E-dictionaries and Phonolexico-graphic Needs of EFL Users*

Włodzimierz Sobkowiak, *School of English, Adam Mickiewicz University, Poznań, Poland (sobkow@amu.edu.pl)*

Abstract: The phonetic aspect of (EFL) dictionaries is among the most seriously underrated and underdeveloped in (meta)lexicography. Pertinent bibliography is scant and even the best learner dictionaries are found wanting on a number of counts. This contribution is both a summary of my thirteen-year-long research into (pedagogical) phonolexicography and a look ahead. The current state-of-the-art in phonolexicography is presented with particular attention paid to how the leading pedagogical EFL e-dictionaries relate to the actual and potential phonolexicographic needs of their users, both students and teachers.

The main themes are: (a) the specificity of phonolexicographic needs of (Polish) EFL users, (b) phonetic representation, both graphic and acoustic, in dictionaries and its problems, (c) phonetic access, i.e. querying the contents of the dictionary via the phonetic code, (d) didactic aspects of phonolexicographic information, i.e. its use in teaching and learning (EFL) pronunciation.

Keywords: E-DICTIONARIES, EFL, ENGLISH, PHONOLEXICOGRAPHY, PHONETIC ACCESS, PRONUNCIATION IN DICTIONARIES

Opsomming: E-woordeboeke en fonoleksikografiese behoeftes van EVT-gebruikers. Die fonetiese aspek van (EVT-)woordeboeke is van die ernstigste onderskatte en onderontwikkelde in die (meta)leksikografie. Tersaaklike bibliografiese besonderhede is skaars en selfs die beste aanleerderswoordeboeke blyk gebrekkig in 'n aantal opsigte te wees. Hierdie bydrae is sowel 'n opsomming van my dertien-jaar-lange navorsing oor (opvoedkundige) fonoleksikografie as 'n blik vorentoe. Die huidige stand van sake in fonoleksikografie word aangebied met besondere aandag gewy aan hoe die toonaangewende opvoedkundige EVT-e-woordeboeke verband hou met die werklike en potensiële fonoleksikografiese behoeftes van hul gebruikers, sowel leerders as onderwysers.

Die hoofonderwerpe is (a) die spesifieke aard van die fonoleksikografiese behoeftes van (Poolse) EVT-gebruikers, (b) die fonetiese weergawe, sowel grafies as akoesties, in woordeboeke en

---

die probleme daarmee, (c) die fonetiese toegang, d.w.s. die raadpleging van die inhoud van die woordeboek deur middel van die fonetiese kode, en (d) die didaktiese aspekte van die fonoleksiko-logiese inligting, d.w.s. die gebruik daarvan in die onderrig en leer van (EVT-)uitspraak.

Sleutelwoorde: E-WOORDEBOEKE, ENGELS, EVT, FONOLEKSIKOGRAFIE, FONETIESE TOEGANG, UITSPRAAK IN WOORDEBOEKE

1. Phonolexicography

1.1 State of the art

In a series of publications spanning the last thirteen years (see http://ifa.amu.edu.pl/~swlodek/public.htm), I have repeatedly pointed out that phonolexico-graphy is almost completely ignored, both linguistically and lexicographically. This has been independently noticed (but not dealt with scientifically) by other researchers, such as Landau (2001: 126), who approvingly summarized Hul-bert's 1955 views in the following way: "Dictionaries are less satisfactory in pronunciation than in spelling, meaning, or etymology. The record of the spo-ken language is difficult to acquire, difficult to transcribe accurately and unam-biguously, difficult to represent understandably in a dictionary transcription, and in most cases of less interest to the user than other kinds of information."

Also lexicographically aware phoneticians like Gimson (1973: 115) have had uncomplimentary views on how pronunciation is treated in dictionaries: "Today, when we no longer regard speech as a degraded form of writing, the pronunciation entry in dictionaries […] should be accorded much greater im-portance." "Unfortunately, the theory underlying the pronunciation component in a dictionary is too frequently difficult to discern."

This, rather sad, state of the art in phonolexicography has not changed noticeably over the last two decades, which have seen an unprecedented growth of lexicography and metalexicography alike. It is enough to browse the abundant proceedings of the biannual Euralex congress, for example, to notice the almost complete absence of any deeper phonolexicographic reflection. It is almost as if there were a consensus among researchers that the few questions that arise at the interface between dictionaries and phonetics have long since been answered, e.g. by Gimson 1973, Abercrombie 1978 or Wells 1985. As a matter of fact, however, these 'classic' treatments of phonolexicography have, doubtless by design, raised more questions than answers. Some of these ques-tions have again been approached in my two books on phonolexicography (Sobkowiak 1999 and 2006), which are now — to the extent that I can ascertain — the only such full-length monographs in existence. In the following section these issues will be considered briefly.

1.2   Main issues

1.2.1   The design of phonetic transcription

One of the most vigorously debated issues in phonolexicography (and beyond: EFL, for example) has been the choice of phonetic transcription to represent the pronunciation of headwords. It is paradoxical that this rather technical question should raise so much controversy. After all — phonolexicographically speaking — there are many other, apparently more substantial, problems than whether chronemes (ː), tildes (~) and bars (-) should be allowed or not in dictionary transcriptions, and why. It is fair to note, of course, that these typographic decisions may reflect much deeper preoccupations with the structure and function of phonetic representation (for example, the classic one: whether vowel timbre in English is best treated qualitatively or quantitatively in such pairs as /i/ and /iː/). By and large, however, the graphical salience of typography has tended to dominate the scene to the disadvantage of other issues.

