On the Problems in EFL lexicography[*]

Anna Włodarczyk-Stachurska

Uniwersytet Technologiczno-Humanistyczny, Poland
Malczewskiego 29, 26600 Radom
PhD
E-mail: a.stachurska@uthrad.pl

Abstract
The aim of the present paper is to analyse the problems in today's EFL lexicography. Taking into consideration the fact that any dictionary is aimed at paving the path to better understanding and – at the same time – to the easier acquisition of the target language, the question that arises pertains to the kind of information a good dictionary should provide? Hence, it is important to investigate the type of information sought and, as a consequence, the way lexicographers deal with the task.
Keywords: lexicography, EFL.

Regarding the EFL student and his reference needs, it has been found that:

1) the dictionary is used mainly for decoding, traditionally understood – from Richards et al. (1985:73) – as the process/act of trying to understand the meaning of a word, phrase or sentence (Béjoint 1981, Hartmann 1983),

2) the vast majority of EFL students use dictionaries to look up meanings (Tomaszczyk 1979, Béjoint 1981, Hartmann 1983),

3) EFL students find bilingual dictionaries more useful than those of a monolingual nature (Tomaszczyk 1979),

4) the use of dictionaries by EFL students decreases as their language proficiency increases (Tomaszczyk 1979),

5) dictionaries are used more competently by the most linguistically proficient users (Tono 1991, Neubach and Cohen 1988),

6) the look-ups are mainly motivated by the following reasons: spelling and meaning to a larger extent, existence, synonymy, grammar, register, collocation and – to a lesser extent – inflection (Harvey and Yuill, 1997).

Obviously enough, the ability of the dictionary user to find the information being sought for depends on his reference skills. Researchers in the field consequently agree on two general stages of the dictionary search; prior to location of the sought word (macrostage), and after location of the sought word (microstage). It is fair to add that the process of finding the right meaning requires a complex set of processes. To be more precise, as Schofield (1982:186-193) rightly indicates, macro strategies demand the following technical skills:

1) locating the word(s) or phrase(s) which the learner does not understand,

2) recovering the canonical form[†] or inflected unknown word,

3) searching for an unknown word in the alphabetical list,

4) taking the following procedural steps if at least one main entry for the unknown cannot be found:

[†] By definition, canonical form is the form of a linguistic item, wchich is usually show as the standard form. For example, the plural morpheme in English is usually shown as -s, even though it may appear as -s, -es, -en.

a) if the unknown seems to be a set phrase, idiom or compound word, look up each element,

b) if the unknown seems to have a suffix, look up the entry for the stem,

c) if the unknown appears to be an irregularly inflected form or spelling variant, scan nearby entries,

d) if there is an addendum, search there.

Note that when the target word has been successfully located, there is a series of strategies that have to be used at a micro-level. Scholfield (1982) identifies the following ones:

1) reducing multiple senses or homographic* entries by elimination but scanning all of the definitions in the entry before making any decision about which fits the meaning that has been decoded from the context,

2) understanding the definition and integrating it into the context where the unknown was met,

3) inferring one appropriate sense that fits the context from the senses entered if none of these senses seems to fit. If more than one sense fits, seek further contextual clues in the source text to disambiguate. Obviously, many statements are ambiguous in isolation but either clear in context or are amenable to logical analysis (see McArthur 1992).

Teaching practice shows that the application of the aforementioned strategies offers a number of challenges for EFL students. Yet, EFL students can face a number of problems as well. In short, one may say that teaching experience shows that all of these problems are merely consequences of a lack of dictionary-using skills. On the other hand, there are several problems of dictionary compilation that should be considered in close connection with the needs of EFL students. The needs are the basis on which the dictionary editor must determine the type of information to include. To what extent, then, can the editor answer the questions of EFL students? One of the major questions that must be answered is: which words should be entered and how should they be treated?

Among issues that seem to call for the utmost attention is determining problems the compilers have to deal with. According to Cowie (1990:685), „[...] the learner`s dictionary has had a number of central concerns". There has been the development of controlled vocabulary that would allow the adequate, precise defining style, yet simple enough to be understood by a language learner, the provision of detailed syntactic, grammatical and inflectional information and, finally, the provision of collocational information. Firstly, commercial considerations have always played an important role in lexicography. As Hanks (2005:249) rightly points out dictionaries are involved in the „[...] pursuit of spiraling marketing claims". It is the function of the EFL dictionary to answer the questions that the user of the dictionary asks and – as a consequence – dictionaries on the commercial market will be successful in the proportion to the extent to which they can answer these questions of the buyer. Landau (1989) is commonly credited with being amongst the first to successfully draw our attention to the fact that a dictionary is a commodity, designed not only to sell but make a profit as well. As a consequence, he indicates „the manner of financing" as a criterion, according to which modern dictionaries can be classified, which is either scholarly or commercial. Whereas the former category takes years to complete – as plainly formulated by the author – „[...] commercial dictionaries are done at a much accelerated rate" (Landau 1989:11).

