

---

# Displacement of Senses, Cluster Equivalence and Bilingual Dictionaries

Barbara Lewandowska-Tomaszczyk

State University of Applied Sciences in Konin

e-mail: blt@uni.lodz.pl

## Abstract

Basing on the data from the BNC and National Corpus of Polish, and relevant parallel corpora, the concepts of *lexical displacement of senses* and *cluster-for-cluster equivalence* are discussed and exemplified on the materials from English and Polish. Direct relevance to bilingual lexicography is further presented and inter-lingual correspondences involving two main types are analysed - lexical-conceptual *L1-cluster-for-L2 cluster mapping* and a *higher syntactic schematization* of the Target as contrasted with the more fine-grained construed Source language version. In order to receive a more complete picture of the *cross-language cluster mapping*, the two-directional translational corpora are consulted and the clusters referring to the mental space of *possibility* generated in both languages, analysed and juxtaposed for the qualitative and quantitative properties. The present project aims at extending the depth of cluster presentation in both languages and the refinement of a broad sense of equivalence between them. A special emphasis in bilingual lexicography project presented here is put on *Part-of-Speech cross-linguistic re-categorization* to recover more complex (asymmetric, incommensurable) inter-linguistic relations. The Principle of *Cluster Equivalence* is shown to be observed in a variety of styles and registers both in the written as well as in spoken mode.

**Keywords:** bilingual lexicography; cluster equivalence; collocational profile; construal; dictionary; displacement of senses; meaning approximation; parallel corpora; re-conceptualization; translation

## 1 Semantic Approximation in Meaning Sharing

Communication generally involves using similar and not identical meanings by interactants. Due to partially indeterminate nature of meanings, communication is approximative and language users accommodate the meanings to one another in terms of a *multi-directional accommodation* (Lewandowska-Tomaszczyk 2012).

Basing on the data from monolingual corpora (BNC and National Corpus of Polish) and relevant parallel corpora, used with corpus tools involving collocation generator (pelcra.pl/hask\_pl, Pęzik 2014) and aligned concordancer (<http://clarin.pelcra.pl/Paralela/#>), the concepts of *lexical displacement of senses* and *cluster-for-cluster equivalence* (Lewandowska-Tomaszczyk forthcoming) are discussed and exemplified on the materials from British English and Polish. Direct relevance of this kind of equivalence and relations to bilingual lexicography is further presented.

A distinction between types of meaning approximation is made here, particularly between that rooted in *ontological vagueness* on the one hand and *epistemological vagueness* on the other (Lewandowska-Tomaszczyk 2016). Ontological vagueness involves natural, existential absence of criterial properties of concept identification, in simpler terms, it corresponds to those objects which are originally indeterminate in real world. Epistemological vagueness on the other hand occurs in the situations in which there exist necessary/sufficient meaning criteria for a linguistic sense definition but, due to various social, psychological, linguistic, or contextual reasons, they are not fully exploited in particular contexts. Both types of vagueness are magnified when meanings from different languages are contrasted, which is particularly evident in the materials from bilingual dictionaries and translation.

## 2 Equivalence and Displacement of Senses

*Equivalence*, which is a topic researched in a very large body of multidisciplinary literature and received numerous philosophical, logical, linguistic and translational interpretations, is understood here in a broad sense as (typically partial) meaning resemblance between concepts. There are basically three major classes of equivalence postulated in the present work. *Generalized equivalence* captures meaning correspondences by shifting them towards higher categorization levels. *Parallel*, more closely *aligned*, equivalence typically occurs in domain-specific texts, and *particularized equivalence*, engaging fine-grained, more explanatory specification of word meanings, goes either deeper into meaning *precision*<sup>1</sup> or else covers comparative semantic levels of a wider range, i.a., inter-categorical extensions (Lewandowska-Tomaszczyk 2012a, 2016). In none of the three types, are equivalent forms in different languages fully *commensurable*.<sup>2</sup> It is rather *a whole array of forms and senses* with relevant concepts in L2 that correspond to a lexical form in L1. Furthermore, the relationship is more branched in fact as it is rather a cluster of concepts in L1 that bear some degrees of similarity (on a number of varying criteria) to those diverse forms in L1, as presented in Table 1 for the form *RUN* in English with its extended cluster of *displaced senses*, (partly) equivalent to a similar cluster of senses in Polish. The English verb *RUN* has a number of possible equivalent forms in Polish (a small part of which is presented in Table 1). Each of the Polish equivalents in turn corresponds to more than one English form, which opens up again a new range of possible Polish equivalents. The equivalence patterns are getting more and more complex with a number of new conceptual areas which function as further points of elaboration on which to build new equivalence forms and meanings in the subsequent Target Language.

