

POLSKA AKADEMIA NAUK · ODDZIAŁ W KATOWICACH

LINGUISTICA SILESIANA Vol. 47/1

Editor

Rafał Molencki, Institute of Linguistics, University of Silesia in Katowice

Co-editors

Artur Kijak, Institute of Linguistics, University of Silesia in Katowice

Anna Majer, Institute of Linguistics, University of Silesia in Katowice

Editorial Board

Paola Attolino, University of Salerno, Italy

Wiesław Banyś, University of Silesia, Katowice, Poland

Bogusław Bierwiazzonek, Jan Długosz University, Częstochowa, Poland

Michael Bilynsky, University of Lviv, Ukraine

Krzysztof Bogacki, University of Warsaw, Poland

Jan Čermák, Charles University, Prague, Czech Republic

Bożena Cetnarowska, University of Silesia, Katowice, Poland

Magdalena Charzyńska-Wójcik, John Paul II Catholic University, Lublin, Poland

Danuta Gabryś-Barker, University of Silesia, Katowice, Poland

Piotr Kakietek, University of Silesia, Katowice, Poland

Marcin Krygier, Adam Mickiewicz University, Poznań, Poland

Andrzej Łyda, University of Silesia, Katowice, Poland

Elżbieta Mańczak-Wohlfeld, Jagiellonian University, Kraków, Poland

John G. Newman, University of Texas Rio Grande Valley, USA

Jerzy Nykiel, University of Bergen, Norway

Marc van Oostendorp, Radboud University, Nijmegen, The Netherlands

Tadeusz Piotrowski, University of Wrocław, Poland

Czesława Schatte, Adam Mickiewicz University, Poznań, Poland

Piotr Stalmaszczyk, University of Łódź, Poland

Wydanie publikacji dofinansowane przez Polską Akademię Nauk

ISSN 0208-4228

© Copyright by Polska Akademia Nauk Oddział w Katowicach, 2026

Linguistica Silesiana has been indexed in SCOPUS database since 2011.

Linguistica Silesiana was approved on 2017-07-13 according to ERIH PLUS criteria for inclusion

LOGOScript

Przygotowanie do druku:
LogoScript Sp. z o.o.
ul. Dembowskiego 4/54, 02-784 Warszawa
e-mail: logoscript@logoscript.pl

CONTENTS – SOMMAIRE – INHALTSVERZEICHNIS

MARCIN KUCZOK

- A contrastive analysis of *sweat/pot* and *tears/lzy* in English and Polish from
a cognitive-semantic perspective 7

ZHOU YUYING AND MUHAMMAD RIDHA ALI HUDDIN

- A cultural comparison of impoliteness strategies in Chinese and English on
Taobao and Amazon 25

MAGDALENA PERZ

- La synonymie en question: du lexique au discours 49

PAWEŁ GOLDA, AGNIESZKA JEDZINIAK-DANOWSKA, MARLENA MOSKAL AND
JOANNA RYSZKA

- Naming the nameless: a preliminary study of bynames for the deceased
from an onomastic perspective 65

PIOTR A. OWSIŃSKI

- Die Wellen der Worte: Eine explorative Studie zu Übersetzungsstrategien
und Herausforderungen im internationalen Seehandelsrecht 87

RENÁTA GREGOVÁ AND RENÁTA PANOCOVÁ

- Do distinctive features play a role in speech errors? Evidence from English 105

TATIANA SZCZYGLÓWSKA

- A corpus stylistic approach to Ernest Hemingway's fiction: keywords and
recurrent phrases in *The Sun Also Rises* 119

MARZENA BĘDKOWSKA-OBLĄK

- Rezensionsbeitrag: Interdisziplinäre Zugänge zu der modernen
Forschung der Sprachkommunikation. 147

MARCIN KUCZOK

University of Silesia in Katowice, Poland

Institute of Linguistics

ORCID: 0000-0002-6921-1995

marcin.kuczok@us.edu.pl

A CONTRASTIVE ANALYSIS OF *SWEAT/POT* AND *TEARS/ŁZY* IN ENGLISH AND POLISH FROM A COGNITIVE-SEMANTIC PERSPECTIVE

This study compares the English words *tears* and *sweat* with their Polish counterparts *łzy* and *pot* from a cognitive-semantic perspective. Beyond their prototypical meanings as bodily fluids associated with the eyes and skin – especially of the face – these words develop overlapping figurative senses. The analysis identifies both shared and language-specific conceptual metaphors and metonymies underlying their semantics, as well as patterns of metaphor-metonymy interaction. The study draws on data from the COCA corpus for English and the NKJP corpus for Polish. The findings show that while similar cognitive mechanisms motivate the figurativeness of the examined concepts in both languages, they result in distinct semantic extensions and complex metaphonymic structures.

Keywords: sweat and tears, pot i łzy, metaphor, metonymy, metaphonymy

1. Introduction

Sweat and tears are bodily fluids that often appear on a person's face – either separately or together – and typically manifest as physiological reactions to various emotions. The face is that part of the body which is almost constantly displayed to others, and for that reason it constitutes an important medium for expressing our emotions and communicating our thoughts, feelings, and attitudes (Wierzbicka 1999: 172-175; Bogdanowska-Jakubowska 2010: 75-82). Notably, the names for those two bodily fluids form strong collocations in both English and Polish. For example, *sweat and tears* occurs 234 times in the *Corpus of Contemporary American English* (COCA), while its Polish counterpart *pot i łzy* appears 159 times in the *Narodowy Korpus Języka Polskiego* (NKJP). Further-

more, dictionary evidence shows that both *sweat/pot* and *tears/lzy* are polysemous, functioning with multiple meanings, including a variety of figurative senses. Some of these senses even overlap, as in the collocations mentioned above, where they collectively convey notions of effort and hard work.

Although numerous body-part terms have been widely examined in cognitive linguistics (e.g. Enfield and Wierzbicka 2002; Kövecses 2005; Sharifian, Dirven, Yu and Niemeier 2008), bodily fluids have received comparatively little scholarly attention. Following the principles of cognitive semantics (Lakoff and Johnson 1999; Lewandowska-Tomaszczyk 2007; Geeraerts 2010), one may argue that the meanings of polysemous words are frequently motivated by metonymic or metaphorical mappings. This leads us to the following research questions: Which conceptual metaphors and metonymies underlying the semantics of *tears/lzy* and *sweat/pot* are shared in English and Polish? What patterns of metaphor-metonymy interaction emerge in their meanings across the two languages? And which of these conceptual mappings are unique to either English or Polish?

The methodology adopted in this study combines the traditional, or lexical, approach to investigating conceptual metaphor (Kövecses 2011: 27) with a corpus-based perspective (Deignan 2015: 145). First, dictionary definitions enable us to compile a list of the potential polysemous senses of *sweat/pot* and *tears/lzy*, which in turn allows us to determine which of these senses are motivated by metaphorical or metonymic processes and should therefore be included in our semantic analysis. Second, authentic examples drawn from language corpora allow us to observe how metaphors and metonymies involving bodily fluids operate in actual usage. Because our aim is to examine the subtleties of the conceptualization patterns encoded in the semantics of the two lexical items, the study is qualitative in nature; quantitative questions fall outside the scope of this paper. For English, the data come from sources representing American usage: the *Merriam-Webster Dictionary* (MWD), one of the oldest and most extensive dictionaries of American English, accessible at <https://www.merriam-webster.com/>, and the *Corpus of Contemporary American English* (COCA), a one-billion-word corpus comprising both written and spoken texts, accessible at <https://www.english-corpora.org/coca/>. For Polish, the materials include Müldner-Nieckowski's *Wielki słownik frazeologiczny języka polskiego* (2003), which provides definitions along with numerous collocations and fixed expressions, and the *Narodowy Korpus Języka Polskiego* (NKJP), a corpus of more than one and a half billion words drawn from diverse text types, available at <http://nkjp.pl/>.

In the first stage of the analysis, the sources listed above were examined in detail to identify the existing meanings and uses of *sweat/pot* and *tears/lzy*. Subsequently, those senses motivated by conceptual metaphors, metonymies, or metaphonymies were isolated, and a representative example of each type of conceptual mapping was selected for discussion. It should be stressed that the

goal of this study is not to present and catalogue all existing collocations, phrases, or expressions containing *sweat/pot* and *tears/lzy*. Rather, our focus lies on the conceptual mappings underlying the semantics of the names of these bodily fluids in English and Polish, with particular attention to the patterns of interaction that arise between them.

The paper begins by examining the role of conceptual metonymy in the semantic extensions of *sweat/pot* and *tears/lzy* in both languages, noting the processes of metonymic expansion and reduction involved in these developments. This is followed by a discussion of the metaphorical extensions of the figurative meanings of the two bodily fluids. The final section explores the various patterns of metaphor-metonymy interaction found in the use of *sweat/pot* and *tears/lzy*, demonstrating the ways in which metonymy may serve as a source for metaphor in certain senses of these words.

2. Metonymic motivation behind the figurative meanings of *sweat/pot* and *tears/lzy*

In cognitive semantics, *conceptual metonymy* is understood as “a cognitive process in which one conceptual entity – the vehicle – grants mental access to another conceptual entity – the target – within the same idealized cognitive model (ICM)” (Radden and Kövecses 1999: 19-21). The metonymic process is perceived as an extremely productive way of extending the meanings of specific lexemes, and various languages can be said to possess their own “metonymy-generators” – lexical items with particularly rich polysemous meanings (Bierwaczek 2013: 268).

According to Ruiz de Mendoza and Galera (2014: 170), the source and target of a metonymic mapping within a given ICM or conceptual domain may be related in two distinct ways. In a *source-in-target* metonymy, the source concept forms a constituent part or subdomain of the target concept. Conversely, in a *target-in-source* metonymy, the direction is reversed: the target is conceptualized as a subdomain of the source. *Source-in-target* metonymy leads to cognitive expansion, whereby the amount of conceptual content accessible from the initial point of reference increases. This metonymy can be illustrated by the sentence *The piano has the flu today*, where the word *piano* stands for a musician, following the INSTRUMENT FOR A MUSICIAN metonymy: here, “piano”, the source for the metonymic link, is a subdomain of “musician”, the metonymic target. *Target-in-source* metonymy, by contrast, results in cognitive reduction, as it highlights only a selected portion of the overall conceptual structure. This type of metonymy can be found in *The Picasso is in the lounge*, where the word *Picasso* stands for “painting”, but “painting”, which is the target of this meto-

nymy, constitutes a subdomain of the domain “Picasso”, the source of this mapping (Ruiz de Mendoza 2000: 124).

Our analysis has shown that the figurative senses of *sweat/pot* and *tears/lzy* are predominantly motivated by the source-in-target metonymy. In fact, the target-in-source mapping could only be identified in those figurative senses of the English *sweat* that are motivated by the more complex double metonymy.

2.1. The source-in-target metonymy in the figurative uses of *sweat/pot* and *tears/lzy*

There are two types of the source-in-target metonymy leading to the cognitive expansion of *sweat/pot* and *tears/lzy*: EFFECT FOR CAUSE or PHYSIOLOGICAL EFFECTS OF EMOTIONS FOR EMOTIONS. The first type of mapping emerges when *sweat* refers to effort, hard work, or exertion. Since individuals who engage intensely in a task – particularly a physical one – typically begin to sweat, the bodily effect becomes linguistically associated with the cause. This is reflected, for example, in the English expression *by the sweat of your brow* (1a) and its Polish counterpart *w pocie czoła* (1b). This expression ultimately traces back to the biblical story, where Adam is condemned by God to toil on the earth as punishment for the original sin: “By the sweat of your brow you will eat your food until you return to the ground, since from it you were taken; for dust you are and to dust you will return” (Gen 3:19).¹ Through the influence of the Bible on the cultures of European countries, the expression has spread to numerous languages, including English and Polish (Kuczok 2023: 107).

- (1a) What happens if you have to move it up a step? Well, then you would do it the old-fashioned way – **by the sweat of your brow**.² (COCA)
- (1b) Pracował przez całą noc **w pocie czoła**, przenosząc przez rzekę sosnę po sośnie, buk po buku, brzozę po brzozie. (NKJP) [He **worked by the sweat of his brow** all night, carrying pine by pine, beech by beech, birch by birch across the river.]

In a similar vein, *tears* and *lzy* can also stand metonymically for effort and hard work, and they may collocate with *sweat/pot* when used in this sense (2a-b). They can additionally appear on the human face as a result of putting significant effort into performing a difficult task. Interestingly, when used in this

¹ All biblical quotations used in this paper come from the English translation called the New International Version (2011) and available from the Bible Gateway website at: <https://www.biblegateway.com/>.

² In all the quotations from language corpora, we use the bold font to emphasize those words and phrases that are lexical correlates of the discussed metonymic and metaphorical mappings.

meaning, the two bodily fluids – sweat and tears – are often combined with *blood/krew*, which may likewise appear on one's face: when a person is working physically hard or fighting, they may easily get wounded and shed blood (Kuczek 2025: 194). The collocation *blood, sweat and tears / krew, pot i łzy* is also very common in both languages, accounting for 156 occurrences (67%) out of the total number of 234 uses of *sweat and tears* in COCA and 98 occurrences (62%) out of the 159 appearances of *pot i łzy* in NKJP. Moreover, *blood* and *sweat* co-occur in this metonymic sense in the English expression *to sweat blood* and the Polish phrase *krwawy pot* ("bloody sweat"), which refer to intense work and great effort. The latter can also be understood literally, as medicine recognizes the phenomenon of hematomatosis, in which individuals under extreme stress may excrete blood through their skin (MWD).³

- (2a) They steal our inheritance left by the **blood, sweat and tears** of our forefathers. (COCA)
- (2b) Fałszywi mesjasze polityczni to specjalność XX w. Przynieśli milionom **krew, pot i łzy**, niewyobrażalne zniszczenia i niepowetowane straty. (NKJP) [False political messiahs are a specialty of the 20th century. They have brought millions **blood, sweat and tears**, unimaginable destruction and irreparable losses.]

Apart from that, *tears* often appear in both English and Polish as a metonymic effect of suffering, as they may signal that the crying person is experiencing intense physical or emotional pain. This meaning of *tears/łzy* can be illustrated with the English collocation *vale of tears* (3a) and its Polish counterpart *padół łez* (3b).

- (3a) And look – hasn't the world always been this way? Vale of sorrow, **vale of tears**. (COCA)
- (3b) Ci, którzy uważali ziemię za **padół łez**, a żywot za pasmo cierpień wszelakich, nie mieli złudzeń, że to dopiero początek, pierwsze zwiastuny tego, co nadejdzie. (NKJP) [Those who considered the earth to be a **vale of tears** and life a series of all kinds of suffering had no illusions that this was only the beginning, the first signs of things to come.]

When it comes to the metonymy PHYSIOLOGICAL EFFECTS OF EMOTIONS FOR EMOTIONS, it is one of the most pervasive ways of conceptualizing emotions. Such metonymic patterns can, in turn, give rise to so-called *phy-*

³ Perhaps the best-known example of that phenomenon can be found in the Bible, where Jesus sweats blood while praying on the Mount of Olives before his passion: "And being in anguish, he prayed more earnestly, and his sweat was like drops of blood falling to the ground" (Luke 22: 44).

siological metaphors, which are common in the language of emotions (Kövecses 1990: 198-199; Kövecses 2014: 17-19; Kövecses 2020: 37-38), for instance ANGER IS HEAT (e.g. *to boil with anger*) or FEAR IS COLD (e.g. *to freeze with terror*). This metonymy can also be identified in a number of uses of *sweat/pot* and *tears/lzy*. First, *sweat* and *pot* may stand for fear – in both languages under comparison, the reference to sweat is connected with the conceptualization of fear as a health problem which makes people weak and feverish (Rewiś-Łętkowska 2019: 134). This meaning appears, for example, in the English expression *in a cold sweat* (4a) and the Polish phrase *oblać się zimnym potem* (literally: “to be covered with a cold sweat”) (4b).

- (4a) I woke up at 4:30 a.m. **in a cold sweat**, fearful that I was wrong, but again could not, for the life of me, dredge up the specifics. (COCA)
- (4b) Była to chwila nad wyraz przykra dla kapitana: nie mógł przypomnieć sobie rysów twarzy własnego synka. Doznał zabobonnego strachu. **Oblał się zimnym potem**. (NKJP) [It was a very unpleasant moment for the captain: he could not recall the facial features of his own son. He was overcome with superstitious fear. He **was covered with a cold sweat**.]

Additionally, the concepts of *sweat* and *pot* can be used metonymically to refer to the physiological effects of anxiety and distress. For example, in English, *being in a sweat* (5a), and in Polish, having sweat on one’s forehead (*pot na czole*) (5b), may function as symptoms of such emotions. In fact, in English there are a number of fixed, idiomatic expressions which are motivated by this metonymy, for instance, *to sweat about something*, *to sweat bullets*, *to sweat it out*, or *don’t sweat it*, and *no sweat*, all of which refer to the feelings of anxiety and worrying situations.

- (5a) He was **in a sweat** and finally he agreed to make that call to the White House. (COCA)
- (5b) Najpierw poczuł ból w dolnej części brzucha. Potem poczuł suchość w gardle i **pot na czole**. Powtarzało się to za każdym razem, gdy myślał o zamieszkaniu z Lone. (NKJP) [First he felt pain in his lower abdomen. Then he felt a dry throat and **sweat on his forehead**. It happened every time he thought about moving in with Lone.]

In line with the same metonymic mapping, *tears/lzy* in both English and Polish stand for a range of negative emotions, especially sadness (Cichmińska 1999: 15), but also regret, helplessness, sorrow, and rage. This can be illustrated, for example, with the English phrase *to be reduced to tears* (6a) and its Polish counterpart *doprowadzać kogoś do łez* (literally: “to bring someone to tears”) (6b).

- (6a) Have you seen scenes of multimillionaires **reduced to tears** because they lose their boat? (COCA)
- (6b) Nieraz **doprowadzała do łez** Polinę wytykając jej bezczynność i dziewczęce romantyczne marzenia. (NKJP) [She often **brought Polina to tears**, pointing out her inaction and girlish romantic dreams.]

However, as Kövecses (1990: 174) writes: “Crying, it should be noted, is not specific to any particular emotion. We can cry as a result of joy, anger, love, and fear, and not just sadness or grief (although prototypically it is associated with these latter emotions).” Thus, in both languages, *tears/lzy* may also metonymically represent positive emotions, such as joy or happiness, as seen in the expressions *tears of joy* (7a) and *lzy szczęścia* (7b).

- (7a) Tears of sorrow quickly turn to **tears of joy** for a young boy fighting to understand why God chooses some and leaves others behind. (COCA)
- (7b) Prawdziwy szal zaczął się po meczu. Na stadionie deszcz szampana, **lzy szczęścia** i trzy rundy honorowe wokół boiska. (NKJP) [The real frenzy began after the match. There was a shower of champagne at the stadium, **tears of joy** and three laps of honor around the pitch.]

2.2. Double metonymy with sweat in English

In some English uses of *sweat*, we can observe a more complex metonymic motivation for meaning, involving two metonymic mappings. This phenomenon is described by Ruiz de Mendoza and Galera (2014: 117) as *double metonymy* or *metonymic chain*. In example (8), the phrase *old sweat* refers to an experienced soldier, and its meaning can be explained through two metonymies connected with the effort and hard work characteristic of military service. Conceptually, these mappings operate within the ACTION ICM (Radden and Kövecses 1999: 19-21). First, the noun *sweat* stands metonymically for an action, following the mapping RESULT FOR ACTION, a subtype of the EFFECT FOR CAUSE metonymy (Panther and Thornburg 2000: 216). This is a source-in-target metonymy that expands the meaning of *sweat*. Subsequently, the action itself stands metonymically for the soldier, through the ACTION FOR AGENT mapping – an instance of metonymic meaning reduction involving a target-in-source metonymy. This mechanism of double metonymy underlying the meaning of *sweat* in this example is illustrated in Figure 2. In fact, the phrase *old sweat* can be used in English in a similar way to refer to experienced people representing other professions, not only to soldiers.

Nevertheless, such a sense of *pot* is absent from Polish. A possible translation of the example would be: *Kompania nr 3 zbuntowała się. Byli to głównie starzy*

wyjadacze, którzy mieli już dość, where instead of *old sweats*, the expression *starzy wyjadacze* (literally: “old eaters”) is used. Although it also has a figurative meaning, it is based on an entirely different semantic motivation.

- (8) No 3 company mutinied. They were mostly **old sweats** and they’d had enough. (COCA)

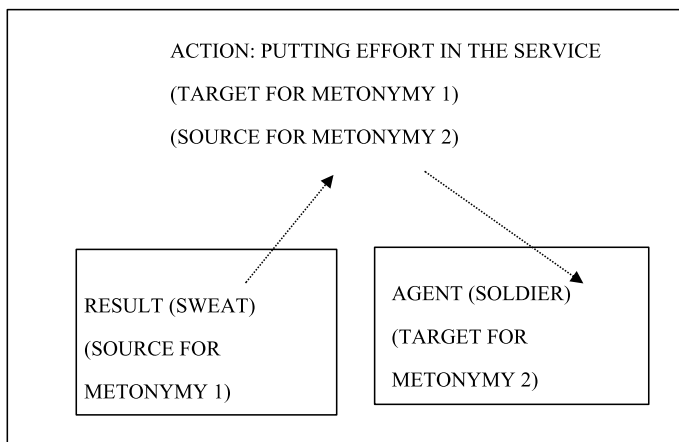


Figure 1. Double metonymy in *old sweat*: the metonymic expansion with the source-in-target metonymy (1) RESULT FOR ACTION (sweat for putting a lot of effort) combined with the metonymic reduction with the target-in-source metonymy (2) ACTION FOR AGENT (putting effort for a soldier)

In the same vein, and only in English, the plural noun *sweats* appears with the meaning “pants or an outfit made of cotton used as sportswear,” as illustrated in example (9). The presence of the bodily fluid *sweat* in the name of this garment seems to be motivated by two metonymies within the ACTION ICM. First, the RESULT FOR ACTION metonymy applies, where *sweat* stands for performing sports; next, the ACTION FOR INSTRUMENT metonymy occurs, as the sports activity explains the use of *sweat* to refer to the special pants used as the gear – or a sort of indispensable instrument – necessary to do the sport. The former is a source-in-target mapping, resulting in an expansion of the meaning of *sweat*, whereas the latter is a target-in-source metonymy, reducing the meaning of the action to the sports gear. The mappings involved in this double metonymy are presented in Figure 2. Moreover, the same metonymic motivation appears to underlie the meaning of a number of English compound nouns, such as *sweat-pants*, *sweatshirt*, or *sweatsuit*, which are often worn by athletes while exercising. In Polish, this type of metonymic motivation based on the bodily fluid *pot* seems to be absent in the names for these sports clothes: *sweats* are *spodnie dresowe*, *sweatsuit* is *dres*, and *sweatshirt* is *bluza*.

- (9) I don't get dressed up all the time either. In fact right now I have a tee and **sweats** on. (COCA)

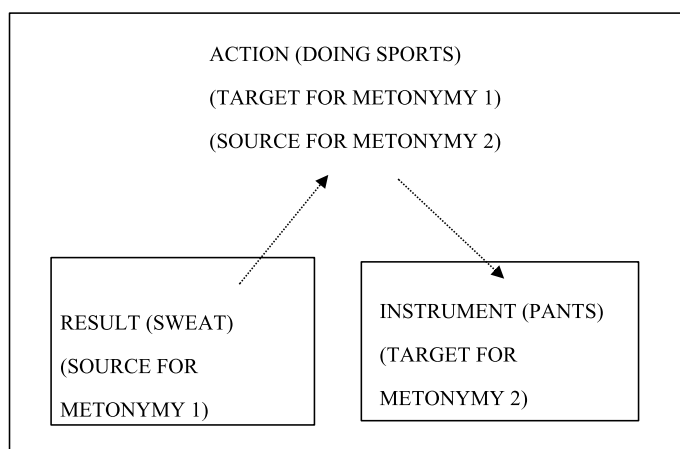


Figure 2. Double metonymy in *sweats*: the metonymic expansion with the source-in-target metonymy (1) RESULT FOR ACTION (sweat for putting a lot of effort) combined with the metonymic reduction with the target-in-source metonymy (2) ACTION FOR INSTRUMENT (putting effort for pants)

3. Metaphorical extensions of *sweat/pot* and *tears/lzy*

In cognitive semantics, *metaphor* is defined as a mapping between two distinct conceptual domains: the source (or donor) domain is partially projected onto the target (or recipient) domain, which is consequently understood in terms of the source (Barcelona 2000: 3-4). This distinguishes metaphor from metonymy, which operates within a single conceptual domain. According to Lakoff and Johnson (2003: 5), a conceptual metaphor involves “understanding and experiencing one kind of thing in terms of another,” as people rely on metaphors to conceptualize and describe abstract concepts across various aspects of life, including art, emotions, and time.

