



UDC 81'25

Original scientific paper

Accepted for publication on 26th September 2003

Combining dictionary and corpus evidence - a lexical study of the English adjective *appropriate* and its translation equivalents in Croatian

Boris Pritchard,
Faculty of Maritime Studies
University of Rijeka

Evidence obtainable from dictionaries and textual corpora can be a very useful tool in applied linguistics, in particular bilingual lexicography and translation. The work deals with dictionaries and corpus material as the two major sources of information in lexicological and translation studies. The English adjective 'appropriate' (and other lexical units belonging to the same lexical set) has been chosen to exemplify the use of traditional and modern tools in discriminating word senses and uses. In addition to general language usage, the paper also studies corpus evidence derived from special-purpose texts (institutional, technical, etc.), mainly restricted by subject-matter or genre-related features, in order to draw parallels that might prove useful in bilingual lexicography and in the process of translation.

1. Introduction

Lexical studies traditionally rely on dictionaries as a source of information. More recently dictionary information has been combined and verified by means of information derived from text corpora.

Lexicography and translation have been in the focus of computational linguistics since its beginnings three decades ago. Computer experts, translators and linguists, in particular those engaged in empirical and applied linguistic studies, have provided and continuously improved both the theoretical framework and practical computer appliances capable of dealing with various aspects of lexical studies and machine translation. Numerous corpus projects have made possible the research in natural language usage. Language corpora today offer the possibility of retrieving extensive information (occurrence, frequency, co-text and context, collocational picture, measures for mutual expectancy or information, etc.) on any given search word within seconds.



Since the beginnings of the computer era concordances have been employed as a useful tool in applied linguistics (e.g. CALL, concordance-aided learning; natural language processing, etc.) as well as in the linguistic description of language (syntactic distribution and lexical phenomena such as collocation, disambiguation, sense discrimination etc.).

Computer-aided translation using bilingual and multilingual concordances obtained from parallel (translation) and comparable corpora are the most recent development. They appear in the form of matching lines of a node (word or phrase) in SL and TL. Parallel concordances are also obtainable on-line and retrievable from e.g. Multiconcord for Windows (King-Wools 1997; web.bham.ac.uk/johnstf/l_text.htm), the Chemnitz English-German Translation Corpus (<http://dict.tu-chemnitz.de/phil/InternetGrammar>), English-Norwegian Parallel Corpus, etc.

This work principally deals with the study of textual corpora and using concordances (a) as a check on lexicographic description in dictionaries and (b) as one of the possible means of disambiguating and discriminating senses of a word or lexical set in the decoding process of translation. The work attempts to show whether corpus evidence can be used in studying the minute differences (shades of meaning) of the adjective *appropriate* and its co-hyponyms forming part of the lexical set lexicalizing the concept of 'the various degrees of adaptableness to a particular use, purpose, or situation'.

Studying corpora, in particular the results of statistical measures and concordances, is valuable in the selection of headwords and their sub-entries and contextualising the meaning and usage of dictionary entries (i.e. definitions and usage notes within them) to disambiguate the various senses of the word (*appropriate*) or the lexical set this word belongs to (*appropriate, suitable, correct, right, proper, fit, fitted, fitting, adapted* etc.). The corpus-based analysis will serve to make an attempt at determining the various uses of these words and their most appropriate translation equivalents (TE) in Croatian.

While a great number of monolingual English corpora is available today (cf. COBUILD, BNC, LOB, etc.), there is only one comprehensive Croatian corpus of a larger size or significance (the 30 million word Croatian National Corpus). Its design and representative character, however, ensures plausible computational insights into the Croatian language and offers the possibility of obtaining Croatian concordances.

Although there is a variety of bilingual corpora available today on the internet or commercially, apart from the Croatian-Slovene Parallel Corpus and the Croatian-English Parallel Corpus (under construction, cf. Tadić 2000) there exist no sizeable bilingual corpora involving Croatian language except for some individually compiled parallel texts or files from which no statistically relevant inferences can be made. Nevertheless, the author's corpus of a number of parallel maritime institutional and technical texts has been used to study the contrastive potential of bilingual concordances (English-Croatian) in order to determine the most appropriate translation equivalents or the degree of substitutability among the members of the lexical set above.



2. Dictionary entries

2.1 For the purpose of this work, i.e. the study of the English adjective *appropriate* and other items of the same lexical set (*acceptable, correct, fitting, proper, right, suitable*), English monolingual dictionaries may be divided according to the degree they exemplify the various senses and uses of headwords. This places desk and encyclopaedic dictionaries (OED, W3, RHD) on the lower end of such a scale, the best known learners' dictionaries (LDOCE, OALD, CIDE) occupying positions from central to higher levels, whereas the COBUILD English Dictionary ranks highest on the scale. For comparison, here are some exemplifications of *appropriate* for each class of the dictionaries:

(a) 1. specially suitable: FIT, PROPER: *sit down anywhere and the appropriate waiter comes up; gift packages are likewise appropriate for the girls you regularly remember; by any means appropriate to our use* (W3)

(b) correct or suitable for a particular time, situation, or purpose: *At an appropriate moment I'll offer the visitors some coffee. Your clothes are hardly appropriate for a job interview. Objectives and strategies which are appropriate to the markets.* (LDOCE)

I haven't got anything appropriate to wear. (CIDE)

(c) 3. *The institutional structure appropriate to each country ...*

a general test for English appropriate for the particular level.

It seemed appropriate to end with a joke. (COBUILD)

2.2 The study of definitions for the word *appropriate* in the dictionaries consulted (i.e. finding the meaning that is "central or common to the set of words being discriminated" (Di Marco - Hirst 1993), or the most neutral, usual and widest sense, shows that these dictionaries mostly agree:

(a) in the choice of synonyms or near-synonyms used to define the word, i.e. <correct>, <suitable>, <acceptable>, <proper>, which at the same time represent the main semantic features of the entry, and

(b) in determining the class of nouns such adjectives qualify: '*purpose*', '*occasion*', '*situation*'

Although a certain degree of circularity can be noted in the explanatory vocabulary used in the definition part of the entries for the set studied, the above synonyms are not always interchangeable. However, such defining vocabulary can be used as evidence of (near)synonymy or substitutability, primarily on the syntactic but also on the semantic level. This is revealed in tables No. 1 and No. 2 showing the adjective *appropriate* and its possible semantic substitutes obtainable from the definitions of the word (i.e. expressed in the form of both defining vocabulary and synonyms) across a number of desk (Table No. 1) and learners' dictionaries (Table No. 2):