Even without serious study, it is easy to notice that there is little unanimity as to the choice of phonetic transcription for lexicography across the Atlantic, Americans sticking to their long tradition of respelling (Paikeday 1993), only occasionally modified by the introduction of schwa and a few other IPA borrowings. The plethora of EFL dictionaries published in European countries, on the other hand, tend to try all kinds of non-IPA experiments in transcriptional user-friendliness, which are bound to lead to widely differing results, depending on the L1 of the target users (see Sobkowiak 1999: 183-200 for a discussion of Polglish (Polish English) transcriptions). There are few experimental studies of dictionary user reactions to different transcription systems (Fraser 1997), and much more research is needed, especially in the area of EFL pedagogical dictionaries, before solid applicational conclusions can be drawn. All in all, the design of an effective phonetic transcription for lexicography remains an open issue.

1.2.2   The scope of phonetic transcription

Which dictionary text should be transcribed: headwords, run-ons, phrases, definitions, examples, equivalents (see Bogaards 2005: 23)? This question has traditionally been answered to minimize the scope of phonetic transcription to head entries only, doubtless for reasons of typographic ease and economy of space. Both considerations must be revised now in the era of electronic dictionaries, leaving pedagogical issues on stage. For example, would learners benefit from having definitions in a monolingual dictionary fully phonetically transcribed? How about examples of usage? After all, "the phonological behaviour of words in context and its representation deserves equal attention" (Magay

1979: 103). Would it be useful to learners to be able to view phonetic transcription of either at different phonostylistic levels, say: 'careful speech' and 'fast speech'? While such solutions could now easily be implemented in electronic pedagogical dictionaries, I am not aware of any research in this direction. Indeed, the only book-length treatment of dictionary definition phonetics (of *Macmillan English Dictionary for Advanced Learners* (MEDAL)) which has ever appeared is Sobkowiak 2006. No empirical user research was attempted here, however. The results are, nevertheless, very interesting, showing, among others, that there are ways to make such definitions phonetically more user-friendly to EFL learners by a judicious choice of defining vocabularies, phonetic control over phraseology and sandhi, audio recordings or synthesis, and others. As before, and much despite the common metalexicographic belief, the issue of the microstructural scope of phonetic transcription remains open.

1.2.3   The choice of default accent and phonostylistic level

These are two of those phonolexicographic issues which have received most attention. All of the 'classic' treatments mentioned above addressed them. There seems to be a fair amount of agreement to the effect that (a) the accent of choice is 'Received Pronunciation' (RP) for British English and 'General American' (GA) for American English, and (b) the citation form of a given word is chosen for phonetic representation. There are solid arguments behind these choices, which mostly reduce to the conviction that offering the EFL users more variety in either of the two sociophonetic dimensions would not only be counterproductive, but virtually confusing, not to mention the ubiquitous BC (before computers) concerns about precious printing space.

Notice, however, that the implicit assumption in the avoidance of sociophonetic variation in EFL dictionaries has been that the proper target users of EFL pedagogical dictionaries are *learners* only. They are — so the argument goes — bewildered enough by other aspects of dictionarese to better spare them this additional worry of having to choose the right phonostylistic form for the occasion. Besides, research has shown that little dictionary use is conducted for purposes of speaking, so in this situation there is little to be gained from phonostylistic advice. These are strong arguments, but they seem to completely ignore the phonolexicographic needs of EFL *teachers* and very advanced learners, who may well need phonetic representation of headwords (a) in other accents than the canonical RP and GA, (b) on a variety of phonostylistic levels. The teacher's dictionary will be treated at some more length below.

1.2.4   Phonolexicographic consistency

Ensuring the consistency of phonetic representation across the entries, i.e. in

the macrostructure of the dictionary, has been one of the preoccupations of (pedagogical) lexicographers: "It goes without saying that, whatever method of indicating pronunciation is adopted, it should be consistently and correctly used in the body of the work" (Abercrombie 1978: 124). And yet, even in this electronic era, with dictionaries compiled with sophisticated software tools from the very inception, there are still too many inconsistencies to ascribe to pure chance. Many examples are listed in my phonolexicographic review of the first edition of MEDAL (Sobkowiak 2003).

Dictionary makers have problems with phonolexicographic consistency on many levels, from "conformity of representations with the established phonological rules of the language" (Sobkowiak 1997a: 98) to aligning headword transcription with its audio representation. Such errors could be spotted quite easily at the design and production stage, should phonolexicography attract more attention than it now does in lexicographic circles.

