

From *Cheap* to *Affordable* – Corpus-Assisted Exploration of Near-Synonymy in English, German, and Slovak Economic Discourse¹

This study explores the lexical-semantic behaviour of near-synonyms in economic discourse across English, German, and Slovak. It focuses on high-frequency lexical pairs such as *cheap*–*affordable*, *billig*–*preiswert*, and *lacný*–*výhodný*. The analysis examines how these items differ in semantic prosody, collocational behaviour, and sentiment function within their respective languages. The study is grounded in two central hypotheses. First, polysemous adjectives *cheap*, *billig*, *lacný* are expected to occur more frequently in neutral or negative sentiment contexts. Second, less polysemous adjectives *affordable*, *preiswert*, *výhodný* are assumed to appear predominantly in positive contexts. To test these assumptions, economic subcorpora from the TenTen family were analysed via Sketch Engine. The strongest nominal collocates were extracted using logDice and subsequently grouped into semantic domains. Cross-linguistic translation issues were also taken into account. Sentiment analysis was performed on concordance lines using AI-assisted tools and Microsoft Excel. Each instance was labelled as positive, neutral, or negative. Chi-square tests were applied to assess whether the observed sentiment distributions deviated significantly from a uniform expectation. The statistical analyses do not confirm either of the proposed hypotheses. However, the data reveal observable tendencies. Less polysemous adjectives occur more frequently in contexts with positive connotations, whereas polysemous adjectives show a broader sentiment range that includes neutral and negative contexts. These tendencies remain descriptively noticeable but do not reach statistical significance, and the distributions show deviations that prevent clear confirmation of the hypotheses. Together, the collocational and sentiment-based analyses contribute to a detailed reconstruction of semantic prosody and illustrate how synonymic variation functions within the economic register of each language.

Keywords: near-synonymy, corpus linguistics, specialised language, economic discourse

Von *cheap* zu *affordable* – Korpusgestützte Untersuchung der nahen Synonymie im englischen, deutschen und slowakischen Wirtschaftsdiskurs

Diese Studie untersucht das lexikalisch-semantische Verhalten von Nahe-Synonymen im wirtschaftlichen Diskurs in Englisch, Deutsch und Slowakisch. Der Schwerpunkt liegt auf häufig vorkommenden Wortpaaren wie *cheap*–*affordable*, *billig*–*preiswert* und *lacný*–*výhodný*. Die Analyse untersucht, wie sich diese Begriffe in ihrer semantischen Prosodie, ihrem Kollokationsverhalten und ihrer Sentimentfunktion innerhalb ihrer jeweiligen Sprachen unterscheiden. Die Studie basiert auf zwei zentralen Hypothesen. Erstens wird

¹ The research has been elaborated within the projects I-26-101-00 *Language as a Sales Tool: Modelling Trust and Effectiveness in AI-Driven Brand Communication*, A-25-106/3020-18 *Artificial intelligence: challenges for linguistics and marketing* and the project APVV 24-0302 *Harmonization of Slovak and English terminology of territorial self-government in SR*.

erwartet, dass die polysemischen Adjektive *cheap*, *billig* und *lacný* häufiger in neutralen oder negativen Sentimentkontexten vorkommen. Zweitens wird angenommen, dass weniger polysemische Adjektive wie *affordable*, *preiswert* und *výhodný* überwiegend in positiven Kontexten vorkommen. Um diese Annahmen zu überprüfen, wurden wirtschaftliche Teilkorpora aus der TenTen-Familie mit Sketch Engine analysiert. Die stärksten nominalen Kollokationen wurden mit logDice extrahiert und anschließend in semantische Domänen gruppiert. Dabei wurden auch sprachübergreifende Übersetzungsprobleme berücksichtigt. Die Sentimentanalyse wurde anhand von Konkordanzzeilen mit KI-gestützten Tools und Microsoft Excel durchgeführt. Jede Instanz wurde als positiv, neutral oder negativ gekennzeichnet. Mit Chi-Quadrat-Tests wurde geprüft, ob die beobachteten Sentimentverteilungen signifikant von einer einheitlichen Erwartung abweichen. Die statistischen Analysen bestätigen keine der vorgeschlagenen Hypothesen. Die Daten lassen jedoch beobachtbare Tendenzen erkennen. Weniger polysemische Adjektive kommen häufiger in Kontexten mit positiver Konnotation vor, während polysemische Adjektive ein breiteres Spektrum an Emotionen abdecken, das auch neutrale und negative Kontexte umfasst. Diese Tendenzen sind deskriptiv erkennbar, erreichen jedoch keine statistische Signifikanz, und die Verteilungen weisen Abweichungen auf, die eine eindeutige Bestätigung der Hypothesen verhindern. Zusammen tragen die kollokations- und sentimentbasierten Analysen zu einer detaillierten Rekonstruktion der semantischen Prosodie bei und veranschaulichen, wie synonymische Variation innerhalb des wirtschaftlichen Registers jeder Sprache funktioniert.

