

LENKA PTAK

University of Wrocław, Faculty of Languages, Literatures and Cultures

lenka.ptak@uwr.edu.pl

ORCID: 0000-0002-9691-3298

## Polish and English Equivalents of Czech Negative Adjectives Ending in *-telný* in the Multilingual Parallel Corpus InterCorp

### Abstract

This paper explores the use of a multilingual parallel corpus in the linguistic analysis of a specific category of deverbal adjectives in English, Czech, and Polish that convey the notion of possibility or potentiality. These adjectives exhibit high morphological productivity and semantic compositionality, particularly in Czech, where their meanings are systematically encoded through specialised derivational suffixes. Previous research suggests that the number of Polish adjectives ending in *-alny*, as documented in lexicographical sources, is lower than their corresponding counterparts in Czech and English. As a result, these adjectives in Polish often have distinct equivalents. The selected adjectives are examined using the Czech multilingual parallel corpus InterCorp, with the aim of highlighting the innovative potential and limitations of parallel corpora in comparative linguistic research. The study specifically investigates the usefulness of multilingual corpora for extracting collocational patterns from textual data.

**Keywords:** comparative linguistics, corpus linguistics, word formation, parallel corpora, deverbal adjectives, derivational suffixes, translation studies

In this paper, I focus on an interlingual comparison of deverbal adjectives denoting possibility and impossibility in Czech, Polish, and English. This constitutes a well-defined semantic category, encompassing adjectives that express whether a given entity can or cannot undergo the action denoted by the verbal root. My research builds on the findings presented in *Przymiotniki o znaczeniu potencjalnym w językach czeskim, słowackim i polskim* (2009), which analyses linguistic data on the basis of academic

dictionary entries<sup>1</sup>. However, in this article, my primary data source is the parallel corpus InterCorp, which has been predominantly compiled by faculty and students at Charles University in Prague.

I aim to focus on several key issues, including the possibility of extracting collocations from specific texts, the use of corpus data to determine optimal translation equivalents, and the analysis of linguistic patterns. Thanks to linguistic corpora, this material can be acquired more efficiently and, most importantly, selected, sorted, or classified at the preliminary stage. This article provides a brief presentation of an automated method for extracting specific language units from diverse texts and determining their equivalents in Czech, Polish, and English translation materials. Corpora provide a large amount of clear and organised data, allowing many hypotheses and theories to be verified more reliably and efficiently, thereby enhancing linguistic research. The thematic scope of this article is intentionally limited to material derived from a parallel corpus, in order to showcase the innovative possibilities of corpus linguistics.

In analysing the linguistic material, I rely on Polish methodological traditions, as presented in academic grammars such as *Zarys słowotwórstwa polskiego* (Grzegorzczkowska [1979] 1979), *Gramatyka współczesnego języka polskiego* (Kallas 1998), and *Zarys gramatyki polskiej (ze słowotwórstwem)* (Nagórko 2000). The term “potential adjectives” is used in accordance with the terminology found in the syntheses of Polish word formation by Alicja Nagórko (1987) and Krystyna Kallas (1998). By conducting a detailed analysis of the lexical material excerpted from the corpus in Czech, Polish, and English, I considered the word-formation parameters outlined in existing studies of word formation in these languages, which were subsequently used in the contrastive analysis. I primarily draw on the theoretical concepts of Miloš Dokulil (1962, 1963) and Mark Aronoff (1976).

### The Multilingual Parallel Corpus InterCorp

Czech, Polish, and English adjectives expressing possibility (or impossibility) exhibit a high degree of correspondence in their functions within texts and translations. To analyse the Polish and English equivalents of Czech adjectives with the suffixes *-telný/-itelný*, I used the parallel corpus InterCorp version 16 in this article, which forms part of the Czech National Corpus and was released in 2023. The latest version of InterCorp contains a total of 387 million words in the aligned foreign-language texts of the core part and 4,506 million words in the collections. The number of words in the Czech texts is 125 million in the core part and 273 million in the collections (Rosen, Vavřín, Zasina 2022). The current structure of the Czech National Corpus is illustrated in Table 1.

1 In the analysis of Czech adjectives, the primary resource was the academic dictionary *Slovník spisovného jazyka českého* (Praha 1960–1971), which contains over 192,000 entries. For Slovak adjectives, the *Slovník slovenského jazyka* (Bratislava 1959–1968), comprising more than 100,000 entries, served as a foundational source. Additionally, Jozef Mistrík's *Retrográdny slovník slovenčiny* (Bratislava 1976) provided a valuable repository of potential adjectives ending in *-telný*. The *Krátky slovník slovenského jazyka* (edited by Ján Kačala and Mária Pisárčiková, Bratislava 1997), which includes approximately 60,000 entries and is also available in an electronic version, was extensively used as well. The material for Polish adjectives was sourced from Witold Doroszewski's *Słownik języka polskiego* (1958–1969), employing the reverse index (*Index a tergo*, edited by Jadwiga Puzynina and Renata Grzegorzczkowska, Warszawa 1973) of this dictionary. Examples were further supplemented with data from Stanisław Dubisz's *Uniwersalny słownik języka polskiego* (Warszawa 2003) and additional studies by various scholars.

Table 1. Structure of the Czech National Corpus.