Table No. 1

<i>APPROPRIATE</i>	W3 ⁱ	EWD	OED	CEDT	RHD
CORRECT	-	-	-	-	-
FITTING	+	+	+	+	+
PROPER	+		+	-	-
RIGHT	-	-	-	+	-
SUITABLE	+	+	+	+	+

Table No. 2

<i>APPROPRIATE</i>	OALD	LDOCE	CIDE	COBUILD	WNW
CORRECT	-	+	-	+	-
FIT(TING)	-	-	-	-	+
PROPER	+	-	-	-	+
RIGHT	+	-	+	+	+
SUITABLE	+	+	+	+	+

The two tables show that, in all the above dictionaries, the word *suitable* (used as a part of the defining vocabulary, but not as the defining term expressing ‘genus proximum’) is substitutable for *appropriate*. *Right* (in only one of its numerous senses), is the second-ranked substitute, particularly in learners’ dictionaries, probably because of more stringent rules on the choice of defining vocabulary in such dictionaries. Certain words in the tables (cf. *proper*, *correct*) are not invariably used as substitutes but have been chosen because of a high degree of their substitutability for the members in the set other than *appropriate*.¹ Though it may seem that ‘appropriate’ in the above use is replaceable with its synonymous counterparts, the substitution test applied to the examples given in some of the above entries for *appropriate* does not always seem to prove satisfactory, cf.:

I didn't think his comments were very appropriate at the time.

~*suitable*²

~*fitting*

+*acceptable*

**proper*

**right*

+ *correct*

¹ The listed co-hyponyms of *appropriate* are: **right, proper, correct, acceptable, approved.**

² ~ (substitutable for)

+ (additional meaning)



The teacher can then take appropriate action.
 ~suitable
 proper
 ?right
 +correct
 +acceptable
 *fitting

It should also be added that the adverbs (e.g. ‘specially’, ‘particularly’) denoting the extent to which a *situation/purpose/occasion* is ‘suitable’ or ‘appropriate’, significantly participate in assigning the distinctive semantic feature to the lexical set by restricting the notion of ‘suitability’ to a specific ‘situation/purpose/occasion, or (in terms of the definition of *appropriate* in LDOCE) by “exactly answering what is suitable”. Such adverbs should therefore be included as obligatory part of the lexicographic definition of the adjective *appropriate* as they are the most prominent lexical expression of the prominent distinctive semantic feature for this word.

A number of the dictionaries mentioned above also emphasize some pragmatic features of the use of these lexical units (e.g. connotation) as well as genre and register, because these also represent important elements for disambiguation or discrimination of senses; cf. explanatory/usage notes or comments such as ‘socially acceptable’, ‘in conformity with the social ethics’ (e.g. OED for *proper*), ‘socially suitable’ (OALD for *right*), ‘acceptable’ (COBUILD for *socially appropriate*) (Longman Activator for *proper*).³

2.3 In bilingual English-Croatian dictionaries the following TE’s for *appropriate* are listed:

- *primjeren, shodan, and zgodan* (Lochmer 1906, the first English-Croatian dictionary)
- *prikladan, zgodan, primjeren, odgovarajući* (Filipović 1996).
- *podesan, uputan; primjeren; doličan; odgovarajući* (Bujas 1999a).

Primjeren is the only translation that appears in all the dictionaries, though the modern dictionaries show a higher degree of match (*prikladan, primjeren, odgovarajući*). The above dictionaries also fail to agree on the sequencing of translation equivalents and their semantic subdivision. It is strange that *odgovarajući*, the most frequent translation equivalent in Croatian, is listed last. This can be justified by formal reasons (rules of word formation and restricted syntactic use in Croatian, i.e. attributive use only of the present participle *odgovarajući*, cf. Ivir 1983:18), and by the relative negligence of the frequency of use in traditional dictionaries.

³ Other senses of *appropriate* (such as ‘belonging to = own, as in ‘they had their *appropriate* methods, CEDT) are not dealt with here.



However, back-translation of the English TE for Croatian headwords found in Croatian-English dictionaries, yields the following counterparts on the English side of Croatian-English dictionaries (Bujas 1999b):

- **podesan** cross-referenced to *prikladan*
- **uputan** = *advisable, expedient, opportune*
- **primjeren** = *suitable, adequate, appropriate, proper, due*
- **doličan** = *fitting, befitting, becoming, seemly*, Lit. *beseeming; proper, appropriate, suitable*
- **odgovarajući** = *corresponding, equivalent, comparable, matching*; and (in the same lexical entry) for the superordinate <prikladan>: *appropriate, suitable, pertinent, adequate; relevant*; with exemplifications such as *u odovarajuće vrijeme* (at the appropriate time), *u odgovarajućim okolnostima* (in appropriate / convenient circumstances)

All the translation equivalents (except *uputan*) contain the word *appropriate* and can therefore also be considered as formal correspondents. The translation equivalents not found in Bujas 1999b show the following back-translations:

- **prikladan** = *suitable, (well) suited (to), lending itself to, serviceable; convenient; fitting, appropriate, proper; adequate*
- **zgodan** = (<prikladan>) *convenient, opportune, suitable, appropriate, proper, fitting, apt*
- **pravi** = *just, right, proper; fit; pure, genuine; correct, accurate.*

In Benson (1989) *appropriate* is listed among the translation equivalents for the Croatian adjectives *pravi* (*real, actual, genuine, appropriate, right*), *odgovarajući* (*1 corresponding, 2 suitable, appropriate*: e.g. *appropriate equipment, appropriate measure*). *Appropriate* is also listed for the adjectives *prikladan* (*appropriate to the situation*), and *zgodan* (*1 suitable, appropriate; convenient*). There is however no headword *prim(j)eren* in the wordlist of this dictionary.

3. A corpus-assisted study of the adjective *appropriate* and its Croatian translation equivalents

3.1 One of the advantages of corpus-based studies is that, within the source material, the corpus shows 'all the contexts in which a word occurs' (Biber et al. 1998: 26). Monolingual corpora (English and Croatian) can be used in the first phase of contrasting the adjective *appropriate* and its equivalents in Croatian.

In the lexical study of English as a source language text, the use of a monolingual corpus can best be made in three stages: (a) by producing collocational pictures of a given word, (b) by retrieving concordances for that key-word, (c) evaluating corpus-driven material.