Ruiz de Mendoza and Galera (2014: 43-44) argue that metaphors generally arise either from resemblance or from experiential correlation. Resemblance metaphors establish a correspondence between two entities based on perceived similarity. For example, in *His nose is an elephant's trunk*, the resemblance is structural, whereas in *Her skin is silk* or *Her eyes could stir the dead*, the similarity lies in certain non-structural attributes (ibid.: 166-168). Metaphors based on experiential correlation, by contrast, link concepts through shared experiences. An example is *warm greetings*, which metaphorically conveys affec-

tion via temperature: since affection is often experienced as physical closeness, it is associated with the bodily warmth of human contact (ibid.: 159-160). However, according to Kövecses (2013: 89), in correlation-based metaphors there must exist a metonymic link between the non-figurative experience and the metaphorical stage of conceptualization. As he explains, “Given a frame in which the correlations apply between two elements, there must be a logically prior metonymic relationship between the elements before one element is moved out of the initial frame to establish a new and independent frame, which will form either a metaphoric source or target” (Kövecses 2020: 49).

When it comes to resemblance metaphors in the analyzed material, they include the MOISTURE / LIQUID IS SWEAT mapping. It can be identified in those examples from the corpora where moisture appears on various objects, such as glass (10a and 11a), cheese (10b and 11b), windows (10c), or walls (11c), though some of them, like *spocony ser* (“sweaty cheese”) in (11b), can be examples of creative metaphors rather than common, conventionalized expressions.

- (10a) An icy-cold Coca-Cola appeared – you could see the **glass sweating**. (COCA)
- (10b) In the glow of the dying night his body is waxy and pale, like a **sweating cheese**. (COCA)
- (10c) The **windows sweated** with steam. (COCA)
- (11a) Można było się także dowiedzieć, że **szkło się poci** lub pokrywa pajęczyną drobnych spękań i że są to zmiany nieodwracalne. (NKJP) [You could also learn that **the glass sweats** or becomes covered with a web of tiny cracks and that these are irreversible changes.]
- (11b) Turnus lipcowy, śniadanie z półmiskiem fajansowym, a na nim wylegują się **spocony ser** przytulony do wysuszonego pasztetu, wtuleni w siebie plaster w plaster. (NKJP) [A July session, breakfast with a faience platter, and on it lying **sweaty cheese** cuddled up to a dried-out pate, snuggled up together slice by slice.]
- (11c) Zjawisko „**pocenia się ścian**” jest naturalnym procesem przy zastosowaniu technologii ścian szczelinowych. (NKJP) [The phenomenon of “**wall sweating**” is a natural process when using diaphragm wall technology.]

Additionally, there are figurative uses of the verb *to sweat* in English that follow the mapping EXUDING SOMETHING IS MAKING IT SWEAT. In example (12a), vegetables can be *sweated*, meaning they release juices when heated. In example (12b), robbing someone’s garage is described colloquially as *sweating* the place, and in sentence (12c), *sweating* is used metaphorically

to mean joining pieces of metal by melting them. None of these figurative meanings of *sweat* occur in Polish. The English *to sweat vegetables* can be translated as *zeszklić warzywa*, a figurative verb literally meaning “to vitrify the vegetables,” perhaps because hot onions and garlic covered in their juices appear as if coated with glass. The translation of *to sweat a garage* might be *oskubać*, which is also metaphorical and literally means “to pluck.” The Polish equivalent of *to sweat the pipe* is *lutować*, which corresponds directly to the English verb *to solder*.

- (12a) In a saucepan melt butter; **sweat onions and garlic**; add shrimps. (COCA)
- (12b) Someone broke the window and wanted **to sweat my garage**. (COCA)
- (12c) Tim, we’ll be **sweating** the pipe. (COCA)

When it comes to *tears*, in both English and Polish we can observe the metaphor X IS A TEAR / TEARS, which is based on the perceived resemblance between the shape or qualities of the target domain and tears. Examples from English include rain (13a), dew (13b), and wax (13c).

- (13a) Thunder jars the windows. We go and stand outside in the first **tears of rain**. (COCA)
- (13b) They wish they could be plants and lie very still near the ground all night and in the morning be covered with **tears of dew**. (COCA)
- (13c) Two **waxen tears** appear. # Inside the chapel – bitter cold. (COCA)

Examples from Polish include referring to something clean, such as water, which is described as “as clean as a tear” (14a). Similarly, the same Polish expression can be used to describe a person who is morally pure (14b). The English equivalent of this Polish simile might be *as clean as a whistle*, which appears to convey the same meaning and usage.

- (14a) Woda wypływająca z filtrów jest **czysta jak łza** – zapewnia zastępcę kierownika zakładu wodociągu centralnego. (NKJP) [The water flowing out of the filters is **as clean as a whistle** – assures the deputy manager of the central water supply plant.]
- (14b) Gen. Komornicki jest **czysty jak łza**. Jest to bardzo solidny i porządny generał, który dobrze wykonuje obowiązki służbowe. (NKJP) [Gen. Komornicki is **as clean as a whistle**. He is a very solid and decent general who performs his duties well.]

4. Metaphtonymies in the figurative meanings of *sweat/pot* and *tears/lzy*

Metaphor and metonymy can interact to form more complex conceptual patterns, known as *metaphtonymies* (Goossens 2002: 366-369). Ruiz de Mendoza and Diez (2002: 522) note that “(...) because of its domain-internal structure, metonymy is always subsidiary in conceptual interaction to metaphor, i.e. metonymy always takes place within the source or the target domain of the metaphor.” Building on the notions of source-in-target and target-in-source metonymy, Ruiz de Mendoza and colleagues (Ruiz de Mendoza and Diez 2002: 489-532; Ruiz de Mendoza 2011: 103-124; Ruiz de Mendoza and Galera 2014: 107-116) identify four fundamental patterns of metaphor-metonymy interaction. The first involves metonymic expansion of the metaphorical target domain (or parts of it) via source-in-target metonymy, as in the expression *to knit one's brows*, where bringing the eyebrows together stands for “being angry” in line with the EFFECT FOR CAUSE metonymy; this metonymy, however, occurs within the target domain of the metaphor PUTTING EYEBROWS TOGETHER IS KNITTING CLOTHES. The second pattern is metonymic reduction of the metaphorical target domain via target-in-source metonymy, exemplified by the authors with *to break someone's heart*, where *heart* represents “love” (the PART OF THE HUMAN BODY FOR THE CULTURALLY ASSOCIATED FUNCTION metonymy), and the abstract concept of the heart is metaphorically reified into an object that can be broken. The third type listed in the sources mentioned above, metonymic expansion of the metaphorical source domain via source-in-target metonymy, occurs in the expression *to beat one's breast*, often used to express apology, where the gesture stands for showing sorrow (EFFECT FOR CAUSE) and simultaneously maps metaphorically onto the act of apologizing (APOLOGIZING IS BREAST BEATING). Finally, the fourth pattern, metonymic reduction of the metaphorical source domain via target-in-source metonymy, is illustrated by the quoted authors with the sentence *Humboldt is the Shakespeare of travelers*, where *Shakespeare* stands for “superior ingenuity or excellent skills in writing poetry” (ENTITY FOR ITS SALIENT PROPERTY) while also invoking the metaphorical mapping ALEXANDER HUMBOLDT IS WILLIAM SHAKESPEARE as a way of conceptualizing Alexander Humboldt's remarkable attributes in the realm of traveling and exploring the world.

When it comes to the extensions of *sweat/pot* and *tears/lzy*, the metaphtonymies identified in our analysis are based on the metonymic expansion of either the source domain or the target domain of a metaphorical mapping.

4.1. Metonymic expansion of the metaphorical source domain

Sweat/pot and *tears/lzy*, when used metonymically to conceptualize hard work and unstinting effort, as discussed above, can give rise to a correlation metaphor, resulting in metaphonymy. As noted by Charteris-Black (2001: 282-284), the collocation *blood, sweat and tears* is generally used figuratively rather than in its literal sense. Thus, in examples (15a) from English and (15b) from Polish, the effort and hard work involved in writing a book or a movie script are conceptualized through the bodily fluids sweat and tears, in line with the mapping MENTAL OR NON-PHYSICAL EFFORT IS PHYSICAL EFFORT. This constitutes an example of metonymic expansion of the metaphorical source domain, with the EFFECT (blood) FOR CAUSE (effort) metonymy, as illustrated in Figure 3.

- (15a) Written in the blood, **sweat and tears** of the author and her family, and featuring expert clinical perspectives, *Chasing Silhouettes* blazes a trail. (COCA)
- (15b) Scenariusz filmowy to nie płwocina, tylko krew, **pot i lzy**. (NKJP)
[A movie script is not spit, but blood, **sweat and tears**.]

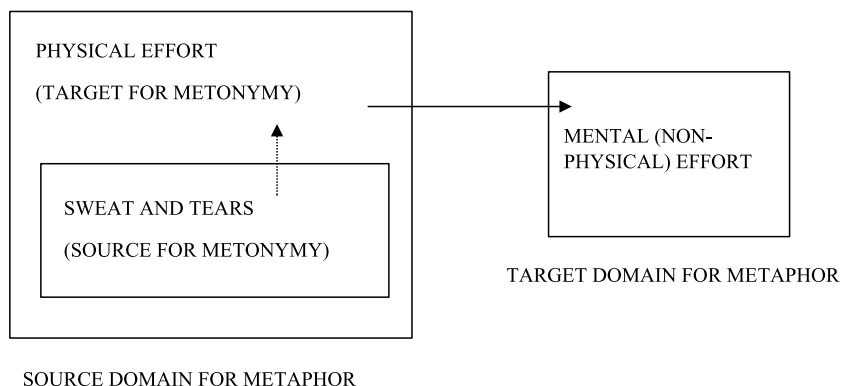


Figure 3. Metonymic expansion of the metaphorical source domain (mental (non-physical) effort is physical effort) with the source-in-target metonymy (effect for cause (tears / sweat for physical effort)) in *sweat and tears/pot i lzy*

Apart from that, the same metaphor-metonymy interaction can be observed in the English phrase *without tears* (16), where it conveys that a given activity can be performed easily, without much effort or difficulty. The Polish translation of this example might be *Metody maszynowego uczenia się bez bólu – podręcznik dla ekologów*, where *without tears* is rendered as *bez bólu* (“without pain”), or alternatively *Metody maszynowego uczenia się dla opornych – podręcznik dla*

ekologów, although, in this version, the phrase *dla opornych* (literally: “for the reluctant”) is based on a different conceptualization pattern.

- (16) Machine learning methods **without tears**: A primer for ecologists.
(COCA)

4.2. Metonymic expansion of the metaphorical target domain

With respect to the metonymic expansion of the metaphorical target domain, both English and Polish provide examples of the resemblance-based personification of *tears*, meaning they are described as if they possess human qualities or behave like people, in line with the mapping TEARS ARE PEOPLE. In English, for instance, tears can *run down someone's face* (17a) or *stand in someone's eyes* (17b), while in Polish they can crowd in one's eyes (*cisnąć się do oczu*) (17c), run to one's eyes (*nabiegły do oczu*) (17d), or stand in one's eyes (*stanęły w oczach*) (17e). Regardless of the human-like activity attributed to them, tears also metonymically represent emotions, following the mapping PHYSIOLOGICAL EFFECTS OF EMOTIONS FOR EMOTIONS. This metonymy occurs within the target domain of the metaphorical mapping, as illustrated in Figure 4.

- (17a) The **tears run** silently down his face. (COCA)
 (17b) **Tears stood** in her eyes. (COCA)
 (17c) I nagle poczułem, że **lzy cisną mi się** do oczu. (NKJP) [And suddenly I felt **tears crowding** in my eyes.]
 (17d) **Lzy nabiegły** matce do oczu. W pierwszym odruchu chciała syna przytulić, ale raptem uciekła do łazienki. (NKJP) [**Tears ran** to the mother's eyes. At first, she wanted to hug her son, but suddenly she ran to the bathroom.]
 (17e) Panna Urszula pobladła, **lzy stanęły** jej w oczach – ostatnie słowa Stadnickiego wydały się jej zwróconymi wprost do niej... (NKJP) [Miss Urszula turned pale, and **tears stood** in her eyes – Stadnicki's last words seemed to be addressed directly to her...]

5. Conclusions

Our analysis of the possible patterns of metonymic and metaphorical motivation behind the senses of *sweat/pot* and *tears/lzy* in English and Polish has shown that both bodily fluids exhibit metonymies, metaphorical extensions of their meanings, and interactions between metaphor and metonymy. In terms of metonymy, in both languages *sweat/pot* and *tears/lzy* may stand for hard work and effort, following the EFFECT FOR CAUSE mapping. The same metonymy also

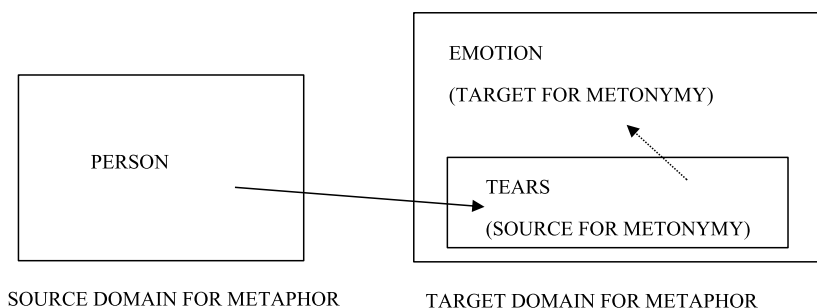


Figure 4. Metonymic expansion of the metaphorical target domain (TEARS ARE PERSONS) with the source-in-target metonymy (PHYSIOLOGICAL EFFECT (tears) FOR EMOTION (sadness)) in personifications of *tears/lzy*

motivates the use of *tears/lzy* to denote suffering. Moreover, in both English and Polish, these bodily fluids represent a range of emotions in line with the metonymic mapping PHYSIOLOGICAL EFFECTS OF EMOTIONS FOR EMOTIONS. Accordingly, *sweat/pot* can stand for fear, anxiety, or distress, while *tears/lzy* can represent either negative emotions, such as sadness, sorrow, rage, and regret, or positive ones, such as joy and happiness.

In English only, we have identified examples of double metonymy, as in *old sweat*, meaning an experienced soldier, and *sweats*, meaning pants worn for sports. In these cases, the RESULT FOR ACTION metonymy co-occurs with either the ACTION FOR AGENT or ACTION FOR INSTRUMENT metonymy.

There are also two common resemblance metaphors of *sweat/pot* and *tears/lzy*: the MOISTURE/LIQUID IS SWEAT metaphor and the X IS TEARS metaphor. The former describes drops of liquid on surfaces, such as glass, or damp on a wall, while the latter is found in English for objects taking the shape of tears, for instance rain, dew, or wax. In Polish, the X IS TEARS metaphor can describe something as clean as a tear, such as water or a person's character. Additionally, in English only, there are metaphorical extensions of *sweat* based on the EXUDING SOMETHING IS MAKING IT SWEAT metaphor, where vegetables, a place, or metal are described as sweating.

Regarding metaphor-metonymy interaction, our study has shown examples of metonymic expansion of either the metaphorical source or target domain in so-called correlation metaphors of *sweat/pot* and *tears/lzy*. Through the EFFECT FOR CAUSE metonymy, both bodily fluids extend the metaphor NON-PHYSICAL EFFORT IS PHYSICAL EFFORT, as illustrated by the phrase *blood, sweat and tears/krew, pot i lzy*. These findings support the view that in experience-based correlation metaphors, it is necessary to include metonymies as conceptual links between experiences and metaphorical extensions of meaning. Metonymic expansion of the metaphorical target domain can also be observed in the resem-

blance-based personification of *tears* (TEARS ARE PERSONS), which metonymically stand for emotions via the PHYSIOLOGICAL EFFECTS OF EMOTIONS FOR EMOTIONS mapping. This metaphonymy is found in both languages, where tears are described as behaving like people.

In conclusion, the contrastive analysis of *sweat/pot* and *tears/lzy* shows that both English and Polish rely on shared embodied experience to motivate figurative meanings through metonymy, metaphor, and their interaction. These bodily fluids systematically conceptualize effort, emotion, and suffering, while language-specific differences reveal the influence of cultural conventions. Further research could examine how metaphor and metonymy motivate the meanings of words denoting other bodily fluids in English and Polish, and potentially in other languages as well.

References

- Barcelona, A. 2000. The cognitive theory of metaphor and metonymy. In A. Barcelona (ed.), *Metaphor and Metonymy at the Crossroads: A Cognitive Perspective*, 1-28. Berlin: Mouton de Gruyter.
- Bierwiaczonek, B. 2013. *Metonymy in Language, Thought and Brain*. Sheffield: Equinox.
- Bogdanowska-Jakubowska, E. 2010. *FACE. An Interdisciplinary Perspective*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- Charteris-Black, J. 2001. Blood, sweat and tears: A corpus-based cognitive analysis of blood in English phraseology. *Studi Linguistici Teoretici e Applicativi* 30: 273-287.
- Cichmińska, M. 1999. Towards an analysis of the concept of sadness in Polish: A cognitive view. *Acta Neophilologica* 1: 11-23.
- Deignan, A. 2015. MIP, the corpus and dictionaries: What makes for the best metaphor analysis? *Metaphor and the Social World* 5: 145-154.
- Enfield, N.J., and A. Wierzbicka (eds.). 2002. *The Body in the Description of Emotion: Pragmatics and Cognition*. Amsterdam: John Benjamins.
- Geeraerts, D. 2010. *Theories of Lexical Semantics*. Oxford: Oxford University Press.
- Goossens, L. 2002. Metaphonymy: The interaction of metaphor and metonymy in expressions for linguistic action. In R. Dirven and R. Pörings (eds.), *Metaphor and Metonymy in Comparison and Contrast*, 349-377. Berlin: Mouton de Gruyter.
- Kövecses, Z. 1990. *Emotion Concepts*. New York: Springer.
- Kövecses, Z. 2005. *Metaphor in Culture: Universality and Variation*. Cambridge: Cambridge University Press.
- Kövecses, Z. 2011. Methodological issues in conceptual metaphor theory. In S. Handl and H.-J. Schmid (eds.), *Windows to the Mind: Metaphor, Metonymy and Conceptual Blending*, 23-39. Berlin: Mouton de Gruyter.

- Kövecses, Z. 2013. The metaphor-metonymy relationship: Correlation metaphors are based on metonymy. *Metaphor and Symbol* 28: 75-88.
- Kövecses, Z. 2014. Conceptualizing emotions: A revised cognitive-linguistic perspective. *Poznań Studies in Contemporary Linguistics* 50: 15-28.
- Kövecses, Z. 2020. *Extended Conceptual Metaphor Theory*. Cambridge: Cambridge University Press.
- Kuczok, M. 2023. The semantics of bodily fluids in Biblical Hebrew and contemporary English: A contrastive study from the cognitive-linguistic perspective. In M. Banaś and G. Wlaźlak (eds.), *Language, Culture, Identity: Between Ethnolinguistics and Ethnomethodology*, 93-113. Göttingen: Vandenhoeck and Ruprecht Unipress.
- Kuczok, M. 2025. Between metaphor and metonymy: A cross-linguistic study of *blood* in English and *krew* in Polish from a cognitive-semantic perspective. *Metaphor and Symbol* 40: 189-205.
- Lakoff, G., and M. Johnson 1999. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.
- Lakoff, G., and M. Johnson 2003. *Metaphors We Live By*. Chicago: University of Chicago Press.
- Lewandowska-Tomaszczyk, B. 2007. Polysemy, prototypes, and radial categories. In D. Geeraerts and H. Cuycens (eds.), *The Oxford Handbook of Cognitive Linguistics*, 139-169. Oxford: Oxford University Press.
- Merriam-Webster Dictionary*. n. d. Available at <https://www.merriam-webster.com/> (accessed 10 September 2025).
- Müldner-Nieckowski, P. 2003. *Wielki słownik frazeologiczny języka polskiego*. Warszawa: Świat Książki.
- New International Version*. 2011. Available at <https://www.biblegateway.com/> (accessed 21 September 2025).
- Panther, K.-U., and L. Thornburg 2000. The EFFECT FOR CAUSE metonymy in English grammar. In A. Barcelona (ed.), *Metaphor and Metonymy at the Crossroads*, 215-232. Berlin: Mouton de Gruyter.
- Radden, G., and Z. Kövecses 1999. Towards a theory of metonymy. In K.-U. Panther and G. Radden (eds.), *Metonymy in Language and Thought*, 17-59. Amsterdam: John Benjamins.
- Rewiś-Łętkowska, A. 2019. The anatomy of fear: interplay between metaphor, metonymy and image schema in the conceptualization of fear/strach in English and Polish. *Studia Pigioniana* 2: 119-143.
- Ruiz de Mendoza, F. 2000. The role of mappings and domains in understanding metonymy. In A. Barcelona (ed.), *Metaphor and Metonymy at the Crossroads: A Cognitive Perspective*, 109-132. Berlin: Mouton de Gruyter.
- Ruiz de Mendoza, F. 2011. Metonymy and cognitive operations. In R. Benczes, A. Barcelona and F. Ruiz de Mendoza (eds.), *Defining Metonymy in Cognitive Linguistics: Towards a Consensus View*, 103-124. Amsterdam: John Benjamins.
- Ruiz de Mendoza, F., and O. Díez 2002. Patterns of conceptual interaction. In R. Dirven and R. Pörings (eds.), *Metaphor and Metonymy in Comparison and Contrast*, 489-532. Berlin: Mouton de Gruyter.

- Ruiz de Mendoza, F., and A. Galera 2014. *Cognitive Modeling: A Linguistic Perspective*. Amsterdam: John Benjamins.
- Sharifian, F., R. Dirven, N. Yu and S. Niemeier (eds.). 2008. *Culture, Body and Language: Conceptualizations of Internal Body Organs Across Cultures and Languages*. Berlin: Mouton de Gruyter.
- Wierzbicka, A. 1999. *Emotions Across Languages and Cultures*. Cambridge: Cambridge University Press.

ZHOU YUYING

Universiti Malaya, Malaysia

Faculty of Languages and Linguistics

22090849@siswa.um.edu.my

MUHAMMAD RIDHA ALI HUDDIN

Universiti Malaya, Malaysia

Faculty of Languages and Linguistics

ORCID: 0009-0003-6862-7539

mardha67@um.edu.my

A CULTURAL COMPARISON OF IMPOLITENESS STRATEGIES IN CHINESE AND ENGLISH ON TAOBAO AND AMAZON

This study investigates the underexplored influence of cultural differences on impoliteness strategies within negative online consumer reviews (OCRs). Conducting a comparative analysis of 300 low-rated reviews from Taobao (Chinese) and Amazon (English), the research examines how Chinese and English consumers strategically use language to express dissatisfaction. The findings reveal that while both groups exhibit a strong preference for direct, on-record impoliteness, qualitative differences in rhetorical style persist, reflecting deeper cultural-communicative norms. The study contributes to cross-cultural pragmatics and offers practical insights for managing global e-commerce communication by highlighting the interplay between genre conventions and cultural specificity in online discourse.

Keywords: pragmatics, impoliteness strategies, online consumer reviews, intercultural communication, e-commerce

1. Introduction

Impoliteness has become an increasingly common phenomenon in digital communication, particularly in the context of e-commerce. When consumers encounter unsatisfactory online shopping experiences, they often use impolite language in negative reviews, which can significantly impact brand reputation

and purchasing decisions (Culpeper et al. 2010). While substantial research exists on impoliteness strategies in OCRs, few studies have systematically examined how cultural differences specifically shape their manifestation across linguistic boundaries. The rapid development of e-commerce platforms has changed the way that consumers and businesses interact, as well as the way that consumers interact with each other. Consumers from different backgrounds adopt different expressions when leaving comments. Platforms like Taobao and Amazon allow customers to express their dissatisfaction through negative reviews, which often contain impoliteness strategies. There is growing evidence that cultural differences between Chinese and English-speaking consumers can significantly affect the way they use impoliteness strategies (Culpeper et al. 2010).

However, while there is a large literature on the use of impoliteness strategies embodied in negative OCRs, there is a lack of focused research on how cultural differences specifically influence the use of impoliteness strategies in OCRs. If this issue is not addressed, it could lead to multiple consequences. First, users from different cultural backgrounds may experience conflicts due to misunderstandings of impolite behavior, which could reduce their satisfaction and trust in the platform. Second, businesses may be unsuccessful to effectively address customer issues because they cannot correctly interpret the cultural perspectives in negative reviews, impacting their business decisions and sales performance. Research shows that if impolite behavior is not properly managed, it may also lead to user attrition and affect the long-term development of the platform.