Taking into consideration the fact that innovation does not guarantee a subsequent commercial success, reviewers rarely undertake a detailed analysis of the content of the work, as an average user does not simply know what a good dictionary should contain. The problem is that each new edition is only cosmetically changed. Let us resort to an author who argues along the following lines:

„[...] in spite of showy graphics and ballyhooed usage notes, there have been very few meaningful changes in commercial American lexicography in the past twenty years. American dictionary publishers are afraid to take risks because of the intense competition and because, being in the main publicity owned corporations, they must show constant growth in revenue. Really innovative works almost always take years to develop, and the investment period is therefore

---

* Traditionally, by homographic entries we mean words which are written in the same way but which are pronounced differently and which may have different meanings, for example, English *lead* /li:d/, as in „This road leads to Warsaw" and *lead* /led/ in the context „His boots are as heavy as lead".

greatly protracted. Moreover, there is no guarantee that the innovative work will be an immediate success: it may take years of expensive promotion campaigns to win back the market share that might have been earned by a conventional revision" (Landau 1989:x).

Also Jackson (1988:172) claims that „[...] innovation is not possible because the public and the publishers have such a fixed idea of what a dictionary should look like, deriving from a tradition developed over centuries". In addition to this, Hartmann (2001:130) stresses that the development of lexicography is not possible without political, economic and other non-lexicographic factors. Also, the author remembers that the state of technology is what matters and boosts lexicographic tradition, too (see Landau 2001:40). On the other hand, the introduction of computer corpora has led to technology becoming greatly involved in the process of dictionary production. The quotation given below shows some of the dangers that arise at the intersection of information technology and lexicography.

„ [...] I confess to some disappointment when I learned that a first step towards the 'New Oxford English Dictionary' project was to be some market research to find out what the consumers of dictionaries want from the product [...] but I trust that they will have a through idea of what might be accomplished to supplement the predictable demands of those who will respond to the questionnaire. Too often the tendency is for the bad dictionaries to drive out the good ones, and for frequently consulted components to drive out the ones rarely used. Commercial considerations like these seem inevitably to shape – or deform – the slow evolutionary growth of our dictionaries" (Bailey1986:123-125).

There seems to have been obtained sufficient evidence that technology affects all aspects of dictionary production. All of the technological limitations must have stood in contradiction with the user's reference needs, at least sporadically. It is worth stressing that – as a result of the attractiveness of this devise for the user – it does actually facilitate editorial work.

Another issue that must be borne in mind is the cultural load in EFL dictionaries. Zgusta (1989b:3-4) stresses the importance of cultural information to pedagogical dictionary users. The significance of the problem of culture in dictionaries is beyond any conceivable doubt, as every dictionary is a snapshot of the society's life reflecting the culture (as the system of values existing in the society). Rey (1987) has attempted to characterize those features of content and organization that can be assumed to convey a cultural load. The author also discusses the issues of internal organization and the range of arrangement conventions. At the same time, Rey (1987:4) admits that the pedagogical dictionary is one of a number of dictionary types with a low cultural content and it stands in direct contrast to such dictionaries as Room's (1986) *Dictionary of Britain* or Crowther's (2000) *Oxford Guide to British and American Culture* the sole task of which is to familiarize the potential users with cultural facts related to Anglo-Saxon countries.

As follows from this short exchange there arises the question of whether Rey`s (1987) claims are open to challenge? It seems that is extremely difficult, if not utterly impossible, to find an equivalent with exactly the same meaning as the lemma as far as culture is concerned, providing that dictionaries should never fail to explain cross-cultural differences[*]. „Definitions of lexical items might, for instance, consist of two parts; a semantic paraphrase of the meaning of the word and then an additional comment of a cultural type" (Stein 2002: 140). Here, a brief illustration may help us indicate the kind of difference discussed at this point:

CCAD (1995)

---

[*] Sometimes it happens that – despite apparent correspondence between lemma and equivalent – the two may refer to different realities Hartmann (1983:122). What is more, the problem of double equivalence may appear when finding the target language equivalent.