From the user perspective, the results of phonolexicographic (or, for that matter, any other) inconsistency may be confusing or positively misleading. While dictionaries are not prototypical resources used for the acquisition of foreign pronunciation, they do convey phonetic information in a supposedly orderly manner. In their craving for rules, learners have the right to expect that also the phonetic content of their dictionaries is rule-governed, and that by frequent use of the dictionary they may expect assistance in acquiring it, be it explicitly or implicitly. If some word-final nasals in the dictionary are transcribed syllabically (to take my favourite example; see Sobkowiak 1999: 94-112) and others — in analogous phonetic contexts — are not, phonolexicographic consistency is compromised and the process of acquiring the pertinent phonological rule of English interfered with. To take another example, in MEDAL1, I discovered twenty headwords starting with /ænti-/ (e.g. *antibiotic, anticlerical, anticyclone, antidepressant, antifreeze*, etc.) and seven with /æntɪ-/ (*antibody, anticlockwise, antidote, antigen, Antipodean, the Antipodes, antithetical*). If there is any linguistic motivation for this variation, it is certainly too subtle for learners to discern.

1.2.5 Phonetic access

Traditional paper-based dictionaries could not be accessed phonetically. Just about the only access path in semasiological dictionaries was through alphabetically ordered headwords. Sometimes indexes of different types were provided to aid users having less orthodox needs. It was quite recently that phonolexicographic needs of EFL users were finally recognized, and pedagogical e-dictionaries started to offer access to the lexicographic content through phonetic transcription, technically mediated by the so-called 'phonetic keyboards'. One such keyboard is reproduced below.

Figure 1:   MEDAL1 *SoundSearch* phonetic keyboard

This is certainly a step in the right direction. Learners can now scan the dictionary for problematic homophones or phoneme clusters; teachers have an excellent tool for material and test preparation. As with many such new developments in contemporary lexicography, it remains to be seen just how user-friendly and useful these new access methods are in the actual EFL practice. Needless to say, there is practically no research yet in this area.

1.2.6   Audio sound representation

Another recent development of potentially significant impact on phonolexicography is good-quality Text-to-Speech Synthesis (TTS). Some systems now offer artificial speech virtually indistinguishable from natural. With the exception of some experimental applications, however, TTS has not yet been used in pedagogical lexicography. The potential benefits go far beyond economy of reader cost and space on the CD-ROMs, now taken up by the memory-hungry recordings of human speech. As I explained in Sobkowiak 1999, chapter 3.10, TTS could be used to audialize (a) any desired stretch of dictionary micro-structure, apart from the headword itself, (b) at any desired phonolectal and phonostylistic level, (c) with the desired gender and voice quality, (d) with or without the desired L1 accent, and (e) for mere representation or as part of a suite of pronunciation exercises built into the dictionary. This phonolexico-graphic flexibility would be an enormous asset to both learners and teachers, who must currently resign themselves to recordings made by a handful of speakers, sometimes with their own phonetic idiosyncrasies. The one-word-

one-phonetic-representation principle of current lexicography is, of course, a serious misrepresentation of linguistic reality.

1.2.7   The typographic design of phonetic representation

Finally, another phonolexicographic issue which has received by far too little serious interest in the relevant literature is the screen typography of phonetic representation. This includes not only phonetic transcription proper, but also the accompanying symbols and flags, fonts with their attributes, scalability, GUI (Graphical User Interfaces) design, platform and copy-and-pasting robustness, and many others. Some examples of common problems in this area, with their consequences for the dictionary use by learners, are provided below (section 3.3). Needless to say, this is again a phonolexicographic field in urgent need of empirical user-directed research.

2.   Phonolexicographic needs of EFL users

2.1   Are there any?

There is now abundant research to show that many EFL dictionary users have rather modest needs when it comes to phonetic representation (see Sobkowiak 1999: 115-121 and Lew 2004: 21-24 for overviews). Obviously, some of the largest groups of EFL dictionary users — translators, businessmen, secretaries, scientists — seldom need pronunciation in their lexical resources for many of their job-related activities. In heavy dictionary use for professional purposes, where the word's pronunciation is irrelevant and can actually be completely ignored, phonetic transcription can be worse than superfluous — it can be obtrusive. (Notice in this context that there is no option, in e.g. MEDAL1, to switch transcription off completely for screen display.)

Also some EFL learners express their disinterest in matters phonetic in dictionaries. This can well be understood in a school situation where reference resources are mostly used to practice the so-called 'passive' skills: listening and reading, with emphasis on the latter. There are clear logistic problems in using dictionaries for speech: listening and speaking alike. All this acknowledged, it is interesting that Lew (2004: 111-112) found as many as 28% (Polish) learners in his study claiming that they were consulting their dictionaries for pronunciation *often* and *most often*; see Table 1 below. Additionally, the phonetic use of the dictionary went up sharply in frequency at the most advanced proficiency level (Lew 2004: 123).

Table 1:   Phonetic consultations of dictionaries in Lew 2004

never	rarely	often	most often	missing
179 (25%)	336 (47%)	160 (23%)	35 (5%)	2 (0%)

All these results show that there is a stable minority of EFL learners, many of them advanced, for whom phonetic representation in dictionaries does matter. As before, an important proviso needs to be mentioned: nobody has yet studied the phonolexicographic needs of EFL *teachers*. Lew's proficiency-stratified results suggest that these may be quite extensive.

My own research in phonolexicographic needs of Poznań university students (Sobkowiak 1999, part 2) points in a similar direction (although this sample may not have been quite representative of the population of advanced EFL learners because of the heavy phonetic bias of the School of English EFL programme). The following is a subset of questions asked in my questionnaire returned by more than six hundred students of English, at the Adam Mickiewicz University, Poznán, Poland.

