

Appendix 4. Index of semantic proximity

For each word pair, an index of semantic proximity was computed as the proportion of the associations shared by both members of this pair (plus the number of mutual associations between the pair members) to the overall number of associations (shared or not) given for them. To illustrate how the index was computed, let us calculate it for the *egzystować – istnieć* pair, which is reproduced below from Appendix 3:

egzystować	istnieć	8	istnieć	egzystować	4
		1		człowiek	2
		1		funkcjonować	2
	być	3		być (byt)	15
	filozofia	3		filozofia	2
	żyć (życie)	12		żyć (życie)	9
	trwać (trwanie, przetrwanie)	5			0
	wegetować (wegetacja)	4			0

The first association common to both members of this pair is *być (byt)*, its frequency being 3 for *egzystować* and 15 for *istnieć*. The lower of these two figures is the number shared by both members of the pair, and so we take 6 (i.e. 3 multiplied by 2) as the pair's number of shared associations with *być (byt)*. This is the first element of the pair's total sum of shared associations, in the numerator of the fraction to be computed. The second element is 4 (the lower of the two frequencies given for *filozofia*, multiplied by 2), the third and the last element is 18 (the lower of the frequencies given for *żyć (życie)*, again multiplied by 2). To this number of shared associations we then add the number of mutual associations, i.e. 8 occurrences of *istnieć* for *egzystować* and 4 occurrences of *egzystować* for *istnieć*. This is because the semantic similarity of any two words manifests not only in the associations they share with other words, but also in the mutual associations they have with each another. In fact, it is the mutual associations that indicate their direct similarity, while the associations shared with other words indicate their similarity only in an indirect way, according to the common rule that if two objects are similar to a third one, they are very likely to be similar to each another.

The total number of associations shared by the members of the *egzystować – istnieć* pair, plus the number of mutual associations between this pair's members, is thus $6 + 4 + 18 + 8 + 4 = 40$. The overall number of associations (shared and not shared) obtained for this pair is 71, which is taken as the denominator of the fraction. Dividing the first number by the second, we arrive at the semantic proximity measure between the elements of the *egzystować – istnieć* pair: $40/71$, which is approximately 0.56.

The measures computed this way are listed below:

absurdalny – niedorzeczny	33/59	0.56	camping – kemping	60/90	0.67
agresor – napastnik	22/78	0.28	chat – czat	53/81	0.65
autonomiczny – niezależny	36/63	0.57	chips – czips	46/92	0.50
businessman – biznesmen	36/84	0.43	cyklista – kolarz	42/72	0.58

dealer – diler	51/83	0.61	konsultant – doradca	32/72	0.44
dewastować – niszczyć	31/69	0.45	korumpować – przekupywać	43/73	0.59
doktor – lekarz	39/83	0.47	kuriozalny – osobliwy	16/67	0.24
e-mail – mejl	71/96	0.74	lobbysta – lobbista	20/49	0.41
egzystować – istnieć	40/71	0.56	manager – menedżer	38/69	0.55
ekspert – znawca	32/87	0.37	menstruacja – miesiączka	61/93	0.66
eksplozja – wybuch	50/77	0.65	partycypować – uczestniczyć	23/61	0.38
fan – miłośnik	3/62	0.05	puls – tętno	60/106	0.57
fenomenalny – nadzwyczajny	26/72	0.36	smartphone – smartfon	48/76	0.63
football – futbol	54/97	0.56	stagnacja – zastój	32/85	0.38
helikopter – śmigłowiec	58/98	0.59	strofa – zwrotka	19/94	0.20
infekcja – zakażenie	50/91	0.55	stymulować – pobudzać	32/68	0.47
jazz – dżez	22/74	0.30	symptom – objaw	52/81	0.64
jeansy – dżinsy	58/89	0.65	toksyczny – trujący	40/95	0.42
komfort – wygoda	52/87	0.60	tryumf – tryumf	62/93	0.67
kompleksowy – całościowy	17/72	0.24	unikalny – wyjątkowy	19/75	0.25
komponent – składnik	41/83	0.49	voodoo – wudu	54/84	0.64

The following list includes the same word pairs, but in descending order of indices:

e-mail – mejl	71/96	0.74	manager – menedżer	38/69	0.55
camping – kemping	60/90	0.67	chips – czips	46/92	0.50
tryumf – tryumf	62/93	0.67	komponent – składnik	41/83	0.49
menstruacja – miesiączka	61/93	0.66	doktor – lekarz	39/83	0.47
chat – czat	53/81	0.65	stymulować – pobudzać	32/68	0.47
eksplozja – wybuch	50/77	0.65	dewastować – niszczyć	31/69	0.45
jeansy – dżinsy	58/89	0.65	konsultant – doradca	32/72	0.44
symptom – objaw	52/81	0.64	businessman – biznesmen	36/84	0.43
voodoo – wudu	54/84	0.64	toksyczny – trujący	40/95	0.42
smartphone – smartfon	48/76	0.63	lobbysta – lobbista	20/49	0.41
dealer – diler	51/83	0.61	partycypować – uczestniczyć	23/61	0.38
komfort – wygoda	52/87	0.60	stagnacja – zastój	32/85	0.38
helikopter – śmigłowiec	58/98	0.59	ekspert – znawca	32/87	0.37
korumpować – przekupywać	43/73	0.59	fenomenalny – nadzwyczajny	26/72	0.36
cyklista – kolarz	42/72	0.58	jazz – dżez	22/74	0.30
autonomiczny – niezależny	36/63	0.57	agresor – napastnik	22/78	0.28
puls – tętno	60/106	0.57	unikalny – wyjątkowy	19/75	0.25
absurdalny – nienormalny	33/59	0.56	kompleksowy – całościowy	17/72	0.24
egzystować – istnieć	40/71	0.56	kuriozalny – osobliwy	16/67	0.24
football – futbol	54/97	0.56	strofa – zwrotka	19/94	0.20
infekcja – zakażenie	50/91	0.55	fan – miłośnik	3/62	0.05