<b>Corpora of the Czech National Corpus</b>	
<b>Written synchronic corpora</b>	<b>Spoken synchronic corpora</b>
<b>General corpora</b> SYN series corpora	<b>General corpora</b> ORATOR ORTOFON ORAL
<b>Web corpora</b> ONLINE NET	<b>Specialised corpora</b> BMK DIALEKT LINDSEI CZ PMK SCHOLA2010 SPEECHES PARLCORP
<b>Learner corpora</b>	
<b>Author corpora</b>	
<b>Specialised corpora</b>	
<b>Foreign language corpora</b> Parallel corpora Comparable corpora Specialised foreign language corpora	
<b>Written diachronic corpora</b> DIACORP OnomOs	

The multilingual parallel corpus InterCorp can be considered referential, which means that all its previous releases remain available in their original form. The corpus consists of a core section, primarily composed of fiction with manually checked alignments and collections. The collections contain texts acquired in multiple languages that are processed and aligned automatically. They do not always include all texts from the sources, particularly those without a Czech counterpart. Each text in the corpus has a Czech counterpart, establishing Czech as the pivot language. For every text, there is a single Czech version (either original or translated), aligned with one or more foreign-language versions (Čermák, Rosen 2012).

### **Passive Potential Adjectives**

The adjectives under analysis are referred to in some Czech studies as “adjectives with the meaning of possibility” or “passive adjectives denoting possibility,” while in English they are commonly termed “passive adjectives denoting possibility.” In the Polish linguistic tradition, these adjectives are identified as “potential adjectives.” The notion of potentiality is a polysemous concept with various terminological applications in word-formation processes. In Polish linguistics, the term encompasses a range of interpretations and has a broad and rather fluid meaning (see Puzynina 1966: 337; Buttler 1962: 239; Chruścińska 1978: 72, 78; Nagórko 2000: 174).

Krystyna Kallas, in her *Gramatyka współczesnego języka polskiego* (1998), categorises potential adjectives as a subclass of deverbal derivatives and defines them as predispositional. According to Kallas, these adjectives encode additional information about predisposition; specifically, a tendency or inherent capacity to perform or undergo an action – alongside an indication of a particular state. For instance, Kallas uses the derivative *uleczalny* (“curable”) as an example, highlighting the inclusion of a modal predicate that conveys information regarding the potentiality or feasibility of the action (Kallas 1998: 471).

Isabel Oltra-Massuet identifies several types of *-ble* adjectives:

- (1) active adjectives, such as *agreeable* and *perishable*, meaning “that V”;
- (2) passive adjectives, such as *modifiable* and *realizable*, meaning “that can be Ved”;
- (3) adjectives denoting possibility, again exemplified by *modifiable* and *realizable*, meaning “that can be Ved”;
- (4) adjectives expressing evaluative judgement, such as *admirable* and *envious*, meaning “that should be Ved”; and
- (5) causative adjectives, such as *horrible* and *terrible*, meaning “that causes N” (Oltra-Massuet 2013: 13).

In Czech, the suffixes *-itelný/-itelný* function as productive morphemes in the formation of potential adjectives. These suffixes are specifically used in the derivational processes of adjectives that retain a verbal base, which distinguishes their application from analogous formations in Polish or English. Unlike in those languages, where *-alny* and *-ble* adjectives may also derive from non-verbal roots, the Czech system restricts *-itelný/-itelný* to the verbal domain.

In Polish, deverbal adjectives expressing potentiality are primarily formed using the suffix *-alny*. Unlike in Czech, where this suffix is largely restricted to verbal roots, in Polish *-alny* is also employed in the formation of desubstantive adjectives, often derived from nouns of foreign origin. Examples include *czensualny* ← *czenzura* (“censorship”), *eksperymentalny* ← *eksperyment* (“experiment”), *kontynentalny* ← *kontynent* (“continent”), and *sentymentalny* ← *sentymet* (“sentiment”) (Ptak 2009: 87). Among Polish adjectives ending in *-alny*, a significant subset does not convey potentiality but instead denotes the direct execution of, or subjection to, an activity. For instance, *grzebalny* (“służący do grzebania zmarłych, związany z obrzędem grzebania” – “serving to bury the dead, related to the burial ritual”) refers to a functional or relational role rather than indicating any potentiality.

Additionally, certain adjectives ending in *-alny* can shift between a predispositional and a non-predispositional interpretation, depending on the context. For example, *jadalny* can denote “one that can be eaten” (*taki, który można zjeść*) when referring to a mushroom (*jadalny grzyb*), thereby conveying potentiality. In contrast, in the phrase *jadalny pokój* (*taki, w którym się jada* – “a room where one eats”), the adjective expresses a functional relation without implying potentiality (see *Gramatyka współczesnego języka polskiego* 1998: 471).

English passive adjectives expressing possibility or potentiality, which can be paraphrased as “fit for being V-ed” or “liable to be V-ed,” are predominantly formed using the suffixes *-able* and *-ible*. According to Bogdan Szymanek (1985), several synchronic processes in English word formation complicate the regularity of *-able* derivatives. He observes that a limited number of *-able* adjectives are derived from nouns

rather than verbs, as in *fashion* → *fashionable*. These derivatives are relatively rare, and their semantics can be unpredictable. However, for certain adjectives of this type, an implicit sense of either activeness (e.g., *profitable*) or passiveness (e.g., *clubbable*) can be identified (Szymanek 1985: 97).