Table 2:   Some phonolexicographic needs of EFL Poznań university students (Sobkowiak 1999)

Statement	yes	no	?
English dictionaries should show pronunciation in some way	601	24	20
I check pronunciation of some words when I read	547	95	3
I have recently checked pronunciation in an English dictionary	611	32	2
In listening comprehension tasks, I look up some words I hear	458	170	17
International Phonetic Transcription (IPA) should be used in all English dictionaries	499	61	85
Learners' dictionaries should address their common pronunciation problems	528	51	66
Separate pronunciation dictionaries are nonsense	90	519	36
In a multimedia computer dictionary, phonetic transcription is useless	33	372	240

2.2   Learner-friendliness

With the advent of computers, especially after GUIs had been introduced, there has been a lot of preoccupation with user-friendliness ("the elegance and clarity with which the interaction with a computer program or a web site is designed" — Wikipedia). In the present context, the issue is (EFL) learner-friendliness, of course, but the design criteria remain the same.

There are many different respects in which dictionaries can be phonolexicographically learner-friendly. If lexicographers are aiming at global 'user-friendliness' of their defining vocabulary, they should certainly also make it phonetically friendly by avoiding notoriously phonetically difficult words, for example. As I showed in Sobkowiak 2006, there are few signs that this aspect of defining vocabularies comes under deliberate lexicographic control in the leading EFL dictionaries now on the market. And yet, phonetically difficult words in dictionary definitions will tend to impede the reading and understanding

process, particularly in those learners who continue to vocalise or articulate subvocally in silent reading, as many do (cf. e.g. Gibson and Levin 1975: 342, Grodziński 1976: 52ff).

EFL electronic dictionaries could dynamically adjust their definitions to the learner's needs and requirements, not only in terms of their lexical scope (defining vocabulary) and syntax, but also in terms of pronunciation (see De Schryver 2003 for this and other lexicographers' dreams). For example, if *thorough* is among the phonetically hardest lexical items for EFL learners, why not use a substitute in definitions (e.g. *complete*) adjusted for pre-intermediate learners, where (syntactically, stylistically and pragmatically) appropriate? Or at least why not reduce the definition incidence of *thorough* (it now stands at 22 in MEDAL1)?

2.3   Teacher's dictionary?

Google search (in website title) for the phrase "teacher's dictionary" currently (7 October 2006) yields 11 hits, none of which is actually a pedagogical EFL dictionary for teachers in the sense which has been of concern here. "Learner's dictionary", on the other hand, yields 21 900 web pages. Why this difference? Do EFL teachers not use EFL dictionaries? Of course, they do: they use their learners' dictionaries, which are in many respects unfit for teachers' needs. What would, then, be the expected features of an EFL teacher's e-dictionary, from the phonolexicographic perspective? For lack of space (and the issue quite clearly deserves an in-depth treatment), they will briefly be listed below. A teacher's dictionary must offer:

—   liberal deployment of phonetic transcription, both for representation and indexing,

—   powerful multicriterial phonetic access mechanisms,

—   phonetically streamlined word-list generation functionalities,

—   robust copy-and-paste facilities, and

—   indication of accentual, phonostylistic and phonolectal variation, with

—   advice on the preferred phonetic form in various situations of use,

—   guidance on likely phonetic problems with the given entry/definition/ example, and

—   remedial drills and exercises, generated automatically from phonolexico-graphic context.

This is certainly a far-from-exhaustive list of desiderata. The lexicographic theory (let alone practice) of the teacher's dictionary does not exist. I believe that, with half the globe learning and *teaching* English as a lingua franca, it is now long overdue.

3.   Phonetic representation

3.1   Transcription vs. others

Doubts have sometimes been voiced as to the desirability of phonetic transcription in dictionaries for learners (Kernerman 1996: 409), in view of the now pervasive use of recorded speech as phonetic representation of headwords. Again, little pertinent research has been conducted in this area. It is intuitively obvious that learners are indeed quite fond of clicking on the headword to listen to the model pronunciation. It is less obvious how many of them would want to keep the transcription as well. The answer, which I obtained in my 1999 questionnaire, listed in Table 2 above ("In a multimedia computer dictionary, phonetic transcription is useless": yes – 33, no – 372, don't know – 240), is hardly indicative of the needs of 'ordinary' EFL learners, i.e. ones of non-academic proficiency, who do not study to become teachers of English. Notice also, that while most subjects claimed that even in a multimedia dictionary phonetic transcription is useful, quite a few (37%) were uncertain.

And yet, there are pedagogically clear advantages of phonetic transcription displayed in addition to the audio representation. While the latter provides a receptive and productive model to be used for immediate mimicry, the former offers categorization, systematization and 'accuracy anchor', to borrow Grabe and Stoller's (1997: 112) term.

