. *Longman Pronunciation Dictionary*. London: Longman.