As shown above, common to all instances of *appropriate* is the fact that it is used as a word qualifying the class of nouns lexicalizing some of the most basic concepts, i.e. SITUATION, PURPOSE, TIME. It appears from the examples above that the differences in the usage among the members of the lexical set under study may be attributable to field or subject-matter, or even to sociolinguistic or pragmatic features of the text (in e.g. *proper*, *correct*, *acceptable*, *fitting*). The corpus studied, however, does not entirely confirm such minute differences of usage and suggests interchangeability of at least some of the above terms (e.g. *appropriate* vs. *suitable*, *appropriate* vs. *proper*) in similar contexts. On the other hand, the substitutability test does not even allow for complete substitution in the same or similar contexts. This, of course, causes problems in deciding on the right (appropriate!) sense of the word *appropriate* and then, to choose the most “appropriate” TE for the adjective *appropriate*. The appropriacy problem is central to the encoding stages of the translation process, especially if decoding allows for different interpretations as shown above. It may, therefore, be stated that the members of the set are substitutable, to a certain extent, for *appropriate* and *suitable* reaching the highest degree of substitutability.

To help solve the problem, pictures of collocates for the adjective *appropriate* obtained from the COBUILD Bank of English (t-scores, MI-scores, frequency ratios)⁴ may prove useful before studying concordances, because by applying them we can restrict the selection of concordances (lines matching the node item) to the more relevant ones and thus save a lot of time.

In the case of adjectives such as *appropriate* and other members of the lexical set studied, the decoding translation procedure principally requires (a) searching for the nouns that are most likely to be qualified by such adjectives (the pattern ADJ + N), (b) looking up the instances of predicative uses of these adjectives (linking verb + ADJ (+ to/for)), and (c) finding concordances for elliptic nominal predicates including such adjectives (i.e. a formulaic pattern of predicative use in: AS/IF/WHEN/WHERE + ADJ)⁵. This work mainly deals with the first pattern.

It is normally recognized that the lexical study (syntactic, semantic, and pragmatic) of a noun qualified by an adjective can be used to arrive at possible clues for disambiguating or discriminating senses of the adjective. This is why, for example, the study of the collocational meaning of a word is “crucial at the first stage of translation” (Baker 1992: 53). The analysis of the COBUILD corpus comparing statistically relevant probability of occurrence versus chance of occurrence) reveals the following nouns (shown in *italic*) to be typically qualified by the set of adjectives under study, i.e. :

⁴ i.e. scores or measurements obtained by some of the most frequent mathematical procedures practised in modern lexicography relying on databases

⁵ It is obvious that a tagger and parser must be previously used to discriminate between the verbal and adjectival form (part of speech) of a word, e.g. *appropriate*.



Table No. 3 ADJ (*appropriate*) + N
(picture by t-score)

it	is	an	APPROPRIATE	for	for	the
would	it	the	APPROPRIATE	to	to	be
is	be	more	APPROPRIATE	<i>time</i>	of	should
be	not	be	APPROPRIATE	<i>action</i>	<p>	be
may	would	is	APPROPRIATE	<i>response</i>	the	this
that	at	most	APPROPRIATE	that	this	your
might	was	where	APPROPRIATE	<i>way</i>	your	particular
think	with	not	APPROPRIATE	<p>	if	action
whether	to	was	APPROPRIATE	<i>level</i>	relevant	age
to	may	are	APPROPRIATE	<i>treatment</i>	and	each

Examples:

Tory voters. It is a particularly appropriate time for this book to appear since
 r General of the BBC, it seems an appropriate time to compile a similar list of
 ct. However she added that at the appropriate time, Mrs Marcos would have to st

te procedures has been and appropriate way of linking 'practitioners' and 'academics'
 Yeah.<S> that would be an appropriate way to go.<S> Yeah. Yeah.<S> So and also
 I'm not sure if that's an appropriate way of describing this work erm it's the

Other nouns (from the rest of the same table) which the adjective *appropriate* significantly modifies are *box, measures, technology, place, moment, steps, use, word, amount, authorities, sentence, setting, responses, course, ways, skills, form, training, size, balance, method, forum, manner, title, name, words, means, choice, care, cases, number, venue, circumstance, times, resources*, e.g.:

and it was met instantly with an appropriate response." Last night Mr Major
 ty - Democrats say the bill is an appropriate response to the changing world en
 the United States would make the appropriate response if the threat was carri⁶

Note that in the COBUILD Bank of English *appropriate* occurs 15,601 times and that it is most frequent in the BREPHEM sub-corpus (British ephemera: miscellaneous texts from every-day life, ranging from instructions on medical prescriptions, via administrative advice to warnings in city traffic, etc.), with the ratio of 168.6 occurrences per each million words. The word is also highly frequent in the USEPHEM sub-corpus (92.4), and the New Scientist Journal (90.4). It is least frequent in spoken English (28.0) i.e. in transcribed BBC transmissions (28.0 per million).

⁶ Concordances chosen from the COBUILD English Collocations on CD-ROM, 1997



Furthermore, some measurements in corpus linguistics can be particularly useful. Among those *t*-score is most valuable, i.e. the probability or likelihood that a word (e.g. noun) will co-occur with a node word (e.g. adjective), as such scores indicate not only the most typical but also the most likely collocates. Thus, under the *t*-score measurement, the nouns such as *level*, *measures*, *age*, *technology*, *circumstances*, *form*, *manner*, *care*, *way* are most likely to co-occur with *appropriate* as their pre-modifier. Combined with the other nouns shown above, a number of semantic classes of nouns can be expected either to follow (attributive use) or precede *appropriate* (predicative use). These include the abstract, “situation” nouns (*circumstances*, *conditions*), nouns expressing degree, intensity or measurement (*level*, *measures*, *amount*, *moment*, *age*, *size*, *times*), concrete nouns (*form*, *box*), nouns related to public administration (*forum*, *authorities*), or abstract nouns referring to human behaviour (*manner*, *care*, *steps*, *responses*), place nouns (*place*, *venue*), words referring to education and training (*skills*, *course*, *training*) and, typically, “clothes/wear” nouns (*suit*, *outfit*, *dress*, *footwear*, etc.). Also very important for the translator is the semi-lexical or sometimes delexical group of nouns used to describe ways or methods of doing something or of describing paronymic (part-whole; Cruse 1986) relationships and other taxonomies (*ways*, *size*, *means*, *use*, *times*).