Also pertinent to this discussion is Isabel Oltra-Massuet's observation that certain *-ble* adjectives are not exclusively derived from verbal bases, and as a result, the subject may not correspond to the direct object of an underlying verb. Furthermore, some *-ble* adjectives, such as *perishable*, do not appear to convey a passive meaning. Additionally, certain adjectives express a modality of necessity rather than possibility, as seen in *admirable* or *deplorable*. In some cases, adjectives like *suitable* have even been argued to lack the expected modality altogether (Oltra-Massuet 2013: 24–25).

Although the languages under discussion employ highly productive and semantically and structurally regular suffixes in the derivation of adjectives denoting potentiality, each language also uses secondary suffixes in the formation of such adjectives. These secondary suffixes frequently extend beyond the basic expression of possibility or impossibility, encompassing meanings related to habitual actions, quantitative aspects, or evaluative judgements. In Polish, for instance, these formative suffixes include *-liwy*, *-ony*, *-any*, and *-ny* (see Nagórko 1987; Gramatyka 1998). As an example, we can examine the Polish adjectives *przeziąkalny* and *przeziąkliwy*. These examples illustrate subtle semantic differences between the two adjectives, with *przeziąkalny* emphasising potentiality, while *przeziąkliwy* conveys an inherent tendency or characteristic: *przeziąkalny materiał* – “taki, który może (łatwo) przeziąkać” (“one that can (easily) permeate”); *przeziąkliwy materiał* – “taki, który łatwo przeziąka” (“one that easily permeates”).

In terms of semantics and morphological derivation, adjectives with the suffix *-telný/-itelný* in Czech exhibit a close paradigmatic relationship with adjectives ending in *-ný* (e.g., *kujný* “malleable,” *ohébný* “flexible,” *podplátný* “bribeable,” *rozbitný* “breakable,” *rozpustný* “soluble,” *únosný* “bearable,” *výsuvný* “extendable”). These forms frequently possess an active or potential meaning, denoting a simple relation to an associated verbal action or process. Furthermore, they demonstrate significant morphological affinity with adjectives bearing the suffix *-lý* (e.g., *jedlý* “edible,” *nevystihlý* “ungraspable,” *neochvělý* “unshakable”), as well as *-ivý* (e.g., *netříštivý* “non-shattering”), *-ací* (e.g., *posunovací* “shiftable,” *rozbírací* “dismantlable,” *zasunovací* “insertable”), and *-avý* (e.g., *otvíravý* “openable”). In certain instances, these adjective types can form synonymous or near-synonymous pairs with *-telný/-itelný* adjectives, reflecting overlapping semantic domains.

As examples of complete synonyms, we can reference several pairs noted in the Czech dictionary *Slovník spisovného jazyka českého*: *dobyitelný* // *dobytný* – for instance, *snadno dobyitelný* // *dobytný hrad*, meaning “a castle that can be easily conquered”; *neotřesitelný* // *neotřestný* – as in *pevná a neotřesitelná nebo neotřesná vůle*, meaning “a firm and unshakeable will”; *nezkrotitelný* // *nezkrotný* – for example, *nezkrotitelná, nezkrotná síla*, meaning “a force that cannot be tamed or subdued”; and *rozbitelný* // *rozbitný* – such as *snadno rozbitelné* // *rozbitné sklo*, meaning “glass that can easily be broken.” In rare cases, synonymic triplets also occur in Czech, formed with the suffixes *-telný/-itelný* // *-ný* // *-lý*. Examples include: *nepostižitelný* // *nepostižný* // *nepostihlý*, *nestižitelný* // *nestižný* // *nestihlý*, and *nevystižitelný* // *nevystižný* (rare) // *nevystihlý* (literary).

## Analysis of Linguistic Material: Polish and English Equivalents of Czech Adjectives with the Suffix *-telný/-itelný* in InterCorp

As previously mentioned, the primary source of linguistic data for this article is the parallel corpus InterCorp version 16. The linguistic material under examination consists of examples from this corpus, specifically pairs of equivalents automatically generated using word segment alignment methods, along with their frequency lists. Additionally, data from the Czech–Polish and English–Polish bilingual dictionaries will be used. The linguistic data retrieved from the parallel corpus has been deliberately restricted to original texts. This is necessary because, although each parallel corpus contains identical texts in two languages, they are not always direct translations; they may also be translations from a third language. While the InterCorp parallel corpus, part of the Czech National Corpus, supports bidirectional searches, it does not guarantee that translations from Czech into other languages are always based on original Czech texts.

To address this, we employed the CQL (Corpus Query Language) command to search within corpus sources. This method allows for complex searches involving various combinations of words, forms, and character strings with precisely defined syntax. Using the concordances (contextual extracts of the search results) obtained from the relevant corpus, we created frequency lists of the searched lemmas (base word forms that encompass occurrences of all their inflected forms). We employed the CQL command in the form [word=" +teln.+ " & upos="ADJ"], which allowed us to search for Czech adjectives containing the suffix *-telný* in the corpus, specifically targeting adjectives (indicated by the tag "ADJ"). This query was then refined to focus on forms containing the negative morpheme *ne-*, using the modified command [word=" ne.+teln.+ " & upos="ADJ"], thereby narrowing the results to adjectives expressing negation, and the original texts, and aligning the three languages together. The query constructed in this manner also retrieves instances of the unnegated superlative form, such as *nejhmatatelnější* "the most tangible," *nejviditelnější* "the most visible," *nejzrozumitelnější* "the most comprehensible," which were subsequently removed through manual filtering.