Although existing research has focused on the impact of cultural differences on language communication, there are still many unknowns regarding the strategies of impoliteness. The impolite phenomenon in network communication has gradually become the new focus of scholars. Lorenzo-Dus et al. (2011) studied the impoliteness strategies used in YouTube comments. However, the study on impoliteness behavior in languages other than English and in bilingual contexts is still relatively lacking. Therefore, this paper chooses the negative consumer reviews of Amazon (American Edition) and Taobao (China Edition) as the research and analysis object, because Amazon and Taobao are the world's largest e-commerce platforms. Understanding these differences is essential for businesses to effectively manage customer feedback and strengthen cross-cultural communication strategies. Moreover, when comparing the use of impoliteness strategies to express dissatisfaction between Chinese consumers on Taobao and English-speaking consumers on Amazon, this gap is also quite evident.

To address this gap, the present study sought to answer the following research questions:

- 1) To what extent do Chinese and English consumers differ in the frequency of impoliteness strategies used in negative OCRs?

2) How do Chinese and English consumers differ in the pragmatic realization of these impoliteness strategies specifically and how is impoliteness linguistically and rhetorically framed across the two cultural contexts?

3) What role do genre conventions of OCRs play in shaping the use of impoliteness strategies, and how do these genre conventions interact with cultural norms?

By analyzing 300 reviews across parallel product categories and addressing these research questions, this study investigates how cultural communication patterns influence the realization of impoliteness, contributing to both cross-cultural pragmatics and e-commerce communication strategies.

2. Literature Review

2.1. Impoliteness strategies

Impoliteness is defined as "a negative attitude toward specific behavior occurring in a specific context" (Culpeper 2011: 23) and involves deliberate face-threatening acts (Bousfield 2008). According to Brown and Levinson (1987), Negative consumer reviews can harm a supplier's positive reputation because the customer's need for acknowledgment is unmet. In certain cases, these reviews may also pose a threat to the supplier's reputation if the complainant demands compensation, thus restricting the supplier's ability to act freely. The impoliteness strategy involves attacking someone's face, which can result in social conflict and discord (Culpeper 2005).

The development of linguistic impoliteness strategy was intermittent and accelerated in the mid-1990s, and reached its climax in 2008 (Culpeper and Hardaker 2017). Researchers believe that all concepts of impoliteness involve at least one participant perceiving the behavior as emotionally negative. In addition to physical actions, verbal behaviors are also involved. The study of linguistic impoliteness covers a broader range of social situations and relationships. It primarily focuses on communicative behaviors, especially those where language is used entirely or primarily to achieve impoliteness. Examples include text messages, blogs, emails, Twitter, and other forms of digital media.

A number of scholars examined impoliteness in various contexts, such as daily communication, workplace, CMC communication, etc. Among them, many scholars have provided their own insights into the framework of the impoliteness model. Culpeper's model of impoliteness includes five super-strategies as follows: 1) Bald on-record Impoliteness, 2) Positive Impoliteness, 3) Negative Impoliteness, 4) Sarcasm or Mock Politeness, and 5) Withhold Politeness. Impoliteness strategy is the research focus of impoliteness theory. Bousfield (2006) proposed an impoliteness model consisting of five strategies: open impolite,

positive impolite, negative impolite, informal impolite and concealed impolite. Later, Kienpointner (2008) improved Bousfield's theory and divided impoliteness strategies into On-record impoliteness and Off-record impoliteness.

2.2. Impoliteness research in OCRs in CMC (Computer-mediated communication) contexts

OCR is a common research topic, especially in cross-cultural studies. For example, Chen (2023) concluded through a survey of online store representatives on Taobao, one of China's largest e-commerce platforms, that on this particular platform, the use of impoliteness strategies provides e-store representatives with an opportunity to renegotiate their image and reconstruct customer perceptions in response to critical comments. Lai's (2019) study, by analyzing the expression of impoliteness in English and Chinese of negative online reviews in the catering industry, found that English reviewers tended to use mixed forms of approval and criticism, as well as "staff oriented" sharp complaints, while Chinese reviewers used more indignant exclamations and "staff oriented" sharp complaints.

Although some studies have analyzed impolite strategies for OCR comparability, the number is still insufficient (Feng and Ren 2020). Existing research on OCRs primarily focuses on English. Many researchers have called for more studies on a wider variety of languages in order to better understand the linguistic features and the localization of social practices within this genre. In addition, Taobao and Amazon are the e-commerce platforms chosen by more consumers for online shopping.

2.3. Cultural differences in communication

Culture has multiple definitions. Culture includes patterned ways of thinking, feeling, and reacting, primarily acquired and transmitted through symbols, and constitutes the distinctive achievements of human groups, including their embodiment in artifacts. Culture for a human group is like personality for an individual. Personality is defined by Guilford (1959) as "the total interaction of the personal characteristics that influence an individual's response to the environment." Culture can be defined as the total interaction of shared characteristics that influence a human group's response to the environment. Culture determines the identity of a human group, just as personality determines the identity of an individual. Previous research has shown that politeness and impoliteness are influenced by multiple factors, such as culture and gender (Mills and Kadar 2011).

The work of Spencer-Oatey (2008) introduces the importance of relationship management and the variability of facial sensitivity across cultures. Studies

focusing on the comparison between East and West, such as Gu's (1990) study on Chinese politeness, emphasize that politeness plays a normative role in restricting the order of conversation and communication of individual speech behaviors in the Chinese context. Politeness is a phenomenon belonging to the social level, which supports the normative constraints of each of its individuals.

Cross-cultural studies of language often focus on differences in how politeness and incivility are perceived and enforced in different societies. Edward Hall (1976, 2000) proposed a popular cultural framework in which he suggested that all cultures can be positioned relative to one another based on their communication styles. In some cultures, such as those of Scandinavians, Germans, and Swiss, communication primarily occurs through explicit statements in texts and speech, and thus these cultures are categorized as low-context cultures. In other cultures, such as Japan and China, communication includes additional cues like body language and the use of silence. These dimensions are crucial to understanding why different impoliteness strategies may be employed in Chinese and English OCRs.

3. Methodology

3.1. Data collection procedure

This study adopted a stratified random sampling method to construct a comparable and representative corpus of negative OCRs. Data were collected from two analogous platforms which were Taobao (for Chinese) and Amazon.com (for English) for ensuring cross-cultural comparability. The corpus was categorized to three key volume, standardized product categories to regulate for variability in consumer expectations namely electronics, clothing, and personal care. For sampling procedure, all reviews were acquired within a defined period of time frame of the comments which was from 1st January 2023 until 30th June 2023. While for product identification on each platform, the search function was employed with neutral keywords namely 'USB-C cable', 'cotton t-shirt', and 'facial moisturizer'. The results were then sorted by 'Best Sellers' or 'Bestselling' in identifying the most popular items, allowing the reviews to illustrate common consumer experiences. The reviews which were ten most recent 1-star and 2-star reviews were selected from the product page of the top five best-selling items in each category which resulted in a total of 150 reviews per language were collected (5 products x 3 categories x 10 reviews). All the data were then filtered and omitted in the sense of the relevancy like irrelevant to the product namely commenting (only on shipping speed), insufficient linguistic data (less than five words long), and written by verified users who received the product through a free promotion (can be biased sentiment). The final balanced corpus eventually

consisted of 300 reviews, which 150 in Chinese and 150 in English. Those data were then fairly distributed across the three product categories. All Chinese data was analysed in its original language; examples were translated post-analysis solely for presentation with careful attention to preserving pragmatic force.

3.2. Data analysis

The data analysis proceeded in two phases: (1) a quantitative coding phase to identify and classify impoliteness strategies, and (2) a qualitative discourse analysis phase to interpret the pragmatic functions and cultural resonances of these strategies. A critical methodological decision regarding the unit of analysis was made prior to both phases.

3.2.1. Unit of analysis and coding protocol

A fundamental consideration in (im)politeness research is the specification of the unit of analysis, as this directly affects frequency counts and the interpretability of quantitative findings (Bousfield 2008; Culpeper 2011). OCRs vary considerably in length and complexity, ranging from single-sentence expressions of dissatisfaction to multi-paragraph narratives containing multiple distinct impoliteness acts. Therefore, a consistent and transparent unit of analysis is essential for reliable coding and meaningful cross-cultural comparison. This study adopted the **speech act** as the primary unit of analysis. Following the framework of Culpeper (2011), a speech act was operationalized as a unit of communicative meaning that performs a specific pragmatic function namely accusing, threatening, warning, insulting, or ridiculing toward a target (the seller, product, or service). Practically, a single speech act could span from a single clause to multiple sentences, provided that the utterance collectively realizes a single coherent pragmatic intention. The identification of speech act boundaries was guided by shifts in pragmatic function which is a transition from product criticism to a threat, for example, was treated as marking a new speech act unit. The coding protocol for handling reviews of varying lengths and complexity was operationalized as follows:

- a) **Single-strategy reviews**: Reviews containing only one identifiable impoliteness speech act contributed **one count** to the total frequency column for that strategy type.
- b) **Multi-strategy reviews**: Reviews containing two or more distinct impoliteness speech acts contributed **multiple counts** which was one count for each identifiable speech act to the respective strategy categories.
- c) **Co-occurring strategies**: In cases where strategies overlapped or were embedded within one another namely sarcasm used within a threat, the

dominant pragmatic function determined the primary classification, with secondary strategies noted in the coding commentary for qualitative reference but not double-counted.

Illustrative example of multi-strategy coding:

"This product is absolute garbage. The seller clearly doesn't care about customers. I will never buy from this store again unless I get a refund"

This review contains three distinct impoliteness speech acts namely product insult, accusation of seller negligence and conditional threat. The phrases "this product is absolute garbage" and "the seller clearly doesn't care about customers" were coded as positive impoliteness, while "I will never buy from this store again unless I get a refund" was coded as negative impoliteness.

Therefore, this single review would contribute two counts to positive impoliteness and one count to negative impoliteness, reflecting the multi-functional nature of the complaint. Because individual reviews may contain multiple impoliteness speech acts, the total frequency of strategies (summed across all categories) will naturally exceed the number of reviews in the corpus. This discrepancy is not an error but a meaningful empirical observation: it indicates the degree of *impoliteness density* within the corpus and reflects the prevalence of multi-strategy reviews wherein reviewers employ multiple impoliteness strategies to intensify dissatisfaction or achieve multiple pragmatic goals. To ensure coding consistency, both coders underwent training in identifying speech act boundaries using pragmatic criteria. A coding guide was developed with explicit decision rules, including illustrative examples of single-strategy, multi-strategy, and borderline cases. Coders were instructed to err toward splitting rather than lumping when pragmatic boundaries were ambiguous, as this approach yields more granular data and reduces the risk of conflating distinct pragmatic functions.

3.2.2 Establishing coding reliability

In Phase I, two trained coders with high fluency in pragmatics and near-native proficiency in the respective languages independently coded the entire dataset. For inter-coder reliability reinforcement, a double-blind coding of a 20% randomized subset (60 reviews) was conducted. Agreement was measured adopting Cohen's Kappa (κ), which yielded a score of $\kappa = 0.85$, demonstrating high agreement beyond chance (Landis and Koch 1977). All differences were resolved through discussion, leading to a revised coding guide that was then applied to the full corpus.

Subsequently, shifting beyond frequency accounts, in Phase II, this study adopted a qualitative discourse analysis (Androutsopoulos 2008) to interpret the pragmatic function and cultural resonance of the strategies. This comprised examining how impoliteness was linguistically realized, its interactional goals, and how it reflected or contested broader cultural communication norms.

Impoliteness in OCRs	Working Definitions
On-record impoliteness <i>Positive impoliteness</i>	Reviews that conducted impoliteness straightforwardly, threatening the online retailers' positive or negative face directly. a. Reviews that questioned the authenticity and quality of the purchased products; b. Reviews that blamed the products, services, after-sales or logistics directly; c. Reviews that used taboo words.
<i>Negative impoliteness</i>	a. Reviews that threatened or expressed feelings such as disappointment, anger, annoyance b. Reviews that requested for further services, questioned the business, or made orders.
Off-record impoliteness <i>Implicated impoliteness</i>	Reviews that conducted impoliteness implicitly, through the implicature of the negative OCR. a. Reviews that advised for cautious purchase of potential customers; b. Reviews that showed doubts and unsolved problems; c. Reviews that blamed the products, services, after-sales or logistics indirectly; d. Reviews that compared the purchased product or service with those sold elsewhere.
<i>Sarcasm</i>	a. Reviews that made unfriendly or unkind joke about the purchased products in form of remarks or rhetorical questions; b. Reviews that appeared to be appropriate on the surface, but were meant to be taken as face-attack and critical in terms of face-management.

Ethical consideration

This study adhered to established ethical guidelines for research involving publicly available online data (Markham and Buchanan 2012; Franzke et al. 2020). All data collected were derived from publicly accessible consumer reviews on Taobao and Amazon. Because the reviews are publicly available and do not involve interaction with human subjects, formal institutional review board approval was not required. Nevertheless, the following ethical safeguards were implemented to protect user privacy and ensure responsible data use.

Data source and accessibility. All reviews included in this study were publicly available on the Taobao and Amazon platforms at the time of data collection (January 1, 2023, to June 30, 2023). No login credentials or special access privileges were required to view these reviews. The study did not involve any direct interaction with reviewers or platform users.

Anonymization and privacy protection. To safeguard user privacy, no personally identifiable information was collected or stored. Username, profile names, and any other direct identifiers were excluded from the dataset. There were no attempts made to contact reviewers or trace their identities, unless the textual content of reviews was retained for analysis. While in the presentation of findings, reviews are cited without any identifying information, ensuring that individual reviewers cannot be identified.

Data usage and purpose limitation. The collected data were used exclusively for academic research purposes related to cross-cultural pragmatics and impoliteness strategies in online consumer reviews. The data were not used for commercial purposes, nor were they shared with third parties. All data were stored securely and accessible only to the research team.

Compliance with platform policies. The data collection procedures were designed to comply with the respective platforms' terms of service and data usage

policies. Only publicly accessible content was collected, and no automated scraping tools or methods that violated platform restrictions were employed. The study's data collection involved manual retrieval of reviews in accordance with standard academic practices for publicly available online data.

Data integrity and transparency. In addition to privacy safeguards, the study ensured data integrity by excluding irrelevant or invalid comments (e.g. reviews commenting only on shipping speed, reviews with insufficient linguistic data). The content of the reviews was not modified except for necessary translations of Chinese comments into English for presentation purposes; translations were conducted with careful attention to preserving pragmatic force.

By implementing these measures, the study sought to balance the objectives of empirical research with the ethical responsibilities of respecting user privacy and adhering to platform norms.

4. Results and Finding

4.1. Impoliteness in Chinese vs. English OCRs

Because individual reviews may employ more than one impoliteness strategy, the total frequency of strategies (172 for Chinese, 161 for English) exceeds the number of reviews. Percentages in Table 2 are calculated based on the total strategy count per language, ensuring that subcategory percentages sum to 100% within each column.

The most significant finding of this study is not a difference, but a striking similarity. As depicted in Table 2, the distribution of impoliteness strategies between Chinese and English reviews is nearly identical, with a strong preference for **on-record impoliteness** in both corpora (Chinese: 62.21%; English: 62.11%). This quantitative parallel depicts a theoretical challenge to determinis-

Table 2: Impoliteness strategies across Chinese vs. English

Language	Chinese(n =150)		English (n = 150)	
	Frequency	Percentage	Frequency	Percentage
Impoliteness Strategies				
On-record impoliteness	107	62.21%	100	62.11%
Positive impoliteness	88	51.16%	79	49.07%
Negative impoliteness	19	11.05%	21	13.04%
Off-record impoliteness	65	37.79%	61	37.89%
Implicated impoliteness	60	34.88%	54	33.54%
Sarcasm	5	2.91%	7	4.35%
Total	172		161	

tic models of cross-cultural communication. It suggests that the context of the online review genre itself imposes a powerful, homogenizing effect. The main *pragmatic functions* of a negative OCR were to issue a firm warning, require for refinement, and perform dissatisfaction for a public audience which basically privilege directness and clarity. In the form of anonymous followed by non-relationship-based setting, the efficiency of bald-on-record impoliteness outweighs the normative cultural preference for indirectness that might govern face-to-face interactions. This finding is parallel to the concept of context collapse (Kini et al. 2022) and online disinhibition (Meriläinen and Ruotsalainen 2024), where the specific affordances of the digital platform including anonymity, asynchronicity, and the public nature of reviews can temporarily suspend or reshape default communicative norms. These platform-specific factors complicate any straightforward application of cultural frameworks derived from face-to-face interaction.

While the strategic choice is quantitatively similar, a qualitative analysis reveals that the linguistic realization and pragmatic framing of these strategies are shaped by culturally influenced communicative preferences though importantly, these preferences function as tendencies instead of deterministic patterns. The cultural divergence, hence, lies not in whether to be impolite, but in how to be effectively and meaningfully impolite within one's cultural framework, as mediated by the specific context of the OCR genre. Both groups frequently adopt positive impoliteness, but the nature of the attack differs in ways that reflect broader communicative tendencies.

In the Chinese OCRs, this strategy often functions as a communal sanction. This can be seen when the reviewer acts as a guardian of the collective, leveraging shared values to amplify the face-attack. This pattern aligns with tendencies toward collectivist orientation, though it is not a universal or invariant feature of all Chinese reviewers. *Example 2*: "大家不要买·超级不好用·谁买谁后悔" ("Everyone, don't buy this; it's super unusable, whoever buys it will regret it"). By analyzing pragmatically, the adoption of "大家" (everyone) and the generic warning "谁买谁后悔" frames the serious complaint of public service announcement instead of solely on personal grievance. The impoliteness is amplified by invoking collective solidarity against the seller.

Example 2A: "兄弟姐妹们听我一句劝, 这家店的东西千万别买! 我买了三件, 一件都没法穿, 料子差得要命, 洗一次就缩水变形。这不是坑人吗? 大家擦亮眼睛, 别像我一样上当受骗!" ("Brothers and sisters, listen to my advice. Don't ever buy from this store! I bought three items, not one of them is wearable. The fabric is terrible; it shrank and deformed after just one wash. Isn't this a scam? Everyone, keep your eyes open, don't be fooled like I was!"). Here, the address "兄弟姐妹们" (brothers and sisters) invokes familial solidarity, framing the reviewer as a protective elder warning kin against harm. The rhetorical question "这不是坑人吗?" (Isn't this a scam?) invites shared moral judgment

rather than simply reporting individual grievance. The impoliteness is not directed at the seller in isolation but is amplified through the construction of an imagined community of consumers whose welfare the reviewer champions.

Example 2B: "各位买家注意了！这家店卖假货还不承认，客服态度极其恶劣。我已经投诉到平台了，希望大家一起抵制这种无良商家！不能让这种人继续坑害消费者！" ("Attention all buyers! This store sells fake products and won't admit it. The customer service attitude is extremely poor. I've already reported them to the platform. I hope everyone will join together to boycott these unscrupulous merchants! We cannot let such people continue to harm consumers!").

This example demonstrates collective warning through explicit calls to collective action. The phrase "无良商家" (unscrupulous merchants) invokes a moral category that positions the seller as violating communal ethical standards. The call to "一起抵制" (join together to boycott) transforms individual dissatisfaction into a collective movement, weaponizing collective identity to amplify the face-attack.

In English OCRs, forensic and individualistic accusation more prevalently reflects positive impoliteness. The reviewer builds a case based on personal observation and evidence. It can be seen in Example 4 where the utterance "the first sign was the misplaced label.... The product did not have a seal, only a sticker illustrating the product was new. By analyzing pragmatically, through a meticulous approach and evidence-based destruction of the seller's credibility can attain impoliteness. The strong influence of the complaint derives from its logical coherence and the presentation of factual and observable details. This aligns with low-context, individualistic norms that value personal autonomy and persuade through reasoned argument and objective proof.

The finding of directness in Chinese reviews may appear counter-intuitive given the frequent characterization of Chinese communication as high-context (Hall 1976) and oriented toward indirectness (Gu 1990). However, rather than viewing this as a contradiction, it can be understood through the concept of contextual salience. In the context of OCRs, the primary objective is not to maintain interpersonal harmony (*rapport-management*) but to publicize a breach of trust and warn other consumers. The cultural script shifts from 'maintaining face' to 'rectifying a wrong' and 'asserting consumer rights.' The directness observed is therefore not a rejection of cultural norms but an adaptation of them to a specific genre context. The reviewer employs culturally potent, collective-oriented directness to achieve a contextually appropriate goal. This illustrates that culture does not prescribe fixed communicative behaviors but rather provides a repertoire of resources that individuals strategically deploy according to context.

4.2. Analysis of individual impoliteness strategies

Positive impoliteness targets the addressee's positive face namely the desire to be liked, respected, and socially validated. When taboo language or insults are directed at the product, quality, or identity of the seller such as calling a product "garbage" or questioning the seller's honesty, these are classified as positive impoliteness because they attack the seller's reputation and social standing. In contrast, negative impoliteness targets the addressee's negative face namely the desire for autonomy, freedom from imposition, and freedom of action. When taboo language or insults are used to express anger that demands action, threatens consequences, or restricts the seller's operational freedom such as *"I will never buy from this company again unless they change"*, these are classified as negative impoliteness. The presence of taboo words may intensify the emotional force, but the classification hinges on whether the primary pragmatic function is to attack the seller's reputation (positive impoliteness) or to impose upon the seller's autonomy (negative impoliteness).

This distinction is summarized in Table 3 below.

4.2.1. Positive impoliteness

As one of the most common impoliteness strategies in negative OCRs, positive impoliteness strategy mainly involves questioning the authenticity and quality of the purchased product, such as concerns about the product's genuine-

Table 3: The difference between positive and negative impoliteness

Impoliteness Type	Pragmatic Target	Definition	Linguistic Indicators	Example
Positive Impoliteness	Positive face (reputation, social standing, identity)	Attacks the seller's credibility, product quality, or social image	– Direct product insults (e.g. "garbage," "trash") – Accusations of dishonesty (e.g. "fake," "scam") – Ridicule of seller competence	"这纸谁买谁后悔 · 假的假的" (<i>"Anyone who buys this paper will regret it; it's fake, it's fake"</i>) attacks product authenticity and seller credibility
Negative Impoliteness	Negative face (autonomy, freedom from imposition)	Threatens the seller's operational freedom; demands action; imposes restrictions	– Purchase withdrawal threats (e.g. "never buy again") – Conditional ultimatums (e.g. "unless you...") – Expressions of anger that demand redress	<i>"I am never ever buying from this company again unless they decide to start caring more about living creatures"</i> restricts seller's future autonomy through conditional threat

ness. Additionally, consumers directly accuse the product, service, logistics, or after-sales service, which also falls under the use of positive impoliteness strategies. The use of taboo words such as 'garbage' or 'trash' can appear in both positive and negative impoliteness, depending on the pragmatic target. When such terms attack the product quality or seller credibility, they function as positive impoliteness; when they intensify demands or threats that restrict seller autonomy, they function as negative impoliteness (see Table 3 above for the classification framework). Examples 1 to 4 illustrate how positive impoliteness works in the Chinese and English OCRs.

Example 1: 使用效果非常差, 遇水都化成纸屑了, 没办法擦手, 第一次买到感觉像假冒伪劣的产品。非常不好用。

Translation: The use effect is very poor. It disintegrates into paper scraps when it gets wet, making it impossible to wipe hands. First time buying, and it feels like a fake, counterfeit product. Very unusable.

Example 2: 这纸抽都抽不出来, 说三层的跟人家一层的差不多厚, 这纸谁买谁后悔, 假的假的。大家不要买, 超级不好用, 谁买谁后悔...□纸超薄, 还抽不出来, 都粘在袋子上 很不好用, 假清风

Translation: You can't even pull the tissues out. They claim it's 3-ply, but it's about as thick as 1-ply. Anyone who buys this will regret it. It's fake, it's fake. Don't buy; it's super unusable, and anyone who buys it will regret it. Super thin, and you can't pull it out; it sticks to the bag. Very bad to use. Fake Qingfeng.

Example 3: I never write reviews but this needs to be said. This is fake Keihl's. The instant I put this on my face my face burned and stung horribly. My eyes watered and stung until I washed it off. This is a dangerous product so be cautious.

Example 4: Do not buy this product! It is fake! The first sign was the misplaced label, which was placed so low that it curled at the bottom. The product did not have a seal, only a sticker indicating the product was "new". I opened the product to compare it to my original Kiehls and the seal was broken. DO NOT BUY!

In Example 1, the consumer repeatedly mentions poor performance and difficulty of use, which reflects positive impoliteness. In addition, the example also highlights specific details, such as pointing out that the product is a counterfeit, as the paper turned into scraps when it came into contact with water. This directly calls attention to the poor quality of the product.