Among nouns, the word *level* has the highest *t*-score (9.95), followed by the plural noun *measures* (9.75), *age* (9.26), *technology* (8.39), *circumstances* (8.37), *form* (7.74), *box* (7.53), *manner* (7.40), *care* (7.37), *skills* (7.22), etc. On the other hand, the most likely linking verbs taking *appropriate* as nominal predicate are: *seem* (9.37) and *was* (8.01), with the latter showing much higher frequency (1215 matching lines) than the former (123). Other verbs forming the same syntactic pattern are *be*, *is*, *are*, *seems*, *seemed*, *consider*, *deemed*, *find*, *thought*. The most likely adverbial collocates of *appropriate* are *more* (25.19), *most* (18.15) and *not* (16.20), followed by *entirely* (*t*-score 10.79), *particularly* (7.20), *only*, *less*, *quite*, *perfectly*, *highly*, *especially*, *culturally*, *wholly*, *hardly*, etc. It is also interesting to show how *appropriate* behaves as compared statistically to other adjectives of the same lexical set occurring in the pattern: ADJ + N (picture by *t*-score):



Table No. 4

Noun collocates shared by *appropriate* and other adjectives of the lexical set⁷

appropriate	acceptable	correct	proper	right	suitable
TIME (318)	10	43	99	147	30
WAY (216)	83	7	183	1100 ⁸	17
ACTION (150)	0	0	6	35	0
RESPONSE (113)	4	4	32	15	11
LEVEL (84)	86	28	20	120	6
MEASURE (78)	1	4	12	15	3
PLACE (70)	6	38	198	1042	63
STEPS (48)	0	1	15	21	1
USE (55)	3	0	109	20	2
WORD (42)	3	36	9	290	9

The table above may therefore offer the lexicographer and the translator the noun-collocates that are statistically most likely to be qualified by *appropriate* and its possible substitutes from the set studied, cf.:

- appropriate / right / proper⁹ way
- appropriate / right / proper / correct / suitable time
- appropriate / right / proper / suitable place
- appropriate / right / acceptable / correct / proper level
- appropriate / correct / suitable word, etc.

T-score information is a valuable source of lexical choice, especially in analysing the SL text, before starting the encoding phase of translation. Together with absolute frequency, it is also useful to foreign learners, as they may not be aware of the most natural or appropriate collocates. Table No. 4 can be used to prepare a list of TE in the target language for any ADJ + N combination, fitting a particular topic or subject-matter.

On the other hand, MI-score¹⁰, a statistical measure indicating strength of a collocation, i.e. the non-randomness present when two words co-occur (Hunston 2002: 71), may prove even more valuable to the translator as MI-scores reveal uncommon but strong (lexical) patterns (Stubbs 1996), cf. the following collocates of *appropriate* in the Bank of English:

⁷ Figures in the brackets show frequency order against the frequency of *appropriate*.

⁸ Extremely high frequency caused by the highly polysemous nature of *right*.

⁹ Adjectives are listed in ascending order.

¹⁰ Mutual Information - a measure that shows how strongly two words are related, i.e. the likelihood with which the presence of one word indicates the presence and co-occurrence of another word.



Table No. 5 - *appropriate* (Picture by MI-score):

deems	APPROPRIATE	safeguards
deem	APPROPRIATE	epiphet
culturally	APPROPRIATE	modification
aesthetically	APPROPRIATE	headings
entirely	APPROPRIATE	remedy
wholly	APPROPRIATE	approvals

Any MI-score higher than 3 significant (e.g. *deem appropriate*, *culturally appropriate*, *appropriate safeguards*) can be taken to be significant and therefore more informative than other, simpler statistical measures. After this the concordance lines can be used to reveal the most central and typical syntactic patterns and the associated meanings (attributive vs predicative use, prepositional phrases *appropriate to*, *appropriate for*) of the collocates, to observe meaning distinctions (types of nouns used as headwords), observing the interdependence of meaning and pattern, phraseology (*as / when appropriate*), etc.

3.2 Croatian National Corpus: Croatian TE candidates for the adjective *appropriate*

Croatian National Corpus (www.hnk.ffzg.hr) has been used in order to study the syntactic and semantic properties of possible translation equivalents for *appropriate* and other members of the respective lexical set. The corpus shows that in Croatian there is a larger number of possible candidates for the concept and typical noun headword collocates for the adjective *appropriate* than usually listed in bilingual dictionaries (e.g. *adekvatan*, *nadležan*). Furthermore, the concordances for *odgovarajući*, *adekvatan*, *primjeren*, *nadležan* largely outnumber those for *prikladan*, *podesan*, *doličan*, which are usually listed among the first equivalents in the dictionaries studied. The typical syntactic structure shared by the two languages is 'adj + N' for all the members of the *appropriate* set (adj + *way, time, place, level, word*, etc.), thus corresponding in Croatian to *odgovarajući / adekvatan / primjeren + način, vrijeme, mjesto, mjere, pomoć, tijelo, organ*, etc.). The structure 'appropriate to + N' corresponds to 'adj + N_{dative}' in Croatian (*primjeren / prikladan potrebama, okolnostima, uvjetima, situaciji, vremenu, stilu, tradiciji*). This use excludes the adjectives *odgovarajući* and *nadležan* and is rare with *adekvatan*. However significant, predicative use of the adjectives studied is much less frequent than the attributive one (*primjerena je i usporedba s, suradnja primjerena zajedničkim zadaćama*). In this corpus the typical sets of nouns qualified by the adjectives above are as follows:

Table No. 6

1a prikladan	<i>način, položaj, oblik; sustav; sadržaj, oprema, prostor, vrijeme</i> (used both ATTRIBUTIVELY and PREDICATIVELY)
1b primjeren	<i>način, odnos, odgovor; izraz, rješenja; naknada, plaća; dar; kazna; terapija, zaštita</i> (ATTRIB. & PREDIC.)
1c pogodan	<i>trenutak; način, oblik; osoba; instrument; mjesto, prostor, sredina</i> (ATTRIB.) <i>terapija, tehnika, materijal, tvari, spojevi, sastav</i> (PREDIC.)
2a odgovarajući	<i>mjere, rješenja, zaključak, odluka, uvjeti; tijela, institucije, zakoni, zakonodavstvo, vijeće; zaštita, pomoć; program, mjesto</i> (ATTRIBUTIVE ONLY)
2b adekvatan	<i>način, zaštita, kontrola; uvjeti, stavovi; mjesto; pomoć, zamjena, cijena, usluge; primjeri</i> (ATTRIB. & PREDIC.)
3a zgodan	<i>prilika, situacija</i> (ATTRIB. & PREDIC.)
3b podesan	<i>oblik, primjer, institucija, prostorija, oprema</i>
3c doličan	<i>stan; način</i>
4a ispravan	<i>dijagnoza; odnos, stavovi, politika; odluka, odgovor; način, put; instalacija</i> (ATTRIB. & PREDIC.)
5b pravilan	<i>način, odgovor; izbor; zaštita; razvoj; odnos, stav, odluka, zaključak</i> (ATTRIB) <i>tretman, djelatnost; način, rješenje; instrument, sredstvo, predmet; mjesto, prilike, vrijeme, prostor; predio</i> (PREDIC.)
6 nadležan	<i>tijelo, organ, ministarstvo, vlast, služba, sud, institucija, sektor; odbor, uprava, odjel; činovnik, službenik, sudac</i> (MOSTLY PREDICATIVE)

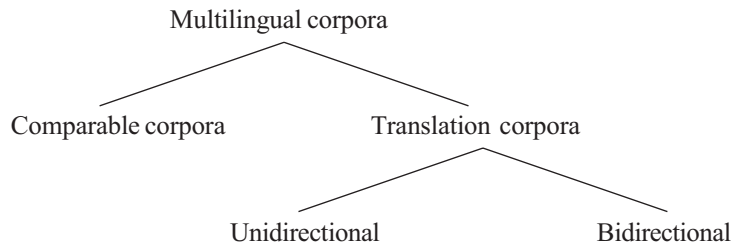
The above list also represents a systematic view of possible Croatian translation equivalents for the various senses of the English adjective *appropriate* and its lexical set. There is a certain degree of correspondence between the Bank of English and the Croatian National Corpus in the selection of some noun headwords qualified by *appropriate* (*way* – *način*, *measures* – *mjere*, *place* – *mjesto*), but a great number of ‘adj + noun’ combinations is language-specific, and these are probably more significant to the lexicographer/translator. Also interesting and useful is a list of nouns that the predicative adjective *primjeren* takes as its subject and dative complement (e.g. *način primjeren okolnostima* = *a method appropriate to the circumstances*): subject (*način, sustav, odgoj, model, naknada, liječenje*) + *primjeren* + dative complement (*vremenu, stilu, svom dobu, prilikama, trenutku, uzrastu, tadašnjim sredstvima*). The dative complement is invariably an expression of time. Similar structure and lexical distribution is also shared by the adjectives *prikladan*, *pogodan*, *adekvatan*). These lexical items and structures can therefore be used as the most likely candidates for selecting translation equivalents in an English-



Croatian dictionary and as headwords and sub-entries in a Croatian-English dictionary.

3.3 *Parallel corpora*

Parallel or translation corpora is another source of lexical information. The development of corpora, particularly the multilingual ones, has brought about the revival of contrastive studies, now frequently termed contrastive linguistics (Altenberg & Granger 2002). Johansson (1998:4-7) has set out a typology of multilingual corpora as shown in the table below:



Though monolingual, the COBUILD Collins Bank of English and the Croatian National Corpus can be taken as an example of comparable corpora on account of their representativeness. For advantages and disadvantages of parallel or translation corpora see Teubert 1996. Recent trends in cross-linguistic studies focus heavily on lexis, contrastive lexical semantics and translation equivalence (Salkie 2002) via bilingual lexicography, parallel corpora and, predominantly, parallel concordancing.

Salkie (2002) argues that translation corpora offer a new perspective on the old issue of translation equivalence. In a monolingual corpus the translator/lexicographer is interested in the patterns which ‘usually involve a phenomenon occurring more frequently than expected’ whereas in a translation corpus (bilingual or multilingual) the interesting cases are those ‘where the correspondence is less frequent than anticipated’ (cf. also Chesterman 1998).

For illustration purposes, the results for the query ‘appropriate’ from the English-German Dictionary (dict.tu-chemnitz) and from the on-line bi-directional English-German Parallel Corpus (cf. Chemnitz InternetGrammar, dict.tu-chemnitz.de/phil/InternetGrammar) are shown:



Entry *appropriate* in the corpus-based English-German Dictionary
(University of Chemnitz: <http://dict.tu-chemnitz>)

Deutsch	English
14 Ergebnisse	14 results
<u>angemessen</u> ; <u>angebracht</u> ; <u>entsprechend</u> ; <u>dementsprechend</u> {adj}	appropriate
<u>passend</u> ; <u>geeignet</u> ; <u>richtig</u> {adj}	appropriate
<u>zutreffend</u> ; <u>zugehörig</u> {adj}	appropriate
<u>artgerecht</u>	appropriate to the species
<u>eigen</u> ; <u>zugehörend</u>	appropriate (to)
<u>geeignete Schritte</u>	appropriate action
<u>Rücklagen auflösen</u>	to appropriate reserves
http://dict.tu-chemnitz.de/dings.cgi? o=3001&service=de-en&query=sich <u>sich etw. aneignen</u> ; <u>sich bemächtigen</u>	to appropriate
<u>gegebenenfalls</u> {adv}	where appropriate
<u>gegen angemessenes Honorar</u>	for an appropriate fee
<u>verwenden</u> ; <u>bereitstellen</u> ; <u>bewilligen</u> (Geld) (für)	to appropriate (for)
Ich halte es für angebracht, ... - : <u>Angehörige des zutreffenden Geschlechts</u>	I feel it is appropriate ... MOTAS : <u>members of the appropriate sex</u>
http://dict.tu-chemnitz.de/dings.cgi?o=3001 &service=de-en&query=Behörde http://dict.tu-chemnitz.de/dings.cgi?o=3001 &service=de-en&query=Behörde {f};	<u>authority</u> ; <u>authorities</u> ; <u>administrative bodies</u>
<u>Behörden</u> {pl}	
<u>nächsthöhere Behörde</u> {f}	<u>superior authority</u>
<u>zuständige Behörde</u> {f}	<u>responsible authority</u> ; appropriate authority

The example above indicates that a translation corpus can offer many more instances of word sense discrimination and use senses than usually encountered in traditional dictionaries (e.g. ODG: *angemessen*, *angebracht* appropriate for or to, *passend für*, *geeignet für*). The on-line parallel corpus also shows patterns and uses specific to the English – German query pair ‘appropriate – angemessen’ and statistics as related to text type.