Eng. (boys) *run* (to school)

**Eng. *RUN*** [ Pol. *biec/biegać*]

Pol. czas biegnie lit. 'time runs' ] Eng. (time) *passes* [

- Pol. *ubiegać się* lit. 'run oneself' [*run for* (presidency)]
- Pol. *działać* lit. 'function' [(machines stop) *running*]
- Pol. *kierować* lit. 'supervise' [*run* (business)]
  - [Eng. *go* (by bus (vehicle))]
    - [Pol. *jechać/jeździć* (iterative) [
      - [Eng. *drive* [Pol. *prowadzić samochód*]
      - [Eng. *ride* [Pol. *jeździć na rowerze/koniu*]]
- Eng. *pass* > Pol. wszystko biegnie lit. 'run'/idzie 'walk' dobrze ] Eng. (all) *GO* (fine)]
  - [Eng. *go* (to school)]
    - [Pol. *iść/chodzić* (iterative)]
      - Eng. *walk* [Pol. *spacerować*]
- Pol. *mijać* [
  - Eng. *leave behind*
  - Eng. *miss*
  - Eng. *pass*

Table 1: Displacement of senses (Eng. *RUN* and Polish and Eng. equivalence clusters).

<sup>1</sup> The term *precision* is used in the logical sense of *truth clarification* as proposed in Grimm (1953/55).

<sup>2</sup> See Lakoff 1987 for the concept of (*in*) *commensurability between languages*.

### 3 Typology of Equivalence and Meaning Clusters

It was proposed in previous studies (Lewandowska-Tomaszczyk 2016) that *meaning clusters* identified in language are mental areas structured around *similar content*. *Similarity* (Lewandowska-Tomaszczyk (2012:1) is determined by multi-peaked radial category spaces (Lakoff 1987), regulated in terms of a number of *tertia comparationis*, or points of reference, which serve as similarity conditioning parameters. The relationship between cluster equivalents is that of semantic *resemblance*. Semantic resemblance can be established in terms of a number of shared properties, i.e., *Shared Qualitative features* [phonetic, morphological, syntactic, semantic (shapes, topology, function, etc.), pragmatic] and *Shared Quantitative features* [near sets], i.e., the number of object feature values in common. Degrees of resemblance between the L1 and L2 and possible inter-substitutions involve the following major dimensions as presented in Table 2:

PERSPECTIVE  
 EVENT (*literal/figurative*) & Participants  
 Event CONSTRUAL<sup>3</sup> linked to Syntactic Selections and Collocation patterns  
 (Object/Event) Granularity/Schematicity  
*Prototypical/Peripheral category members* & their (intra- and inter-category extensions)  
*Form* → *Definition*  
*Form* → *Examples (list)*

Table 2: Semantic resemblance dimensions.

Although meanings and their semantic spaces are not fully predictable outside of context, they retain their core, prototypical senses, which can serve as a point of reference to compare degrees of similarity with other concepts. *Sufficiently similar concepts*<sup>4</sup> can be predicted to belong to the same *cluster*. This membership, as is always the case with categorization, is dynamic and subject to modulation.