Schlüsselwörter: Nahe Synonymie, Korpuslinguistik, Fachsprache, Wirtschaftsdiskurs

Authors: Filip Kalaš, Bratislava University of Economics and Business, Dolnozemská cesta 1, 851 04 Petržalka, Slovakia, e-mail: filip.kalas@euba.sk
Pavol Lipták, Bratislava University of Economics and Business, Dolnozemská cesta 1, 851 04 Petržalka, Slovakia, e-mail: pavol.liptak@euba.sk

Received: 1.8.2025

Accepted: 10.1.2026

1. Theoretical background

Synonymy as a central concept in lexical semantics denotes the relationship between words that share the same or similar meanings. This refers to a semantic or formal-semantic oppositional relationship between lexemes, in which the differentiating feature is either neutralised or weakened. As a result, the lexemes are in the closest relationship within the onomasiological lexical paradigm (Dolník 2007). It follows that in the traditional lexical semantics, there are two types of synonymy: absolute and near-synonymy. The absolute synonymy is characterised as a relationship where two lexical items can substitute for one another in all contexts without affecting the utterance's meaning, truth conditions, stylistic register, or connotation (Cruse 1986, Edmonds/Hirst 2002). Interestingly, absolute synonymy can only exist under strict conditions, which requires perfect substitutability across all syntactic, pragmatic, and stylistic levels (Cruse 1986). Taylor (2002: 263) adds that "while perfect synonyms are rare, 'near-synonyms' are especially numerous", because this shows their centrality in natural language use. Therefore, it is very unlikely to encounter an absolute synonymy in natural language. Lipka (1992: 142) corroborates this that „synonymy is rarely absolute but usually partial i. e., we are dealing with near-synonyms which differ in their collocational and stylistic restrictions or in their connotations“. This is evident in the pair *big* and *large*, often presented as synonyms but they exhibit distinct collocational preferences (*big mistake* vs. *large number*). In this sense, their mutual substitution is highly limited.

As a result of the limitations of absolute synonymy, lexical semantics focuses more on near-synonymy, which is a phenomenon involving words with substantial semantic overlap yet differing in connotation, stylistic value, collocational patterns, or distributional properties (Taylor 2002, Divjak 2010). Some authors use the term “plesionyms” to describe near-synonyms, emphasizing that these words, while semantically close, exhibit “non-trivial differences in collocational ranges and stylistic value” (Cruse 1986: 266, Leech 1974, Hanks 2013). Yet, we do not use this term in this study.

2. Corpus-based studies on near-synonyms

Corpus studies have significantly improved our understanding of how words with similar core meanings function in real-word communication. These studies move beyond the concept of absolute synonymy and focus on the ways how near-synonyms deviate in collocational patterns, semantic prosody, and discourse functions. The corpus-based approach helps scholars gain insights into subtle distinctions that are often missed by traditional lexicography.

Among a number of published studies there are several notable similarities that need to be mentioned. Smith (2021), Ikonen (2022), Taylor (2002), and Louw (1993) regard near-synonymy as a context-dependent, gradual phenomenon, demonstrating that words with similar denotational meanings are distinct in terms of collocational behaviour and stylistic value. When it comes to the methodological approach, many examined studies adopt corpus-driven and corpus-based approaches, making use of tools like Sketch Engine, the Corpus of Contemporary American English (COCA), and the British National Corpus (BNC). Commonly, researchers employ techniques including collocational profiling, frequency counts, concordance analysis and examinations of semantic prosody (Louw 1993, Gries 2006). While dictionaries are useful supplement to research, corpus data typically offer a more accurate understanding of words in authentic language setting (Ikonen 2022). Taken together, the studies underscore the profound importance of context and vividly illustrate how even small variations in meaning can influence pragmatic role of a word within discourse.

Despite certain shared approaches, there are studies differing in both theoretical and methodological orientation. Taylor (2002) and Gries (2006, 2012) work within cognitive linguistics and distributional semantics frameworks, while focusing on conceptual categories and how words are used in different contexts. In contrast, Louw (1993) and Polizzi et al. (2024) prioritise semantic prosody and collocational patterns. Furthermore, there are corpus-based studies analysing near-synonyms in specialised domains e. g., military (Smith 2021) or medical English (Maci et al. 2019). Tomášková and Borovská (2021) concentrate on the semantic and collocational comparison of *lockdown* and *shutdown* within German pandemic discourse, employing Sketch Engine to build a targeted corpus. Next, some studies address near-synonymy within general English or computational applications using broad corpora such as COCA and BNC

(Ikonen 2022, Yang 2016, Hu/Yang 2015, Inkpen/Hirst 2006). Moznárová (2018) and Inkpen and Hirst (2006) further demonstrate the practical application of near-synonym studies within lexicography and computational linguistics.

Altogether, corpus-based studies of near-synonyms demonstrate the value of empirical linguistic methods in revealing the nuanced semantic and stylistic distinctions between closely related words. These insights strengthen theoretical linguistic research and contribute to language pedagogy, lexicography, translation studies, and specialised discourse analysis.

3. Semantics of the chosen near-synonyms

In this part, we examine the recorded semantics of the selected near-synonymous pairs. To ensure consistency and reliability, the provided overview is based on online accessible lexicographic resources in each language.

For English, the individual meanings along with their collocates were extracted from the Cambridge Dictionary and the Oxford Dictionary, for German, the DWDS and the Duden Dictionary were used, and for Slovak, the headword *lacný* was examined using the SSSJ, while the headword *výhodný* was analysed using the KSSJ. The reason is that the SSSJ, albeit the most up-to-date dictionary, has been processed just up to the letters Pn so far.

cheap (EN)	billig (DE)	lacný (SK)
low or reduced in price, inexpensive		
<i>cheap flight</i> <i>cheap food</i> <i>cheap imports</i>	<i>billiges Obst</i> ‚cheap fruits‘ <i>billige Waren</i> ‚cheap goods‘ <i>billige Arbeitskräfte</i> ‚cheap labour‘	<i>lacný tovar</i> ‚cheap goods‘ <i>lacný obed</i> ‚cheap lunch‘ <i>lacný dar</i> ‚cheap gift‘
low quality; of poor quality		
<i>cheap wine</i> <i>cheap shoes</i> <i>cheap jewellery</i>	<i>billiger Ramsch</i> ‚cheap rubbish‘ <i>billiger Schnapps</i> ‚cheap liquor‘ <i>billiges Kleid</i> ‚cheap dress‘	
unkind or unpleasant		
<i>cheap jokes</i> <i>cheap tricks</i> <i>cheap laughs</i>	<i>billige Ausrede</i> ‚lame excuse‘ <i>billiger Trick</i> ‚cheap trick‘ <i>billiger Trost</i> ‚cheap consolation‘	<i>lacná zábava</i> ‚cheap entertainment‘ <i>lacná výhovorka</i> ‚lame excuse‘ <i>lacná útecha</i> ‚cheap consolation‘
low status		
<i>cheap crook</i>		
ungenerous (AmE)		
<i>be cheap</i>		
	legitimite, justified	
	<i>billige Forderung</i> ‚equitable claim‘ <i>billiger Vorschlag</i> ‚reasonable proposal‘	
		easy; effortless
		<i>lacný úspech</i> ‚cheap success‘ <i>lacná sláva</i> ‚cheap fame‘ <i>lacná popularita</i> ‚cheap popularity‘