Since by the Croatian-English Corpus (Tadić 2000) is still under construction the results (i.e. parallel concordances) for *appropriate* are not illustrated here.

4.

This paper finally claims that lexicographic and translation procedures should make use and combine all the sources of information discussed above: dictionary data, the corpora avail-



able (monolingual corpora for both languages, comparable corpora, translation corpora), translation competence (translator's experience, translator's memory), and lexicographer's intuition.

Though highly synonymous with other members of the same lexical set, neither *appropriate* nor *odgovarajući* are fully substitutable for the other members of the set (*acceptable, correct, proper, right, suitable, fitting* in English or *primjeren, shodan, zgodan, prikladan, odgovarajući, podesan, uputan, doličan* in Croatian).

Substitutability versus frequency of the adjectives of the studied lexical set (*appropriate* etc.), must therefore be treated with caution. This is partly revealed in Table No. 4 (occurrence of the adjectives *acceptable, correct, suitable* with the noun *action*; or very low rate of occurrence of *correct* with *way, correct* or *suitable* with *steps, acceptable* or *suitable* with *use* etc.). The syntactic patterning must therefore be compared against the semantic properties of a word as well as its discourse or pragmatic environment, in order to decide on the choice of the right TE. This is where human analysis and decision-making takes over, cf.:

- | | |
|--|--|
| <p>(1) But he added mid-season was probably not the <i>appropriate time</i> to do it.</p> <p>(2) What is the most <i>appropriate way</i> to restore the northeast wing of Windsor Castle?</p> <p>(3) We must promise that <i>appropriate action</i> will be taken whenever possible.</p> <p>(4) Democrats say the bill is an <i>appropriate response</i> to the changing world environment.</p> <p>(5) ... and the situation demanded that the government took <i>appropriate measures</i>.</p> <p>(6) And it just didn't seem to be the <i>appropriate place</i> to have an argument.</p> <p>(7) It is unlikely the public would consider this is an <i>appropriate use</i> of resources ...</p> <p>(8) ... justices were in doubt as to the <i>appropriate level</i> of sentence for his particular class of case.</p> <p>(9) The Dublin summit is to decide on the most <i>appropriate form</i> of assistance for the Greek economy.</p> <p>(10) A deed registered with the <i>appropriate authority</i> usually constituted sufficient proof.</p> <p>(11) 'empathy' is not quite the <i>appropriate word</i> for this process, but it is close.</p> <p>(12) the Commissioner will take <i>appropriate action</i> to prevent a multiplicity of suits in various jurisdictions</p> <p>(13) ... times of the year when the Sun is at the <i>appropriate angle</i> above the horizon, ...</p> <p>(14) This is a major obstacle to <i>appropriate care</i>.</p> | <p>(1a) Međutim, dodao je, srce sezone vjerojatno nije <i>pravo vrijeme</i> za to.</p> <p>(2a) Koji je <i>najbolji način</i> da se obnovi sjeveroistočno krilo Windsorskog dvorca?</p> <p>(3a) Moramo obećati da će se, kad god to bude moguće, poduzeti <i>odgovarajuće mjere</i>.</p> <p>(4a) Demokrati kažu da prijedlog zakona predstavlja <i>primjeren odgovor</i> na promijenjeno međunarodno okruženje.</p> <p>(5a) ... a to je zahtijevalo da vlada poduzme <i>odgovarajuće mjere</i>.</p> <p>(6a) A činilo se je da to i nije <i>pravo mjesto</i> za prepirku.</p> <p>(7a) Javnost će vjerojatno držati da se sredstva ne koriste <i>pravilno</i>.</p> <p>(8a) ... suci nisu dvojili o tome koliko je <i>presuda primjerena</i> takvoj vrsti slučaja.</p> <p>(9a) Dubliski skup na vrhu treba odlučiti o <i>najprikladnijem obliku</i> pomoći grčkoj privredi.</p> <p>(10a) Dokument upisan kod <i>nadležnog organa vlasti</i> obično bi bio dovoljnim dokazom.</p> <p>(11a) 'suosjećanje' nije sasvim <i>prikladna riječ</i> za to, ali je ta riječ prilično blizu pravom značenju.</p> <p>(12a) Inspektor će poduzeti <i>odgovarajuće radnje</i> kako bi spriječio pokretanje više sudskih postupaka u različitim zakonodavstvima.</p> <p>(13a) doba godine kada se sunce nalazi pod <i>odgovarajućim</i> kutom iznad horizonta</p> <p>(14a) to je veća prepreka <i>ogovarajućoj skrbi</i>.</p> |
|--|--|



Even a superficial analysis points out some specific syntactic and semantic properties of the adjective *appropriate*:

First, combining all the sources above one may note the following patterns in English:

(a) *appropriate* + N + to + inf. (e.g. 1,2,6); (b) *appropriate* + N + for + NP (11); (c) take + *appropriate* + N (3,5); (d) ADV + *appropriate* + N (2,11); (e) NEG-S1 + S2 (NP = *appropriate* + N) (6,7,8,9,11). All these patterns call for separate semantic analyses of the syntactically conditioned senses of *appropriate*, which also produce different grammaticalizations and lexicalizations in TL.

Second, the concordances in the more general corpus (examples. 1-11 from the COBUILD corpus) do not always allow for invariable substitution of *appropriate* for other members of the lexical set without considerable change in the meaning, cf.:

(7b) It is unlikely the public would consider this is an *appropriate* use of resources

- a *proper* use
- a *suitable* use
- ?a *right* use
- a **fitting* use
- an *acceptable* use

Only *proper* and, partly, *suitable*, can be full lexical substitutes for *appropriate* in the above sentence. *Correct* and *acceptable*, though matching the grammatical pattern, produce slight changes in meaning (possible connotations arising from pragmatic elements of the context of situation, Kalogjera 1995). *Right* is syntactically questionable in its use with the indefinite article, although semantically it fits the *correct* and *acceptable* group (<morally acceptable>). There are semantic restrictions on the selection of *fitting*, which is confirmed by the absence of *fitting* used as a node in the corpus.

The same can be proved by the possible (author's) renditions in translation:

(7c) Javnost će vjerojatno držati da se sredstva ne koriste *pravilno*..