Similarly, a visual waveform representation of spoken words, increasingly more often provided as a thrilling multimedia extra, may offer multisensory feedback for practice, but is risky because "the *visual* comparison of the two sound waves, the model's and the learner's repetition [...] is at best inconsequential, and at worst thoroughly misleading and frustrating. The graphic representation of a waveform has a very complex relationship to its acoustic basis, and the latter to both its articulation and perception" (Sobkowiak 1997b: 335). Consider, for example, my own best rendition of *Good morning* compared with the native model in Figure 2. Should I not have known better, I would have been very frustrated by the dissimilarities between the two waveforms, apparently testifying to my complete mispronunciation of this phonetically easy phrase.

Figure 2:   A waveform of *Good morning*; model left, learner right
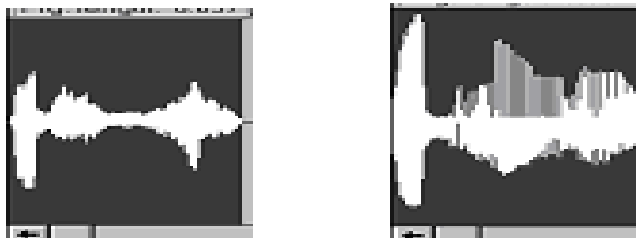
3.2   Simplified transcription?

To foreign learners there are no self-pronouncing words. For a quarter of a century, or more, my own phonetic image of the word *desultory* has been parallel to *compulsory*, i.e. */dəˈsʌltərɪ/. *Instead* may look self-pronouncing to native speakers of English, but is far from this, as many beginning/intermediate Polglish errors testify, e.g. */ɪnˈstiːd/. *All* headwords in an EFL pedagogical dictionary should have phonetic representation, however obvious it may seem to the (native) dictionary makers.

As mentioned in section 1.2.1, while phonetic transcription has definitely been normalizing during the whole 20th century, there is still wide diversity, especially cross-Atlantically. Respelling continues to be quite popular in America, where the concept of an EFL monolingual learner's dictionary with the 'international' IPA transcription is a relatively recent graft from Great Britain. From the point of an EFL dictionary user, however, respelling can be thoroughly confusing, as it is of course heavily target-language-dependent. The following praise, voiced on http://dictionary.reference.com/help/faq/language/s36.html, can then only refer to native speakers of English: "A respelling pronunciation system is fairly practical. No special characters or diacritics are used. No pronunciation guide must be relied upon. Examples are: *accident* (AK-si-dunht), *diamond* (DIE-muhnd), *garage* (gah-RAHZH, guh-RAHZH, GA-rahzh)." A Polish EFL learner would be uncertain as to the correct phonetic value of /AK/, /dunht/ or /RAHZH/. The bottom line is, then, that in the EFL situation any such 'simplifications' are very risky and should be thoroughly tested empirically before deployment. It is not surprising that the IPA standard, despite much criticism directed at it from EFL circles (especially in the 1970s) holds on quite staunchly in Europe.

An important proviso in this context is that in e-dictionaries many types of transcription are in principle mutually convertible and selectable. There is thus nothing to stop EFL e-dictionary makers from offering learners a transcription of their choice, including the many subtle varieties of IPA on the one hand, and some L1-sensitive derivative simplifications on the other (see, e.g. Sobkowiak 1997c).

3.3   Typography

As mentioned briefly in 1.2.7, the whole issue of typographic design of phonetic transcription for screen rendering has so far received very little attention. For example: "the whole of CD-ROM MEDAL is set in sans serif Arial-like font [...]; the only exception being the non-Roman IPA symbols [...] which are serifed. Depending on the graphics card's selected screen resolution, both the shape and the size of the transcription field symbols may differ: /ˈkɒntrəvɜːsi/" (Sobkowiak 2003: 427). This is indeed a common weakness of e-dictionaries, including the bilingual English–Polish dictionaries which I have seen. The con-

sequences for the user/learner are twofold: (a) font inconsistencies show that phonetic representation is somehow inherently sloppier than the rest of the microstructure, hence (maybe) of less lexicographic importance; and (b) subtle differences in symbol shape may be taken to imply phonetic contrasts (after all, in transcription, unlike in orthography, every little bit counts).

Another weakness of e-dictionaries is that transcription does not survive copy-and-pasting, a common enough procedure, especially with young computer users, who usually type better than they handwrite from kindergarten onwards. The following is the phonetic fragment of the *controversy* entry copied from MEDAL1 on CD-ROM into MS Word 2000: / B `k\197ntrEv3:si, B kEn`tr\ 197vEsi /. While the mapping with IPA may be one-to-one, it would certainly be too much to expect learners to translate from this SAMPA-like transcription to anything closer to standard.

3.4   Choice of model

As mentioned above, the traditional monolectal solutions to the problem of pronunciation model choice appear to work best for most, especially lower-proficiency, learners. On the other hand, real phonetic variation caused by dialect, phonic style, lexical frequency, or indeed the so-called 'free variation' should not be suppressed by the overwhelming force of the "monostylistic curse" (Bailey 1986: 26). Advanced learners and teachers may need this information to follow their aims. The solution appears to be to customize display of phonetic content to the needs of the user. Just as different transcription types can be switchable in an e-dictionary, so can different phonolects. To the extent that they are phonologically rule-governed, the switching can be done automatically across the whole macrostructure of the dictionary. Idiosyncrasies would need to be programmed individually for each word, i.e. the way that all dictionary transcription is prepared nowadays.