Tab. 1. Semantic fields of the adjective *cheap* and its equivalents in German and Slovak

The table 1 reveals that the first three meanings overlap across all three languages. English, however, has two additional meanings not paralleled in the other languages, which are indicated in orange low status and ungenerous. The latter one is mostly uttered in American English, whereas the British English uses the adjective *mean* in this sense. German has one additional meaning used in specialised juridical discourse, signifying *legitimate* or *justified*. In Slovak, one additional meaning appears, as well. This is marked in green, where *lacný* extends metaphorically to *easy*, *effortless*. This nuance is not shared with English or German.

Overall, while *cheap/billig/lacný* share a core denotation of low cost and reduced price across languages, they reveal differences in semantic extension and sentiment prosody. The shared meanings (the first three ones) often carry negative connotations that imply inferior quality or lack of worth. This aligns with Cruse's (1986) and Lyon's (1977) observations that near-synonyms frequently overlap in denotational core while they deviate or show variations in sentiment and stylistic load. Thus, these synonyms share initial semantic similarities across languages, however, their additional meanings reveal language-specific extensions influenced by cultural and discourse-specific norms.

affordable (EN)	preiswert (DE)	výhodný (SK)
not expensive, good value, cheap		
<i>affordable prices</i> <i>affordable housing</i> <i>affordable transportation</i>	<i>preiswertes Angebot</i> ,good-value offer' <i>preiswerte Schuhe</i> ,affordable shoes' <i>preiswertes Zimmer</i> ,affordable room'	<i>výhodná pozícia</i> ,advantageous position' <i>výhodná kúpa</i> ,good deal' <i>výhodný produkt</i> ,good-value product'

Tab. 2. Semantic field of the adjective *affordable* and its equivalents in German and Slovak

The data show that all three adjectives (table 2) share the core meaning of “not expensive, good value, or cost-effective”, which reflects positive sentiment prosody across languages. In English, the adjective *affordable* commonly collocates with nouns that emphasize accessibility without implying low quality. The German equivalent similarly creates collocations that convey the meanings of good value, and as such they maintain a neutral to positive connotations, mostly in marketing discourse. The Slovak adjective aligns with this semantic core but introduces a slight shift in sentiment nuance, frequently emphasising advantageousness or benefit rather than mere affordability.

From the viewpoint of synonymy, these adjectives can thus be treated as near-synonyms; yet, they exhibit subtle differences in focus and collocational behaviour from the overall semantic point of view.

4. Methodology

The methods used in this study are grounded in corpus linguistics. The distributional analysis of the three near-synonymous adjectives is based on their occurrence and

co-occurrence with nominal collocates in economic discourse. For this purpose, three comparable web corpora were selected, specifically their subcorpora focused on economic and financial texts. These are operated by Sketch Engine (Kilgarriff et al. 2014).

English Web 2021 (enTenTen21)	52,268,286,493	German Web 2023 (deTenTen23)	16,666,049,248	Slovak Web 2023 (skTenTen23)	898,031,101
Economy & finance & business	822,858	Economy & finance & business	60,676	Economy & finance & business	89,428

A partial collocational profile was constructed for each adjective, measured by absolute frequency and logDice score. In most cases – except for the adjective *preiswert* due to its lower frequency – the 20 strongest nominal collocates (ranked by logDice) were extracted and categorized into semantic domains. For German and Slovak data, particular attention was given to contrastive aspects, especially to how the adjectives under investigation are rendered in English depending on context.

In addition, the study sought to examine the sentiment context in which each adjective typically appears. To this end, sentiment analysis was carried out using Microsoft Excel Sentiment Analysis add-in powered by Microsoft Azure Cognitive Services. The sentiment values generated by the tool were subsequently reviewed manually on a sample of concordance lines to ensure interpretative plausibility in the context of economic discourse. Neutral sentiment was operationalised as instances where the adjective occurred in a descriptive or informational context without explicit positive or negative sentiment value, as indicated by near-zero polarity scores and corroborated by contextual reading. The sentiment analysis used in this study is intended as a supportive and exploratory method. While the AI-based tool ensures consistency and replicability, its results may be affected by context-dependent meaning and polysemy, particularly in economic discourse. For this reason, automated classifications were supplemented by manual inspection of concordance lines, and the results are interpreted in combination with collocational and statistical analyses rather than as standalone evidence.

In this study, two hypotheses were formulated:

- H_1 : Polysemous adjectives (*cheap*, *billig*, *lacný*) are significantly more likely to occur in neutral or negative contexts than in positive ones.
- H_0 : There is no significant difference in the distribution of polysemous adjectives across sentiment contexts;
- H_2 : Less polysemous adjectives (*affordable*, *preiswert*, *výhodný*) are significantly more likely to occur in positive contexts than in neutral or negative ones.
- H_0 : There is no significant difference in the distribution of less polysemous adjectives across sentiment contexts.

The quantitative component of the study is grounded in descriptive statistics (absolute frequency, relative frequency, and logDice), complemented by inferential statistical testing and sentiment analysis.