### 4 Cluster Equivalence – Parallel Corpora

To exemplify semantic clustering processes on authentic language materials selected parallel English-to-Polish and Polish-to-English corpus data is presented in the present paper to support the claim that a translated text involves firstly equivalent structures from the core as well as successively more and more peripheral areas of the Source Language (SL) - Target Language (TL) similarity space. It is proposed that there eventually exists a certain *approximation threshold* (Gärdenfors 2000), which confines the area of *permissible* as opposed to a cline of *risky* equivalence patterns, some of which can be *nonce formations*, some others are permissible although rarely used in authentic discourse.

The data used in this section are generated from monolingual national corpora (BNC and National Corpus of Polish nkjp.pl) and parallel (Polish-to-English & English-to-Polish) corpora. The parallel corpus as well as the corpus tools used such as a collocation generator (pelcra.pl/hask\_pl) and aligned concordancer (<http://clarin.pelcra.pl/Paralela/#>) (Peżik 2014) have been built by the PELCRA team at the University of Lodz, Poland. As an example the top-frequency collocational profiles and parallel concordance-based equivalent clusters of items involving the conceptual space of *possibility* in the two languages are presented in Table 3:

<sup>3</sup> The concept of *construal* is understood in terms of Cognitive Linguistics (Langacker 1987) and refers to the way how linguistic structure portrays different semantic content.

<sup>4</sup> The concepts of *sufficient similarity* and *allowable substitution* are accounted for in Lewandowska-Tomaszczyk (2016).

#	Collocate	POS	A	TTEST
1	distinct	AJ%	54.0	6.51
2	remote	AJ%	27.0	4.14
3	real	AJ%	80.0	4.12
4	exciting	AJ%	27.0	3.98
5	ever-present	AJ%	10.0	3.06
6	endless	AJ%	14.0	2.94
7	infinite	AJ%	10.0	2.59
8	intriguing	AJ%	8.0	2.42
9	alternative	AJ%	20.0	2.27
10	mere	AJ%	15.0	2.22
11	theoretical	AJ%	14.0	2.20
12	future	AJ%	27.0	2.05
13	strong	AJ%	44.0	2.04
14	realistic	AJ%	10.0	2.02
15	very	AJ%	21.0	1.91
16	alarming	AJ%	5.0	1.71

Table 3: Top adjectival collocations of *possibility*.

#	Collocate	POS	A	TTEST
1	equal	AJ%	440.0	20.05
2	educational	AJ%	107.0	8.53
3	earliest	AJ%	82.0	8.41
4	unique	AJ%	90.0	8.02
5	golden	AJ%	86.0	7.97
6	ideal	AJ%	89.0	7.85
7	ample	AJ%	64.0	7.67
8	excellent	AJ%	77.0	6.35
9	miss	AJ%	33.0	5.08
10	lost	AJ%	25.0	4.43
11	rare	AJ%	41.0	4.11
12	perfect	AJ%	44.0	3.99
13	wonderful	AJ%	39.0	3.84
14	unrivalled	AJ%	12.0	3.28
15	tremendous	AJ%	20.0	3.05

Table 4: Top adjectival collocations of *opportunity*.

#	Collocate	POS	A	TTEST
1	better	AJ%	241.0	11.16
2	best	AJ%	182.0	7.20
3	good	AJ%	370.0	6.16
4	realistic	AJ%	35.0	4.9
5	only	AJ%	124.0	4.59
6	fat	AJ%	35.0	4.47
7	fair	AJ%	61.0	4.40
8	pure	AJ%	35.0	4.10
9	slightest	AJ%	21.0	3.97
10	even	AJ%	15.0	3.45
11	slim	AJ%	18.0	3.35
12	off	AJ%	12.0	3.23
13	outside	AJ%	28.0	3.09
14	sport	AJ%	13.0	2.80

15	treble	AJ%	8.0	2.64
----	--------	-----	-----	------

Table 5: Top adjectival collocates of *chance*.

Verbal collocates of these three items signal both semantic differences and similarities between senses such as stronger meaning links between *opportunity* and *chance* as juxtaposed to *possibility*, which is summarised in Table 6:

Top Verbal Collocates of *possibility*: *consider, explore, discuss, raise, offer*

Top Verbal Collocates of *opportunity*: *provide, take, offer, give, seize*

Top Verbal Collocates of *chance*: *get, win, take, stand, give*

Table 6: Verbal collocates.