CALD (2005)

> **suburb** Ⓐ /ˈsʌb.ɜːb/ ⓤⓢ /-ɜːb/ *noun* [C] an area on the edge of a large town or city where people who work in the town or city often live: *Box Hill is a suburb of Melbourne.* ○ *We drove from middle-class suburbs to a very poor inner-city area.*

LDCE (2005)

> **sub·urb** /ˈsʌbɜːb $ -ɜːrb/ *n* [C] an area where people live which is away from the centre of a town or city: *a London suburb* | [+of] *a suburb of Los Angeles* | *a kid from* **the suburbs** | **in a suburb** *Don't you get bored living out here in the suburbs?*

As may be concluded from the examples given above, there is no comprehensive, highly informative cultural note. The information provided does not indicate that the word „suburb" has many connotations in English. Note that it is sometimes used derogatorily to refer to a kind of middle-class way of life, socially respectable, yet definitely as dull as can be. Sometimes its use alludes to the orderliness of the neatly laid-out, semi-detached houses with front gardens that characterize many suburban areas, the connotative element that is specified in none of the dictionaries under consideration.

It seems obvious enough that most of the vocabulary is culture-specific. That means that the lexicon reflects the particular and unique way of life of its speakers. It is fair to say at the same time that – while there are degrees of culture-specificity – some items are more culture-bound than the others, there is very little in the lexicons of different languages that is truly universal (cf. Hartmann, 1983). In the words of Zgusta (1989:3): „ [...] since language is embedded in culture, cultural data are important to the learner not only for steering his linguistic behaviour but frequently for choosing the correct lexical equivalent. Such cultural information can be understood in a broad way, so that it can pertain to political and administrative realities of the country or countries whose language is being learned, and so on. Undoubtedly a good part of this information is of encyclopaedic character; be this as it may, it belongs to what the learner has to learn".

In general, however, it appears that compilers do have problems with the culture-bound words. Subsequently, it goes without saying that pronunciation labelling in learners' dictionaries poses some problems in lexicography. Non-native speakers of English expect EFL dictionaries to describe the standard language, for the purpose of communication between non-natives. Sobkowiak (2002) is of the opinion that the phonetic aspect of EFL dictionaries is „[...] among the most seriously underrated and underdeveloped in (meta)lexicography".

Along similar lines is Hulbert (1955), quoted in Landau (1991:97), who states: „[...] Dictionaries are less satisfactory in pronunciation than in spelling, meaning, or etymology. The record of the spoken language is difficult to acquire, difficult to transcribe accurately and unambiguously, difficult to represent understandably in a dictionary transcription, and in most cases of less interest to the user than other kinds of information".

Also Gimson (1973:115) stresses that „[...] Today, when no longer recorded speech as a degraded form of writing, the pronunciation entry in dictionaries [...] should be accorded much greater importance". The same author goes on to add that „[...] unfortunately, the theory is too frequently difficult to discern".

At this point it seems reasonable to dedicate more time and space to the state of the art. The OALD (2005:1540) specifies the model in the following manner: „[...] The British pronunciations given are those of younger speakers of General British. This includes RP (Received Pronunciation) and a range of similar accents which are not strongly regional. The American pronunciations chosen are also as far as possible the most general (not associated with any particular region). If there is a difference between British and American pronunciations of a word, the British one is given first, with AmE before the American pronunciation".

CALD (2005:x) seems to clarify the situation by saying that: „[...] British and American pronunciations of a word are shown after the headword. These are written using the International Phonetic Alphabet (IPA)".

LDCE (2005:xii) characterizes the language it describes as „[...] Pronunciation is shown using the International Phonetic Alphabet. If the British and American pronunciations are different, the British pronunciation is shown first and the American pronunciation has a dollar sign $ in front of it".

The latest, CCAD (1995: xxxviii) focuses on the following assumption: „[...] the basic principle underlying the suggested pronunciation is 'If you pronounce it like this, most people will understand you'. The pronunciations are therefore broadly based on the two most widely taught accents of English, RP or Received Pronunciation for British English, and GenAm or General American for American English".

Significantly, all of the big four dictionaries currently employ some versions of IPA to indicate pronunciation, which seems logical from a pedagogical point of view, as appealing to an international patent in EFL dictionaries. One may say that for the majority of learners this practice is far too demanding. It seems that the statement that no pedagogical dictionary of English would be marketable without reference to the IPA pronunciation is obvious, but – at the same time – it must be borne in every lexicographer's mind that such a system needs replacement, or at least some major supplementation. It is for precisely this reason, for learners who bring little or no literacy skills in their L1, that it is particularly difficult to take advantage of the IPA system that bears insignificant similarity to anything they read in English.