3.5   Transcription vs. audio

The main problems with the recorded audio component of many dictionaries are: (a) the list-reading effects (e.g. contrastive stress, sustained intonation contours), (b) the occasional mismatch between recording and transcription, (c) limiting the recordings to headwords only (the recent edition of *Longman Dictionary of Contemporary English* is an exception here: http://www.longman.com/ldoce/about_cd.html). As an example of (b), listening to all British headword recordings in MEDAL1 within the ⟨a⟩-letter range, I found about seventy cases of mismatch in the few thousand entries, e.g. *academician* is transcribed as /əˌkædəˈmɪʃn/ but pronounced as /æˌkədəˈmɪʃn/ in MEDAL1. Or consider the non-coalesced recording of *aperture* /ˈæpətʃə/. Good quality TTS, as mentioned above in section 1.2.6, would solve all of these problems.

4.   Phonetic access

In the traditional paper dictionaries, transcription fulfils but one function —
representational. Learners refer to it to find out about the pronunciation of the
(orthographically) located word. This is very different from the indexical
(query) function, which transcription additionally has in a phonetic-access
enabled e-dictionary, such as MEDAL. Here learners enter phonetic transcrip-
tion for a word or word-list which they search. This function of transcription
simply did not exist in traditional dictionaries. In e-dictionaries, both functions
must be smoothly integrated. As it appears, this is far from trivial. Consider the
question of grapho-phonemic biuniqueness for data entry, for example. From
the point of view of mere representation, it does not matter whether English
affricates or diphthongs are coded mono- or bisegmentally. In their representa-
tional function /tʃ/ and /t ʃ/, they do not differ at all. Bisegmental coding for
querying (indexical function) will not do, however. The search for /*ʃ/ words
will yield: *aitch, approach, arch, attach,* etc., as it does in MEDAL1, for instance.
As I described in Sobkowiak 2003, there are many other such problems, con-
cerning stress marks, wildcards and brackets, for example.

　　　Some EFL-wise useful queries, such as one for aspirated plosives, are quite
simply impossible, not only in the electronic versions of standard learners' dic-
tionaries, such as MEDAL or Longman, but even in the dedicated *pronouncing*
dictionary, such as the *Cambridge English Pronouncing Dictionary* (CEPD) on CD-
ROM (see my review Sobkowiak 2005). As can be seen in the screenshot of the
CEPD phonetic keyboard in Figure 3, such EFL-wise exoticisms as the Welsh
voiceless laterals can be entered, but there is no way to type aspiration or sono-
rant syllabicity, both phenomena of obvious interest to EFL learners. The con-
clusion is that phonetic access has yet some way to go before it can be fully and
profitably integrated in EFL learners' e-dictionaries.

Figure 3:   CEPD phonetic keyboard
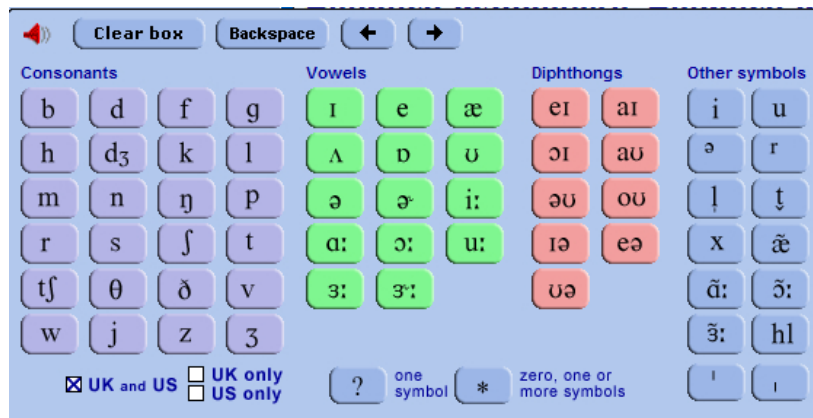
5.   Didactic aspects of phonolexicography

5.1   Dictionaries as teaching/learning resources?

Despite some voices to the contrary: "It cannot realistically be seen as part of the dictionary's function to teach the sound system" (Brazil 1987: 161), EFL dictionaries are gradually transformed into one-stop learning resources, with grammatical, pragmatic and phonetic advice on board, a variety of self-study facilities, multiple suites of exercises and hyperlinks to even more support on the world-wide web. In Sobkowiak 1998, I argued that despite the traditional idiographic perspective of lexicography, EFL MRDs (machine-readable dictionaries) can indeed teach foreign pronunciation. This is all the more true of e-dictionaries, with their phonolexicographic query potential, which I have been unveiling in this article. For example, which EFL resource other than properly coded e-dictionary could (semi-)automatically generate exercises spun around: (a) those English animal names which are (b) relatively common in colloquial English, but (c) relatively difficult (grapho)phonetically to Polish (foreign?) learners: *calf, lamb, sow, bison, donkey, giraffe, leopard, monkey, reindeer*?

5.2   Techniques

The only type of pronunciation exercise currently used in EFL e-dictionaries is "repeat after me", as in the following screenshot from MEDAL1.