5. Results

5.1 Collocational analysis and semantic distributions of the near-synonyms *cheap-affordable*

The data in table 3 depicts the nominal collocates of the adjective *cheap* drawn from Sketch Engine. This data is crucial since it enables a systematic analysis of the semantic and distributional behaviour of this adjective. Moreover, it highlights its dominant usage patterns across different domains.

Nr.	collocate <i>cheap</i>	AF	RF _{100,000}	logDice	Nr.	collocate <i>cheap</i>	AF	RF _{100,000}	logDice
1.	<i>gasoline</i>	107	13,00	5,25	11.	<i>booze</i>	17	2,07	3,68
2.	<i>labour</i>	338	41,08	5,14	12.	<i>LNG</i>	19	2,31	3,62
3.	<i>crude</i>	79	9,60	4,8	13.	<i>scam</i>	113	13,73	3,53
4.	<i>labor</i>	467	56,75	4,69	14.	<i>tracker</i>	113	13,73	3,52
5.	<i>valuation</i>	193	23,45	4,57	15.	<i>coal</i>	75	9,11	3,5
6.	<i>electricity</i>	189	22,97	4,31	16.	<i>borrowing</i>	40	4,86	3,43
7.	<i>import</i>	278	33,78	4,12	17.	<i>money</i>	1108	134,65	3,29
8.	<i>SMPP</i>	11	1,34	3,99	18.	<i>feedstock</i>	19	2,31	3,23
9.	<i>yen</i>	27	3,28	3,93	19.	<i>alternative</i>	400	48,61	3,19
10.	<i>liquidity</i>	49	5,95	3,88	20.	<i>financing</i>	133	16,16	3,18

Tab. 3. Nominal collocates of *cheap* extracted from enTenTen21

Building on this quantitative foundation, the subsequent analysis explores how these collocates cluster around distinct semantic domains, while they also reveal less frequent but semantically significant associations with low quality or questionable value. Among 20 extracted collocates we can distinguish the following semantic clusters:

- **Energy domain** – collocates such as *gasoline*, *crude*, *electricity*, *LNG*, *coal*, *feedstock* build a tight cluster around commodity markets. In this case the adjective *cheap* here aligns with price volatility, affordability for consumers, and market reporting contexts.
- **Labour Market** – the collocate *labour* and its American variant *labor* reach the highest AF/RF what underscores the prominence of discussions around cheap labour in economic corpus. Here, the adjective *cheap* aligns with cost-efficiency and productivity. Moreover, it may be connected also to discussions around labour exploitation, migration and/or ethical concerns.
- **Finance** – collocates like *money*, *liquidity*, *financing*, *borrowing*, *valuation* reflect a specialised financial sense of cheap referring to easy credit conditions, low-interest rates, and accessible financing.
- **Others** – collocate *import* and its adjectival base may denote affordable substitutes or low-cost foreign goods. In the collocations with *scam*, *tracker*, *booze* the adjective *cheap* implies poor quality, unreliability, or disrepute.

The table 4 represents the noun collocates of the near-synonym *affordable*. Unlike *cheap*, which requires semantic segmentation, this adjective consistently carries a single sense of something financially accessible while retaining acceptable quality.

collocate <i>affordable</i>	AF	RF _{100 000}	logDice	collocate <i>affordable</i>	AF	RF _{100 000}	logDice
<i>(patent) attorney</i>	1290	156,77	5,41	<i>minimum</i>	36	4,37	2,38
<i>housing</i>	1596	193,96	5,24	<i>flagship</i>	9	1,09	2,13
<i>plumbing</i>	57	6,93	3,88	<i>smartphone</i>	50	6,08	2,05
<i>EV (electric vehicle)</i>	10	1,22	3,69	<i>housing</i>	18	2,19	2,01
<i>price-point</i>	6	0,73	3,67	<i>indulgence</i>	8	0,97	1,95
<i>homeownership</i>	7	0,85	3,63	<i>price</i>	902	109,62	1,91
<i>luxury</i>	56	6,81	3,33	<i>pricing</i>	57	6,93	1,9
<i>childcare</i>	16	1,94	3,09	<i>electricity</i>	34	4,13	1,88
<i>healthcare</i>	68	8,26	3,01	<i>mortgage</i>	32	3,89	1,53
<i>rent</i>	71	8,63	2,8	<i>repayment</i>	8	0,97	1,21

Tab. 4. Nominal collocates of *affordable* extracted from enTenTen21

Thus, the analysis focuses primarily on domain-based distribution and contextual patterns:

- **Housing, living and essential services** – the strongest collocational pattern appear with *affordable housing* accompanied by *rent*, *mortgage*, *homeownership*, *plumbing*, *electricity*, *healthcare* and *childcare*. It is evident that these collocates align with policy, economic and social discourse, where *affordable* signals accessibility to essential needs.
- **Professional services** – a notable collocate with high frequency and logDice is *(patent) attorney* which indicates the professional service domain. Here, the adjective *affordable* is employed to discuss reducing barriers to access legal or specialised professional service and as such it frames the concept within contexts of justice and entrepreneurship.
- **Consumer goods and technology** – collocates such as *smartphone*, *EV*, *flagship*, *luxury* or *indulgence* demonstrate the usage in the marketing and consumer discourse, framing high-value or premium goods as financially attainable for a broader consumer base.
- **Market discourse** – the collocates *price*, *pricing*, *price-point* and *minimum* implies market strategy and pricing discourse. In this context, the adjective is used to describe strategies designed to appeal to price-sensitive consumers. The focus tends to remain on financial accessibility without denoting diminished quality.
- **Financial Transaction** – the only collocate *repayment* suggests its role in discussions on loan and credit management. This may align with discourse in personal finance and banking.