When we move further, we see that part of speech coding and grammatical information is another problematical area in lexicography. It goes without saying that grammatical information is very important for the learners of English. To pick a random example, LDCE has, for instance, traditional grammatical coding.

LDCE (2005)
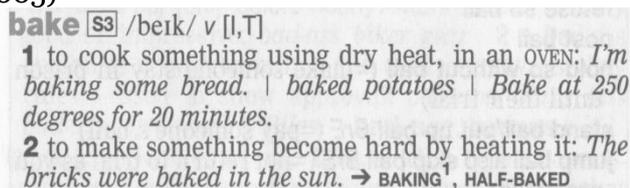
It requires the learner to thumb frantically back to find simple details. On the other hand, *CCAD* includes grammatical information in an extra column, a narrow column alongside each column of entries, defining which part of speech the particular words are.

CCAD (1995)
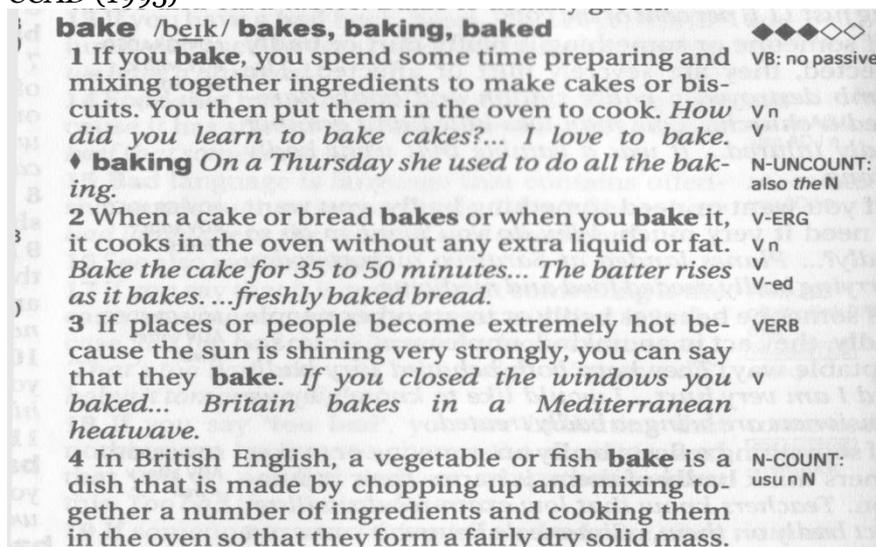
Observe that here some terminology seems rather confusing. For example, the code *N UNCOUNT* stands for uncountable noun[*]. It is fair to say that the abbreviation is cryptic or – at least – highly confusing. In any event, the majority of learners find these patterns particularly frustrating and difficult, leading to a consequent neglect of dictionary usage. Evidently, then,

---

[*] The obvious user association here would be non-countable.

learners should be exposed to such grammatical coding readily available to them, in clear-cut language or non-obscure symbols, preferably by means of adequate, well-selected illustrative material*. The use of abbreviations *per se* is not the case; obscurity, obfuscation and confusion are issues that must be avoided†.

Another problematic aspect of EFL dictionaries compilation is language phraseology. And our aim here is to cast some light on the question of how lexicographers encode the evidence of phraseological patterning. Yet, before looking *at EFL dictionaries*, it is important to consider why idiomatic information should be recorded at all. Obviously, idioms and other fixed phraseological expressions must be taken into consideration, as the primary role of a dictionary is to list and account for the lexical items of a given language‡. There is also a need to show phraseology when senses or items are restricted co-textually (for example, when verbs are followed by exact prepositions or related to particular kinds of objects).

What is more, phraseology has a purpose in clarifying sense differentiation, if the information appears as part of the definition or contained by illustrative example (it can – at the same time – clarify the definition itself). Another reason for including phraseological information is linguistic, or – to put it differently – there exists the ultimate objective to create a record of lexical behaviour as a part of an entire and incorporated description of a language. However, only very large-scale dictionary projects with unlimited funding would be in position to do this for all words. What is more, average users are unlikely to find the information useful enough to be worth the extra work, while interactive corpus/tools provide the information both more economically and effectively.