Figure 4:   "Repeat after me" exercise in MEDAL1

This limitation, however, is not one of technology or pedagogy, but of imagination. For example, pronunciation practice could easily be combined with flashcards. Why is it that the only elements of an entry's microstructure used in flashcards are its headword and definition? Why not let the learner "type the mystery word" in response to its phonetic transcription or recording (dictation), or both, as well as to its definition? Why not flash the headword and ask the learner "Do you remember how to pronounce this word?", as well as "Do you remember what this word means?", the only option now built in?

      But one could, of course, go much further. The wildcard and regular-expression queries of phonolexicographic content yield lists of items from which the original criterion could be inductively reconstructed by the learner. For the beginner: "What do these words have in common, as far as pronunciation goes: *dough, go, know, sew, toe*?" (a selection from forty-nine MEDAL1 /ʔəʊ/

headwords). For the advanced: *aphrodisiac, Chianti, cordiality, piano, react*?" (a selection from ninety-nine /*iæ*/ headwords, containing a particularly recalcitrant case of vocalic hiatus).

Of course, it is easy to transform these exercises into binary- or multiple-choice format, matching, selection or minimal pairs. It is possible to combine them with audio and/or transcription, as well as part-of-speech information, lexical frequency tagging, dialectal and stylistic stratification, illustrations, etc.

5.3   L1-sensitivity

Both theoretical and applied phonolapsology must be L1-sensitive because phonetic interference from L1 is the amplest source of L2 pronunciation errors on most levels of proficiency. Contemporary EFL dictionaries are not L1-sensitive, or only superficially so (e.g. false-friends lists), mostly due to the overwhelming commercial factors. For example: the EFL dictionaries currently available on the Polish market do not take account of one interference-prone Polish graphophonemic rule: ⟨c⟩ → /ts/. To Polglish beginners, *romantic* is /ro'mantits/. There is an abundant bibliography of sources on common Polglish phonetic errors at different proficiency levels. All that needs to be done is to build this knowledge into EFL e-dictionaries for Polish learners, and then use it to post cautionary flags or construct phonetically L1-sensitive exercises using the phonolexicographic content of the dictionary as input. This is already happening on other levels of lexical structure in dictionaries: morphological, syntactic, pragmatic. Phonetics is obviously lagging behind.

5.4   The undiscovered potential of definitions

In the final chapter of Sobkowiak 2006, I unveil some of the so-far undiscovered pedagogical potential of monolingual EFL learners' dictionary definitions. If treated as a properly phonolexicographically annotated corpus of text, rather than separate text bits attached to headwords, they can furnish an excellently rich resource for learners, providing opportunity for pronunciation practice along different phonetic dimensions: from segmentals, through sandhi, to stress, rhythm and intonation.

Consider an example. A learner who has problems with linking-r, dental(ized) clusters or fast-speech /d+j/ affrication (and many invariably do), such as in this MEDAL1 definition of *melt – if you melt into o<u>r against someone you relax a<u>s they hol<u>d you close in a romantic way</u>*, could obtain some advice and ample material for practice taken from other phonetically treated definitions, where such phonetically troublesome phenomena would be suitably coded and could be retrieved in the form of a phonetic concordance focused on the relevant difficulties, such as in the following example:

—   exactly: *in every way o<u>r e</u>very detail*
—   intense: *very great o<u>r e</u>xtreme*

—   severe/ly: *very strict o**r e**xtreme*

—   to have one foot in the grave: *to be very old o**r i**ll and likely to die soon*

6.   More (phono)lexicographic dreams

6.1   Phonetically treated definitions

I dream of dictionaries which would redress the anti-phonetic bias of current lexicography. In such dictionaries, not only the phonetic representation of the headword would be carefully thought over, but the entire entry would likely receive phonolexicographic attention. Definitions would be designed and written according to some phonetic guidelines, just as they are written according to strict syntactic and stylistic guidelines today. In consequence, such definitions would be easier to read, both as meta-text and text, and the incidental learning of vocabulary would get a boost. Properly annotated, they could be used as a (phonetic) electronic corpus resource in simple lookup as well as in a variety of sophisticated queries informing word-list generation, test preparation, materials design, etc. (see 5.4 above). Audio-recorded or speech-synthesized, they could make an e-dictionary a yet more functional didactic aid, with all the multimedia involvement expected by contemporary learners (Sobkowiak 2006).

6.2   Articulatory animation

The so-far lexicographically unimplemented type of potentially pedagogically useful phonetic representation is articulatory animation. Realistic avatars of the *Baldi* kind (http://mambo.ucsc.edu/pdf/WilliamsRDD.pdf) could be used, with speech animated in real time. Different perspectives, zooms, transparencies and tempos could be used (see Figure 5 below). The animation can be coordinated with a human recording or with TTS output. The technology is now mature for use in e-dictionaries. The benefits for e-dictionary EFL users are too obvious to elaborate on.

Figure 5:   *Baldi* going transparent

7.   Instead of conclusions

As was alluded to above, "indicating pronunciation is often under-estimated by the critics of dictionaries as being a derivative business" (Magay 1979: 99). I hope to have shown that phonolexicography, while far from being "a derivative business", can and should become a thriving scholarly field, as well as a fully-fledged discipline of applied lexicography. Many users use their EFL dictionaries for pronunciation, starting from the modest one-off lookup of the word's phonetic transcription, through phonetic-access wildcard searches for troublesome clusters, to pronunciation exercises built into some e-dictionaries for learners. Despite its enormous pedagogical potential, and unlike other sub-disciplines of learners' lexicography, theoretical and applied phonolexicography has hardly been developing in the last decades. This is one more attempt to start to change this unsatisfactory situation.