5.2 Collocational analysis and semantic distributions of the near-synonyms *billig-preiswert*

The adjective *billig* in German exhibits a broad distribution while preserving its core association with low price and affordability. Nonetheless, it frequently extends into sentiment dimensions and implies inferiority, insufficiency, or superficiality. In terms of domain distribution, the collocates (Table 5) can be classified into the following categories:

- **Financial and economic domain** – this domain covers a significant number of collocates like *Baugeld*, *Kredit-Zins*, *Kredit-Angebot*, *Zentralbankgeld*, *Geld*, *Kredit* and *Aktie*. In this domain, the adjective typically indicates financial accessibility through low prices or low interest rates. From a contrastive point of view, the suitable English equivalents often include *low-interest*, *inexpensive* or *cheaply priced*.
- **Energy and commodities** – this sector is represented by two nominal collocates, *Erdgas* and *Massenware*, both emphasising low prices within trade and consumer markets.
- **Labour** – collocates such as *Arbeitskraft* and *Hilfskraft* highlight this semantic domain, where the adjective *billig* denotes low-wage or inexpensive labour. It often carries a potentially negative sentiment layer.
- **Legal context** – this domain is described by *Ermessen*, *Kündigungsgrund*, and *Entschädigung* and the adjectival node takes on a sentiment sense implying arbitrariness, insufficiency, or inadequacy.
- **Other** – the collocates *Polemik* and *Ausrede* together with the adjective denote superficiality or poor quality of argumentation.

collocate <i>billig</i>	AF	RF _{100,000}	logDice	collocate <i>billig</i>	AF	RF _{100,000}	logDice
<i>Baugeld</i> ,building loan capital ^c	16	26,37	8,35	<i>Arbeitskraft</i> ,labour ^c	31	51,09	3,65
<i>Kredit-Zins</i> ,low-interest loan ^c	6	9,89	7,47	<i>Geld</i> ,money ^c	203	334,56	3,3
<i>Kredit-Angebot</i> ,loan offer ^c	4	6,59	6,75	<i>Kredit</i> ,credit ^c	36	59,33	2,92
<i>Ermessen</i> ,arbitrary discretion ^c	140	230,73	6,43	<i>Ausrede</i> ,lame excuse ^c	5	8,24	2,72
<i>Zentralbankgeld</i> ,central bank money ^c	7	11,54	6,4	<i>Aktie</i> ,stock ^c	22	36,26	2,53
<i>Kopierpapier</i> ,copy paper ^c	5	8,24	6,4	<i>Erdgas</i> ,natural gas ^c	4	6,59	2,38
<i>Entschädigung</i> ,meager compensation ^c	34	56,04	4,83	<i>Import</i> ,import ^c	4	6,59	2,33
<i>Massenware</i> ,mass merchandise ^c	4	6,59	4,68	<i>Hilfskraft</i> ,auxiliary worker ^c	4	6,59	2,31

collocate <i>billig</i>	AF	RF _{100 000}	logDice	collocate <i>billig</i>	AF	RF _{100 000}	logDice
<i>Kündigungsggrund</i> ,flimsy/weak reason for dismissal'	4	6,59	4,19	<i>Einkauf</i> ,purchase'	4	6,59	0,69
<i>Polemik</i> ,polemics'	5	8,24	3,76	<i>Kopie</i> ,copy'	4	6,59	0,68

Tab. 5. Nominal collocates of *billig* extracted from deTenTen23

Table 6 shows that the collocational profile of the adjective *preiswert* is much narrower than that of its near-synonym. This results in a smaller number of collocates. Another evident fact is the low occurrence of these collocations within the given subcorpus, which is, of course, also due to the limited size of the subcorpus itself.

collocate <i>preiswert</i>	AF	RF _{100 000}	logDice	collocate <i>preiswert</i>	AF	RF _{100 000}	logDice
<i>Anschlussfinanzierung</i> ,follow-up financing'	2	3,30	5,03	<i>Darlehen</i> ,loan'	6	9,89	1,42
<i>Kreditangebot</i> ,loan offer'	2	3,30	3,29	<i>Kredit</i> ,credit'	8	13,18	0,76
<i>Wohnlage</i> ,residential area'	2	3,30	2,08	<i>Aktie</i> ,undervalued credit'	5	8,24	0,4
<i>Aktienmarkt</i> ,undervalued stock market'	2	3,30	1,79	<i>Unterkunft</i> ,accommodation'	8	13,18	0,33
<i>Übernachtungsmöglichkeit</i> ,accommodation option'	3	4,94	1,48	<i>Umzug</i> ,relocation'	2	3,30	0,24

Tab. 6. Nominal collocates of *preiswert* extracted from deTenTen23

The German adjective *preiswert* consistently carries a single core meaning “good value, affordable or reasonably priced. It denotes financial accessibility while retaining acceptable quality, similar to English *affordable*.

The collocates in the Table 6 fall into two domains:

- **Finance** – the adjective together with its collocates *Anschlussfinanzierung*, *Kreditangebot*, *Darlehen*, *Kredit*, *Aktienmarkt* and *Aktie* signal attractive financial conditions. In terms of financial domain, it may be linked with discourses such as banking, personal finance, and investment markets.
- **Housing** – here, the adjective implies affordability in the housing and living domain. It is accompanied by the nouns *Wohnlage*, *Übernachtungsmöglichkeit*, *Unterkunft* and *Umzug*.