Additionally, to be classified as monolingual, a lexicographic work of reference must display the feature explained; in the words of Hartmann and James (1998:95) „[...] the words must be explained by means of the same language". No matter whether it is done by means of synonymous equivalents, a definition, antonyms in negation or a combination of these, all are relatively space consuming. That means that the space left for other information categories is scarce. As a consequence, the compiler may be forced to reduce the amount of phraseological information to the bare minimum. Another thing is that monolingual definitions are more difficult to process than native language equivalents. When dictionary consultation repeatedly involves finding the meaning relatively fast, and the students' assignments concern many new vocabulary items, such difficulties may result in the learner switching back to a bilingual dictionary.

The present period of EFL dictionaries, that is the corpus era which began with CCAD (1987), led to a special focus on corpus evidence and the typological lexico-grammatical patterns revealed§. revealed§. The truth is that within the body of EFL dictionaries one may find merely limited reference to phraseological phenomena other than collocation. Yet, from even this narrow focus, there are clearly important points to consider, apart from the quality, range and information provided. It seems that of essential importance is the function of phraseological information in relation to the needs and interests of the target users. The compiler's task here is to estimate what learners might want to know about the phraseology of an individual lemma, form or sense, as well as identifying which patterns to record.

Another aspect here is the challenge of the move from the position where the release of phraseological information is considered from the perspective of linguistic research, to the

---

* Illustrative materials are here understood as example sentences that follow the definition. They are useful as they provide extra denotative and connotative information, what is more they can convey or reinforce grammatical information by exemplifying its syntactic behaviour.

† The grey area refers also to the pattern of indication of the inflected forms. There is a constant need for clearly indicating irregularly formed words; additionally non-transparent inflected forms of a main entry should be spelled out.

‡ Phraseology is a domain of linguistic study which illustrates the correlation between language and culture. An important reason why cultural information of this kind should be included in an account of EFL dictionaries concerns the needs of lexicography today. For the practical purposes of dictionary making, cultural markedness must be taken into account.

§ Landau (1989) presents the lexicographical methodology, Klotz (1999), Mittmann (1999) the methodical comparisons of the treatment of complementation and collocation in the 1990s editions of *EFL* dictionaries. Siepmann (2006) discusses the presentation of collocational information in dictionaries. Recently, Osuchowska (2007) examines the collocational value of *EFL* dictionaries in a great detail.

situation when the needs of the user become the primary objective. It appears that particularly crucial is the function of phraseological information in relation to the needs and interests of the EFL students. The lexicographer's task here is to second-guess what users might want to know about the phraseology of individual lemma, form, or sense, as well as identifying which patterns to record.

Here, the discussion concerning electronic lexicographic products inevitably emerges. Of course, the challenge here has been to move from the position where the retrieval and delivery of phraseological information is designed from the perspective of linguistic research, including the provision of data for lexicography, to one where the users' needs are prioritised. Yet, it seems arguably even more difficult to identify what these are than in the case of traditional printed dictionaries. In the past, dictionaries simply provided raw corpus data, encouraging users to work empirically, observing patterns for themselves. Nevertheless, there are disadvantages, including time factors, and difficulties with interpreting the evidence found. Furthermore, extensive corpora are too large to be used effectively; small corpora are subject to skewing from constituent texts especially relevant where phraseological patterning varies according to genre.

It seems obvious enough that tools should be dynamic and provide filtered data, organized in terms of significance, word class, syntagmatic positioning, genre and meaning, but overly filtered data may be misleading and may become underinformative entries in printed dictionaries at the same time. The major conclusion that seems to be emerging from the above considerations is that lexicography – although the science has been recently developing at an unprecedented pace – still suffers from numerous problematic issues. It sets up a number of indispensable requirements which any lexicographical description is to observe if it is hoped to be somehow satisfactory.

Abbreviations:
CALD – Cambridge Advanced Learner's Dictionary
CCALD – Collins Cobuild Advanced Learner's Dictionary
LDCE – Longman Dictionary of Contemporary English
OALD – Oxford Advanced Learner's Dictionary

References:
Dictionaries:
1. Collins Cobuild Advanced Dictionary 6th edition. (2009) New York: Harper Collins.
2. Cambridge Advanced Learner`s Dictionary 3rd edition. (2008) Cambridge: Cambridge University Press.
3. Longman Dictionary of Contemporary English. (2005) London: Pearson Education Limited.
4. Oxford Advanced Learner`s Dictionary. (2005) Oxford: Oxford Univeristy Press.