From the lexicographic material of *Słownik czesko-polski* ([1991] 2010) and *Velký česko-anglický slovník* (Fronek [2000] 2013), 84 negative deverbal adjectives ending in *-telný/-itelný* were extracted as primary entries, alongside 172 negative deverbal adjectives ending in *-telný/-itelný* extracted from *Velký česko-anglický slovník* ([2000] 2013). For practical reasons – namely the inability to exhaustively cover all derivatives in this article – only selected negative adjectives ending in *-telný/-itelný* and their Polish or English equivalents were analysed. These were restricted to those recorded simultaneously in the Czech–Polish dictionary, the Czech–English dictionary, and the InterCorp corpus, and whose frequency in the corpus was at least two occurrences.<sup>2</sup> The absence of certain Polish equivalents for Czech adjectives ending in *-telný/-itelný* in Tables 2 and 2a does not imply that they are nonexistent in modern Polish. The list of derivatives extracted from the InterCorp corpus, as well as from the Czech–Polish dictionary,

---

2 In this article, we reference the research presented in the paper "O polskich odpowiednikach czeskich przymiotników potencjalnych w korpusie InterCorp oraz w słownikach dwujęzycznych" (Ptak 2015), which demonstrated that the parallel corpus *InterCorp* version 8 provides a significantly broader range of translation equivalents for source words than traditional bilingual dictionaries. The study underscores the potential of *InterCorp* as an invaluable tool for identifying and translating analogous formations across the 38 languages available in the corpus at the time the research was conducted, showcasing its utility in offering richer linguistic insights and translation possibilities than the dictionaries.

could be expanded with examples drawn from authoritative academic dictionaries of the Polish language, national reference corpora, and personal annotations. For instance, Doroszewski's *Słownik języka polskiego* (SJP Dor) lists 106 negative deverbal adjectives ending in *-alny*. However, as previously noted, the scope of this article was deliberately confined to data sourced from the InterCorp corpus and bilingual dictionaries, with the aim of illustrating the innovative potential of corpus linguistics.

Polish equivalents of Czech negative adjectives ending in *-telný/-itelný* that appear solely in the InterCorp corpus are indicated with an asterisk (\*) preceding the word. In Tables 2 and 2a, the frequency of the lexical items is presented both in the original texts and in their corresponding translations within the corpus. Frequency analysis of lexical items within a defined text corpus is one of the most frequently used methodologies in corpus linguistics. Indeed, word frequency lists often serve as foundational data for lexicographic work and dictionary compilation.

Table 2. Czech negative adjectives with the suffix *-telný/-itelný* and their Polish and English equivalents in the parallel corpus InterCorp and in the dictionaries *Słownik czesko-polski* (Basaj and Siatkowski [1991] 2010) and *Velký česko-anglický slovník* ([2000] 2013).

Czech negative adjectives with the suffix <i>-telný/-itelný</i> (frequency in InterCorp)	Polish and English equivalents of Czech negative adjectives with the suffix <i>-telný/-itelný</i>				
	Adjectives with the suffix <i>-alny</i> and <i>-ble</i> (frequency in InterCorp)	Polish adjectives with the suffix <i>-ny</i> (frequency in InterCorp)	Polish adjectives with the suffix <i>-liwy</i> (frequency in InterCorp)	Polish adjectives with the suffix <i>-ly</i> (frequency in InterCorp)	Polish adjectives with the suffix <i>-ony, -ty</i> (frequency in InterCorp)
<i>nedělitelný</i> (4)	<i>indivisible</i> (4)	<i>niepodzielny</i> (4)			
<i>nedocenitelný</i> (2)	<i>invaluable</i> (1) <i>inestimable</i> (0)				<i>nieocenyony</i> (0)
<i>nedotknutelný</i> (4)	<i>nietykalny</i> (2) <i>nienaruszalny</i> (1) <i>integralny</i> (0) <i>untouchable</i> (1) <i>inviolable</i> (1) <i>*inalienable</i> (1)	<i>niezmienny</i> (0)			<i>nienaruszony</i> (1)
<i>nenahraditelný</i> (7)	<i>*niepowtarzalny</i> (1) <i>irreplaceable</i> (7) <i>irretrievable</i> (0) <i>irredeemable</i> (0) <i>irreparable</i> (0)				<i>niezastapiony</i> (5)
<i>nenapodobitelný</i> (7)	<i>*niepowtarzalny</i> (5) <i>inimitable</i> (6)				
<i>nenapravitelný</i> (11)	<i>incorrigible</i> (8) <i>irreparable</i> (2) <i>irretrievable</i> (0) <i>irredeemable</i> (0)	<i>niepoprawny</i> (9)			