- adekvatno*
- na odgovarajući način*
- ispravno*
- na prihvatljiv način* .
- ?*na prikladan/primjeren način*

The analysis above yields more or less the same results across various specific translation corpora (texts) in different registers or contents by subject-matter (e.g. from INCOTERMS, Paris Memorandum on Port State Control, Collision Rules, etc.):



(15) The Seller has to load the waggon or container in the *appropriate* manner. INCOTERM p. 12)
 (16) The Authorities will seek satisfaction whenever they will deem this *appropriate*. (Paris Memorandum)

(15a) Prodavalac je dužan izvršiti utovar u vagon ili kontejner na *odgovarajući* način.

(16a) Vlasti mogu obavljati preglede uvijek kada to smatraju *prikladnim*.

Third, attributive *appropriate* preceded by an adverbial of intensity or degree (*more, most, quite, very*) excludes *odgovarajući*, the most neutral and, therefore, the most likely TE from the list of possible candidates for the translation equivalents in Croatian (**najdogovarajući, ?najviše odgovarajući, *vrlo odgovarajući*), and selects *prikladan* as the most appropriate TE. The same holds for the instances of predicative use, where *odgovarajući* (most frequently quoted TE in English-Croatian dictionaries) is ruled out as ungrammatical, e.g.:

(18) and in conformity with *appropriate* administrative and financial practices

(18a) i u skladu s *odgovarajućim* upravnim i finansijskim praksama (Paris Mem.)

(19) With regard to *appropriate* training, reference is made to ...

(19a) U pogledu odgovarajuće obuke upućuje se na ... (Paris Mem.)

(20) He shall also ... use and may, if necessary, detain the ship until appropriate corrective action is taken (Paris Mem.)

(20a) On će također ... te može, ako je potrebno, zadržati brod dok se ne poduzmu *odgovarajuće* mjere ispravljanja ... (Paris Mem.)

(21) to proceed to the nearest *appropriate* repair yard available ...

(21a) da isplovi do najbližeg raspoloživog *odgovarajućeg* brodogradilišta (Paris Mem.)

The second most frequent translation equivalents are *prikladan* and *primjeren*. It should be noted that this corpus too shows that the use of *prikladan* is indicated when the meaning of *appropriate* is emphasized by adverbs of degree or intensity (*more, most, entirely, very, only, quite, highly, wholly, hardly*):

(22) by the most *appropriate* means of telecommunication

(22a) *najprimjerenijim* sredstvima telekomunikacije (Off. Gaz.)

(23) provide the most *appropriate* assistance available

(23a) radi osiguranja *najprikladnije* raspoložive pomoći (Paris Mem.)

Exceptionally, the corpus also reveals the use of some other TE, such as *određen* ('equipment required by an *appropriate* convention' = 'opremi koja zahtijeva *određena* konvencija'), and *podesan* ('codes of practices or other *appropriate* means' = 'kodeksima rada ili na drugi *podesan* način', Official Gazette).

In one parallel text (COLREGS), the attributive *appropriate* has the following TE in Croatian: *nadležan* (*appropriate* authority = *nadležne* vlasti), *odgovarajući* (*appropriate* signal = *odgovarajući* signal, *appropriate* action = *odgovarajući* manevar), *određen* (in the *appropriate* traffic lane = po *određenom* plovidbenom putu) and *djelotvoran* (may send an *appropriate* whistle signal = mora davati druge *djelotvorne* signale).



In another maritime institutional parallel text *appropriate* has been used as an English translation equivalent for the following Croatian words: *odgovarajući* (odgovarajući pravilnik = appropriate Regulations; odgovarajući kapaciteti = appropriate capacity, na odgovarajućoj razini = on the appropriate level), *ispravan* (u ispravnom stanju = in appropriate condition), *propisan* (na propisanom obrascu = on the appropriate form, propisana tehnička dokumentacija = appropriate technical documentation), *svrsishodan* (postupati kako je najsvrsishodnije = take the most appropriate action), *prikladan* (na najprikladniji način = operate in the most appropriate manner).

Inappropriate, an inherent negation form of the adjective *appropriate*, shows a much higher restriction on the selection of the headword noun than its positive counterpart. The most typical nouns accompanied by this adjective belong to the set marked semantically by the feature “human behaviour or moral”, e.g. *behaviour*; *manner*; *treatment* and *relationships*. The same holds for some other members of the set, e.g.: *unacceptable behaviour* (*interference*, *thoughts*, *infringement*); *incorrect advice* (*diagnosis*, *behaviour*, *assumption*, *views*, *assessment*); *improper conduct* (*behaviour*, *dealings*, *interference*, *manner*, *inducement*, *treatment*, *compulsion*); *unsuitable marriage* (*behaviour*, *policies*). In the COBUILD corpus the predicative use of *inappropriate*, however, is far more frequent than the attributive one, a feature that is less pronounced in the case of its co-hyponyms (*unacceptable*, *incorrect*, *ill-fitting*, *improper*, *unsuitable*, and the suppletive form *wrong*). The predominant pattern is: ‘it + BE/SEEM + *inappropriate* + to-inf/that-clause’ or ‘it + BE + *inappropriate* + for + NP + to-inf’:

(29) Scotland Yard said it would be ‘*inappropriate*’ at this stage to comment on the judge’s ruling.

(30) It seemed *inappropriate* that the court should decline to exercise its authority ...

(31) It would be *inappropriate* for me to comment on the outcome of that review.

(29) U Scotland Yardu su izjavili da u ovoj fazi ne bi bilo *primjereno* (*dobro/pametno*) osvrtni se na odluku suca.

Činilo se *neprijmerenim* da će sud odustati od primjene prava iz svoje nadležnosti.

Ne bi bilo *dobro/pametno* da ja dadem osvrt na rezultat recenzije/prikaza.

As indicated by the translation (near)equivalents the choice in Croatian is also very limited, primarily by the kind of subject-matter of the text or register, by syntactic reasons (predicative use), and by the requirements of the communicative effect.

Other very frequent, predicative patterns or elliptic (indented) sentences such as: *as* + *appropriate*, *if* + *appropriate*, *where* + *appropriate*, *when* + *appropriate*, or similar patterns involving other members of the lexical set, which cause major problems both in the syntactic/semantic interpretation and, consequently, in the translation into Croatian are not the subject of this paper.