Bibliography

Abercrombie, D. 1978. The Indication of Pronunciation in Reference Books. Strevens, P. (Ed.). 1978. *In Honour of A.S.Hornby*: 119-126. Oxford: Oxford University Press.

Bailey, C.-J.N. 1986. Even Dialectologists Get the Blues: Inadequate Phonetic Transcription of English. *Papiere zur Linguistik* 35(2): 3-38.

Bogaards, P. 2005. Dictionaries and Productive Tasks in a Foreign Language. *Kernerman Dictionary News* 13: 20-23.

Brazil, D. 1987. Representing Pronunciation. Sinclair, J.M. (Ed.). 1987. *Looking Up. An Account of the COBUILD Project in Lexical Computing and the Development of the Collins COBUILD English Language Dictionary*: 160-166. London/Glasgow: Collins ELT.

De Schryver, G.-M. 2003. Lexicographers' Dreams in the Electronic-Dictionary Age. *International Journal of Lexicography* 16(2): 143-199.

Fraser, H. 1997. Dictionary Pronunciation Guides for English. *International Journal of Lexicography* 10(3): 181-208.

Gibson, E.J. and H. Levin. 1975. *The Psychology of Reading*. Cambridge, Mass.: The MIT Press.

Gimson, A.C. 1973. Phonology and the Lexicographer. McDavid, R.I. Jr and A.R. Duckert (Eds.). 1973. *Lexicography in English*: 115-121. International Colloquium, New York 1992. Annals of the New York Academy of Sciences 211. New York: Academy of Sciences.

Grabe, W. and F.L. Stoller. 1997. Reading and Vocabulary Development in a Second Language. Coady, J. and T. Huckin (Eds.). 1997. *Second Language Vocabulary Acquisition: A Rationale for Pedagogy*: 98-122. Cambridge: Cambridge University Press.

Grodziński, E. 1976. *Mowa wewnętrzna* [Inner speech]. Wrocław: Ossolineum.

Hulbert, J.R. 1955. *Dictionaries: British and American*. London: Deutsch.

Kernerman, L. 1996. English Learners' Dictionaries: How Much Do We Know about Their Use? Gellerstam, M. et al. (Eds.). 1996. *Euralex '96 Proceedings*: 405-414. Göteborg: Göteborg University, Department of Swedish.

Landau, S.I. 2001. *Dictionaries: The Art and Craft of Lexicography*. Second edition. Cambridge: Cambridge University Press.

Lew, R. 2004. *Which Dictionary for Whom? Receptive Use of Bilingual, Monolingual and Semi-bilingual Dictionaries by Polish Learners of English*. Poznań: Motivex.

Magay, T. 1979. Problems of Indicating Pronunciation in Bilingual Dictionaries with English as the Source Language. Hartmann, R.R.K. (Ed.). 1979. *Dictionaries and Their Users*: 98-103. Proceedings of the 1978 BAAL Seminar on Lexicography. Exeter Linguistic Studies 4. Exeter: Exeter University Press.

Paikeday, T.M. 1993. Who Needs IPA? *English Today* 9(1): 38-42.

Sobkowiak, W. 1997a. Consistency in EFL Dictionary Phonetics. Waniek-Klimczak, E. (Ed.). 1997. *Teaching English Phonetics and Phonology II. Accents '97*: 95-102. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Sobkowiak, W. 1997b. Speech in EFL CALL. Strykowski, W. (Ed.). 1997. *Media a edukacja*: 329-339. Poznań: eMPi[2].

Sobkowiak, W. 1997c. Radically Simplified Phonetic Transcription for Polglish Speakers. Hickey, R. and S. Puppel (Eds.). 1997. *Language History and Linguistic Modelling. Festschrift for Jacek Fisiak on his 60th Birthday*: 1801-1830. Berlin: Mouton.

Sobkowiak, W. 1998. Can EFL MRDs Teach Pronunciation? Fontenelle, T. et al. (Eds.). 1998. *Euralex '98 Proceedings*: 271-277. Liège: University of Liège, English and Dutch Departments.

Sobkowiak, W. 1999. *Pronunciation in EFL Machine-Readable Dictionaries.* Poznań: Motivex.

Sobkowiak, W. 2003. Pronunciation in *Macmillan English Dictionary for Advanced Learners* on CD-ROM. *International Journal of Lexicography* 16(4): 423-441.

Sobkowiak, W. 2005. Review of *Cambridge English Pronouncing Dictionary* (CEPD) on CD-ROM. *TESL-EJ* 8.4. http://writing.berkeley.edu/TESL-EJ/ej32/m1.html.

Sobkowiak, W. 2006. *Phonetics of EFL Dictionary Definitions.* Poznań: Wydawnictwo Poznańskie.

Wells, J.C. 1985. English Pronunciation and its Dictionary Representation. Ilson, R. (Ed.). 1985. *Dictionaries, Lexicography and Language Learning*: 45-51. London: Pergamon.