Czech negative adjectives with the suffix <i>-telný/-itelný</i> (frequency in <i>InterCorp</i> )	Polish and English equivalents of Czech negative adjectives with the suffix <i>-telný/-itelný</i>				
	Adjectives with the suffix <i>-alny</i> and <i>-ble</i> (frequency in <i>InterCorp</i> )	Polish adjectives with the suffix <i>-ny</i> (frequency in <i>InterCorp</i> )	Polish adjectives with the suffix <i>-liwy</i> (frequency in <i>InterCorp</i> )	Polish adjectives with the suffix <i>-ły</i> (frequency in <i>InterCorp</i> )	Polish adjectives with the suffix <i>-ony, -ty</i> (frequency in <i>InterCorp</i> )
<i>neodolatelny</i> (6)	<i>irresistible</i> (3) <i>uncontrollable</i> (0) <i>*irrepressible</i> (1)	<i>*niepohamowany</i> (1) <i>*niezłomny</i> (1)			<i>nieodparty</i> (1) <i>nieprzewyciężony</i> (0)
<i>neotřesitelný</i> (3)	<i>*unshakable</i> (2)	<i>niezachwiany</i> (2) <i>niezłomny</i> (0)			
<i>nepochopitelný</i> (22)	<i>incomprehensible</i> (17) <i>*inconceivable</i> (1) <i>*inexplicable</i> (3)			<i>niezrozumiały</i> (12)	<i>niepojęty</i> (8)
<i>nepopsatelný</i> (8)	<i>indescribable</i> (7) <i>unspeakable</i> (0) <i>indefinable</i> (0) <i>ineffable</i> (1)	<i>nieopisany</i> (4) <i>niewymowny</i> (0)			
<i>neviditelný</i> (48)	<i>niewidzialny</i> (34)	<i>niewidoczny</i> (10)			
<i>nevysvětlitelný</i> (16)	<i>*niewytłumaczalny</i> (8) <i>unexplainable</i> (0) <i>inexplicable</i> (10)	<i>niejasny</i> (0)		<i>niezrozumiały</i> (1)	<i>*niepojęty</i> (1) <i>*niewyjaśniony</i> (1)

Table 2a. Czech negative adjectives with the suffix *-telný/-itelný* and their Polish and English equivalents in the parallel corpus *InterCorp* and in the dictionaries *Słownik czesko-polski* ([1991] 2010) and *Velký česko-anglický slovník* ([2000] 2013).

Czech negative adjectives with the suffix <i>-telný/-itelný</i> (frequency in <i>InterCorp</i> )	Polish and English equivalents of Czech negative adjectives with the suffixes <i>-telný/-itelný</i>	
	The construction <i>do+genitive</i> of a deverbal noun (frequency in <i>InterCorp</i> )	Adjectives ending with other suffixes, descriptive phrases, or adjectives that do not express the full content of the given structure (frequency in <i>InterCorp</i> )
<i>nedělitelný</i> (4)		
<i>nedocentelný</i> (2)		<i>*bezcenny</i> (1) <i>*priceless</i> (1)
<i>nedotknutelný</i> (4)		<i>sacrosanct</i> (0)
<i>nenahraditelný</i> (7)		<i>*niedający się zastąpić</i> (1) <i>*bezcenny</i> (1)

Czech negative adjectives with the suffix <i>-telný/-itelný</i> (frequency in <i>InterCorp</i> )	Polish and English equivalents of Czech negative adjectives with the suffixes <i>-telný/-itelný</i>	
	The construction <i>do+genitive</i> of a deverbal noun (frequency in <i>InterCorp</i> )	Adjectives ending with other suffixes, descriptive phrases, or adjectives that do not express the full content of the given structure (frequency in <i>InterCorp</i> )
<i>nenapodobitelny</i> (7)	<i>nie do naśladowania</i> (0) * <i>niemożliwy do naśladowania</i> (1) * <i>nie do podrobienia</i> (1)	
<i>nenapravitelny</i> (11)	<i>nie do naprawienia</i> (2)	<i>hopeless</i> (1)
<i>neodolatelny</i> (6)		<i>zniewalający</i> (0) * <i>przemóżny</i> (1) * <i>tajemniczy</i> (1) * <i>overwhelming</i> (1) <i>overpowering</i> (0) <i>compelling</i> (0) <i>magnetic</i> (0)
<i>neotřesitelny</i> (3)		* <i>taki, który nie daje się zbić z tropu</i> (1) <i>steadfast</i> (0) <i>unswerving</i> (0) * <i>unshaken</i> (1)
<i>nepochopitelny</i> (22)		<i>przekraczający możliwości zrozumienia</i> (1) <i>tajemniczy</i> (1) <i>baffling</i> (0) <i>perplexing</i> (0) <i>beyond all understanding</i> (0) <i>it defies reason</i> (0) * <i>hard to understand</i> (1)
<i>nepopsatelny</i> (8)	<i>nie do opisania</i> (1)	* <i>niezapomniany</i> (1) * <i>ogłuszający</i> (1) <i>beyond all description</i> (0)
<i>neviditelny</i> (50)		* <i>niesprecyzowany</i> (1)
<i>nevysvětlitelny</i> (16)	* <i>nie do wytłumaczenia</i> (1)	<i>nie dający się wyświecić (wyjaśnić)</i> (0) * <i>nie dający się wytłumaczyć</i> (1) * <i>bezsensowny</i> (1) * <i>no longer explicable</i> (1) * <i>without explanation</i> (1) * <i>incredible</i> (1)

The use of adjectives with the negative morpheme **ne-** is well represented in Czech. The ease with which these adjectives can be negated is likely attributable to the syntactic role that adjectives fulfil as a part of speech. This phenomenon is also connected to the modal nature of the adjectives under investigation, particularly in expressing the impossibility of an action's occurrence. In academic dictionaries of Czech, Polish, or English, word-formation paraphrases are used to describe these derivatives. Czech examples

include phrases such as *takový, který nelze, takový, který se nedá* and *takový, který nemůže být*; Polish equivalents include *niedający się, niemożliwy do* and *niemogący być*; while English equivalents include *not fit for being V-ed, not liable to be V-ed*.<sup>3</sup>