In Example 2, the consumer questions the authenticity of the product, stating twice that the product is fake. Additionally, the author directly criticizes the product by stating that the tissue cannot be pulled out, it's inconvenient to use, and that anyone who buys it will regret it. This is a direct accusation of the product.

In Example 3, the consumer claims that using the product caused a painful burning sensation on their face, but after washing it off, everything was fine. This leads them to directly state that the product is fake. By accusing the product of being fake, the consumer is directly attacking the product's credibility, which

also threatens the reputation of the seller. Therefore, Example 3 also falls under positive impoliteness.

In Example 4, the consumer directly points out that the product is fake and provides specific details. They note that the product lacks a seal and the label is placed incorrectly, further supporting the claim of it being counterfeit.

As demonstrated in Examples 1-4, in the anonymous CMC (Computer-Mediated Communication) environment of OCRs, reviewers use a variety of expressions to convey their dissatisfaction. By highlighting the poor quality of the purchased products and questioning their authenticity, they attack the quality and reputation of retailers on platforms like Taobao and Amazon. Moreover, when looking at the frequency of strategies, positive impoliteness strategies account for the largest proportion among all impoliteness strategies, making them the most common in OCRs. Browsing through unpleasant real user experiences and explicit recommendations not to buy can have a significant impact on potential customers, influencing their purchasing decisions.

It should be noted that while taboo words such as 'garbage' or 'disgusting' may appear in these examples, their classification as positive impoliteness is determined by their pragmatic target: they attack the product's quality and, by extension, the seller's reputation. The same taboo words, if used to intensify a threat or demand that restricts the seller's autonomy, would be classified as negative impoliteness (see Section 4.2.3.).

4.2.2. *Implicated impoliteness*

Implicated impoliteness refers to face-threatening acts that are conveyed indirectly, rather than through direct attacks on someone's face. Implicated impoliteness often involves indirect criticism of the product, after-sales service, or logistics. It is the second most common type of impoliteness strategy among all impoliteness strategies. Examples 5-8 demonstrate how implicated impoliteness is reflected in both Chinese and English negative OCRs.

Example 5: 看上成分才买的, 但有股油漆味, 假滑

Translation: Purchased for ingredients, but smells like paint and feels greasy.

Example 6: 买过很多次了, 这次的新包装味道跟之前一点也不一样, 想一股发霉的味道, 打出的泡沫也没有之前的多, 很水

Translation: Bought multiple times, but the new packaging smells not the same as before, like a musty taste and has less foam than before.

Example 7: Package was thrown and damaged. The product leaked through the box it's leaked all over the concrete

Example 8: I've never had acne issues before. Just stubborn chin blackheads so I figured it was time to switch it up. This review is long overdue and late so I have scars now rather than pimples. I saw this brand had great reviews and it worked for a lot of people so I was thinking I'd try it. This made my face

breakout so much. It's 5 months later and my scars are still very prominent. Hoping they heal soon with my new skin care brand that has been working so far. In the pictures are both my cheeks.

In Example 5, the reviewer points out that the product has a paint-like smell and feels slippery, which clearly indicates dissatisfaction with the product. However, the reviewer does not directly say that the product is bad or that they dislike it. By highlighting certain quality issues, the reviewer indirectly criticizes the product, which is an implicit form of impoliteness.

In Example 6, the consumer mentions that the packaging of the product is different from the previous ones. In addition, the consumer points out that the product has a moldy smell and that the foam produced is less than before. Clearly, the consumer is implying doubts about the product's authenticity. However, the consumer does not explicitly state that the product is counterfeit; instead, they use phrases like "different from before" and "less foam" along with the mention of a moldy smell. These clues suggest that the consumer is indirectly criticizing the product's quality and expressing doubts about its authenticity.

In Example 7, the review mentions that the packaging was damaged and the product leaked, which indirectly suggests a quality issue with the product. However, the reviewer does not directly attack the seller's reputation or claim that the product is of poor quality, making this an example of implicated impoliteness. Similarly, in Example 8, the consumer states that her face has scars and was severely damaged after using the product. However, the reviewer does not directly assert that the product is bad or has quality problems. They only express their personal experience after using the product without directly attacking the seller's face.

In conclusion, as demonstrated by these examples, implicated impoliteness strategies in negative OCRs can be realized in various ways. For example, they can be expressed through indirect criticism of the product, after-sales service, logistics, or customer service. Implicit reminders not to buy the product can also serve as a form of impoliteness. Additionally, comparing the purchased product to other products can also be an implicated form of impoliteness.

4.2.3. Negative impoliteness

Negative impoliteness strategies attack the addressee's negative face which is the desire for autonomy and freedom from imposition. While taboo language or emotionally charged expressions may accompany such strategies, the defining feature is not the presence of taboo words but the pragmatic function of restricting the seller's operational freedom, issuing threats, or imposing demands. In the following examples, note that the classification as negative impoliteness stems from these functional criteria instead of the use of taboo language alone.

Among all impoliteness strategies, negative impoliteness is the third most common type of impolite behavior. Negative impoliteness strategies account for

11.05% of all impoliteness strategies in Chinese, and 13.04% in English. Examples 9-12 demonstrate how negative impoliteness strategies are expressed in negative OCRs.

Example 9: 恶心 美白效果扯淡 什么破玩意□这么贵价格还没用, 不买了以后 保湿控油没效果 吸收效果无用

Translation: Disgusting. whitening effect bullshit. what shit, so expensive. It's useless. Won't buy again. The moisturizing oil control has no effect. Absorption effect is useless

Example10: 整体评价 差 质地手感 差假货 吸收效果差 一塌糊涂 产品质量差得不行 用了起痘烂脸, 退了货, 卖家收到货, 赖着不退款, 说没收到货, 倒了十八辈子 血霉买他家东西找气受

Translation: Overall evaluation: Poor. Texture and feel: Poor, fake product. Absorption effect: Poor, a complete mess. Product quality: Terrible. Caused breakouts and ruined my face, returned the product, but the seller refused to refund, claiming they didn't receive the item back. It was a terrible experience buying from them, nothing but frustration.

Example 11: I've used this over the last couple of days for the first time, this morning my finger has popped an air bubble which now shows the true amount of contents in the tub!!! That is the bottom of the tub you can see in the middle!!! I am absolutely disgusted I have spent £20 on an item with half of product! Not that I can afford to waste this kind of money on skin care in this climate! I am absolutely fuming

Example 12: I just found out that this company tests on animals!!! I will be using the Inkey List SA cleanser from now on because they are cruelty free!! Not to mention this cleanser is literally cheaper AND better. I will be finishing the bottle I have now then I am never EVER buying from this company again unless they decide to start caring more about living, breathing, creatures, instead of adding more money to the billions of dollars they already have!!

In Example 9, the reviewer says "disgusting" mentions the lack of whitening effect, and refers to the product as "garbage" which clearly indicates the reviewer is extremely angry. Additionally, the statement "I will never buy again" is a direct threat to the seller's reputation, which is a form of negative impoliteness.

In Example 10, the consumer states that buying the product was a way of "seeking trouble" showing that they are very upset. The phrase "this is bad luck for generations" further emphasizes the consumer's frustration and anger. This is an expression of negative impoliteness.

In Example 11, the consumer uses numerous exclamation marks to stress their dissatisfaction. The use of words like "fuming" and "disgusted" further illustrates the consumer's anger and disappointment toward the seller, again reflecting negative impoliteness.

In Example 12, the consumer says they will "never buy from them again unless they do something about it" This is a threat to the seller, demanding

a change in their actions. The statement "I am never ever buying... again, unless.." restricts the seller's freedom to act, which is an example of negative impoliteness. This represents a challenge to the retailer's autonomy and threatens their reputation.

In summary, unlike positive impoliteness, negative impoliteness attacks the operational freedom of Taobao and Amazon. Consumers express threats, disappointment, and anger, and sometimes demand further services or actions before making another purchase. These behaviors affect the rights and interests of online retailers on platforms like Taobao and Amazon (Spencer-Oatey 2008).

4.2.4. *Sarcasm*

This study found that sarcasm is the least frequently used impoliteness strategy, both in Chinese and English OCRs. Examples 13-16 demonstrate how sarcasm is expressed in reviews.

Example 13: 实物与宣传不符 谁买谁是大冤种**商家虚假宣传 还在看评价的赶紧退出去别拍！！自己看图 量了尺寸不够 客服竟然让我用20cm的尺子 自己家产品质量心里不清楚吗 刁难我？我带着外包装说小客服不承认让我打开量里面的纸 我打开了又不给我退款说影响二次销售 恶心至极 清风一生黑

Translation: The product does not match the advertisement. Whoever buys this is a sucker. The seller is false advertising. If you're still looking at reviews, quickly exit and don't buy! Look at the pictures. I measured, and it's too small. Customer service told me to use a 20cm ruler. Don't they know the product's quality? I showed them the packaging, and they told me to open it to measure the inside. After opening it, they refused to refund, claiming it affects resale. It's absolutely disgusting. Qingfeng is blacklisted forever.

Example 14: 刚买还没用就空一大块，缺斤少两，想用常规版的包装，结果发个花里胡哨的，那么个大品牌做这样的事，也是绝绝子了。干皮擦了还是一样干，一点用都没有，不再回购系列。

Translation: I just bought it but haven't used it yet, and there's a huge empty space inside – it's missing product. I wanted the regular packaging, but they sent me some flashy one instead. For such a big brand to do this is really wonderful. I have dry skin, and it still feels dry after using it – no effect at all. I won't be repurchasing from this line.

Example 15: I've always had great skin and I haven't had a pimple in over 20 years. Since I started using this product, I look like a 15 years old kid going thru puberty. Even when I was 15, I never struggled with acne and always had clear skin. I do not have sensitive skin at all. Bottom line is my pores aren't any smaller and my face looks like a slice of pepperoni pizza.

Example 16: If you're like me and deal with oily skin around your T-zone, DO. NOT. BUY. THIS. PRODUCT. It has corn oil listed as an ingredient, which

is a complete joke in and of itself. The amount of oil my face produced after using this product was scary. You would've thought I ran a 5k in the amazon rainforest by the looks of my face when I woke up in the morning. If you're going to use a CeraVe cleanser, use the Foaming Facial Cleanser for normal to dry skin and save yourself from nightmare-ish oily skin.

In Example 13, the reviewer says that the actual product does not match the advertisement, accusing the seller of false advertising, which implies that the product's quality is poor. The reviewer then sarcastically asks the seller if they are unaware of the quality of their own product. In most cases, if the product were of good quality, the seller would be well aware of it. This rhetorical question is clearly a form of sarcasm, mocking the seller's lack of honesty or awareness about their product's poor quality.

In Example 14, the reviewer uses the phrase "绝绝子" (which generally means "fantastic" or "awesome") in a sarcastic way. Normally, "绝绝子" would be a compliment, but from the context of the review, it's clear that the product is subpar, with issues such as it being sold with a shortage of product. The consumer sarcastically uses this phrase to express disappointment in the brand's actions, not to praise it. This is a form of sarcasm, mocking the brand's poor performance.

In Example 15, the reviewer mentions that they have had clear skin for over 20 years without ever suffering from acne, but after using the product, their skin became like that of a teenager going through puberty. This sarcastic remark suggests that the product caused the reviewer to break out in acne. Furthermore, the reviewer compares their face to an "Italian sausage pizza" implying that their skin became unhealthy after using the product. This sarcastic comparison not only criticizes the product but also attacks the seller's reputation by suggesting the product had a negative effect on their appearance.

In Example 16, the reviewer calls the product a "joke" sarcastically implying that it is of poor quality. In addition, the reviewer humorously describes how the oil on their face in the morning was like running 5 kilometers in the Amazon rainforest, exaggerating the excessive oil production caused by the product. This humorous exaggeration is a sarcastic way of highlighting how ineffective the product is, thus indirectly mocking its performance.

In conclusion, reviewers often use unfriendly jokes or metaphors to attack the face of Taobao and Amazon, and these jokes are typically disguised as rhetorical questions. In negative OCRs, using sarcasm as an impoliteness strategy is more persuasive and makes the reviewer's expression more humorous. Although sarcasm is the least frequently used impoliteness strategy, it is a unique form of impoliteness and serves as a distinctive rhetorical strategy. By incorporating sarcasm, consumers can deliver their criticisms in a way that both undermines the product or seller and entertains or engages the audience, making it a particularly effective but less common approach.

5. Discussion

5.1. Frequency of Impoliteness Strategies (RQ1)

The quantitative findings presented in Table 2 reveal a striking similarity instead of a significant difference. Both corpora demonstrated a strong preference for on-record impoliteness (Chinese: 62.21%; English: 62.11%), with positive impoliteness emerging as the most frequently adopted subcategory in both languages (Chinese: 51.16%; English: 49.07%). Similarly, off-record impoliteness accounted for comparable proportions (Chinese: 37.79%; English: 37.89%), with implicated impoliteness far outnumbering sarcasm in both datasets. This quantitative parallel challenges overly deterministic interpretations of cultural frameworks such as Hall's (1976) high-context/low-context distinction or Hofstede's (1984) individualism/collectivism dimension. While these frameworks offer valuable insights into broad cultural tendencies, they risk essentializing cultural differences if applied without attention to contextual factors. The findings suggest that the context of the online review genre itself exerts a powerful, homogenizing effect that can override or reshape culturally patterned preferences. This does not mean that culture is irrelevant; in contrast, it indicates that cultural influence is *contextually mediated* and *non-deterministic*. In this specific genre characterized by anonymity, public accountability, and the pragmatic goal of warning others, the preference for directness becomes functionally salient across both cultural groups.

The main pragmatic functions of a negative OCR are to issue a firm warning, demand remediation, and perform dissatisfaction for a public audience privilege directness and clarity. In the form of anonymous, non-relationship-based setting, the efficiency of bald-on-record impoliteness outweighs the normative cultural preferences for indirectness that might govern face-to-face interactions. This finding aligns with the concepts of context collapse (Kini et al. 2022) and online disinhibition (Meriläinen and Ruotsalainen 2024), where the specific affordances of digital platforms can temporarily suspend or reshape default communicative norms. Therefore, in response to RQ1, **no substantial cross-cultural difference was observed in the frequency distribution of impoliteness strategies**, suggesting that genre conventions exert a strong homogenizing force across cultural contexts.

5.2. Pragmatic realization of impoliteness (RQ2)

While the strategic choice was quantitatively similar, the qualitative analysis revealed that the linguistic realization and pragmatic framing of these strategies are profoundly shaped by culture. The cultural divergence, therefore, lies not in whether to be impolite, but in how to be effectively and meaningfully impolite within one's cultural framework. Positive impoliteness, the most frequent strategy

in both corpora, manifested differently across cultures. In Chinese OCRs, this strategy often functioned as a communal sanction. As illustrated in Example 2, the reviewer acts as a guardian of the collective, leveraging shared values to amplify the face-attack: "大家不要买 · 超级不好用 · 谁买谁后悔" ("Everyone, don't buy this; whoever buys it will regret it"). The adoption of "大家" (everyone) and the generic warning frame the complaint as a public service announcement instead of a personal grievance. This reflects a collectivist orientation where social validation serves as a powerful rhetorical tool. In contrast, English OCRs manifested positive impoliteness through forensic and individualistic accusation. As shown in Example 4, the reviewer builds a case based on personal observation and evidence: "The first sign was the misplaced label... The product did not have a seal, only a sticker" Impoliteness is achieved through meticulous, evidence-based destruction of the seller's credibility. This aligns with *tendencies* associated with low-context, particularly individualistic orientations, a preference for explicit reasoning, evidence-based argumentation, and appeals to personal autonomy. It is important to emphasize that these are tendencies, instead of fixed characteristics; individual reviewers may diverge from these patterns, and the online context itself may amplify or suppress certain communicative preferences.

Implicated impoliteness, the second most frequent strategy, also revealed cultural divergence. Chinese reviewers tended to use subtle cues namely comparisons to previous purchases or allusions to unfamiliar smells to imply inauthenticity without direct accusation (Examples 5 and 6). English reviewers, while also employing indirectness, however, often framed implicated impoliteness within detailed personal narratives that documented consequences over time (Example 8), reinforcing the cultural preference for evidence-based argumentation. Sarcasm, though the least frequent strategy, served distinct cultural functions. Chinese sarcasm often involved rhetorical questions that mocked seller incompetence (Example 13) or ironic praise (Example 14), while English sarcasm frequently employed exaggerated metaphors and humorous analogies (Examples 15 and 16) to underscore product failure. Thus, in response to RQ2, the pragmatic realization of impoliteness exhibits meaningful cross-cultural *tendencies* that align with broader cultural orientations identified in the work of Hall (1976) and Hofstede (1984). Chinese reviewers in this corpus demonstrated a greater tendency toward communal, collectivist-oriented framing, while English reviewers showed a greater tendency toward individualistic, evidence-based narratives. However, these findings should not be interpreted as evidence of culturally deterministic patterns. The observed differences are best understood as *probabilistic tendencies* that emerge within the specific context of OCRs as a context that itself shapes how cultural resources are deployed. Moreover, the considerable overlap between the two corpora (e.g. both groups used on-record impoliteness at comparable rates) underscores that culture interacts with, instead of determines and communicative behavior.

5.3. The role of genre conventions (RQ3)

The findings suggest that OCRs have evolved into a distinct genre with its own pragmatic conventions that transcend cultural boundaries. The strong preference for on-record impoliteness in both corpora indicates that the genre prioritizes directness, clarity, and impact features that serve the primary functions of warning potential buyers and compelling seller response. However, genre conventions do not omit cultural specificity; in contrast, they interact with cultural norms in complex ways. The concept of contextual salience helps explain this interaction. In the context of OCRs, the objective is not to maintain harmony (*rappport-management*) but to rectify a perceived wrong and warn others. The cultural script shifts from "maintaining face" to "asserting consumer rights"

For Chinese reviewers, directness is not a rejection of culturally patterned preferences but an adaptation of them to the OCR genre. They draw upon culturally available resources namely collective-oriented rhetoric and appeals to shared values to achieve the genre's functional goals. For English reviewers, the genre's demand for directness aligns more seamlessly with tendencies associated with low-context, individualistic communication norms, potentially resulting in a more straightforward alignment between genre expectations and culturally preferred styles. However, this observation should not be taken to imply that English reviewers are uniformly 'direct' or that Chinese reviewers are uniformly 'indirect'; however, it points to differing *resources* that reviewers draw upon when navigating the generic demands of OCRs. Importantly, the online environment introduces additional mediating factors. Anonymity, asynchronicity, and the public, transactional nature of reviews can alter how cultural norms are enacted (Meriläinen and Ruotsalainen 2024). What might be considered excessively confrontational in face-to-face interaction may be perceived as acceptable even expected in the context of online consumer feedback. Therefore, the interplay between culture, genre, and digital context is dynamic and non-linear; cultural influences are not simply 'reflected' in language but are actively negotiated and sometimes transformed through the affordances of the medium.

This study thus shifts beyond a mechanical application of Culpeper's model by emphasizing the pragmatic function of the strategies. The choice of a specific linguistic realization which was a collective warning *vs.* an evidentiary account constitutes a performance of cultural identity. For the Chinese reviewer, the bald-on-record insult is not solely an emotional outburst; but it is a tool for enacting communal judgment. For the English reviewer, the detailed narrative is not solely a complaint; however, it is a performance of rational individualism. Therefore, genre conventions provide the structural framework within which impoliteness occurs, but culture fundamentally shapes how impoliteness is pragmatically weaponized to be most meaningful and effective within a specific cultural world-view.

6. Conclusion

In response to **RQ2 (pragmatic realization)**, the study revealed that despite quantitative similarities, qualitative differences in linguistic framing persist. Chinese reviewers demonstrated a *tendency* toward collectivist-oriented strategies, framing complaints as communal warnings and leveraging shared values to amplify face-attacks. However, English reviewers, demonstrated a *tendency* toward individualistic, evidence-based narratives that construct impoliteness through logical argumentation and personal accountability. These patterns align with broader cultural-communicative tendencies associated with collectivism vs. individualism and high-context vs. low-context communication styles (Hall 1976; Hofstede 1984). However, it is crucial to interpret these findings as *probabilistic tendencies* instead of deterministic cultural prescriptions. The considerable overlap between the two corpora and the near-identical rates of on-record impoliteness underscores that culture does not rigidly determine communicative behavior. Instead, cultural orientations provide repertoires of resources that individuals strategically deploy in response to genre conventions, platform affordances, and the specific interactional goals of the OCR context.

References

- Androutsopoulos, J. 2008. Potentials and limitations of discourse-centred online ethnography. *Language@ internet* 5(8).
- Bousfield, D. 2008. *Impoliteness in interaction*. Amsterdam: John Benjamins.
- Bousfield, D., and M. Locher (eds.) 2008. *Impoliteness in Language: Studies on its Interplay with Power in Theory and Practice* (Vol. 21). Berlin: Walter de Gruyter.
- Brown, P., and S. Levinson 1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.
- Cenni, I., and P. Goethals 2017. Negative hotel reviews on TripAdvisor: A cross-linguistic analysis. *Discourse, Context & Media* 16: 22-30.
- Chen, J., and D. Feng 2023. Strategic impoliteness in e-commerce communication: How online shops challenge negative customer reviews. In C. Xie (ed.), *Advancing (Im)politeness Studies: Cultural, Digital and Emotional Aspects*, 203-223. Cham: Springer International Publishing.
- Culpeper, J. 1996. Towards an anatomy of impoliteness. *Journal of Pragmatics* 25(3): 349-367.
- Culpeper, J. 2010. Conventionalised impoliteness formulae. *Journal of Pragmatics* 42(12): 3232-3245.
- Culpeper, J. 2011. *Impoliteness: Using Language to Cause Offence*. Cambridge: Cambridge University Press.

- Culpeper, J. 2005. Impoliteness and entertainment in the television quiz show: The Weakest Link. *Journal of Politeness Research: Language, Behaviour, Culture* 1(1): 35-72.
- Culpeper, J., and C. Hardaker 2017. Impoliteness. In J. Culpeper, M. Haugh and D. Kádár (eds.), *The Palgrave Handbook of Linguistic (Im)politeness*, 199-225. London: Palgrave Macmillan.
- Culpeper, J., L. Marti, M. Mei, M. Nevala and G. Schauer 2010. Cross-cultural variation in the perception of impoliteness: A study of impoliteness events reported by students in England, China, Finland, Germany and Turkey. *Intercultural Pragmatics* 4(7): 597-624.
- Culpeper, J. 2011. Politeness and impoliteness. In K. Aijmer and G. Andersen (eds.), *Sociopragmatics (Vol. 5). Handbook of Pragmatics*, 391-436. Berlin: Mouton de Gruyter.
- Feng, W., and W. Ren 2020. Impoliteness in negative online consumer reviews: A cross-language and cross-sector comparison. *Intercultural Pragmatics* 17(1): 1-25.
- Gu, Y. 2011. Modern Chinese politeness revisited. In F. Bargiela-Chiappini and D.Z. Kádár (eds.), *Politeness Across Cultures*, 128-148. London: Palgrave Macmillan UK.
- Guilford, J.P. 1959. *Personality*. New York: McGraw-Hill.
- Hall, E.T. 1976. *Beyond Culture*. New York: Doubleday.
- Hall, E.T. 2000. Context and meaning. *Intercultural Communication: A Reader* 9: 34-43.
- Hofstede, G. 1984. *Culture's Consequences: International Differences in Work-Related Values* (Vol. 5). Beverly Hills, California: Sage.
- Jackson, L.A., and J.L. Wang 2013. Cultural differences in social networking site use: A comparative study of China and the United States. *Computers in Human Behavior* 29(3): 910-921.
- Kini, S., M.P. Shelat and V. Jain 2022. Conceptualizing 'Filter-ing': Affordances, context collapse, and the social self online. *International Journal of Communication* 16: 1573-1593.
- Lai, X. 2019. Impoliteness in English and Chinese online diners' reviews. *Journal of Politeness Research* 15(2): 293-322.
- Landis, J.R., and G.G. Koch 1977. The measurement of observer agreement for categorical data. *Biometrics* 33(1): 159-174.
- Lorenzo-Dus, N., P.G.C. Blitvich and P. Bou-Franch 2011. On-line polylogues and impoliteness: The case of postings sent in response to the Obama Reggaeton YouTube video. *Journal of Pragmatics* 43(10): 2578-2593.
- Meriläinen, M., and M. Ruotsalainen 2024. Online disinhibition, normative hostility, and banal toxicity: Young people's negative online gaming conduct. *Social Media + Society* 10(3): 1-12.
- Mills, S. 2009. Impoliteness in a cultural context. *Journal of Pragmatics* 41(5): 1047-1060.

- Mills, S., and D.Z. Kadar 2011. Politeness and culture. In D.Z. Kádár and S. Mills (eds.), *Politeness in East Asia*, 21-44. Cambridge: Cambridge University Press.
- Spencer-Oatey, H. 2011. Conceptualising 'the relational' in pragmatics: Insights from metapragmatic emotion and (im)politeness comments. *Journal of Pragmatics* 43(14): 3565-3578.