5.3 Collocational analysis and semantic distributions of the near-synonyms *lacný-výhodný*

From a domain perspective, the collocational profile given in the Table 7 can be grouped into three overarching categories:

- **Housing** – this domain includes collocates such as *nehnutelnosť, elektrina, matrac, komoda, rošť, kemp*. Here, the adjective *lacný* retains its literal sense of affordability and low price and it maintains a neutral connotation.
- **Finance** – this domain covers the collocates *peniaze, kryptomeny, benzín, ropa, hypotéka* or *úver* and the adjective refers mostly to low-cost financial assets, and favourable market pricing. Thus, it emphasises economic accessibility in investment and consumer decision-making.
- **Frivolity** – this domain captures collocates like *špás* and *brak* where the adjective implies worthlessness, superficiality or poor quality. In such contexts, equivalents like *frivolous, shoddy* or *trash* could be more appropriate than *cheap*.

collocate <i>lacný</i>	AF	RF _{100 000}	logDice	collocate <i>lacný</i>	AF	RF _{100 000}	logDice
<i>letenka</i> ,plane ticket‘	57	63,74	8,56	<i>peniaze</i> ,easy money‘	26	29,07	5,78
<i>hypotéka</i> ,mortgage loan‘	99	110,70	8,29	<i>kryptomeny</i> ,cryptocurrencies‘	7	7,83	5,78
<i>špás</i> ,frivolous joke‘	14	15,66	7,45	<i>kemp</i> ,campsite‘	10	11,18	5,71
<i>válanda</i> ,daybed‘	6	6,71	6,44	<i>ropa</i> ,oil‘	12	13,42	5,7
<i>Bulharsko</i> ,Bulgaria	7	7,83	6,28	<i>matrac</i> ,mattress‘	26	29,07	5,67
<i>úver</i> ,loan‘	93	103,99	6,03	<i>benzín</i> ,gasoline‘	9	10,06	5,66
<i>elektrina</i> ,electricity‘	16	17,89	5,99	<i>alternatíva</i> ,alternative‘	41	45,85	5,59
<i>komoda</i> ,chest of drawers‘	6	6,71	5,98	<i>brak</i> ,trash / shoddy goods‘	4	4,47	5,57
<i>nehnutelnosť</i> ,property‘	40	44,73	5,96	<i>variant</i> ,version‘	9	10,06	5,53
<i>rošť</i> ,bed slat‘	18	20,13	5,92	<i>vitrína</i> ,showcase‘	6	6,71	5,49

Tab. 7. Nominal collocates of *lacný* extracted from skTenTen23

The adjective *výhodný* consistently exhibits the positive sense of advantageous, beneficial, or attractive. Given the high frequencies and remarkably high logDice values (see Table 8), the adjective demonstrates a high degree of collocability, indicating it is deeply embedded in practical, financial discourse in Slovak.

collocate <i>výhodný</i>	AF	RF _{100 000}	logDice	collocate <i>výhodný</i>	AF	RF _{100 000}	logDice
<i>kúpa</i> ,good buy‘	3843	4297,31	11,42	<i>hypotéka</i> ,mortgage‘	267	298,56	7,59
<i>poloha</i> ,location‘	1815	2029,57	9,44	<i>set</i> ,set‘	318	355,59	7,52

collocate <i>výhodný</i>	AF	RF _{100 000}	logDice	collocate <i>výhodný</i>	AF	RF _{100 000}	logDice
<i>cena</i> ,price'	6420	7178,96	9,24	<i>pozícia</i> ,position'	503	562,46	7,41
<i>ponuka</i> ,attractive offer'	1836	2053,05	8,52	<i>zľava</i> ,discount'	279	311,98	7,4
<i>úver</i> ,loan'	756	845,37	8,24	<i>financovanie</i> ,financing'	241	269,49	7,34
<i>podmienka</i> ,condition'	2444	2732,92	8,13	<i>alternatíva</i> ,alternative'	298	333,23	7,28
<i>pôžička</i> ,loan'	499	557,99	8,13	<i>sadzba</i> ,rate'	396	442,81	7,15
<i>nákup</i> ,purchase'	586	655,28	8,11	<i>ubytovanie</i> ,accommodation'	345	385,79	7,14
<i>úrok</i> ,interest rate'	401	448,41	8,09	<i>balíček</i> ,package'	232	259,43	7,05
<i>balenie</i> ,value pack'	560	626,20	8,06	<i>riešenie</i> ,solution'	824	921,41	7,04

Tab. 8. Nominal collocates of *výhodný* extracted from skTenTen23

When it comes to semantic domains, these can be divided into following groups:

- **Consumer Transactions** – the collocates like *cena*, *kúpa*, *nákup*, *ponuka*, *zľava*, *balenie* or *balíček* indicate along with the adjective a purchase or price that offers good value or is particularly attractive.
- **Financial contexts** – the collocates *úver*, *pôžička*, *hypotéka*, *financovanie*, *úrok* and *sadzba* may here denote favourable financial conditions.
- **Real estate** – the collocates *nehnutelnosť*, *ubytovanie*, *poloha*, *pozícia* demonstrate advantageous features of the accommodation.

6. Evaluation of sentiment analysis

Sentiment analysis was employed to uncover sentiment tendencies associated with the examined adjectives. This sentiment value is essential for understanding the semantic profiles of these adjectives in contexts.

The heatmap (Fig. 1) clearly visualises the sentiment distribution by applying an intuitive colour gradient. Several key patterns emerge:

1. Less polysemous adjectives *affordable*, *preiswert*, *výhodný* exhibit high positive sentiment scores (0.76–0.79). This support the assumption that they tend to have a consistent positive sentiment profile in discourse.
2. Importantly, the adjectives *preiswert* and *výhodný* do not display any negative values.
3. Interestingly, polysemous adjectives *cheap*, *billig*, *lacný* reveal high positive sentiment scores (0.76–0.80) as well. This may reflect their dual sentiment potential depending on context and collocate.
4. Neutral sentiment scores remain stable across all adjectives.