In the parallel corpus InterCorp, we identified **751 occurrences**<sup>4</sup> of adjectives with the suffix *-telný/-itelný*, of which **698 instances** were classified as negative potential adjectives with this suffix. Adjectives with unclear word-formation structures, such as *nesmrtelný* “immortal” (30 occurrences), *nezřetelný* “indistinct” (5 occurrences), *neprožřetelný* “imprudent” (1 occurrence), and *nevyzpytatelný* “unpredictable” (2 occurrences), were excluded from the analysis. After generating frequency lists for the relevant lemmas, we obtained **138 distinct examples** of negative adjectives with the *-telný/-itelný* suffix, each with varying frequencies. The most frequent adjective was *neuvěřitelný* “unbelievable,” which appeared 100 times in the corpus, while **63 adjectives** were observed with a frequency of just one occurrence each: *nedefinovatelný, nehmatatelný, neměřitelný, nenapadnutelný, nenaučitelný, neobsáhnutelný, neobyvatelný, neoddiskutovatelný, neodhalitelný, neodstranitelný, neoholitelný, neosvojitelný, nepopíratelný, neporaditelný, neporušitelný, nepostradatelný, nepotopitelný, nepozorovatelný, neprokazatelný, nepřeložitelný, nepřeslechnutelný, nepřestupitelný, nepřevtělitelný, nerecyklovatelný, nereformovatelný, nerekonstruovatelný, nerozlišitelný, nerozložitelný, nerozlužitelný, nesdělitelný, neslibatelný, nesmazatelný, nesouměřitelný, neuhasitelný, neuhlídatelný, neukojitelný, neukončitelný, neurčitelný, nevnímátný, nevyhodnotitelný, nevyhojitelný, nevyjádřitelný, nevykořenitelný, nevykoupitelný, nevypátratelný, nevypověditelný, nevyvěratelný, nevzkřísitelný, nezahubitelný, nezachytitelný, nezanedbatelný, nezastavitelný, nezbadatelný, nezkratitelný, nezlikvidovatelný, nezměnitelný, neznatelný, neznormalizovatelný, nezodpověditelný, nezrušitelný, nezvratitelný.*

Out of the 63 examples, only 11 instances involved translating a Czech adjective with the suffix *-telný/-itelný* into a Polish adjective with the suffix *-alny* (e.g., *nienaruszalny, nieodwracalny, niepodważalny, niepoznawalny, nieprzetłumaczalny, niereformowalny, nieuleczalny, niewyobrazalny, niezatapialny, niezauważalny, niezniszczalny*), which clearly predominates within this onomasiological category. In other Polish translations, various strategies were used: adjectives with the suffix *-ny* were employed 12 times, the construction *do + genitive* of a deverbal noun appeared in 7 instances, adjectives with other

3 To provide a more comprehensive account of the usage of negative passive potential adjectives in contemporary Czech and Polish, data obtained from academic and bilingual dictionaries may be complemented with evidence drawn from the *Czech language corpus Syn2020* and the *National Corpus of Polish*.

In the *Syn2020* corpus, the following CQL query was employed: [word = “ne.+teln.+” & pos = “A”]. This search yielded more than 30,000 instances of negative adjectives containing the suffix *-telný/-itelný*, offering a substantial empirical basis for further analysis. To obtain the highest possible number of relevant attestations from the National Corpus of Polish, we used the Poliqarp search engine and the CQL query [base = “nie.+aln.+” & pos = adj]. As previously noted, this type of query enables the retrieval of various combinations of words, forms, and other character sequences using precisely defined syntax. From the resulting concordances – ordered sets of citations featuring a given word or construction – we generated a frequency list that included all inflected forms of the retrieved adjectives. Since the suffix *-alny* appears in the derivation of both denominal and deverbal adjectives, the dataset initially included over 500 denominal adjectives with the prefix *nie-*, such as “nieaktualny,” “nieformalny,” “nielegalny,” “niematerialny,” “nienormalny,” and “nieprofesjonalny.” After a thorough verification of the attestations, we identified 420 valid instances of passive potential adjectives with the suffix *-alny*, including examples such as “niepodważalny” (“undeniable”), “nieporównywalny” (“incomparable”), “nieprzewidywalny” (“unpredictable”), “nierozdzielny” (“indissoluble”), “nierozpuszczalny” (“insoluble”), “nieuleczalny” (“incurable”), “niewidzialny” (“invisible”), and “niewybaczalny” (“unforgivable”).

4 The application of individual search filters led to a marked reduction in the range of textual sources, which significantly decreased the number of occurrences.

suffixes were used 8 times, descriptive phrases occurred 9 times, and in 17 cases, adjectives or descriptive expressions that did not fully convey the meaning of the original structure were used, or no appropriate adjectival translation was provided. In English, a different tendency was observed, with the majority of these adjectives (41 instances) being translated using the suffix *-ble*.

From the dictionary material of the *Słownik czesko-polski* ([1991] 2010), we extracted 86 negative potential adjectives with the suffix *-telný/-itelný*, and from the *Velký česko-anglický slovník* ([2000] 2013), we extracted 172 negative potential adjectives with the same suffix. A total of 63 adjectives was recorded simultaneously in the Czech-Polish dictionary, the Czech-English dictionary, and the InterCorp corpus. Some of these adjectives, specifically those with a frequency of 2 or more occurrences, are presented in Tables 2 and 2a.