Other sources:
5. Bailey, R.W. (1986) 'Dictionaries of the next century', [in:] Ilson, R.F. (ed.), pp. 123-137.
6. Béjoint, H. (1981) 'The foreign student's use of monolingual English dictionaries: A study of language needs and reference skills', [in:] *Applied Linguistics*, vol. 2.3, pp. 207-215.
7. Cowie, A.P. (1990) 'Language as words: Lexicography', [in:] Collinge, N.E. (ed.) *An Encyclopaedia of Language*. London: Routledge, pp. 671-700.
8. Hanks, P. (2005) 'Johnson and modern lexicography', [in:] *International Journal of Lexicography*, vol. 18.2, pp. 243-266.
9. Hartmann, R.R.K. (1983). 'The bilingual learner's dictionary and its uses', [in:] *Multilingua*, vol. 2 (4), pp. 195-201.
10. Hartmann, R.R.K. (1991). 'The bilingual learner's dictionary and its uses', [in:] *Multilingua*, vol. 2 (4), pp. 195-201.
11. Hartmann, R.R.K. and James, G. (1989) *Dictionary of Lexicography*. London: Routledge.
12. Harvey, K. and Yuill, D. (1997). 'A study of the use of a monolingual pedagogical dictionary by learners of English engaged in writing', [in:] *Applied Linguistics*, vol. 18 (3), pp. 253-273.

13.   Jackson, H. (1988) *Words and their Meaning*. London: Longman.

14.   Klotz, M. (1999) 'Word complementation in English learners` dictionaries – A quantitative study of CIDE, COBUILD2, LDOCE3 and OALD5', [in:] Herbst, T. and Popp, K. (eds.), *The Perfect Learners` Dictionary (?)*. Tübingen: Niemeyer, pp. 33-43.

15.   Landau, S.I. (1989) *Dictionaries: The Art and Craft of Lexicography*. Cambridge: Cambridge University Press.

16.   Mittmann, B. (1999) 'The treatment of collocations in OALD5, LDOCE3, COBUILD2 and CIDE', [in:] Herbst, T. and Popp, K. (eds.) *The Perfect Learners` Dictionary (?)*, Tübingen: Niemeyer, pp. 101-111.

17.   McArthur, T. (1992) *The Oxford Companion to the English Language.* Oxford: Oxford University Press.

18.   Neubach, A. and Cohen, A.D. (1988) 'Processing strategies and problems encountered in the use of dictionaries', [in:] *Dictionaries: Journal of the Dictionary Society of North America*, vol.10, pp. 1-19.

19.   Osuchowska, D. (2007) In search of the collocational value of EFL dictionaries. Rzeszów: Wydawnictwo Uniwersytetu Rzeszowskiego.

20.   Rey, A. (1987) `Le dictionnaire culturel`, [in:] *Lexicographica*, vol. 3, pp. 3-50.

21.   Scholfield, P. (1982) 'Using the dictionary for comprehension', [in:] *TESOL Quarterly*, vol.16 (2), pp. 185-194.

22.   Siepmann, D. (2006) 'Collocation, Colligation and Encoding Dictionaries. Part II: Lexicographical Aspects', in: *International Journal of Lexicography* 19, pp. 1-39.

23.   Stein, G. (2002) *Better Words: Evaluating EFL Dictionaries*. Exeter: Short Run Press Ltd.

24.   Tono, Y. (1991) 'A good dictionary user: what makes the difference?', [in:] Ito, K. *et al.* (eds.) *Recent studies on English language teaching*. Tokyo: Yumil Press, pp. 229-253.

25.   Tomaszczyk, J. (1979) 'Dictionaries: Users and uses', [in:] *Glottodidactica,* vol. 12, pp. 103-119.

26.   Włodarczyk-Stachurska, A. (2010) *On sociolinguistic variation: the case of English women terms*. Radom: Wydawnictwo Politechniki Radomskiej.

27.   Zgusta, L. (1989a) 'The role of dictionaries in the genesis and development of the standard', [in:] F.J. Hausemann *et al.* (eds) *Wörterbücher. Dictionarics. Dictionaries. An International Encyclopedia of Lexicography*. Berlin: De Gruyter, pp. 70-79.

28. Zgusta, L. (1989b) 'Idle thoughts on an idle fellow; or, vaticinations on the learners` dictionary`, [in:] M.L. Tickoo (ed.) *Learners` Dictionaries: State of the Art*. Singapore: SEAMEO Regional Language Centre, pp. 1-9.