A closer observation of the equivalence patters in the translated materials of the two languages reveals additional properties of the meanings, which provide arguments for the *Principle of Cluster Equivalence* presented here. A number of possible equivalence patterns identified in the parallel texts (6) can be proposed and a sample of the examples presented reveals equivalence clusters as used in the translation practice.

### (1) possibility > możliwość

(1E<sup>5</sup>) the landlord is forgoing the possibility of renting to someone else

(1P) (wynajmujący) nie ma możliwości wynajęcia go komuś innemu

### (2) usability > możliwość

(2E) It is therefore the usability of the house that is the subject of the contract

(2P) Przedmiotem umowy najmu jest więc możliwość korzystania (lit. possibility of using) z mieszkania

### (3) opportunity - możliwość

(3E) look for new opportunities

(3P) poszukują nowych możliwości

### (4) option > możliwość

(4E) This is a better option

(4P) To lepsza możliwość

### (5) chance > możliwość

(5E) the chance to choose something happy or sad.

(5P) możliwość wyboru szczęśliwego lub smutnego zdarzenia.

### (6) potential - możliwość

(6E) jobs with significant learning potential

(6P) a ich zajęcia oferują znaczne możliwości kształcenia się

### (7) capability > możliwość

(7E) We have been very lucky to have had the capability to monitor the polar regions with satellites

<sup>5</sup> The symbol E is used for the English examples, while P – for the Polish examples.

(7P) dysponujemy możliwościami monitorowania regionów polarnych przez satelity

**(8) reach > możliwość**

(8E) categories that are not mega-sized, but are still beyond a single country's reach, need such coordination and support

(8P) kategorie, które, choć nie mają dużych rozmiarów, wykraczają poza możliwości jednego państwa, potrzebują takiej koordynacji i wsparcia.

**(9) possibility > perspektywa**

(9E) its mission and the possibilities it opens for them

(9P) jej misji i perspektyw, jakie przed nimi otworzyła,

**(10) option > opcja**

(10E) Rejecting this compromise is no option

(10P) Odrzucenie tego kompromisu nie jest żadną opcją

**(11) POS Recategorization**

(V>N)

(11E) She hoped to be able to present the results

(11P) Wyraziła ona ponadto nadzieję co do możliwości przedstawienia wyników

**(VP>NP)**

(12E) so that we could observe the storm as soon as possible

(12P) abyśmy mieli możliwość (lit. have a possibility) obserwowania burzy możliwie jak najszybciej

The analysis of the patterns presented in examples (1-12) above makes it possible to posit a set of *SL-TL prototypical equivalence cluster correspondences* between English and Polish conceptual area of possibility with extended Part-of-Speech re-categorisation types as structured in (13):

**(13) PROTOTYPE (State) possibility - możliwość**

(State/Event) *opportunity* > możliwość, sposobność

(Event/Result) *option* > opcja, możliwość

(External conditions) *chance/reach* > możliwość, szansa

(Particularised Event) *usability* > możliwość użycia

(Event) *perspective* > perspektywa, możliwość

(Individual/Organization) *potential/capability/reach* > możliwość/możliwości (Plural)

**Extended POS re-categorization (N>V, V>N, X>Adjective/Adverb/PrepNP** *it is possible for/that, possibly, optionally, perhaps, maybe*

While the equivalence between *possibility* – *możliwość* can be considered a *prototypical case*, *opportunity/option/chance* demonstrates the salient role of external conditioning, and constitutes cases of *extended* equivalence, similarly to the forms such as *potential* and *capability* involving rather the internal predispositions of an Agent. The particularised equivalence of the *usability* - *możliwość* case on the other hand is an instance of an *emergent equivalence* type, conditioned by the discourse situational factors. Thus, *Extended Equivalence* involves areas of presupposition,

causality, and a number of implicational clues, and *Emergent Equivalence* is dynamically set up in discourse and can also act as ‘nonce’ formation in the text. A similar scale can be applied to the example of *RUN* in Table 1 above.