In the case of the corpus, the diversity of Polish equivalents was achieved primarily through the use of descriptive phrases or adjectives, which often fail to fully convey the complete meaning of the original term. As demonstrated by our previous detailed analysis of adjectives denoting possibility (Ptak 2009), Czech adjectives with the suffix *-telný/-itelný* are more numerous than Polish potential adjectives with the suffix *-alny*. In Czech, the total number of derivatives with the suffix *-telný/-itelný*, according to the *Slovník spisovného jazyka českého* (SSJČ), amounts to 625. In contrast, in Polish, based on dictionary data from Doroszewski's *Słownik języka polskiego* (SJP Dor), there is a significant difference, with only 252 adjectives formed with the suffix *-alny*, a disparity noted in Ptak (2009: 134–135).

For example, the Czech adjective *nepopsatelný* is translated in the Czech-Polish dictionary as *nie do opisania*, *nieopisany*, or *niewymowny*. In the *Dictionary of the Czech Language* (*Slovník spisovného jazyka českého*, 1971), this adjective is paraphrased as *takový, který se nedá popsat, vymykající se jakémukoli popisu*. In the parallel corpus InterCorp, we recorded 8 instances of the adjective *nepopsatelný*, 4 of which were translated as *nieopisany* (1a), one as *niezapomniany*, and one using the construction *do* + genitive of a deverbal noun. In other instances, this adjective was translated using descriptive phrases or adjectives that did not fully capture the entire meaning of the term or the modal notion of possibility, or it was left untranslated altogether (1b). In English, it was consistently translated using the suffix *-ble*.

- (1) a. **C** ... *ve vzdálenosti asi dvou set metrů viděla plameny a z jiného místa, blíž u ní, se zvedal z příkopu k temnému nebi stále ten stejný nepopsatelný, strašný křik* – **P** ... *jakieś dwieście metrów dalej zobaczyła płomienie, a z rowu, bliżej niej, wznosił się wciąż ku ciemnemu niebu ten sam przeraźliwy krzyk* – **E** ... *about two hundred yards away she saw flames, and from another spot, closer to her, that same indescribable, terrible screaming still rose;*
- b. **C** *Hajlování přešlo v nepopsatelný řev.* – **P** *Hajlowanie przerodziło się w ogłuszający ryk.* – **E** *The siegheiling disintegrated into indescribable pandemonium.*

A parallel corpus should not be conflated with an “automatic translator” or a translation memory system used in information technology. Achieving more refined results, such as filtering collocations, limiting searches to specific documents or text types, or conducting more advanced queries requires a basic understanding of corpus tools and, above all, patience from the researcher. Not all equivalents are appropriate or contextually suitable:

- (2) **C** *Za jednoznačný a nezpochybnitelný fakt považují pouze rostoucí závislost zemí EU na dovozu energetických surovin.* – **P** ... *zmian [ y ] klimatycz [ e ] ... wywierac będą duży wpływ na ludność.* – **E** *The first point I want to make is that the increasing dependency of EU countries on imports of raw energy materials is the only unambiguous fact, about which there is no doubt.*

- (3) *C Historická příslušnost Turecka k Evropě je **nezpochybnitelná**. – P Patrząc realnie, Turcja jest obecnie członkiem Rady Europy oraz Organizacji Bezpieczeństwa i Współpracy w Europie. – E Turkey's historic affiliation with Europe is **undeniable**.*

The inadequacy of examples (2) and (3) arises from the incorrect alignment of concordances in the Polish language and the improper tagging of individual elements. A key advantage of InterCorp is the ability to analyse query results in both immediate and broader textual contexts, to further filter the identified concordances, and to verify their collocability in both languages. However, a notable drawback is the requirement to manually review the concordances in the “target” language and select the appropriate equivalents.

### Conclusion

Differences between the individual languages are most apparent in the number of derivatives formed using primary suffixes within the examined onomasiological category. As our analysis of adjectives expressing possibility demonstrates, English and Czech adjectives with the suffixes *-ble* and *-telný/-itelný*, respectively, are more numerous than their Polish counterparts with the suffix *-alny*. Consequently, Polish often employs a range of equivalents for these derivatives.

For our research, we relied primarily on linguistic data extracted from the InterCorp corpus. This corpus-based material was further supplemented with lexicographic data and additional evidence drawn from the Syn2020 corpus and the National Corpus of Polish. One of the key advantages of using corpora lies in the vast range of available source texts and, crucially, in the ability to automatically search and analyse linguistic items based on their frequency and distribution within authentic contexts. To ensure consistency and comparability, we deliberately restricted our corpus searches to original Czech texts and their translations into English and Polish, focusing on a dataset of 63 texts translated from Czech into both target languages.

After generating frequency lists for the searched lemmas, we identified 698 instances of negative adjectives ending in *-telný/-itelný* with varying frequencies. Additionally, we extracted 84 negative potential adjectives with the *-telný/-itelný* suffix from Czech–Polish dictionary data and 172 from Czech–English dictionary data. The analysis of this linguistic data supports our initial hypothesis: Czech adjectives with the suffix *-telný/-itelný* and English adjectives with the suffix *-ble* are more numerous than their Polish equivalents ending in *-alny*. In numerous cases, Polish lacks a direct equivalent for these adjectives, which leads to context-dependent translations. As a result, Polish often resorts to alternative descriptive phrases or different adjectives, which may not fully convey the precise meaning of the original form or the potentiality of the action expressed by the adjective.