5. Conclusion

The lexical study of the adjective *appropriate*, and the set of qualifying adjectives semantically related to it, starts with the analysis in the two languages of the different syntactic patterns and the various senses of the word in English and its counterparts in Croatian. Dictionary information in both languages is then combined with corpus evidence in order to be able to contrast the lexical sets covering the shared concept: 'correct or suitable for a particular time, situation, or purpose' or 'the various degrees of adaptability to a particular use, purpose, or situation'. Substitutability test performed on a selection of monolingual and parallel concordances will reveal the degree of match between the units of the lexical sets in English and Croatian. In addition, the choice of the appropriate translation equivalent will depend on the semantic and usage interrelations holding among the adjectives within the lexical set, thus also reflecting the various factors acting in the context of situation. Various corpus linguistics tools, such as statistical measures of mutual expectancy, concordances, parallel concordances etc. can be of great value in contrasting the adjective - noun relations in the two languages and will in the near future offer on-line translation equivalents from general or field-specific bilingual lexical databases.

The benefits of combining dictionary and corpus information are threefold. For the lexicologist it will reveal the principal semantic relations holding among the members of the lexical set (*appropriate, suitable, correct, right, proper, fit, fitting*, etc.). It will help sense discrimination and suggest shared features and ordering of senses versus usage for each lexical unit of the set. For the lexicographer it offers syntactic-semantic properties of the lexical set and enables extraction and classification of statistically relevant collocates to create sub-entries. Such an analysis will also provide the elements for continuous developing of appropriate lexicographic definitions. Finally, bidirectional corpus and lexical analysis will be useful to the translator in arriving at translational equivalents and possible formal correspondents for the lexical set studied on the syntactic, semantic and pragmatic levels.

REFERENCES

- Altenberg, B. & Granger, S. (2002). 'Recent trends in cross-linguistic lexical studies'. In: Altenberg, B. & Granger, S. *Lexis in Contrast*, Studies in Corpus Linguistics, Amsterdam/Philadelphia, J. Benjamins Publ. Co., 3-48
- Baker, M. (1992). *In Other Words. A Coursebook in Translation*. Routledge, London
- Biber, D. Conrad, S. Reppen, R. (1998). *Corpus Linguistics, Investigating Language Structure Use*. Cambridge University Press
- Chesterman, A. (1998). *Contrastive Functional Analysis*. Amsterdam, Philadelphia, J. Benjamins Publ. Co.
- Cruse, D.A. (1986). *Lexical Semantics*. CUP



- Di Marco, C., Hirst, G. (1993). 'Usage notes as the basis for a representation of near-synonymy for lexical choice'. *Making Sense of Words*. UW Center for New OED and Text Research, 33-43
- Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge University Press
- Ivir, V. (1983). *A Contrastive Analysis of English Adjectives and their Serbo-Croatian Correspondents*. University of Zagreb
- Johansson (1998). "On the role of corpora in cross-linguistic research". In: Johansson and Oksefjell (eds), 282-296
- Justeson, J.S., Katz, M.S. (1993). 'Principled disambiguation: discriminating adjective senses with modified nouns'. *Making Sense of Words*. UW Center for New OED and Text Research, 57-74
- Kalogjera, D. (1995). 'Cultural Allusions in the Translation of English Newspaper Headlines'. *Prevođenje: Suvremena strujanja i tendencije*, HDPL, Zagreb, 29-38
- King, P., Wools, D. (1997). *MultiConcord - Parallel Concordancer, Ver 6.9H*,
- Salkie, R. (2002). Two types of translation equivalence", In: Altenberg, B. & Granger, S. *Lexis in Contrast*, Studies in Corpus Linguistics, Amsterdam/Philadelphia, J. Benjamins Publ. Co., 73-96
- Stubbs, M. (1996). *Text and Corpus Analysis*. Blackwell, Oxford
- Tadić, M. (2000). 'Building the Croatian -English Parallel Corpus', ELRA, Paris-Athens, Vol I, 523-530
- Tadić, M. (1999). Procedures in Building the Croatian-English Parallel Corpus, 4th TELRI Seminar, www.hnk.ffzg.hr/mt
- Teubert, W. (1996). "Comparable or parallel corpora?" In Sinclair et al. (eds) *An international project in multilingual lexicography*, 238-264

Dictionaries consulted:

Monolingual dictionaries:

- CEDT - *Collins English Dictionary & Thesaurus*, 1999.
- COBUILD - *Collins COBUILD English Dictionary*, 1996.
- English-German Dictionary*, University of Chemnitz (<http://dict.tu-chemnitz>)
- LDOCE - *Longman Dictionary of Current English*, 1998.
- Longman Language Activator*, 1994.
- OALD - *Oxford Advanced Learners' Dictionary*, 1999.
- ODG - *Oxford-Duden German-English / English-German Dictionary*, SuperLex, OUP, 1996.
- RHD - *Random House Webster's Electronic Dictionary & Thesaurus*, 1992.
- WNW - *Merriam - Webster's New World Dictionary*, 1998.
- W3 - *Webster's Third New International English Dictionary*, 1976.



English-Croatian and Croatian-English dictionaries:

Benson, M. (1989). *Srpskohrvatsko-engleski rečnik*, Prosveta, Beograd

Bujas, Ž. (1999a). *Veliki englesko-hrvatski rječnik*. Zagreb, Globus

Bujas, Ž. (1999b). *Veliki hrvatsko-englesko rječnik*. Zagreb, Globus

Filipović (1996). *English-Croatian Dictionary*. Školska knjiga, Zagreb

Lochmer (1906). *Englesko-hrvatski rječnik*. Naklada I. Pl. Hreljanovića, Senj

KOMBINIRANJE PODATAKA IZ RJEČNIKA I KORPUSA - LEKSIČKA STUDIJA
PRIDJEVA 'APPROPRIATE' I NJEGOVIH EKVIVALENATA U HRVATSKOM

U ovom radu razmatraju se podaci iz rječnika i korpusa kao dva glavna izvora informacija u leksikološkim studijama. Engleski pridjev 'appropriate' (kao i druge leksičke jedinice koje pripadaju istom leksičkom skupu) izabran je kao primjer uporabe tradicionalnih i modernih strategija u razlikovanju značenja i uporabe riječi. Osim općom jezičnom uporabom, ovaj se rad bavi i korpusom podataka dobivenih iz specijaliziranih stručnih tekstova, kako bi se napravile usporedbe koje mogu biti korisne u dvojezičnoj leksikografiji i u prevodenju.