MAGDALENA PERZ

Université de Silésie à Katowice, Pologne

Faculté des sciences humaines

ORCID: 0000-0003-1446-6920

magdalena.perz@us.edu.pl

LA SYNONYMIE EN QUESTION : DU LEXIQUE AU DISCOURS

SYNONYMY INTO QUESTION : FROM LEXICON TO DISCOURSE

The purpose of the present article is to discuss, through adjectival uses, a number of difficulties that affect the description of synonymy relation. The paper reviews various approaches to synonymy by examining a number of fundamental concepts. The article offers a theoretical and applicative dimension devoted to the field of adjectives. The elements of the co-texte are the factor which takes part in the emergence of the meaning – therefore they are involved in the synonymic process. The contextual approach adopted to describe synonymous units can overcome the shortcomings of classical dictionaries of synonyms.

Keywords: adjectival synonymy, distributional classes, object-classes, co-text, idiom principle

1. Introduction

Un regain d'intérêt pour les études relatives à la synonymie – son traitement et son encodage demeure intéressant à plus d'un titre. D'abord, la plupart des résultats théoriques ne semblent pas encore trouver d'écho dans la pratique lexicographique. En effet, la synonymie est souvent présentée davantage comme un domaine de recherche théorique qu'appliqué. L'absence des solutions opératoires se reflète dans des dictionnaires de synonymes qui se présentent, dans la majorité des cas, sous forme d'un listing des termes proches. Le terme de synonymie reçoit des interprétations diverses selon les écoles et les réponses s'avèrent discordantes. Par ailleurs, les approches relatives à la synonymie incarnent de multiples fluctuations théoriques survenues au cours des années.

Dans cet article nous allons nous pencher sur les questions qui se situent à la frontière entre la théorie et la pratique en présentant nos dernières avancées dans le domaine de la synonymie d'adjectifs. Nous le ferons en deux étapes : en explorant les bases théoriques sur la synonymie à la lumière des préoccupations propres à la représentation linguistique. Dans un deuxième temps, nous recourrons à une analyse succincte de l'adjectif *fort* pour prouver que les choix synonymiques ne sont pas libres, mais conditionnés par les mécanismes de restriction lexicale, parfois au point de conduire à une non-substituabilité par tout autre mot.

2. Enjeux pratiques de la synonymie

La description des unités synonymiques constitue un intérêt considérable pour la didactique du lexique (substitutions synonymiques) ainsi que pour un nombre important d'applications en lexicographie comme modélisation des relations lexicales. Comme l'ont souligné A. Lehmann et F. Martin-Berthet (2018 : 54), l'existence de la synonymie relève d'une pratique intuitive, mais la synonymie demeure une notion problématique. Les dictionnaires de synonymes¹ sont, dans la majorité des cas, des dictionnaires du type cumulatif – ils présentent les termes censés être synonymes suivant divers critères. Dans de tels dictionnaires, on a beau chercher des nuances d'emplois des mots ou le degré de leur synonymie. Le problème se pose de savoir comment hiérarchiser cet inventaire des acceptions des unités dites synonymiques. Comment structurer leurs possibilités syntaxiques afin de rendre compte de leur fonctionnement. Le traitement de la synonymie fait ressortir quelques difficultés en sémantique lexicale et nous croyons indispensable de revenir sur quelques aspects que cette notion implique.

Le survol des études sur la synonymie révèle que ce concept a été traité de deux perspectives opposées. L'une qui prend comme point de départ le niveau virtuel (Apresjan 1974 ; Mel'čuk 1984, 1988 ; Culioli 1990) et l'autre qui part du niveau de l'expression linguistique vers la détermination du contenu sémantique. (Ploux et Victorri 1998).

3. Synonymie – de contenu à expression

Ayant sa source au départ dans l'étude des équivalences sémantiques, la synonymie a été d'abord appréhendée comme un fait de discours. En effet, la synonymie-équivalence de sens a été examinée comme une relation de signifié à expression dont le but a été de recueillir la diversité des formes linguistiques renvoyant à la même idée. Comme le remarque G. Doualan (2014 : 413), ce type

¹ Dictionnaire de synonymes de CRISCO

de synonymie, qui ressort du procédé rhétorique nommé *amplificatio*, a connu un vif succès au Moyen-Age et à la Renaissance. Le cadre dans lequel elle est apparue relève donc d'onomasiologie. Les études menées par Ju. Apresjan et I. Mel'čuk s'inscrivent dans cette perspective. Selon les auteurs, la langue regorge des mots synonymes et des phrases inter-paraphrastiques et par conséquent, la synonymie ne limite pas aux unités lexicales. Ju. Apresjan dans ses travaux en matière de la sémantique lexicale assigne à la synonymie une place toute à fait centrale. L'auteur avance que les synonymes forment une catégorie irrégulière, purement lexicale. Dans la langue, ils se présentent sous une forme toute faite et ne peuvent pas être générés selon aucun modèle productif. L'auteur a inventorié trois conditions indispensables pour juger d'une situation de synonymie lexicale (Apresjan 2000 : 211) :

- le chevauchement de leurs définitions, c'est-à-dire les synonymes devraient être définis par la même expression au niveau sémantique.²
- l'identité de leurs structures de valence³
- l'appartenance à la même catégorie (profonde) du discours.

Une telle vision de la synonymie nécessite une analyse assez approfondie des unités en question – une reformulation de leurs définitions y incluant la décomposition sémantique permettant de déterminer le nombre d'arguments impliqués par les unités. Pour être sémantiquement équivalents, les structures de valence des unités doivent coïncider.

Il ressort de ces caractéristiques définitoires que la voie choisie par Ju. Apresjan est assez radicale et peu de couples seraient capables de satisfaire aux exigences évoquées. La synonymie dite exacte prend également ses sources dans l'emprunt. Ju. Apresjan cite des paires de synonymes : *indépendance* - *autonomie*, *base* – *fondement*, *dégradation* - *dégénération*.

Pour l'auteur, c'est dans le domaine du lexique expressif où se manifeste le plus explicitement la disposition de la langue à renouveler ses ressources lexicales par la création des synonymes. Le lexique expressif fait apparaître des mots de sens à peu près équivalents qui se distinguent par quelques nuances de sens ou par le degré d'intensité. Par conséquent, la langue est dominée par, pour reprendre la terminologie de Ju. Apresjan, la quasi-synonymie.

Les unités phraséologiques, grâce à leur forte expressivité, constituent un terrain propice à l'apparition des expressions synonymiques. En témoignent les

² Pour décrire avec précision les équivalences et les coïncidences sémantiques entre des mots, Ju. Apresjan en collaboration avec I. Mel'čuk et A. Žolkovskij ont développé un système sémantique formel – langage de composantes (de multiplicateurs) sémantiques – des sens élémentaires qui sont universels et dont la combinatoire permettrait de reconstruire la représentation sémantique de tous les mots d'une langue naturelle.

³ Le nombre d'actants impliqués par le prédicat et leurs rôles doivent être identiques.

expressions telles que : *avoir peur* = *avoir la trouille* ; *un brouillard dense* = *un brouillard à couper au couteau* ; *casser sa pipe* = *souffler la chandelle* ; *être paresseux* = *avoir un poil dans la main*⁴

Si on accepte que la synonymie soit une relation de contenu à expression, on est amené à admettre que la synonymie peut être poussée très loin et englober un nombre important de phénomènes langagiers, quelles que soient leur forme extérieure et leur fonction grammaticale. C'est la conclusion à laquelle aboutissent les analyses des auteurs cités : « la synonymie des expressions linguistiques ne résulte pas de l'utilisation de synonymes lexicaux au sens strict, mais fait appel à de nombreux autres moyens linguistiques ». (Apresjan 1974/2000 : 49).

4. Synonymie – de l'expression au contenu

Une autre orientation d'appréhender la synonymie a été proposée dans la perspective dite *différentialiste*. Ce cadre prend appui sur deux items de sens équivalents pour en exposer des ressemblances et des différences de sens. C'est cette approche de synonymie qui est adoptée par les dictionnaires de langue et dans beaucoup de travaux linguistiques modernes. (cf. Ducháček 1967; Victorri 1997; Fuchs 2007; Perz 2023). Il faut admettre que la sémantique lexicale a toujours privilégié l'analyse du mot isolé au détriment de l'analyse textuelle. Mais, une telle approche – de l'unité lexicale à son inventaire de significations possibles pour proposer le partenaire synonymique connaît de nombreux obstacles.⁵

Il existe un très large accord selon lequel les items dits synonymes devraient appartenir à la même catégorie morphosyntaxique. Cette condition se trouve confirmée, entre autres par J. Sanchez et P. Cadiot (2007 : 138) qui précisent que « la condition préalable à tout rapprochement synonymique » est « l'appartenance des lexèmes à une même classe grammaticale ». Ce constat est à rapprocher de celui formulé par J. Pustejovsky (1995 : 23) qui écrit : « synonymy (...) is an intra-category relation, e.g. nouns for nouns, verbs for verbs, and so on ». En conséquence, les suites:

Jacques est craintif = *Jacques est un lâche*
un argument solide = *un argument de poids*

⁴ Nous avons donné les exemples en français pour illustrer cette relation synonymique. Ju. Apresjan cite : *brat'v oborot – brat'v rabotu, byt'na veršok ot gibeli – byt'na volosok ot gibeli*. (2000 : 213).

⁵ F. Rastier (1987) dans sa sémantique interprétative condamne ce type d'approche. Selon lui, la démarche dite sémasiologique empêche de rendre compte du fait que les différents emplois polysémiques peuvent apparaître dans les contextes très différents - ils ont chaque leur histoire et souvent n'appartiennent pas au même niveau de la langue.

sont des synonymes au niveau de discours. *Craintif*, *ive* (adj.) et *lâche* (n.), *solide* (adj.) et *de poids* (express.) ne sont pas donnés comme synonymes en langue, c'est-à-dire ne sont pas recensés dans les dictionnaires de synonymes.

Dans *La sémantique* (1988) I. Tamba nous évoque le même critère :

Dans l'usage courant, on appelle synonymes, des mots de forme différente, mais de sens identique ou semblable et de même statut morphosyntaxique (1988 : 80).⁶

Soulignons qu'en pratique, ce critère n'est pas toujours respecté. Ainsi, parmi les synonymes de l'adverbe *cependant* figurent : *n'empêche que*, *malgré tout*, *avec tout cela*.⁷ Le critère du même statut morphosyntaxique n'est pas neutre sur le plan théorique - il constitue l'une des raisons pour laquelle la synonymie est assignée à l'axe paradigmatique - l'axe des choix lexicaux et de la substitution. Ce rapport s'établit *in absentia*, virtuellement, entre les unités de même catégorie grammaticale. Par conséquent, les utilisateurs ordinaires de la langue réussissent facilement à distinguer les mots synonymes des mots non synonymes ainsi qu'à déceler divers degrés de similarité de sens entre deux mots. Notons que, le test de substituabilité, souvent appliqué pour mesurer le degré de synonymie entre deux items en contexte, fait que la frontière entre l'axe paradigmatique et l'axe syntagmatique s'estompe. La substitution peut engendrer des suites phraséologiques que son synonyme ne permet pas, comme *fort* et *puissant*:

être fort en mathématiques vs **être puissant en mathématiques*

et des constructions que le partenaire synonymique n'accepte pas – *fort* peut être employé comme adverbe :

Ils doutent fort qu'il puisse le faire.

Ces familles étaient fort riches.

Sémantiquement, si un tel rapport est possible entre les items lexicaux cela signifie que le mot hors emploi peut être soumis aux procédures d'analyse propre à la langue. La synonymie serait donc une relation sémantique ancrée dans le système de la langue et produite par la langue.

Ainsi, bien que longuement discutée – d'une relation s'établissant au niveau discursif à une relation valable au niveau du système linguistique, la manière de concevoir la synonymie ne fait pas l'unanimité. C. Fuchs (1982 : 22) souligne ce changement de perspective en écrivant que « la synonymie de phrases relevait de la rhétorique, et la synonymie de mots de la grammaire ». On peut se poser la question s'il s'agit de deux visions opposées ou complémentaires de la synonymie ? À une telle question, il n'existe pas de réponse unique, elle touche le problème de la délimitation du champ de l'analyse.

⁶ Soulignons que, dans la deuxième édition de son livre (2005), l'auteure nous fournit la définition un peu différente que celle présentée en 1988.

⁷ <https://www.cnrtl.fr/synonymie/cependant>

La constatation faite par A. Cruse présente un mouvement en avant par rapport à la vision classique. Dans son livre *Meaning in Language*, il souligne que les synonymes “are words whose semantic similarities are more salient than their differences” (Cruse 2004: 156). Les synonymes sont des mots pour lesquels les similitudes sémantiques sont plus saillantes que les différences. L’apport essentiel de A. Cruse aux études des relations synonymiques est que la synonymie est une notion scalaire dont la synonymie totale serait le point final de l’échelle synonymique. L’échelle s’étend donc de la synonymie absolue à la non-synonymie en passant par diverses situations de synonymie regroupées sous l’étiquette de *synonymie cognitive*. Les synonymes cognitifs doivent partager les mêmes restrictions de sélection, mais peuvent différer au niveau des restrictions collocationnelles. C’est pourquoi, certaines paires de mots sont “more synonymous than others” (A. Cruse 1986 : 267). Pour être ressenti comme synonymes les mots doivent afficher plus de ressemblances que des différences d’emploi.

5. Du virtuel à l’actuel

Toute analyse de la synonymie engage une réflexion sur la nature du sens lexical – comment il s’organise et comment il est structuré. B. Victorri, dans sa conception dynamique du sens, pour rendre compte de la dichotomie existant entre la langue et le discours introduit la notion d’énoncé-type et énoncé-occurrence.

[...] il s’agit d’une part de comprendre comment les éléments linguistiques qui composent un énoncé-type se combinent pour lui conférer un sens, et, d’autre part, comprendre les mécanismes par lesquels ce sens peut entrer en interaction avec des données extra-linguistiques pour donner un sens à un énoncé-occurrence.

(Victorri 1996 : 29)

Si un mot hors-contexte peut avoir plusieurs sens, ce sont des sens virtuels. Dans un contexte donné ce n’est qu’un de ses sens virtuels qui se trouve actualisé. Actualiser un mot, c’est réaliser en discours les potentialités dont il est chargé en langue. Les différents sens dont l’unité est composée se dégagent grâce aux éléments du contexte. Afin d’illustrer l’influence de l’adjectif sur le sémantisme du substantif, prenons le cas du nom *voiture* :

- *voiture puissante* – dont le moteur a de grandes capacités
- *voiture rapide* – qui roule rapidement
- *voiture verte* – 1) dont la couleur est verte ; 2) qui est écologique
- *voiture confortable* – qui assure le confort durant le voyage
- *voiture sportive* – qui a l’allure sportive, qui a de grandes capacités

- *voiture écologique* – *qui ne produit pas d'émissions polluantes*
- *voiture spacieuse* – *qui dispose de beaucoup de place*
- *voiture électrique* – *qui a un moteur électrique associé à une batterie*

L'adjectif active certains aspects de l'objet caractérisé – une “zone sémantique”⁸ à chaque fois différente du substantif qu'il accompagne.

Nous pouvons déjà inférer de ces exemples que le contenu sémantique du mot est conditionné par les relations qui s'établissent durant le processus d'actualisation. En effet, au niveau discursif, le mot en question ne retient qu'une partie de son sémantisme qu'il affiche en tant qu'élément en langue. Cela laisse supposer que les différents sens (signifiés) des mots y compris les adjectifs se construisent durant le processus d'actualisation – ils sont dépendants des contextes où ils se manifestent. Le signifiant reste inchangé tandis que le signifié est réduit lors du passage du virtuel à l'actuel.

On peut comprendre, à la lumière de ces observations pourquoi il y a deux directions opposées de voir la synonymie – synonymie au niveau du système linguistique et synonymie au niveau discursif. Les différentes approches relatives à la synonymie tendent à dépasser cette opposition entre deux plans : celui de la sémantique du mot et celui de la sémantique du discours. Vu l'asymétrie que la langue présente entre le niveau lexical et celui du discours, il n'existe pas de cadre formel qui toucherait à ces deux aspects de la synonymie. Nous ne visons pas, dans cet article, de choisir entre les deux directions qui se disputent la matière de synonymie. La question est de savoir comment on peut intégrer ce mouvement de sens.

6. Théorie vs pratique

Deux adjectifs ne donnent lieu à synonymie que pour une partie plus ou moins étendue de leurs acceptions. En effet, les adjectifs *clair* et *compréhensible* représentent des synonymes dits partiels ou approximatifs parce qu'ils ne peuvent être interchangeables que dans certains environnements tels que <écrits> : *article, commentaire, description, paragraphe, texte, information*, <propos> : *énoncé, explication, exposé, discours, langage, réponse* ou <signes> : *avertissement, message, signe, signal*.⁹ Deuxièmement, les revois synonymique ne touchent qu'une seule acception particulière comme : *fort* et *touffu* (barbe) – ce couple passe difficilement comme synonymique dans l'usage intuitif du terme.

Nous essayons de clarifier notre propos à partir des emplois de l'adjectif *fort*. Notre choix porte sur l'adjectif de type qualificatif parce que c'est un champ

⁸ Pour la notion de zone ou “zone active”, nous nous inspirons des travaux de R. W. Langacker. Cette notion est un élément central du dispositif proposé par cet auteur (Langacker 1991 b : 190)

⁹ Pour plus de détail, voir (Perz 2018).

d'investigation riche et complexe. L'adjectif *fort* présente une pluralité d'acceptions, il apparaît avec une fréquence considérable dans le lexique et appartient à un groupe restreint d'adjectifs que M. Borodina appelle de *formation primaire*.¹⁰ (Borodina 1963 : 194, cité par Goes 1999 : 48). Cet adjectif sert à exprimer l'idée générale d'intensité.

L'adjectif *fort, e* se combine avec la catégorie d'humains :

(1) *un homme fort*, (2) *une femme forte*, (3) *un enfant fort*

À ce stade déjà, nous pouvons observer que la relation qui s'établit entre l'adjectif et les substantifs du type *humain* est ambiguë. *Fort* peut signifier – 1) '*de force physique, puissant*', 2) '*qui est corpulent*' ou encore 3) '*qui est capable de supporter la souffrance, la fatigue*'. C'est l'interaction avec le contexte plus large qui déterminera leur valeur sémantique et leur interprétation.

Fort se combine avec les noms désignant des objets :

(4) *un papier fort*, *de fortes chaussures*, (5) *une colle forte*, (6) *une carte forte*, (7) *un café fort* (8) *un fromage fort*, (9) *de fortes lunettes*

Avec des substantifs abstraits :

(10) *une voix forte*, (11) *une odeur forte*, (12) *un caractère fort*, (13) *une forte influence*, (15) *un fort accent*, (16) *une monnaie forte*, (17) *des traits forts*, (18) *le sexe fort*, (19) *une forte discipline*, (20) *une plaisanterie forte*

Avec les substantifs dénotant des phénomènes naturels :

(21) *une pluie forte*, (22) *un vent fort*, (23) *une forte chaleur*

Il n'est pas étonnant que la liste de ses synonymes recensés par le Dictionnaire électronique des synonymes (*DES*)¹¹ est imposante. La consultation de ce dictionnaire demeure problématique dans la mesure où les partenaires synonymiques sont donnés sans commentaires et on tombe sur les renvois pour lesquels on a du mal à restituer le lien synonymique.

Pour révéler le partenaire synonymique d'un adjectif, il faut dresser la liste de ses acceptions et ensuite analyser les rapports qui s'établissent entre l'adjectif et les noms auxquels il se rapporte. La rection nominale des adjectifs constitue l'un des paramètres fondamentaux dans le processus de son interprétation – l'adjectif est sélectionné pour exprimer une propriété donnée, mais toujours sous le contrôle du substantif. Comme l'a précisé M. Riegel, les adjectifs se caractérisent par une "incomplétude" ou "dépendance" référentielle. (Riegel 1985 : 97). Cette

¹⁰ M. Borodina est l'auteur qui a essayé d'inventorier une liste d'adjectifs très fréquents qui proviennent de vieille souche. Ils sont courts et généralement monosyllabiques.

¹¹ DES est un dictionnaire de synonymes du français sous forme électronique conçu par CRISCO (Centre de recherches inter-langues sur la signification en contexte) qui apporte de nombreuses innovations au traitement de la synonymie. Il énumère des synonymes issus de sept dictionnaires compilés par l'INALF. Pour avoir une meilleure compréhension des synonymes, il les organise sous forme d'espaces sémantiques. Une telle représentation du sens permet de visualiser la répartition des sens d'une unité lexicale et par là sa polysémie.

dépendance sémantique de l'adjectif au nom fait que l'adjectif en tant que catégorie se montre beaucoup plus difficile d'accès en termes d'analyse que d'autres items autonomes - les substantifs ou les verbes.

L'adjectif *fort* peut être synonyme de *vif* si on précise que cette équivalence n'existe que pour la classe de <sensations> : *une forte/vive agitation, une forte/vive passion, une forte/vive émotion*. Ce sous-classement de la combinatoire lexicale des unités est connu sous le nom des classes d'objets (Gross 1994, 2012, Banyś 2018). L'idée directrice de cette démarche est « une généralisation des objets qui partagent le même type de relations amène à construire une classe d'objets » (Banyś 2018 : 29).

Pour l'adjectif *fort* et le type des substantifs qu'il qualifie, on peut établir le profil combinatoire en établissant les classes distributionnelles¹².

- fort – SYN : *abondant* <précipitations> : *pluie, averse, neige, crue*
- fort – SYN : *aigu* <douleurs> : *fièvre, douleur, toux, maladie, inflammation*
- fort – SYN : *assourdissant* <bruits> : *bruit, musique, cri, hurlement, tapage, bourdonnement*
- fort – SYN : *autoritaire* <régimes> : *gouvernement, empire, système, pouvoir*
- fort – SYN : *corpulent* <humains> : *homme, gaillard, femme, fille*
- fort – SYN : *gros* <constructions> : *muraille, digue, bâtiment, édifice*
- fort – SYN : *gros* <ensembles> : *volume, quantité, ouvrage, somme*
- fort – SYN : *important* <montées> : *hausse des prix, augmentation, demande, progrès, croissance, développent, stimulation*
- fort – SYN : *important* <relations> : *corrélation, lien, dépendance, rapport, correspondance, position*
- fort – SYN : *intense* <odeurs> : *arôme, parfum, odeur, saveur*
- fort – SYN : *intense* <luminosités> : *lumière, soleil, éclat, luminosité, reflet*
- fort – SYN : *intense* <nourriture_1> : *fromage, boisson, charcuterie*
- fort – SYN : *massif* <parties du corps extérieurs> : *bras, épaule, mâchoire, nez, épaule, jarret, muscles, poitrine, taille*
- fort – SYN : *piquant* <nourriture_2> : *épices, moutarde, vinaigre, radis, sauce, poivre, piment*
- fort – SYN : *puissant* <séismes> : *tremblement de terre, secousse, éruption, secousse*
- fort – SYN : *puissant* <animaux> : *cheval, bœuf, chien*
- fort – SYN : *puissant* <odeurs> : *arôme, parfum, odeur, saveur*
- fort – SYN : *puissant* <substances chimiques> : *acide, électrolyte, désherbant, remède, explosif*

¹² Soulignons que les appellations des classes ne sont que des étiquettes et elles ne doivent pas être assimilées aux classes ontologiques.