## 5 Quantitative Criteria of Re-conceptualization. Collocation Distribution

The quantitative criteria of meaning re-conceptualization (Lewandowska-Tomaszczyk 2010) are applied when, inter alia, collocational strings are compared for their textual distribution as for the Nouns *possibility* and *opportunity* in English presented in (1) and (3) and two of their possible Polish cluster equivalents *możliwość* and *opcja* (8 & 10), generated from the BNC and the National Corpus of Polish respectively. The quantitative distributional criteria reveal additional semantic specificity of the items compared (Table 7).

### (i) *możliwość* Adjectival Collocates

C	Collocate	POS	A	TTEST	
1	taki	adj	6172.0	54.57	‘such’
2	finansowy	adj	1448.0	31.66	‘financial’
3	ograniczony	adj	641.0	23.97	‘limited’
4	swój	adj	2910.0	20.26	‘one’s’
5	techniczny	adj	577.0	20.25	‘technical’
6	nieograniczony	adj	392.0	19.33	‘unlimited’
7	realny	adj	447.0	19.00	‘real’
8	nowy	adj	1797.0	18.62	‘new’
9	dodatkowy	adj	538.0	17.41	‘additional’
10	szybki	adj	445.0	16.84	‘fast’
11	własny	adj	887.0	16.00	‘own’
12	swobodny	adj	269.0	15.34	‘free’
13	szeroki	adj	401.0	15.30	‘wide’
14	prawny	adj	561.0	14.81	‘legal’
15	potencjalny	adj	273.0	14.48	‘potential’

Verbal Collocates of *możliwość*: *mieć* ‘have’, *dawać* ‘give’, *istnieć* ‘exist’, *stwarzać* ‘create’, *rozważyć* ‘consider’

Table 7: Collocations of the Polish equivalent forms (*możliwość*)

### (ii) *sposobność* Adjectival Collocates

#	Collocate	POS	A	TTEST	
1	każdy	adj	89.0	8.34	‘each’
2	ten	adj	135.0	4.58	‘this’
3	pierwszy	adj	39.0	3.99	‘first’
4	doskonały	adj	16.0	3.84	‘perfect’
5	bliski	adj	17.0	3.46	‘close’
6	dobry	adj	29.0	3.28	‘good’
7	dogodny	adj	11.0	3.28	‘comfortable, convenient’
8	żaden	adj	20.0	3.17	‘no one’
9	jedyny	adj	12.0	2.85	‘single, unique’
10	znakomity	adj	9.0	2.82	‘excellent’
11	wyjątkowy	adj	8.0	2.66	‘exceptional’
12	taki	adj	35.0	1.77	‘such’
13	zdarzony	adj	3.0	1.73	‘that occurred’
14	świetny	adj	4.0	1.72	‘very good’
15	niejeden	adj	3.0	1.69	‘more than one’

Verbal Collocates: *mieć* ‘have’, *dawać* ‘give’, *(s)korzystać* ‘benefit’, *nadarzy(a)ć* ‘happen’, *czekać* ‘wait’

Table 8: Collocations of the Polish equivalent forms (*sposobność*).

A comparison between the collocational profiles of *możliwość* (Table 7) and *sposobność* (Table 8) shows that the two Polish forms demonstrate closer cohesiveness and similarity between each other than in the case of the English corresponding equivalents *possibility* and *opportunity* as contrasted in Tables 3 and 4. They also present differentiating properties with respect to their meaning, more specifically the presence of the Verbal collocate *rozważyć* ‘consider’ for *możliwość* rather than for *sposobność*, which makes *możliwość* a closer equivalent to *possibility* than to any of the other members of the presented cluster.