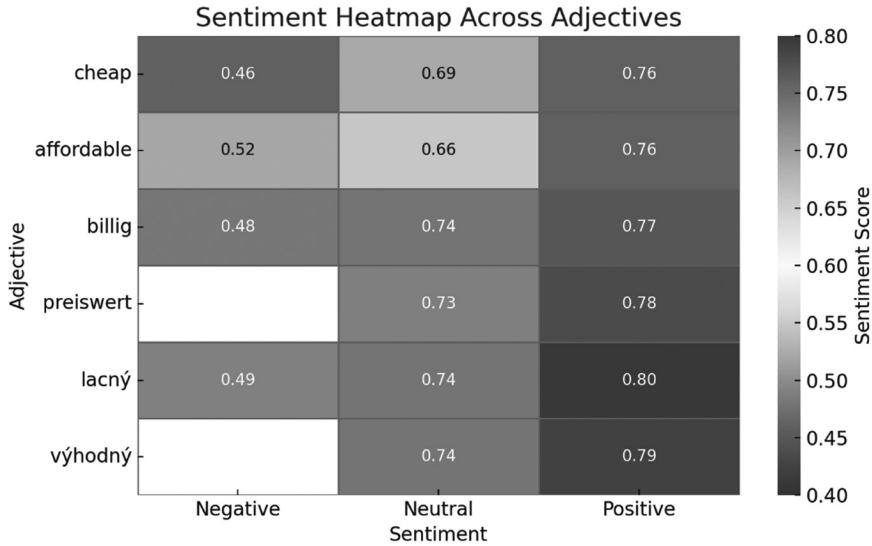


Fig. 1. Sentiment analysis of the examined adjectives

The graph (Fig. 2) illustrates the proportional distribution of sentiment across adjectives, normalized to 100.% for comparability. It is evident that *cheap* and *lacný* occur predominantly in neutral contexts (87.2.% and 74.2.%, respectively) with minimal positive usage. In contrast, *preiswert* and *billig* exhibit a strong positive skew (77.8.% and 66.6.%), while *výhodný* shows a balanced distribution with a notable positive share (45.8%). *Affordable* displays a mixed profile, with neutral contexts (50.6.%) dominating but with a substantial positive component (36.2%).

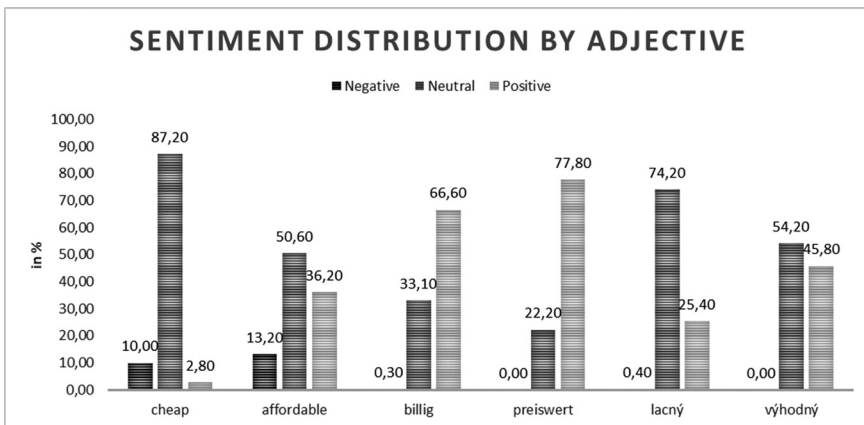


Fig. 2. Distribution of sentiment categories normalised to 100 % across adjectives

The results revealed significant deviations for all adjectives tested. Specifically, *cheap* ($\chi^2=570.9$, $p<0.001$), *affordable* ($\chi^2=25.5$, $p<0.001$), *billig* ($\chi^2=17.4$, $p<0.001$), *preiswert*

($\chi^2=26.9$, $p<0.001$), *lacný* ($\chi^2=85,5$, $p<0.001$), and *výhodný* ($\chi^2=3,53$, $p=0.060$) all demonstrated highly significant differences from the null hypothesis of equal distribution across positive and non-positive contexts. These findings indicate that the sentiment distribution of these adjectives in corpus data is not uniform, supporting the observation that adjectives such as *cheap*, *affordable*, and *lacný* tend to occur predominantly in non-positive contexts, while adjectives like *preiswert*, *billig*, and *výhodný* demonstrate a significant skew towards positive contexts. Nevertheless, statistical significance in the chi-square tests only indicates deviation from a uniform distribution and does not by itself confirm the directional hypotheses, as the direction and consistency of sentiment tendencies vary across individual adjectives and languages.

7. Discussion and conclusion

This study showed that near-synonymous adjectives *cheap-affordable*, *billig-preiswert* and *lacný-výhodný* display distinct sentiment tendencies shaped by polysemy and discourse. Polysemous adjectives seemed to be used predominantly in neutral or negative contexts, while less polysemous ones appeared in positive sentiment discourse. Chi-square tests confirmed deviations from an even sentiment distribution, what supports the hypothesis that lexical polysemy correlates with sentiment variability.

Nevertheless, the tendency for the polysemous adjective *billig* to occur in neutral to negative contexts and, conversely, for the monosemous adjectives *preiswert* and *výhodný* to occur mainly in positive contexts is partially distorted, as the Fig. 2 shows. We can therefore reject both hypotheses since statistically significant deviations do not consistently align with the predicted sentiment directions across all adjectives and languages, although certain tendencies can be observed. In the future, it will be necessary to extend the analysis to include other adjectives such as *inexpensive*, *low-priced*, *economic* or *reasonable*. Their sentiment distributions could be analysed to refine the understanding of lexical variation.

These findings highlight the importance of context, in our case, economic discourse where perceived value and tone are critical. Moreover, the context played a significant role also in translation process. Speaking about translation process, it is important to note that the nominal collocate and the context itself impact on the equivalency of the adjectives in the target languages. The German *preiswert* requires translations such as *affordable* or *good-value*, and thus avoiding the potentially negative tone of *billig*. Similarly, the Slovak *výhodný* demands context-sensitive equivalents like *attractive*, *advantageous* or *well-located*. In contrast, *lacný* is semantically broader, with translations ranging from *cheap* to *affordable* or *low-cost*.