- fort – SYN : *puissant* <régimes> : *gouvernement, empire, système, pouvoir*
- fort – SYN : *perçant* <sons aigus> : *bruit, clameur, cri, sifflement, son*
- fort – SYN : *profond* <blessures> : *blessure, plaie, ride*
- fort – SYN : *solide* <constructions> : *bâtiment, coffre, construction, demeure, digue, édifice, château, matériau, mur, muraille*
- fort – SYN : *solide* <relations> : *amitié, mariage, affection, liaison, rancune*
- fort – SYN : *solide* <classe_S3> : *corde, bottes, chaussures, argument, équipe, base, économie, fondement, raison, structure*
- fort – SYN : *sonore* <sons> : *musique, voix, soupir, ronflement*
- fort – SYN : *significatif* <ampleurs> : *augmentation, pourcentage, nombre*
- fort – SYN : *résistant* <matériaux> : *tissu, ruban, corde*
- fort – SYN : *résistant* <espèces> : *race, génération, espèce, peuple, tribu*
- fort – SYN : *résistant* <constructions2> : *bâtiment, coffre, construction, demeure, digue, édifice, château, matériau, mur, muraille, ville, place*
- fort – SYN : *rigoureux* <ordres> : *discipline, exigence, hiérarchie*
- fort – SYN : *vif* <sensations> : *agitation, désir, joie, émotion, passion, excitation, impression*
- fort – SYN : *violent* <perturbations> : *orage, tempête, ouragan, cyclone*

Au vu de ces classes distributionnelles, nous voyons que l'adjectif *fort* possède une grande extension et peut qualifier un grand nombre de substantifs. Signalons que l'adjectif reçoit une partie de son sémantisme du substantif qu'il qualifie. À titre d'exemple, prenons la classe d'<aliments> où l'adjectif *fort* affiche une autre valeur sémantique en fonction du substantif qu'il accompagne :

- *fromage fort* – '*fromage à un goût très prononcé*'
- *moutarde forte* – '*moutarde piquante*'
- *vin fort* – '*vin à forte teneur en alcool*'
- *café fort* – '*café très concentré*'
- *alcool fort* – '*boisson à degré d'alcool élevé*'

Les lexèmes censés être synonymes ne se rencontrent pas dans les mêmes environnements. Leurs comportements combinatoires et restrictions d'emploi peuvent diverger de façon importante. Comme l'a remarqué G. Kleiber (2009 : 20) la synonymie – et toutes les relations lexicales – n'est pas la propriété d'un mot, mais d'un sens d'un mot, ce qui est souvent source de confusion. La difficulté réside dans le manque d'une description opératoire de l'environnement où l'interchangeabilité peut avoir lieu.

7. Interchangeabilité restreinte

On ne peut pas négliger le fait que certains couples synonymiques nécessitent un traitement lexical différencié – leurs éléments ne vont pas être inclus dans les classes étant donné que le choix de synonymes est contraint. Il s’agit d’un savoir lexical qui tend à passer inaperçu, mais qui représente un niveau avancé dans la maîtrise d’une langue. Les synonymes retenus paraissent sémantiquement explicables, mais ne semblent pas prédictibles sur le plan sémantique. Penchons-nous sur les partenaires synonymiques de *fort* dans les couples suivants:

- fort – SYN : *à couper au couteau* (brouillard)
- fort – SYN : *adhérent* (colle)
- fort – SYN : *épais* (végétation)
- fort – SYN : *grossier* (plaisanterie)
- fort – SYN : *impétueux* (vent)
- fort – SYN : *marqué* (accent)
- fort – SYN : *opaque* (brouillard, brume)
- fort – SYN : *profond* (racines)
- fort – SYN : *suffocant* (chaleur)
- fort – SYN : *torrentiel* (pluie)
- fort – SYN : *tropical* (chaleur)
- fort – SYN : *vif* (concurrence)
- fort – SYN : *vigoureux* (tempérament)
- fort – SYN : *à crever les tympans* (bruit)
- fort – SYN : *bleu* (peur)
- fort – SYN : *compact* (terre)
- fort – SYN : *copieux* (repas, plat)
- fort – SYN : *cuisant* (échec)
- fort – SYN : *de canard* (froid)
- fort – SYN : *de fer* (santé)
- fort – SYN : *de plomb* (soleil)
- fort – SYN : *de poids* (argument)
- fort – SYN : *furieux* (vent)

Ces contraintes ne permettent pas de combiner à volonté les éléments du système. Signalons que les chercheurs ont remarqué l’existence de ce type de stabilisation entre les unités. Ce phénomène a reçu différentes appellations, entre autres *semi-phrasèmes* (Mel’čuk 1998, 2003), *combinaisons phraséologiques* (Hausmann 1989, 2007) ou *collocations* (Sinclair 1991) ou encore *figement* (Mejri 2019, 2023). Nous n’allons pas entrer ici dans le débat sur les collocations – elles dépassent le cadre de cette étude, précisons seulement que la façon de les définir ne fait pas l’unanimité.

Retenons de ces exemples que le sémantisme de l'adjectif *fort*, étant assez générique, se révèle difficilement saisissable. On voit à travers la quantité d'exemples retenus que le discours semble contourner son usage au profit des emplois plus nuancés et plus expressifs. Ce cheminement conduit à (ré)affirmer l'idée déjà formulée par Ju. Apresjan, à savoir qu'il y a dans le discours un espace qui est propice à la spécialisation du sens et par conséquent à la lexicalisation plus ou moins forte des unités.

Le calcul de la combinatoire lexicale entre l'adjectif *fort* et nom conduit à déterminer les séries pour lesquelles la substitution synonymique n'est pas possible. Il y a une spécialisation du sens en contexte qui conduit à une non substituabilité des éléments – de telles expressions doivent être listées séparément. Les suites retenues sont des phraséologismes dont l'interprétation ne pose pas de grands problèmes, mais leur degré de figement est beaucoup plus important – il empêche la substitution synonymique et par conséquent ces suites s'emploient telles quelles. Nous avons retenu les syntagmes suivants :

- *le sexe fort*
- *un café fort*
- *un esprit fort*
- *une âme forte*
- *une carte forte*
- *une consonne forte*
- *une main forte*
- *une monnaie forte*
- *de fortes lunettes*

L'adjectif devient un composant formel de séquences plus larges dont il ne peut pas être détaché. Les constituants ne peuvent pas être commutés sans que le sens soit modifié ou que la construction devienne incorrecte/peu usuelle dans une langue. Ces suites se rapprochent plutôt du nom composé ou de *groupements usuels*.

8. Synonymie – phénomène de discours

Comme nous venons de mentionner plus haut que le sens d'un item dans un énoncé est au moins en partie déterminé par le sens global de l'énoncé. Les phrases, en tant qu'unités communicatives, ont un statut instable et individuel, par opposition à la permanence du système linguistique. Par conséquent, le choix des partenaires synonymiques est dicté par les propriétés sémantiques des mots environnant le lexème analysé. Les éléments du co-texte constituent un facteur prenant part dans l'émergence de l'acception en question – ils sont donc impliqués dans le processus synonymique entre les deux acceptions. Considérons quelques acceptions de l'adjectif *fort* construites en co-texte.

1. (a) *Une soudaine pluie **forte** a inondé les rues en un rien de temps.* – SYN : **torrentielle**
 (b) *Un instant plus tard, il fut aveuglé par une **forte** pluie* – SYN : **cinglante**
 (c) *La période de **fortes** pluies crée des conditions idéales pour l'éclosion de moustiques tout au long de l'été.* – SYN : **continuelles**
2. (a) *Pendant la saison du jasmin, les fleurs fraîchement cueillies sont mêlées aux feuilles soyeuses des bourgeons, ce qui engendre une **forte** odeur.* – SYN : **enivrante**
 (b) [...]développement des bactéries, des champignons, des pollens et d'autres micro-organismes qui génèrent une **forte** odeur. – SYN : **nauséabonde**
 (c) *La forte odeur du marché aux poissons qui lui a donné envie de vomir* – SYN: **fétide**
3. (a) *Si vous avez été exposé à un bruit fort et que vous êtes inquiet à propos de votre audition, devez consulter (...)* - SYN : **perçant**
 (b) (...) les éléphants, soit en provoquant un choc à l'aide d'un bruit fort et inattendu, soit par des bruits spécifiques pour effrayer (...) – SYN : **intense**

Il ressort de cette courte exploration qu'en discours, de nombreux facteurs interviennent dans l'émergence d'une acception et par conséquent guident le choix synonymique. L'interaction des éléments du co-texte spécifie les rôles des unités présentes dans une phrase et permet de voir comment les acceptions se dégagent grâce à la contextualisation. Nous rejoignons ici les idées de S. Cigada (2009 : 22) selon qui la synonymie/*langue* est en effet une synonymie de sémèmes et la synonymie/*parole* est une synonymie de sémès uniques.

En d'autres termes, la synonymie telle qu'elle est perçue en discours est régie par deux principes opposés: d'un côté, le principe de liberté permettant de substituer les unités pourvu qu'on applique les règles assurant la grammaticalité sans modification significative du sens et de l'autre côté, le principe qui restreint cette liberté combinatoire sous le contrôle d'autres paramètres linguistiques. Ainsi, dans le discours les suites **un fort argent*, **un fort respect* ne sont pas recevables, alors que *un argent fou*, *un profond respect* le sont tout à fait. Il convient de souligner que, l'auteur qui a posé l'existence de ces deux principes régissant la combinatoire lexicale est J. Sinclair (1991). L'auteur, dans son ouvrage, oppose le principe de liberté (*open-choice principle*) qui autorise la liberté dans la sélection des éléments dans la langue au principe d'idiomaticité (*idiom principle*) qui restreint cette liberté combinatoire par des habitudes langagières.

9. En guise de conclusion

L'un des objectifs de la lexicographie est de mettre en évidence les capacités combinatoires des unités lexicales. La difficulté majeure à laquelle se heurte la description dictionnaire est due à l'impossibilité d'affiner la description de toutes les combinaisons entre les éléments synonymiques.

Pour rendre compte de la synonymie de deux adjectifs, plus précisément de leur communauté d'emplois, il est nécessaire de les considérer de manière multilatérale :

1/ du point de vue des combinaisons de l'adjectif avec différents substantifs qu'il qualifie afin de rendre compte de ses paradigmes lexicaux. Pour ce faire, certains chercheurs ont forgé la théorie de *classes d'objets* – les sous-classes sémantiquement homogènes capables de rendre compte des paradigmes distributionnels des unités. Mais, il s'avère que la langue/discours présente tant d'irrégularités que les formaliser toutes semble illusoire. La liberté et le caractère optionnel des combinaisons, et en même temps la flexibilité des adjectifs, rendent plus difficile la description des groupes synonymiques clairement distincts. Il est nécessaire aussi de les traiter :

2/ du point de vue des combinaisons de deux adjectifs synonymiques avec le même substantif régisseur pour vérifier s'il n'y a pas de restriction dans la sélection des éléments. Ces contraintes lexicales ne peuvent pas être totalement prédites à partir des seules règles de logique ou de grammaire. Elles sont souvent spécifiques à chaque acception du mot. Il est requis de saisir ces contraintes individuellement pour chaque acception.

Le linguiste se trouve confronté à deux plans qui paraissent opposables – d'un côté la langue est traitée comme un ensemble clos et relativement stable de signes avec tout son potentiel combinatoire et de l'autre côté cet ensemble est mis en emploi et analysé comme manifestation de ce système. Cette confrontation entre ces deux plans – discours vs langue affine la description en permettant de mettre au jour le brouillage sémantique qui se réalise en discours, mais c'est une entreprise très complexe.

Références

- Apresjan, Ju. 1973. *Principles and Methods of Contemporary Structural Linguistics*. La Haye: Mouton.
- Apresjan, Ju. 1974/2000. *Semantyka leksykalna. Synonimiczne środki języka*. Wrocław : Zakład Narodowy im Ossolińskich.
- Bally, Ch. 1909/1921. *Traité de stylistique française*, 2 vol. Heidelberg et Paris: Zinter et Klincksieck.
- Banyś, W. 2018. Nouveau anciens paradigmes: Approche orientée objets, Classes d'objets, Psychologie écologique et Linguistique. *Neophilologica* 30 : 25-41.

- Blumenthal, P. 2006. De la logique des mots à l'analyse de la synonymie. *Langue Française* 150 : 14-31.
- Cadiot, P. 2009. Couleurs des mots ou synonymie. *Pratiques* 141/142 : 26-38.
- Cigada, S. 2009. Le concept de synonymie entre « langue » et « parole ». *Cahiers de l'Association internationale des études françaises* 61 : 13-29. DOI : <https://doi.org/10.3406/caief.2009.2566>.
- Doualan, G. 2014. Éléments pour une lecture de l'histoire de la synonymie. In F. Neveu, P. Blumenthal, L. Hriba, A. Gerstenberg, J. Meinschaefer, S. Prévost (Éd.), *Congrès Mondial de Linguistique Française*, 409-424. Berlin. <http://dx.doi.org/10.1051.shsconf/20140801108>.
- Fuchs, C. 2007. Relation de synonymie entre les polysèmes : le réseau comme-manière - façon. *Le français moderne* 75(1) : 97-113.
- Fuchs, C. 1982. La paraphrase entre la langue et le discours. *Langue française* 53 : 22-33. DOI : <https://doi.org/10.3406/lfr.1982.5113>.
- Goes, J. 1999. *L'adjectif. Entre nom et verbe*. Louvain-la-Neuve : Duculot.
- Grobelak, L. 1990. *Dictionnaire collocationnel du français général*. Varsovie : Państwowe Wydawnictwo Naukowe.
- Gross, G. 1991. Typologie des adjectivaux. In H. Stammerjohan (Éd.), *Analyse et synthèse dans les langues romanes et slaves*, 163-178. Tübingen : Günter Narr Verlag.
- Gross, G. 1994b. Classes d'objets et synonymie. *Annales Littéraires de l'Université de Besançon. Série Linguistique et Sémiotique* 23 : 93-102.
- Hausmann, F.J. 2007. Lexicographie française et phraséologie. In E. Haag (Éd.), *Collocations, phraséologie, lexicographie*, 121-153. Aachen: Shaker Verlag.
- Honeste, M.-L. 2007. Entre ressemblance et différence : synonymie et cognition. *Le français moderne* 75(1) : 160-173.
- Langacker, R.W. 1991. *Concepts, Image and Symbol. The Cognitive Basis of Grammar*. Berlin : Mouton-de Gruyter.
- Lehmann, A., et F. Martin-Berthet 1998. *Introduction à la lexicologie. Sémantique et morphologie*. Paris : Duclot.
- Lehmann, A., et F. Martin-Berthet 2018. *Lexicologie. Sémantique, morphologie et lexicographie*. Paris : Armand Colin.
- Kleiber, G. 2009. La synonymie – identité de sens n'est pas un mythe. *Pratiques* 141/142 : 9-25.
- Mejri S. 2023. Prédicats, sens, polylexicalité : un parcours heuristique. *Neophilologica* 35: 1-40. DOI: <https://doi.org/10.31261/NEO:2023/35/14>.
- Mejri S. 2005. Figement; néologie et renouvellement du lexique. *Linx*, 52 / 163-174. DOI : <https://doi.org/10.4000/linx.231>.
- Mel'čuk I. et al. 1984, 1988, 1992. *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques I, II, III, IV*. Montréal : Les Presses de l'Université de Montréal.
- Moon, R. 1987. Monosemous words and the dictionary. In A.P. Cowie and European Association for Lexicography (Eds.), *The Dictionary and the Language Learner*,

- Papers from the EURALEX Seminar at the University of Leeds*, 173-182. Berlin, Boston: Max Niemeyer Verlag.
- Murphy, L. 2003. *Semantic Relations and the Lexicon*. Cambridge: Cambridge University Press.
- Noailly M. 1999. *L'adjectif en français*. Paris : Éditions Ophrys.
- Pecman, M. 2005. De la phraséologie à la traduction proactive : essai de synthèse des fondements théoriques sous-tendant la recherche en phraséologie. *Meta* 50(4). DOI : <https://doi.org/10.7202/019853ar>.
- Ploux, S., et B. Victorri 1998. Construction d'espaces sémantiques à l'aide de dictionnaires de synonymes. *Traitement automatique des langues* 39 : 161-182.
- Perz, M. 2018. Les emplois et le potentiel informationnel de l'adjectif clair(e). *Lublin Studies in Modern Language and Literature* 42: 184-205.
- Pustejovsky, J. 1995. *The Generative Lexicon*. Cambridge : MIT Press.
- Rastier, F. 1987. *Sémantique interprétative*. Paris :PUF.
- Riegel, M. 1993. Grammaire et référence : à propos du statut sémantique de l'adjectif qualificatif. *L'Information Grammaticale* 58 : 5-10. DOI : "<https://doi.org/10.3406/igram.1993.3148> .
- Sanchez, J., et P. Cadiot 2007. Le cas des substantifs construits. *Le français moderne* 75(1) : 114-140.
- Tamba, I. 1998. *La sémantique*. Paris : PUF.
- Tutin, A., et F. Grossmann 2002. Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif. *Revue Française de Linguistique Appliquée* 7 : 7-26.

PAWEŁ GOLDA

University of Silesia in Katowice, Poland

Faculty of Humanities, Institute of Linguistics

ORCID: 0000-0001-5505-7731

pawel.golda@us.edu.pl

MARLENA MOSKAL

University of Silesia in Katowice, Poland

Faculty of Humanities, Institute of Linguistics

ORCID: 0000-0002-8935-503X

marlena.moskal@us.edu.pl

AGNIESZKA JEDZINIAK-DANOWSKA

University of Silesia in Katowice, Poland

Faculty of Humanities, Institute of Linguistics

ORCID: 0000-0002-1163-4936

agnieszka.jedziniak-danowska@us.edu.pl

JOANNA RYSZKA

University of Silesia in Katowice, Poland

Faculty of Humanities, Institute of Linguistics

ORCID: 0000-0002-5142-3949

joanna.ryszka@us.edu.pl

NAMING THE NAMELESS: A PRELIMINARY STUDY OF BYNAMES FOR THE DECEASED FROM AN ONOMASTIC PERSPECTIVE

The article proposes a new term, *postmortonyms*, denoting names assigned to unidentified deceased individuals. Drawing on an exemplary semantic onomastic analysis of 220 postmortonyms from multiple cultures, the study explores the motivations, structures, and functions of these bynames, distinct from *posthumous names*, *deadnames*, and *necronyms*. The research demonstrates that postmortonyms often incorporate details such as location or distinguishing features, which serve investigative, communicative, and memorial purposes. The analysis highlights the phenomenon of polyonymy, in which temporary names coexist with potential authentic identities, and discusses the broader cultural and legal contexts that influence naming practices for the nameless dead. The research is a preliminary study, and the findings provide a foundation for further research on the conceptualization of death and naming across cultures.

Keywords: onomastics, proper noun, name, onyms, postmortonym, nickname, byname, death

1. Introduction

The urge to name is a fundamental human trait, reflecting our need to categorize, communicate, and make sense of the world. It stems from the fact that “[t]he unnamed is the unnoticed, and the unnoticed is for cognitive and communicative purposes nonexistent” (Algeo and Algeo 2000: 265). Moreover, according to Bright (2003: 674), “the use of personal names, having varying levels of descriptiveness, is a sociolinguistic universal of the human species.” Thus, names can be regarded as social phenomena (Ernst 2010: 61). Consequently, it is deeply human to assign names even to the dead, particularly in the absence of an established identity. Such units fall within the domain of onomastics, the scientific study of proper names (or onyms), which explores their social, cultural, historical, and psychological significance.

Naturally, death is a taboo subject in many cultures; yet, it did not stop these cultures from establishing social rules that often emphasize respect for the deceased, following both legal provisions (cf. British Common Law) and the Latin aphorism *de mortuis nil nisi bonum dicendum est*. Therefore, some legal systems have established guidelines for referring to individuals whose names are unknown or cannot be disclosed. For instance, Order No. 48 of the Chief Commander of Police dated June 28, 2018, stipulates that in Poland, the term *NN zwłoki* (‘NN dead body’) should be used for unidentified bodies. Whereas in English-speaking countries, placeholders such as *John* and *Jane Doe* or *Richard Roe* (e.g. Bryant 1957; Felton 1972) are used as “specified names of unspecified identities” (LeMay 1995: 643). However, these guidelines are not always sufficient, leading to the creation of new, distinct names that circulate in both the mass media and official institutions.

In our study, we refer to the names assigned to unidentified deceased individuals as *postmortonyms* (hereinafter, PMOs) – a term we propose due to the absence of an appropriate one in the onomastic literature. As the analysis shows, these names often incorporate specific details, such as the location of discovery or distinguishing physical features. Furthermore, these names may serve multiple purposes: they help differentiate between cases, aid law enforcement and forensic investigation, and attract media attention through memorability. Thus, in their purpose and use, they are significantly different from *posthumous names* (Nordvall 2020), *deadnames* (Koles 2024), and *necronyms* (Testoni *et al.* 2020; Gałkowski 2024). Moreover, despite general interest in the onomastics (Hough (ed.) 2016), PMOs have not been studied yet, contrary to other investigation-related names, e.g. the aliases of criminals (e.g. Golda and Mężyk 2021) or nicknames (e.g. Leslie and Skipper 1990), which are typologically similar to PMOs.

The following onomastic analysis examines 220 PMOs from various cultures and languages, focusing on their structure and semantic properties. While some discourse analyses have explored the use of *John* and *Jane Doe* (e.g. Klaidman

and Beauchamp 1986; Ehrlich 2021), these studies focus on specific cases rather than on naming systems and the structure of names. This research aims to address the outlined gap by offering a semantic onomastic analysis of PMOs, a new onymic category. Moreover, the present analysis serves as a starting point for a discussion on different approaches to the conceptualisations of death, with the aim of broadening the perspective to include other cultures in future comparative analyses.

With regard to the cross-cultural and cross-linguistic dimensions of this research, in the era of globally shared media and the growing popularity of true crime, cases involving unidentified deceased individuals are often highly mediated across different countries, giving them a cross-cultural and cross-linguistic character. This is reflected in the lack of translations in many instances, as the name of a case already circulates in English. Moreover, death itself is both a cultural and universal phenomenon, independent of language, which justifies the introduction of a new term, *postmortonym*. If the research were conducted on the material from one language, there would be a risk of creating a concept limited to a single linguistic and cultural context. However, by the inclusion of multiple cultures and codes we aim at supporting the legitimacy of introducing this term.

2. Theoretical Framework

The interest in proper names has a long history and has preoccupied philosophers such as John Stuart Mill and Gottlob Frege. Consequently, various scholars have attempted to understand the role of a name – for example, whether it carries both connotative and denotative meaning (as in the Millian theory), and whether different proper names referring to the same entity are truly synonymous (as discussed in sense-based theories of names). Naturally, these questions and theories remain under debate, but many linguists (e.g. Del Pinal 2015; Lederman 2021) continue to engage with them to provide a broader perspective on the complexity of proper names.

In anthropology, names are of paramount importance, as understanding how communities bestow them “provides a grasp of the community’s understanding of identity, kinship, genealogy, and social life” (De Stefani 2016: 54). This relevance is highlighted by linguists who say that

“[p]roper names are not only linguistic signs – they are also social constructs, often strongly ideologized, entering into numerous text and discourse relations, forming both formal and semantic-connotational, or metaphorical-metonymic series, playing in many cases a key role in (re)constructing social reality” (Rutkowski and Skowronek 2019: 49).

Moreover, proper names play a crucial role in the discussions of reference because they fundamentally identify and differentiate, as they refer to a specific person, object, or phenomenon, i.e. a specific referent. Thus, a proper name, as a designation of an unambiguous onymic object, should be considered in a specific onymic context and situation (Sičáková 2002).

The language's anthroponymic system is highly diverse, encompassing various classes of human names. Among these, in addition to the first and last names (i.e. prosoponyms), are pseudonyms and nicknames (cf. *bynames* in Brylla 2016), which serve as their *substitutes* in different communicative situations and social contexts, and are often created for specific purposes. In other words, the “[u]se of certain names under particular circumstances may be forbidden by religious taboo; or then again, such names may be replaced by descriptive nicknames” (Bright 2003: 672). However, the relationship between pseudonyms and nicknames to the first name and last name can vary. As Aleksiejuk (2014: 249) notes

[p]seudonyms do substitute – they are false names that play the role of the real ones, and are used instead. In turn, nicknames can be described as supplementary – they supplement a selection of alternative identification devices – depending on the social situation we can use, for example, a given name, a surname, a title or a nickname.

A pseudonym, a self-given name (Aleksiejuk 2016: 448), is adopted with the intention of hiding one's identity, recognized by the first name and last name, and various motivations are indicated for its creation: personal, ideological, conditioned by custom or prevailing trend, and forced by the political or social situation (Cieślíkowa 1998a: 135-136). On the other hand, according to Cieślíkowa (1998b: 122), the existence of a nickname is supported by the author's need to externalize their attitude towards another person, the bearer. Therefore, it is said that nicknames “tend to have strong emotive connotations” (Van Langendonck and Van de Velde 2016: 32). Moreover, the material for both pseudonyms and nicknames is an abundant repertoire of linguistic forms (cf. Cieślíkowa 1998a, 1998b) and “[w]ithin that identification may be communicated a sense of relationship, purpose, or social characterization” (Leslie and Skipper 1990: 277).

However, research in the field of anthroponomastics lacks findings on the names assigned to deceased individuals when their real identity, indicated by first and last name, is somehow unknown. As Gałkowski (2024) mentions, the term *onym* is extremely productive and actively used as a compositional element combined with various thematic terms that semantically lead to different sections of onomastics and different classes of onyms (Gałkowski 2024: 14). We therefore propose, for the purposes of this study and others of a similar nature, to distinguish a *postmortonym* (from Latin *post mortem* ‘after death’ and Greek *onoma* ‘name’) within anthroponymy as a naming unit used to refer to the unidentified deceased. We assume that a *postmortonym* is a postmortem byname,

fictional, non-authentic onym assigned to a deceased person, whose true and authentic name is (at least initially) unknown. It is given to the individual by themselves, by society, by the media (primarily the internet, press, TV, true crime podcasts, etc.), or by institutions (courts, law enforcement agencies, such as the prosecution and police), formally or informally. Such a name must be comprehensible to the broader linguistic community or at least a part of it, repeatable and reproducible, and should unambiguously refer to that specific deceased individual.