## 6 Conclusion

What is observed in the authentic parallel corpus data are two basic correspondences: first of all, lexical-conceptual *L1-cluster-for-L2 cluster mapping* and secondly, a *higher syntactic schematization* of the Target as contrasted with the more fine-grained construed Source version. To receive a fuller picture of the *L2-to-L1 cluster mapping*, the two-directional translational corpora have been consulted and the clusters generated in both languages are analysed and juxtaposed to each other for the qualitative and quantitative properties and adequately structured to be used as bilingual dictionary English-to-Polish and Polish-to-English cluster entries.

First attempts of a lexicographic application of networked clusters of senses were undertaken in a project on a Bilingual Dictionary Thesaurus (BIT, Lewandowska-Tomaszczyk 1993), which presented the BIT formula and bilingual materials, stored in electronic surrounding, which made it possible for the user to have access to the network of lexical senses in the two languages, and could be used both as an explanatory monolingual dictionary based on the encyclopaedic thesaurus principle and, at the same time, as a reference tool accounting for similarities and contrasts between English and Polish. The software was combined with authentic language corpus data to provide the user with a more developed meaning exposition as well as with the contextual restrictions on the distribution and pragmatics of lexical senses. The present project aims at extending the depth of cluster presentation in both languages and the refinement of a broad sense of equivalence between them.

A special emphasis in bilingual lexicography project presented here is put on *L1-to-L2 Part-of-Speech re-categorization*, in which a headword description needs to be more adequately elaborated on to cover more complex (asymmetric, incommensurable) inter-linguistic relations.

The Principle of *Cluster Equivalence* can be observed in a variety of styles and registers: literary language, general language and restricted-domain LSP. In Lewandowska-Tomaszczyk (2016) wider areas of English - Polish materials from *spoken discourse* are also consulted to broaden the study of the equivalence areas towards neighbouring conceptual domains.

## 7 References

- Gärdenfors, P. (2000). *Conceptual Spaces – The Geometry of Thought*. Cambridge, Mass.: The MIT Press. A Bradford Book.
- Grimm, L. (1953/55). On the Application of the Concept of Precisation. *Synthese* Vol. 9, No.2 104-20.
- Lakoff, G. (1987). *Women, Fire and Dangerous Thing: What Categories Reveal about the Mind*. Chicago: Chicago University Press.
- Langacker, R.W. 1987. *Foundations of Cognitive Grammar* vol. 1. Stanford University Press.
- Lewandowska-Tomaszczyk, B. (2016). Principle of Cluster Equivalence and parallel corpora. In M. Thelen, G-W van Egdom, D. Verbbeck, L. Bogucki and B. Lewandowska-Tomaszczyk (eds.) *Translation and Meaning. New Series, Vol.1*. Frankfurt a. Main: Peter Lang. 229-244.



- Lewandowska-Tomaszczyk, B. (2012). Approximative spaces and the tolerance threshold in communication. *International Journal of Cognitive Linguistics* volume 2, number 2. Nova Science publishers, inc.
- Lewandowska-Tomaszczyk, B. (2012a). Approximative Spaces and the Tolerance Threshold in Communication. *International Journal of Cognitive Linguistics* Volume 2 Issue 2. 2-19.
- Lewandowska-Tomaszczyk, B. (2010). Re-conceptualization and the emergence of discourse meaning as a theory of translation. In B. Lewandowska-Tomaszczyk and M. Thelen *Meaning in Translation*. Peter Lang. 105-148.
- Lewandowska-Tomaszczyk, B. (1993). Bilingual Thesaurus (BIT), Field Theory, and Cognitive Linguistics. In P. R. Lutzeier (ed.), *Studies in Lexical Field Theory*, Niemeyer Verlag.
- Pęzik, P. (2014). Spokes – a search and exploration service for conversational corpus data. [http://pelcra.pl/docs/lib/exe/fetch.php?media=spokes\\_sub\\_2.pdf](http://pelcra.pl/docs/lib/exe/fetch.php?media=spokes_sub_2.pdf)
- Parallel English-to-Polish and Polish-to-English Corpora*, CLARIN Project, University of Lodz.