The fact that the scope and textual diversity of a parallel corpus are essential for effective statistical analysis does not necessarily preclude its suitability for translating specific units. As shown by our analysis, we are able to narrow our corpus search successfully. Overall, it may be concluded that the InterCorp parallel corpus is a tool that can be effectively used for translation and for determining possible equivalents. And although a linguist using parallel corpora encounters a number of methodological problems, their very existence represents a significant advance in the possibilities for linguistic analysis. With its growing size, the InterCorp parallel corpus is becoming increasingly important in contrastive research.

## References

- Aronoff, Mark (1976) *Word formation in generative grammar*. Cambridge, Massachusetts, and London, England: MIT Press.
- Basaj, Mieczysław, Janusz Siatkowski (eds.) ([1991] 2002) *Słownik czesko-polski*. Second Edition. Warszawa: Wiedza Powszechna.
- Buttler, Danuta (1967) "Rekcja przymiotników odczasownikowych." [In:] *Poradnik Językowy*. Vol. 8; 357–370.
- Čermák, František, Alexandr Rosen (2012) "The case of InterCorp, a multilingual parallel corpus." [In:] *International Journal of Corpus Linguistics*. 17 (3); 411–427. Retrieved from <https://doi.org/10.1075/ijcl.17.3.07cer> [date of access: 15 Oct. 2024].
- Chruścińska, Krystyna (1978) "O formacjach potencjalnych i okazjonalnych." [In:] Mieczysław Szymczak (ed.) *Z zagadnień słownictwa współczesnego języka polskiego*. Wrocław: Zakład Narodowy im. Ossolińskich; 69–79.
- Dokulil, Miloš (1962) *Tvoření slov v češtině 1. Teorie odvozování slov*. Praha: Československá akademie věd.
- Doroszewski, Witold (ed.) (1958–1969) *Słownik języka polskiego*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Dubisz, Stanisław (ed.) (2003) *Uniwersalny słownik języka polskiego*. Vols. I–IV. Warszawa: Wydawnictwo Naukowe PWN.
- Grzegorzczkova, Renata ([1979] 1979) *Zarys słowotwórstwa polskiego. Słowotwórstwo opisowe*. Third Edition. Warszawa: Państwowe Wydawnictwo Naukowe.
- Fronek, Josef ([2000] 2013) *Velký česko-anglický slovník*. Second edition. Voznice: LEDA spol. s r.o.
- Grzegorzczkova, Renata, Roman Laskowski, Henryk Wróbel (eds.) ([1984] 1998) *Gramatyka współczesnego języka polskiego. Morfologia*. Warszawa: Wydawnictwo Naukowe PWN.
- Křen, Michal, Václav Cvrček, Jan Henyš, Milena Hnátková, Tomáš Jelínek, Jan Koček, Dominika Kovářková, Jan Krívan, Jiří Milička, Vladimír Petkevič, Pavel Procházka, Hana Skoumalová, Jana Šindlerová, Michal Škrabal (2020) *SYN2020: reprezentativní korpus psané češtiny*. Ústav Českého národního korpusu FF UK, Praha. Available from the web: <http://www.korpus.cz>.
- Nagórko, Alicja (1987) *Zagadnienie derywacji przymiotników*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Nagórko, Alicja (2000) *Zarys gramatyki polskiej (ze słowotwórstwem)*. Warszawa: Wydawnictwo Naukowe PWN.
- NKJP – Narodowy Korpus Języka Polskiego. Available at: <http://www.nkjp.pl/>.
- Oltra-Massuet, Isabel (2013) *Deverbal adjectives at the interface: A crosslinguistic investigation into the morphology, syntax and semantics of -ble*. Boston: De Gruyter Mouton.
- Pęzik, Piotr (2012) *Wyszukiwarka PELCRA dla danych NKJP. Narodowy Korpus Języka Polskiego*. Adam Przepiórkowski, Mirosław Bańko, Rafał Górski, Barbara Lewandowska-Tomaszczyk (ed.). Warszawa: Wydawnictwo PWN.
- Ptak, Lenka (2009) *Przymiotniki o znaczeniu potencjalnym w językach czeskim, słowackim i polskim*. Wrocław: Oficyna Wydawnicza Atut.
- Ptak, Lenka (2015) "O polskich odpowiednikach czeskich przymiotników potencjalnych w korpusie InterCorp oraz w słownikach dwujęzycznych." [In:] *Language and Literary Studies of Warsaw*. Vol. 5. Warszawa; 55–68.
- Puzynina, Jadwiga (1966) "O pojęciu potencjalnych formacji słowotwórczych." [In:] *Poradnik Językowy*. Vol. 8; 329–338.

- Rosen, Alexandr, Bohumil Šimčík, Martin Vavřín, Adrian J. Zasina (2024) *Korpus InterCorp – čeština, verze 16ud z 17. 9. 2024*. Praha: Ústav Českého národního korpusu FF UK. Available from the web: <http://www.korpus.cz>.
- Rosen, Alexandr, Martin Vavřín, Adrian J. Zasina (2022a) *InterCorp, Release 15 of 11 November 2022*. Prague: Institute of the Czech National Corpus, Charles University. Available from the web: <http://www.korpus.cz>.
- Rosen, Alexandr, Martin Vavřín, Adrian J. Zasina (2022b). *The InterCorp Corpus – Czech3, version 16 of 11 November 2022*. Prague: Institute of the Czech National Corpus, Charles University. Available from the web: <https://kontext.korpus.cz>.
- Szymanek, Bogdan (1985) *English and Polish Adjectives*. Lublin: Wydawnictwo KUL.