PMOs may therefore share some features with other classes of anthroponymic names, such as the above-mentioned pseudonyms or nicknames. They are units that exist alongside the proper (usually unknown) anthroponyms identifying the deceased, but they are created in specific social and communicative conditions and for a particular purpose, and therefore, they function as name substitutes.

Moreover, PMOs, due to their sphere of occurrence, are thematically, not denotatively, similar to necronyms (from Greek *necros* ‘dead’ and *onoma* ‘name’), which are either proper names of necropolises or names given to newborns after deceased relatives in order to honor them (Caffarelli and Gagliardi 2018: 37; Gałkowski 2024: 22). The issue of necronyms, in the second sense, is taken up by anthropologists, who draw attention to the fact that in different cultures, “death-names” are treated differently, and are associated with local traditions as well as taboos (Testoni *et al.* 2020). Gałkowski, on the other hand, places necronyms in the group of onymic terms whose status is still debated; they have not gained recommendations in onomastics studies and have not been widely accepted in research works (Gałkowski 2024: 22). Furthermore, as PMOs are mainly used in the mass media and the criminal cases they cover, they should be considered in the thematic vicinity of investigation-related names, e.g. the aliases of criminals (see Golda and Mężyk 2021). However, it would be erroneous to assume that postmortem bynames are exclusively associated with criminal cases. They may also be assigned, for instance, to ancient mummies or bodies recovered in high-altitude mountain regions whose true identities remain unknown.

Lastly, the emergence of a PMO is associated with the multiplication of names indicating one referent, which thus requires an understanding of the problem of the multiplicity of names. The term *polyonymy* originally referred to the use of more than one name within an anthroponymic system (Bijak 2016: 55). Over time, it was extended to other branches of onomastic studies, especially toponomastics, where the phenomenon of multiple names gained various terms, e.g. *polymorphism* (*dimorphism*), *naming parallelism*, *variantness*, *duality of naming acts* or *naming alterations* (Bijak 2016: 55-56). We understand polyonymy similarly to Bijak (2016) and earlier Krško (2003), as any instability and variability in a name referring to one referent, consisting in changes within the

onym itself, its phonetic or grammatical form, as well as in the creation of a new name with a different motivation. The issue of variety primarily falls within the scope of socio-onomastics, the branch of onomastics that acknowledges that onymic variables are mainly shaped by different societal and cultural registers (Nicolaisen 1985: 123). In terms of the analyzed material, the last possibility is vital, i.e. the creation of a new name with a different motivation and in a strictly defined nomination context. The newly created units (postmortonyms) thus enter into a relationship with officially standardized and sanctioned names, i.e. the first and last names of the deceased, which have not yet been identified during the ongoing investigation. Therefore, a PMO exists temporarily alongside an unknown proper anthroponym that establishes the identity of the deceased, while the relationship between these names operates primarily on two levels: the social (official vs. unofficial name) and the codificational (standardized vs. non-standardized name). Thus, we treat the phenomenon of polyonymy as a particular synonymous relation in which the uses of specific onyms are conditioned by the social and institutional situation of use. It is impossible to ignore this phenomenon, especially when the identity of the deceased is finally established during the ongoing investigation. Additionally, it seems that the PMO becomes independent over time, functioning in the media sphere alongside the proper name, such as first and last name, to prove the individual's identity.

3. Methodology

As previously stated, this study aims to validate the existence of the term *postmortonym* by the analysis of specific examples of this phenomenon. Accordingly, we began by collecting 220 units that fulfilled the proposed definition of PMO. The postmortonyms were primarily collected from the Interpol website, Wikipedia, and Unidentified Wiki. The last one is a website “dedicated to unidentified individuals [...] as well as formerly unidentified persons”, which hosts the largest publicly available database of unidentified individuals, including their PMOs. To ensure the validity of the collected (potential) PMOs, each one was checked and verified in other sources. Only the bynames that appeared in at least three other sources were included in the data. Once the data had been gathered, they were organized in a table accessible online (Golda *et al.* 2026) – importantly, the analyses below follow the source spelling.

The data table includes essential information for each analyzed unit: translation into English (if applicable), disambiguation, source, country, gender, and whether the unit refers to a child. In this article, the term *child* refers to a young human being who has not yet reached adulthood; that is, a person in the early stages of physical and mental development, from birth to adulthood. It must be acknowledged that, in many cases of nameless bodies, the lack of identity does

not help determine whether the individual was a child (often, only an approximate age range, e.g. 15-25 years, can be specified). For this reason, we have also adopted the category *possibly a child* in our data table.

It should be noted that the search was not limited to English-language units, although they constitute the majority of our research data – all non-English units have been translated into English. The multilingual and multicultural nature of this research constitutes both an asset and a limitation. On the one hand, it demonstrates that PMOs are a universal phenomenon, appearing in diverse languages and cultures (e.g. English, French, Polish, and Japanese), which aligns with our objective of introducing this new category of proper names. The distribution by country of the sourced units is as follows: Argentina: 1, Australia: 5, Belgium: 11, Brazil: 4, Canada: 4, Czech Republic: 1, Denmark: 2, France: 7, Germany: 11, Indonesia: 1, Ireland: 3, Israel: 1, Italy: 6, Japan: 7, Netherlands: 15, Norway: 2, Pakistan: 2, Poland: 4, South Korea: 1, Spain: 10, Tibet: 1, UK: 10, USA: 111. On the other hand, the research's multilingual material complicates the analysis of the structural patterns presented later in the article.

The analytical part of this study is grounded in two types of analysis: semantic and structural. Semantic analysis allows linguists to explore the principles that govern the construction of meaning, focusing on the relationships between expressions and the concepts they represent (Cruse 2011). This framework was applied to examine the references, possible etymology, and motivations behind PMOs. In this analysis, if possible, we establish the basic connotative meanings according to the four types discussed by Van Langendonck and Van de Velde (2016: 31-32). Importantly, as suggested by Brylla (2016: 250), the analysis of historical data “involves the risk of speculation and misinterpretation [...] since it is not always easy to find naming motives.” This remark also applies to data that cannot be verified by the author. Therefore, in the section *Culture*, we propose some potential references that may not have influenced the origin of the PMO.

The structural analysis, in turn, aims to identify and describe recurring structural patterns observed in these linguistic units. The rationale for this analysis is rooted in the idea that “the semantic status of names is mirrored by certain syntactic (called ‘symbolic’) constructions” (Van Langendonck and Van de Velde 2016: 17). Importantly, due to the pilot nature of the study and its focus on the category of PMOs, the analysis remains fairly simplistic in form and deserves further, more detailed investigation – for instance, through the application of the categories, criteria, and terminology discussed by Aleksiejuk (2016) for the analysis of pseudonyms.

The sources used in the research are all openly and freely accessible on the Internet. We did not have access to any case files and based our research exclusively on information readily available in news articles and other digital sources.

4. Semantic analysis

4.1. Anthroponyms

As the material shows, postmortonyms incorporate anthroponyms, including both first names and last names. However, the anthroponyms observed in PMOs are notably heterogeneous. The most prevalent and significant examples identified within the collected data are *John Doe* and *Jane Doe*. Among the collected data, there are instances of both *John Doe* and *Jane Doe*, with specific variations such as *Bulldog John Doe*, *Bibb County John Doe*, and *San Mateo County John Doe*, as well as *The St. Louis Jane Doe*, *Tujunga Jane Doe*, and *Wexford Jane Doe*.

There are also cases where the names *John Doe* or *Jane Doe* are followed by a number, such as *Jane Doe #40* (also spelled *Jane Doe 40* or *Jane Doe No. 40*). Interestingly, this PMO was used twice: for a female body found on October 1, 1976, in San Francisco (California), later identified, and for a female body found on May 28, 1974, in Long Beach (California), which remains unidentified. The case of *Mary Doe four* – a combination of *Doe* and the numeral *four* – illustrates a similar phenomenon. This designation was assigned to an elderly woman who died in March 1978 at Queenwood East Nursing Home in Illinois. Her true identity had remained unknown for four decades, during which she spent time in various institutions. Throughout her institutionalization, the name *Mary Doe-four* was repeatedly applied and ultimately became her postmortem identifier.

Remaining within the context of full names (i.e. [first name] + [last name]), we also observed examples such as *Jason Doe* (given to a deceased named after a note addressed to *Jason*), *Betty Boop* (given to a deceased found with a T-shirt featuring a picture of Betty Boop on the front), and *Elvis Frank* (given to a newborn found in a shoebox near the Elvis Presley bus stop in Rhode-Saint-Genève). Importantly, despite the presence of an anthroponym in the last PMO, it was derived from a place name (a bus stop) rather than from Elvis Presley himself.

Additionally, some instances of PMOs involve only first names, with female-related examples including *Horseshoe Harriet*, *Wych Elm Bella*, and *Baby Theresa*, and male examples such as *Fred the Head*, *Adam*, and *Coal Train Carlos*, including also diminutives, e.g. *Charlie* (*Charlie* is a short form of *Charles*), *Beth Doe* (*Beth* is a short form of *Elizabeth* or *Bethany*), *Castleberry Kate* (*Kate* is a short form of *Catherine*), and *Palisades Pete* and *Pipeline Pete* (*Pete* is a short form of *Peter*).

Interestingly, there are PMOs that incorporate initials or simply letters, e.g. 女性 X ('Woman X '), *Miss X*, and *Mister S*. The latter two examples incorporate titles such as *Mister* and *Miss*, and similar instances of PMOs can be observed in English, such as *Mister Somerton*, *Ms. Wiggins*, and *Ms. Startex*.

The final point to note is that the collected data allow us to conclude that some PMOs containing anthroponyms were assigned by police, investigators, or

journalists, such as *Eklutna Annie*, *Baby Theresa*, *Millie Doe*, and *Adam*. However, there are cases of names actually used by the deceased prior to their death that later became their PMOs, such as *Mary A. Anderson*, 吉田 慶子 ('Yoshida Keiko'), *Lyle Stevik*, and *Peter Bergmann*. Moreover, there are cases where nicknames given to individuals during their lifetime later became PMOs, such as *Mary Doefour*, *Six Shooter Jim*, or *Funny*. It is worth noting that in one case, a PMO, an anthroponym, was later confirmed to be the deceased's true name: *Monique*.

4.2. Location

In the collected examples of PMOs we can also observe the reference to the location where the deceased was found. Thus, they refer to natural features of the landscape and may include examples of hydronyms (i.e. names of bodies of water), oronyms (i.e. names of the mountains and hills), and urbonyms (i.e. names of the cities). Once again, this category of *nomen falsum* is highly heterogeneous. PMOs may also specify the exact location of discovery, while others provide an approximate reference. For example, *the Bolney Torso* was found near Bolney, while *Smyrna Doe* was discovered near Smyrna.

The PMOs that mention location use county names (i.e. choronyms), e.g. *Walker County Jane Doe*, *Miami County Jane Doe*, and *San Mateo County John Doe*. Other units within this category include city names where the nameless people were found or near which they were discovered, e.g. *the Kambo Man*, *Tempe Girl*, and *Chłopczyk z Cieszyna* ('the Boy from Cieszyn'). Additionally, PMOs can specify particular districts or neighborhoods within cities (i.e. urbanonyms), such as *Tujunga Jane Doe* (Los Angeles, USA), *Setiabudi 13* (Jakarta, Indonesia), and *Ueno John Doe* (Tokyo, Japan).

Several units refer to bodies of water, e.g. *the Girl in the River Main*, *the Woman in the Gaasp River*, *the Woman in the IJ River*, *Otonabee River Man* (a river), *the Woman in the Canal* (a canal), and *Delta Dawn* (a river delta). Moreover, the analyzed PMOs include names of islands, e.g. *the Gentleman of Heligoland* (discovered in the waters surrounding Heligoland), caves, e.g. *Trow Ghyll skeleton*, or are elevation-based, e.g. *Lady of the Hills* (a hill), *the Lady of the Dunes* (sand dunes), *The woman of Mount Artxanda* (a mountain), *Isdalskvinnen* ('The Isdal Woman', a valley). Notably, there are units referring to the beaches, e.g. *La Dama del Camello* ('The Lady of the Camello') and *Ballinamona Woman*, as well as the one referring to a bog, *The Body in the Bog*. PMOs are also derived from forests or park names, e.g. *The Burned Body in the Forest*, *Clocaenog Forest Man*, and *Angel of the Meadow* (based on the Angel Meadow public park in Manchester). However, they may also designate the general vegetation around the body, e.g. *Nude in the Nettles*, or include other common

nouns referring to the surroundings, e.g. *Mężczyzna ze ścieżki* ('The Man from the Path'), *The Woman on the Road*, *The Woman by the Motorway*, and *The Girl in the Parking Lot*. Therefore, PMOs can also refer to buildings, such as hotels, e.g. *Sahara Sue* (named after the Sahara Hotel in Las Vegas), sheds, e.g. *Mill Creek Shed Man*, or industrial structures, e.g. *Pipeline Pete*.

4.3. Gender and stage of life

Postmortonyms may provide information about the gender or stage of life of the deceased. Regarding gender, PMOs, especially those in English, employ the terms *John Doe* and *Jane Doe*. Other examples contain personal names that inherently indicate masculinity or femininity, e.g. *Saja* (male) and *Viola* (female). Additionally, as previously discussed, some PMOs incorporate honorifics that specify gender, such as *mister* or *miss*, e.g. *Mr. X* and *Ms. Wiggins*.

PMOs also incorporate gender-specific nouns. For example, the noun *lady* is included in units such as *Lime Lady*, *Lady of the Lake*, and *The Lady in the Fridge*. Others employ general terms such as *man* or *woman* (or their equivalents in other languages) to indicate gender. Examples of male-associated PMOs include *Mill Creek Shed Man*, *Somerton Man*, and *Tamam Shud Man*, whereas female-associated examples include *Woman in the Trunk*, *Ballinamona Woman*, and *Isdalskvinnen* ('The Isdal Woman'). There are other gender-specific terms that indicate a specific sex. *The Gentleman of Heligoland* refers to a male, while *Princess Blue* and *Princess Doe* refer to a female. Additionally, *Cheerleader in the Trunk* uses the term *cheerleader*, which typically connotes a young female.

One may question the rationale for discussing gender and stage of life within a single section, as some PMOs simultaneously indicate both gender and possible age by incorporating terms such as *boy* or *girl* (or their equivalents in other languages). Examples containing *boy* (or its equivalents) include *Boy Under the Billboard*, *Chłopczyk z Cieszyna* ('The Boy from Cieszyn'), and *The Boy in the Box*. Similarly, examples incorporating *girl* (or its equivalents) include *Das Mädchen aus dem Main* ('The Girl from the Main'), *Lanaye Girl*, and 少女A ('Girl A').

In some cases, the stage of life – particularly young age and the fact that the individual was a child – is indicated through the use of the word *baby*, as in *Baby Jane*, *Baby Jane Doe*, and *Baby Elle*, or through the adjective *little*, as seen in *Little Zion Doe* and *Little Miss Nobody*, or *young*, e.g. *The young woman found in Saint-Denis*. Due to the inherent heterogeneity of PMOs, other creative units also exist, such as *Sweetie Doe*, which, similarly, refers to a young person.

However, it should be noted that a proper analysis of the incorporation of the terms *baby* and *girl* requires a detailed analysis of information about the case, since both of the terms may, by semantic extension, be used to refer to an adult woman to metaphorically denote her fragility.

4.4. Circumstances

As previously discussed, location is a crucial factor in forming PMOs and often remains an integral part of the discovery context. However, PMOs may reflect the circumstances surrounding the discovery of the body or the cause of death, e.g. *Paducah Plane Jumper*, who fell from the wing of a departing airplane. In some cases, even the specific condition in which the body was found is explicitly indicated in PMOs, e.g. *The Woman in the well*, *Nude in the Nettles*, *The woman on the boat*, and *Boy Under the Billboard*. Additionally, PMOs may suggest the enclosed space of body discovery. Examples include *Septic Tank Sam*, *The woman in the cardboard box*, *The woman in the suitcase*, and *The Lady in the Fridge*. Some of these, such as *The Boy in the Box* (applied to three different deceased) and *Woman in the Trunk*, emphasize especially confined spaces. Some units reference vehicles, as seen in *Coal Train Carlos*.

However, the circumstances influencing the creation of PMOs can vary significantly. Some names highlight the condition of the body, particularly when only remains are discovered. For example, *Setiabudi 13* refers to the discovery of 13 bones, while *The Bolney Torso* denotes a case in which only a torso was found. Similarly, *Trow Ghyll Skeleton* and *Skeleton in Sneakers* emphasize the advanced state of decomposition of the remains. There are also cases in which the PMO reflects the order in which body parts were discovered. For example, in the case of *Fred the Head*, the head was found first, whereas in *Theresa Torso*, it was the torso.

Some PMOs indicate the way the body was treated postmortem, e.g. *Lime Lady*, whose body was completely covered in lime to accelerate decomposition, *Tent Girl*, who was wrapped in tent fabric, and *Body in the carpet*, which refers to a body wrapped, among others, in a carpet. Here, we should also mention the case of a body staged to appear as a Persian princess, which is related to two PMOs, namely, *Persian Mummy* and *Persian Princess*. Another case that refers to the circumstances is *The Woman Who Wasn't Alone*, attributed to the body of a victim found in a hotel room. Her hands were bound, and there were visible signs of violence on her body.

Additionally, PMOs can also be inspired by the time of discovery. Examples include *Valentine Sally* and *Valentine Doe*, both named after Saint Valentine's Day – *Valentine Sally* was discovered on February 14, 1982, while *Valentine Doe* was found on February 15, 1991, a day after Valentine's Day. There is at least one example of a PMO that includes the year, namely *Jane Smyrna Doe 2019*.

4.5. Appearance

The collected examples of PMOs also refer to the body's physical appearance. For instance, one PMO is derived from the deceased's general appearance, suggesting a Latino origin. This is reflected in the name *Coal Train Carlos*,

where *Carlos* is a common name in Latin American countries, which aligns with the individual's presumed ethnic background. Similarly, based on the overall appearance, *The Bride* was given to the deceased due to her clean look.

In the material, we can also find *Scarlett Doe*, named for her distinctive red hair, and *The Woman with the Artificial Nails*, which highlights the deceased's appearance by emphasizing her notable nails. Both cases illustrate how unique physical features can be the source of the PMOs. Similarly, tattoos may play a crucial role in identification and thus inspire the name-givers, e.g. *the Girl with the Scorpion Tattoo*, *the Woman with the Flower Tattoo*, *the Woman with Panther and Scorpion tattoos*, *the Woman with Butterfly Tattoos*, *Cherries Doe*, *Bulldog John Doe*, *Daisy Doe*, and *Peaches*. While the first four units are self-explanatory, the others may require further clarification. First, *Cherries Doe* was named after a cherry tattoo. Second, *Bulldog John Doe* derived from a distinctive Marine Corps tattoo on his right arm, depicting a bulldog wearing a First World War helmet. Third, *Daisy Doe* was named as such because the individual had a daisy tattoo on their shoulder. Lastly, *Peaches* was given to an individual who had a tattoo of a heart-shaped peach with a bite taken out of it, accompanied by two falling drops.

Clothing items worn by the deceased can also inspire PMOs. Interestingly, these units do not necessarily refer to the clothing *per se*, but to the prints or logos. For example, *Betty Boop* was named after the cartoon character Betty Boop, who appeared on her T-shirt. Similarly, *Corona Girl* got her name from a T-shirt with a Corona beer logo, which became one of the few clues in her case. Lastly, *Vandy Jane Doe* was named after her hoodie with the Vanderbilt University logo. *Smurfette Jane Doe* was named after the distinctive blue-green shirt she wore, which featured the image of the cartoon character Smurfette and the text "HE SMURFS ME/HE SMURFS ME NOT." Another example of the unit that points to the text printed on the clothing item is *The woman with the 'Little Italy' T-shirt*.

There are also names that point to the textile designs, e.g. *Disco Dan*, attributed to a man found murdered in Texas in 1978. This byname is derived from the deceased's distinctive blue floral-print shirt, whose pattern resembled musical notation. Other cases based on textile designs include *The woman with the flower skirt* and *The woman with the paisley jeans*.

However, the color of clothing can also influence the creation of PMOs. For instance, *Lavender Doe* was named for her purple sweater; similarly, *Blue Pacheco*, was named for her blue clothing, and *Woman in Pink*, was named for her fuchsia pink pants, a blouse with a pink floral print, and black-and-pink sandals.

Interestingly, the type of clothing seems to be crucial to the PMO. *Pyjama Girl* was named after the pajamas she was found wearing, and *Buckskin Girl* after the buckskin jacket she wore. *Orange Socks* refers to a person who was found naked except for a pair of orange socks, which became the source of the PMO.

Another example is *Skeleton in Sneakers*, where the individual's sneakers were visible above ground at the time of discovery. A notable case is *Green Boots*, a name given to the body of an unidentified climber identified by his green boots, which have become a landmark along the main Northeast ridge route of Mount Everest. Various theories exist regarding the identity of the deceased. There is also an interesting case related to the gender of clothing items, namely, *The Woman in Men's Clothing*.

In addition to clothing, jewelry can also inspire the creation of PMOs. For example, *Girl with the Turquoise Jewelry* was named due to the several pieces of turquoise jewelry she had, while *Princess Blue* received her name from a ring with a blue gemstone. The *Woman with the 'Jean & Nelly' Ring* refers to a woman who wore a gold wedding band engraved with laurel leaves on the outside and the inscription on the inside.

Furthermore, clothing may sometimes allude to pop culture, movies, or literature. One example is *Little Lord Fauntleroy*, a name derived from the children's novel by Frances Hodgson Burnett, in which the protagonist, Cedric Errol, is referred to as *Little Lord Fauntleroy*. This PMO was assigned to a boy from Waukesha, whose elegant clothing suggested he came from a wealthy family. Similarly, *Black Dahlia* was given to an individual who was last seen wearing a tight-fitting black dress – a reference to the 1946 film *The Blue Dahlia*, which was popular at the time and known for its noir themes.

4.6. Belongings

As previously discussed, clothing and jewelry can influence the postmortonyms assigned to deceased individuals, as they are essential personal belongings. However, other types of possessions, aside from clothing and jewelry, may also contribute to the creation of PMOs. For instance, two PMOs referring to the same individual – *Jason Doe* and *Grateful Doe* – are both derived from belongings found with the body. A note containing the name *Jason* led to the first byname, while two Grateful Dead concert tickets found with the deceased inspired the second. Similarly, *Tamam Shud Man*, a name derived from the Persian phrase “tamám shud”, meaning “It is over” or “It is finished.” The phrase was printed on a scrap of paper found in the fob pocket of the man's trousers. The note had been torn from the final page of a copy of *Rubaiyat*, a 12th-century poem by Omar Khayyám.

An interesting case is that of *The Globetrotter*, a woman found with several items suggesting international connections: a map of Vancouver, a map of Copenhagen Airport, a scientific American magazine with a subscription request for Payson (Arizona), a dermatology article with an address in Cambridge, magazine pages featuring the Danish Royal Family, and a Scandinavian Airlines luggage

label. Here, the *Christmas Tree Lady* can also be mentioned, since her body was discovered with a backpack and an 8-inch (200 mm) Christmas tree. Similarly, the PMO *Monique*, most likely derived from the empty handbag with the name *Monique* written on the outside, which was found near the body.

Other examples of PMOs derived from the objects include *the Girl with the 10-Pence Coin*, *the Woman with the Cane*, *the Woman with the Watch*, and *the Woman with the German Keys*, all of which may be considered self-explanatory. Some belongings, however, were not immediately visible at the time of discovery, e.g. *the Woman with the Richmond Dental Crown*.

4.7. Culture

Naturally, culture (e.g. popular culture, film, and literature) also influences postmortonyms. Several of the cases already mentioned under appearance fall into this category: *Little Lord Fauntleroy* (here linked to prose), *Tamam Shud Man* (here linked to poetry), *Betty Boop*, *Smurfette Jane Doe*, and *Black Dahlia* (here linked to movies). However, this list is not exhaustive.