Overall, the combination of collocational and sentiment analysis offers a productive method for capturing subtle sentiment distinctions among near-synonyms within and/or across languages.

Literature

- CAMBRIDGE UNIVERSITY PRESS. *Cambridge Dictionary*. <https://dictionary.cambridge.org/>. 4.7.2025.
- CRUSE, D. A. *Lexical semantics*. Cambridge: Cambridge University Press, 1986. Print.
- DIGITALES WÖRTERBUCH DER DEUTSCHEN SPRACHE (DWDS). *DWDS*. <https://www.dwds.de/>. 4.7.2025.
- DIVJAK, Dagmar. „Structuring the lexicon: A clustered model for near-synonymy“. *Quantitative methods in cognitive semantics: Corpus-driven approaches*. Ed. Dylan Glynn and Kerstin Fischer. Berlin, Boston: Mouton de Gruyter, 2010, 47–88. Print.
- DOLNÍK, Juraj. *Lexikológia*. Bratislava: Univerzita Komenského, 2007. Print.
- DUDENREDAKTION. *Duden online*. <https://www.duden.de/>. 4.7.2025.
- EDMONDS, Philip and Graeme HIRST. „Near-synonymy and lexical choice“. *Computational Linguistics* 28.2 (2002): 105–144.
- GRIES, Stefan Th. „Frequencies, probabilities, and association measures in usage-/exemplar-based linguistics: Some necessary clarifications“. *Studies in Language* 36.3 (2012), 477–510.
- GRIES, Stefan Th. „Exploring variability within and between corpora: Some methodological considerations“. *Corpora* 1.2 (2006): 109–151.
- HANKS, Patrick. *Lexical analysis: Norms and exploitations*. Cambridge, MA: MIT Press, 2013. Print.
- HU, Chunyu and Bin YANG. „Using Sketch Engine to investigate synonymous verbs“. *International Journal of English Linguistics* 5.4 (2015): 29–36.
- IKONEN, Noora. *Big and large: A corpus study of near-synonyms in contemporary American English* [Master's thesis, University of Helsinki]. Helda Digital Repository. 2022. <https://helda.helsinki.fi/>
- INKPEN, Diana and Graeme HIRST. „Building and using a lexical knowledge-base of near-synonym differences“. *Computational Linguistics* 32.2 (2006): 223–262.
- JAROŠOVÁ, Alexandra and Katarína BUZÁSSYOVÁ (eds.). *Slovník súčasného slovenského jazyka. H–L*. Bratislava: Veda, 2011. Print.
- KAČALA, Ján, Stanislav KOLESÁR and Mária MISLOVIČOVÁ. *Krátky slovník slovenského jazyka*. 4th ed. Bratislava: Veda, 2003. Print.
- KILGARRIFF, Adam, Vít BAISA, Jan BUŠTA, Miloš JAKUBÍČEK, Vojtěch KOVÁŘ, Jan MICHELFEIT, Pavel RYCHLÝ and Vít SUCHOMEL. „The Sketch Engine: Ten years on“. *Lexicography* 1 (2014): 7–36. Print.
- LEECH, Geoffrey. *Semantics*. Harmondsworth: Penguin, 1974. Print.
- LIPKA, Leonhard. *An outline of English lexicology: Lexical structure, word semantics, and word-formation*. Tübingen: Max Niemeyer Verlag, 1992. Print.
- LOUW, Bill. „Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies“. *Text and technology: In honour of John Sinclair*. Ed. Mona Baker, Gill Francis and Elena Tognini-Bonelli. Amsterdam: John Benjamins, 1993, 157–176. Print.
- MACI, Stefania, Réka JABLONKAI and Rita SALVI. „Evaluation in a cross-linguistic perspective: A corpus-assisted study of English and German medical discourse“. *Linguistica e Filologia* 39 (2019): 65–95. Print.
- MOŽNAROVÁ, Michaela. *Near-synonyms in English and Slovak: A corpus-based study* [Bachelor's thesis, Masaryk University], 2018. MUNI Theses
- OXFORD UNIVERSITY PRESS. *Oxford English Dictionary online*. <https://www.oed.com/>. 4.7.2025.

- POLIZZI, Daniele, Silvia BERNARDINI and Adriano FERRARESI. „Evaluation in a cross-linguistic perspective: Investigating semantic prosody across English and German near-synonyms“. *Journal of Corpora and Discourse Studies* 7.1 (2024): 77–100. Print.
- SMITH, Mike. „A Corpus-Driven Study of Near-Synonymous Command Verbs Used in a Military English Context“. *Kwansei Gakuin University Humanities Review* 25.1 (2021): 69–86.
- TAYLOR, John R. „Near synonyms as co-extensive categories: ‘high’ and ‘tall’ revisited“. *Language Sciences* 25.3 (2002): 263–284.
- TOMÁŠKOVÁ, Simona and Edina BOROVSÁ. „Lockdown vs. Shutdown: Zur Semantik der partiellen Synonyme im Deutschen“. *Jazyk a kultúra* 47–48 (2021): 69–85.
- YANG, Bin. „A corpus-based comparative study of learn and acquire“. *English Language Teaching* 9.1 (2016): 209–216.

Author’s Declaration on the Use of Generative Artificial Intelligence (GenAI) Tools

During the preparation of this work the authors used [ChatGPT] in order to plan the structure/outline of the paper, language editing in all respects. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

ZITIERNACHWEIS:

- KALAŠ, Filip, LIPTÁK, Pavol. „From *Cheap* to *Affordable* – Corpus-Assisted Exploration of Near-Synonymy in English, German, and Slovak Economic Discourse“, *Linguistische Treffen in Wrocław* 29, 2026 (1): 143–158. DOI: 10.23817/lingtreff.29-9.