One notable case of literary inspiration is *Lyle Stevik*, the name (here also a pseudonym) used by an unidentified man who checked into a motel in Amanda Park, Washington, on September 14, 2001. Two days later, he was found deceased in his room, having died by suicide through hanging. The name *Lyle Stevik* was taken from a character – Lyle Stevick (with slightly different spelling) – in Joyce Carol Oates' 1987 novel *You Must Remember This*, and it was self-assigned by the deceased. Another literature-related name is *Mostly Harmless* given to Vance Rodriguez, who died of starvation. He was also known by the false names *Denim* and *Ben Bilemy*, the latter of which he used while staying in various hostels. The name *Mostly Harmless* is likely a reference to a character from Douglas Adams' *The Hitchhiker's Guide to the Galaxy* series.

Here we should also mention *Lady of the Lake*, which presumably refers to a character from Arthurian legend – a mysterious ally of Merlin who took part in various enigmatic missions. In some versions of the legend, only her arm is visible, always adorned in white, shimmering brocade. At other times, she is depicted walking on water or appearing in a castle, embodying a matriarchal deity.

Cinematic influences on PMOs are noticeable. While *Ms. Wiggins* was previously mentioned as an example containing the honorific *Ms.*, it also derives from a television series. *Mrs. Wiggins* was a recurring comedy sketch on *The Carol Burnett Show*, with one installment airing on Carol Burnett & Company. Another example, less straightforward, is *Julie Doe*, a PMO assigned to a transgender woman believed to have been murdered in Clermont, Florida, in 1988. The name *Julie* was chosen by forensic anthropology students examining her

remains. It was inspired by the LGBT-themed film *To Wong Foo, Thanks for Everything! Julie Newmar*.

The last example, *Delta Dawn*, is the PMO given to the body of a female infant. This name possibly originated from a song written by Larry Collins and Alex Harvey, first recorded in 1971. However, the body was actually found near the delta of the Escatawpa River in 1982. Thus, the name *Delta Dawn* may derive from both the location where the body was found and the song.

4.8. Linguistic Effects

The final category of postmortonyms consists of various linguistic effects including two rhyme-based PMOs, *Fred the Head* and *Belle in the Well*.

Furthermore, we have wordplay based on the place of discovery, *Angel of the Meadow*, which refers to remains discovered in 2010 in Angel Meadow, a public park in Manchester.

Finally, PMOs feature alliteration, where at least two words in the PMO begin with the same sound, e.g. *Horseshoe Harriet*, *Theresa Torso*, *Tommy Torso*, *Palisades Pete*, *Pipeline Pete*, *Disco Dan*, *Christy Crystal Creek*, *Castleberry Kate*, *Skeleton in Sneaker*, *Ben Bilemy*, and *Nude in the Nettles*. Notably, these examples often incorporate common first names (e.g. *Pete* or *Christy*), presumably deliberately chosen to create alliteration, paired with nouns that serve to identify the location or condition of the remains upon discovery.

5. Structural analysis

The second analytical section is dedicated to the structural analysis of post-mortonyms and revisits some information presented in the previous one. However, it does so in a different context, emphasizing patterns of formation, i.e. focusing on the structure or syntax.

The most basic structural distinction within the collected PMOs is between monolexical and polylexical units. In the research data, the latter category clearly predominates, with monolexical units accounting for 15 examples. Some of them are commonly used first names or their diminutives, e.g. *Anna* or *Charlie*, while others are common nouns, such as *Peaches* or *Denim*. Moreover, there is also one adjective: *Funny*, or even two examples of single letters, as in the Japanese *N* and *W*. As previously noted (see Methodology), the multilingual character of the research data poses challenges for structural analyses. For instance, the Dutch *Heulmeisje* ('Heul Girl') and the Norwegian *Isdalskvinnen* ('The Isdal Woman') are classified as monolexical since they are closed-form compound nouns in their

respective language. However, when translated into English, they consist of two or three separate words.

Polylexical units in the research data range in length from two to eight words. The frequency distribution shows a clear inverse correlation between the length of PMOs (counted in words) and the number of occurrences in the research data. That is, as the number of words in a PMO increases, the number of such examples decreases. Thus, the most common are two-word PMOs (with over 70 examples), while there is only one example of an eight-word PMO. Two- and three-word PMOs constitute more than half of the research data, suggesting that brevity is clearly favored; yet the small number of monolexical examples might indicate that they are not informative, distinctive, or semantically transparent enough.

In two-word PMOs (73 instances), several structural patterns can be distinguished. Some follow the pattern [attributive noun] + [person-denoting noun], in which the second unit is a common noun referring to a human, while the first one modifies or specifies it, e.g. *Ballinamona Woman*, *Corona Girl*, *Lime Lady*, or *Somerton Man*. A similar pattern is [adjective] + [person-denoting noun], e.g. *Persian Princess*, *Różana dziewczyna* ('Rose Girl'), or *Lil Miss*, which uses an abbreviation of *little* and an honorific. In one example (*Princess Blue*), this order is reversed. Another pattern containing a person-denoting noun is [definite article] + [person-denoting noun], e.g. *the Bride* or *the Globetrotter*. Moreover, as noted in the section on anthroponyms, the PMOs also contain personal names, as in: [attributive noun] + [prosonym], which is used when a personal name constitutes the second unit, e.g. *Castleberry Kate* or *Baby Theresa*. Sometimes the prosonym is used as the first unit: *Theresa Torso* and *Tommy Torso*. Some examples are [first name] + [last name] combinations, e.g. *Peter Bergmann*, *Lyle Stevik*, or 吉田 慶子 ('Yoshida Keiko'). A special case is the use of *Doe*, when the PMO can be created to resemble a [first name] + [last name] sequence, e.g. *Millie Doe* or *Jason Doe*, or the first unit can be something related to the person – a noun, e.g. *Daisy Doe* – or an adjective, e.g. *Grateful Doe*. An interesting case is a PMO with an honorific as the first unit, followed by a proper noun: *Ms. Star-tex*, or by an abbreviation: *Mister S*. There are more cases in the research data, but they are scarcely represented. For instance, two examples follow the pattern [adjective] + [noun], where a noun phrase is headed by a common noun and modified by an adjective: *Green Boots* and *Orange Socks*. There is also one adjective phrase: *Mostly Harmless*. An example from Indonesia features a number: *Setiabudi 13*, and two examples from Japan – 女性X ('Woman X') and 少女A ('Girl A') – include single letters, which might suggest different postmortonym-formation patterns in some languages.

In three-word PMOs (47 instances), are created by adding a word (typically a noun) in front of *Jane Doe* or *John Doe*, e.g. *Vandy Jane Doe*, *Baby Jane Doe*, or *Ueno John Doe*. Another common pattern is the use of a first name or

a person-denoting noun preceded by two other words, typically nouns (common or proper), e.g. *Christmas Tree Lady*, *Coal train Carlos*, or *Wych Elm Bella*, and in one example, the first name is the first unit: *Christy Crystal Creek*. Some follow the pattern [definite article] + [proper noun] or [adjective] + [noun], often person-denoting nouns, e.g. *The Grauballe Man*, *The Nameless Girl*, or *The Bolney Torso*. Moreover, there are also some PMOs that include nouns followed by prepositional phrases, as in *Mężczyzna ze ścieżki* ('The Man from the Path'), *Stranger of Lipari*, or *Skeleton in Sneakers*. Three examples of PMOs for children include the adjective *little*, e.g. *Little Miss Nobody* and *Little Zion Doe*. Rare cases within the research data include, as already said, numbers, e.g. *Jane Doe 40*, as well as [first name] + [last name] combinations with abbreviated middle name: *Mary A. Anderson*, and appositive constructions: *Fred the Head*.

Within the four-word subcategory of PMOs (32 instances), two main formation patterns can be distinguished. The first one consists of a noun (mainly a person-denoting noun) followed by a prepositional phrase, e.g. *Boy Under the Billboard*, *Woman at the border*, *La Dama del Camello* ('The Lady of the Camel'), or *Jack in the box*. The second one is similar to three-word PMOs as it uses *Jane Doe* or *John Doe*, but they are preceded by two words (proper names, typically toponyms), e.g. *Miami County Jane Doe* or *Parker Dam John Doe*. One example, *Mill Creek Shed Man*, does not use *Doe*, but instead uses *Shed Man*, which provides more information. Other formation patterns, although scarcely represented, include two PMOs containing *Little Miss*, e.g. *Little Miss Lake Panasoffkee*, a combination of *Jane Doe*, place name, and year: *Jane Smyrna Doe 2019*, and an appositive structure: *Betty the Bag Lady*.

The remaining PMOs containing five or more words are discussed collectively, as they are rather homogenous structurally, regardless of their varying lengths.¹ Moreover, it should be noted that PMOs containing six or more words are almost exclusively sourced from Interpol. All but one (*Girl with the Turquoise Jewelry*) of these PMOs begin with a definite article, e.g. *la extraña dama de Garde* ('The Strange Lady of Gardey') or *The woman with the cane*. The article is followed by a noun, mainly a person-denoting noun, of which *woman* (or its equivalents) is the most common in the research data. In certain cases, the noun is preceded by an adjective, e.g. *pregnant* or *young*. All such examples include prepositional phrases, of which the most common is *in*, as the preposition of place indicating typically where the body was found, e.g. *The woman in the canal*, and *with*, indicating a characteristic (appearance or belongings), e.g. *The woman with the flower skirt* or *The woman with panther and scorpion tattoos*. Only three examples do not follow this pattern. Firstly, *The St. Louis Jane Doe*, which is a common pattern in the two and three-word PMOs,

¹ The distribution is as follows: five words: 30, six: 16, seven: 6, eight: 1.

The woman who wasn't alone, where the noun phrase also includes a restrictive relative clause, and *The young woman found in Saint-Denis*, with a reduced restrictive relative clause.

6. Discussion and conclusions

Since time immemorial, names have carried symbolic, metonymic significance, as though they embodied the bearer's character or appearance rather than merely being assigned on the basis of such traits. It can also be said that "[t]he meanings of names are the result of complex social negotiations, learned, and reified through socialization" (Leslie and Skipper 1990: 273). In the case of postmortonyms, the name-giver may be less concerned with what a name might offer the individual and more interested in what the name represents. In other words, PMOs typically reflect the deceased's physical condition. It also shows that the deceased are not merely bodies that need to be identified – if it were so, they could be just numbered – but are also people whose stories should be remembered or understood.

As the analysis shows, PMOs can usually be categorised fairly easily due to their relatively repetitive structure. The names examined were predominantly based on: (1) the possessions of the deceased, (2) the place of discovery, and (3) the appearance of the body. However, these categories are neither mutually exclusive nor exhaustive.

Regarding physical appearance, PMOs can, either directly or indirectly, indicate the deceased person's age. For children or infants, terms such as *boy*, *baby*, or diminutives are commonly used, while *John Doe* and *Jane Doe* generally refer to adults and also signal the gender of the deceased. Interestingly, the last name *Doe* on its own appears more frequently in male cases. In terms of age, the word *lady* is sometimes used to suggest an older woman (e.g. *Christmas Tree Lady*). However, in the analyzed research data, *lady* is also employed in other contexts – such as indicating a presumed female (e.g. *Lady of the House*, *Lady of the Hills*, *Lady of the Dunes*), referencing cultural motifs (e.g. *Lady of the Lake*), or simply as a polite form of address for an unidentified woman (e.g. *Lime Lady*, *Lady in the Fridge*). The anonymity of individuals is also highlighted in other PMOs, such as *Stranger of Lipari*, *Little Miss Nobody*, and *L'Inconnue de la Seine* ('The Unknown Woman of the Seine'). Some names also stem from unverifiable information. For instance, *the Frenchwoman* was so named because witnesses heard her speaking French – although this does not necessarily confirm her nationality. Similarly, *the Introvert* was named based on claims that the individual rarely spoke to others.

We also observe the use of generic names such as *John*, *Jane*, and *Jack*. However, it should be noted that the use of *John* and *Jane Doe* may also be

influenced by legal conventions in the respective country. Notably, in the collected material, there are 48 instances of *Doe*, primarily from the USA, with isolated cases from Spain, Japan, and Ireland.

While PMOs may originate from various sources or individuals, tracing their origins is difficult and certainly beyond the scope of this article. In some cases, their emergence in media outlets is apparent or suspected; in others, they originate from or are adopted by investigative organizations, e.g. Interpol or DNA Doe Project. Thus, although the media may play a role in creating or popularizing particularly striking PMOs, these bynames ultimately become established through their use beyond the originating source.

However, the analyses suggest that the form and content of PMOs may be source-dependent. For instance, PMOs sourced from Interpol tend to be longer and emphasize the deceased's belongings or clothing, with one exception: *The woman with the Belgian connection*, which appears more mysterious than informative. In contrast, shorter PMOs, often linked to highly publicized cases, may have been coined by journalists seeking a sensational, memorable name. Nevertheless, structural analysis shows that two- and three-word units are more common, possibly because monolexical units may lack informativeness, while longer ones may be less memorable. The topic of PMO length – and thus their descriptive character – deserves further analysis, since descriptiveness in onomastics is most prevalent in place names (see Bright 2003). Moreover, the presence of valorizing elements (either directly or indirectly) in the structure of PMOs provides a basis for studying the expressive function of these names, which – generally speaking – conveys the name-giver's emotional attitude toward the bearer (see Van Langendonck and Van de Velde 2016). This function is also present in names that include, for example, diminutives of personal names, further indicating their expressive character.

Nevertheless, regarding descriptiveness, it should be noted that the present analysis did not consider the capitalization of the PMOs. This aspect may be relevant when evaluating the nominative or descriptive character of the analyzed units. For instance, there may be a meaningful distinction between *Belle in the Well* and *The woman by the motorway*, with the former functioning as a naming unit and the latter likely serving as a mere description of the case.

Given the cultural and linguistic diversity of the analyzed material, the methodology presented can also be applied in cross-linguistic comparative studies of PMOs. Moreover, the above analysis not only demonstrates the need for a new category of *postmortonyms* within onomastic typology but also serves as a pilot study for future research. First, the results presented here are worth pairing with a questionnaire to assess the units' reception – whether they convey the intended information and are truly memorable. Second, the motivation behind the naming calls for further investigation, for example, whether specific names given to deceased infants are intended to evoke sympathy or possibly reflect the name-

giver's attitude toward the deceased (see *Precious Doe* case). Third, since PMOs used in widely recognized or international cases tend to be translated, a detailed analysis of how these names are transferred across languages would be valuable. Furthermore, it should be noted that the units were sourced from various websites with differing purposes (cf. Interpol vs. Unidentified-awareness.fandom), making a source-based analysis of PMOs another worthwhile direction for future research. In other words, the study could also focus on the name-givers, including not only institutions such as the police but also journalists and even the deceased themselves, who adopted names, i.e. pseudonyms, that later became PMOs. At the same time, it would be worthwhile to examine how these names function across various discursive spaces, such as institutional ones (e.g. judicial, legal, police) and media-related ones (e.g. press, internet, television), and in the texts that actualize them. The use of PMOs across such diverse areas of communication would also enable a deeper analysis of the functions they perform and their significance in social interactions. From a discursive perspective, proper names are treated as social constructs that offer insight into the collective consciousness of language users (Rutkowski and Skowronek 2019: 49) and, according to De Stefani (2016), they are worth further analysis in discursive environments.

Acknowledgement

The authors would like to express their sincere gratitude to Professor Iwona Loewe for her encouragement, support, and guidance throughout this research, as well as to two anonymous reviewers for their valuable comments.

References

- Aleksiejuk, K. 2014. Internet Names as an Anthroponomastic Category. In J. Tort Donada and M. Montagut Montagut (eds.), *Els noms en la vida quotidiana. Actes del XXIV Congrés Internacional d'ICOS sobre Ciències Onomàstiques. Annex 3*, 243-255. Barcelona: Generalitat de Catalunya.
- Aleksiejuk, K. 2016. Pseudonyms. In C. Hough (ed.), *The Oxford Handbook of Names and Naming*, 439-452. Oxford: Oxford University Press.
- Algeo, J., and K. Algeo 2000. Onomastics as an interdisciplinary study. *Names: A Journal of Onomastics* 48(3-4): 265-274.
- Bijak, U. 2016. Polionimia, czyli 'wojna nazw'. *Polonica* 36: 55-66.
- Bright, W. 2003. What IS a name? Reflections on onomastics. *Language and Linguistics* 4(4): 669-681.
- Bryant, M.M. 1957. Names in everyday speech. *Names: A Journal of Onomastics* 5(1): 47-58.

- Brylla, E. 2016. Bynames and nicknames. In C. Hough (ed.), *The Oxford Handbook of Names and Naming*, 237-250. Oxford: Oxford University Press.
- Caffarelli, E., and C. Gagliardi 2018. Terminologia onomastica. *Archivio per il vocabolario storico italiano* 1: 11-54.
- Cieślíkowa, A. 1998a. Pseudonimy. In E. Rzetelska-Feleszko (ed.), *Polskie nazwy własne. Encyklopedia*, 135-141. Warszawa-Kraków: Wydawnictwo IJP PAN.
- Cieślíkowa, A. 1998b. Przewiska. In E. Rzetelska-Feleszko (ed.), *Polskie nazwy własne. Encyklopedia*, 119-134. Warszawa-Kraków: Wydawnictwo IJP PAN.
- Cruse, D.A. 2011. *Meaning in Language: An Introduction to Semantics and Pragmatics* (2nd ed.). Oxford: Oxford University Press.
- Dal Pinal, G. 2015. Dual content semantics, privative adjectives, and dynamic compositionality. *Semantics & Pragmatics* 8(7): 1-53.
- De Stefani, E. 2016. Names and discourse. In C. Hough (ed.), *The Oxford Handbook of Names and Naming*, 52-66. Oxford: Oxford University Press.
- Ehrlich, S. 2021. Semiotic ideologies and trial discourse: Implications for multimodal discourse analysis. In C. Gordon (ed.), *Approaches to Discourse Analysis*, 123-135. Washington, D.C.: Georgetown University Press.
- Ernst, P. 2010. Onomastische Motivationsforschung. Vom Beschreiben zum Erklären. *Onoma* 45: 61-78.
- Felton, G. 1972. Who is the real John Doe? *Names: A Journal of Onomastics* 20(4): 297-300.
- Gałkowski, A. 2024. Kilka spostrzeżeń na temat stanu i interdyscyplinarności terminologii onomastycznej. *Acta Universitatis Lodziensis. Folia Linguistica* 58(1): 9-28.
- Golda, P., and J. Mężyk 2021. English, French, and Polish aliases of criminals: Diversity of inspirations in their creation and typical nicknaming schemes. *Academic Journal of Modern Philology* 14: 173-200.
- Golda, P., A. Jedziniak-Danowska, M. Moskal and J. Ryszka 2026. *Naming the nameless: A corpus of selected postmortonyms*. <https://zenodo.org/records/19178816>.
- Hough, C. (ed.). 2016. *The Oxford Handbook of Names and Naming*. Oxford: Oxford University Press.
- Klaidman, S., and T.L. Beauchamp 1986. Baby Jane Doe in the media. *Journal of Health Politics, Policy and Law* 11(2): 271-284.
- Koles, T. 2024. The semantics of deadnames. *Philosophical studies* 181: 715-739.
- Krško, J. 2003. *Hydronymia povodia Turca*. Banská Bystrica: Matej Bel University.
- Lederman, H. 2021. Fine-grained semantics for attitude reports. *Semantics & Pragmatics* 14(1): 1-45.
- LeMay, E.C. 1995. A biography of the nameless: John and Jane Doe. *The Georgia Review* 49(3): 633-646.
- Leslie, P.L., and J.K. Skipper 1990. Toward a theory of nicknames: A case for socio-onomastics. *Names: A Journal of Onomastics* 38(4): 273-282.
- Nicolaisen, W.F.H. 1985. Socio-onomastics. In E. Eichler, G. Hilty, H. Löffler, H. Steger and L. Zgusta (eds.), *Namenforschung / Name Studies / Les noms*

propres. Ein internationales Handbuch zur Onomastik / International Handbook of Onomastics / Manuel international d'onomastique, 118-132. Leipzig: Karl-Marx-Universität.

Nordvall, C. 2020. The inconsistent translation of posthumous epithets. *Translation Quarterly* 96: 69-79.

Rutkowski, M., and K. Skowronek 2019. The onomastic discourse analysis. *Annales Universitatis Mariae Curie-Skłodowska sectio-FF Philologicae* 37(1): 47-60.

Sičáková, L. 2002. Polyonymy, polysemy, homonymy and antonymy of hydronyms. *Slovenská reč* 67(6): 330-341.

Testoni, I., M. Dorsa, E. Iacona and G. Scalici 2020. Necronym: the effects of bearing a dead little sibling's name. *Mortality* 3: 343-360.

Van Langendonck, W., and M. Van de Velde 2016. Names and grammar. In C. Hough (ed.), *The Oxford Handbook of Names and Naming*, 17-38. Oxford: Oxford University Press.

PIOTR A. OWSIŃSKI

Jagiellonen-Universität in Krakau, Polen

Philologische Fakultät / Institut für Germanistik und Nordistik

ORCID: 0000-0001-7862-3345

piotr.owski@uj.edu.pl

***DIE WELLEN DER WORTE: EINE EXPLORATIVE STUDIE
ZU ÜBERSETZUNGSSTRATEGIEN
UND HERAUSFORDERUNGEN
IM INTERNATIONALEN SEEHANDELSRECHT***

***THE WAVES OF WORDS: AN EXPLORATORY STUDY
ON TRANSLATION STRATEGIES AND CHALLENGES
IN INTERNATIONAL MARITIME LAW***

The article examines the challenges and possible strategies involved in translating maritime law terminology, based on a contrastive analysis of the German provisions on maritime trade in the *Handelsgesetzbuch* and the Polish *Kodeks morski*. German legal language is characterized by high nominalization, compound nouns, and lexically encoded legal roles, which may lead to misinterpretation or unintended shifts in legal qualification if translated literally into Polish. By contrast, Polish maritime legislation tends to favor descriptive constructions, verbalization of duties, and explicit specification of legal effects. As a result, translators are often required to unpack German lexical structures and apply functional equivalence rather than literal translation.

The paper discusses typical translational challenges related to bills of lading, maritime carriage contracts, risk transfer, and shipowner liability, and outlines possible translation strategies that help maintain terminological transparency, systemic coherence, and legal effectiveness. The analysis shows that maritime law can serve as a useful “translation laboratory” for examining terminological and functional equivalence in international legal communication, with implications for cross-border commercial and legal translation, where linguistic precision directly affects legal and financial outcomes.

Keywords: maritime law, HGB Seehandel, Polish maritime law, legal translation, functional equivalence, legal risk

1. Geleitwort

Die Ozeane bedecken über 70 % der Erdoberfläche und stellen damit eine zentrale Lebens- und Handelsader unseres Planeten dar. Sie tragen nicht nur entscheidend zum interkontinentalen Güterverkehr bei und bilden die Grundlage der globalen Kommunikation über Unterseekabel, sondern liefern auch einen erheblichen Anteil der weltweiten Versorgung mit hochwertigem Eiweiß aus Fischerei und Meeresprodukten (FAO 2022; European Commission Knowledge Centre 2025). Für mehr als 3,2 Milliarden Menschen machen diese maritimen Produkte mindestens 20 % des täglichen Konsums an tierischem Eiweiß aus (FAO 2022). Gleichzeitig haben jahrzehntelange Umweltbelastungen die Meere als fragile Ökosysteme hervorgehoben. Vor diesem Hintergrund verfolgt das internationale Seerecht als Teilordnung des Völkerrechts das Ziel, ein kohärentes und nachhaltiges System für die Nutzung der Weltmeere zu etablieren. Es reguliert nicht nur die kommerzielle Nutzung der Meere, sondern schützt deren ökologische Integrität und gibt Leitlinien für Forschung, Schutzmaßnahmen und ressourcenschonende Meeresbewirtschaftung vor. Während das internationale Seerecht einen umfassenden Regelungsansatz verfolgt und die Meere als Ganzes erfasst, richtet sich das Seehandelsrecht als Teilgebiet des Privatrechts gezielt auf die besonderen rechtlichen Beziehungen zwischen privaten Rechtssubjekten. Dazu gehören Regelungen zu Schiffsbetrieb, Frachtverträgen, Konnossementen und Haftungsfragen. Gerade diese Spezialisierung macht das Seehandelsrecht zu einem besonders spannenden, aber auch anspruchsvollen Feld für linguistische und translationswissenschaftliche Analysen. Im Vergleich zu den zahlreichen Studien im Erb- und Strafrecht, wie sie beispielhaft von Owsinski und Paluch durchgeführt wurden (bspw. Owsinski/Paluch 2020, 2022a, 2022b; Paluch/Owsinski 2020), sind sprachwissenschaftliche Untersuchungen des Seehandelsrechts im polnisch-deutschen Kontext bislang deutlich seltener. Während im Erb- und Strafrecht die Terminologie weitgehend etabliert und die Übersetzbarkeit der Termini relativ eindeutig ist [z. B. dt. (der) *Erbe* ↔ poln. *spadkobierca* oder dt. *Täter* ↔ poln. *sprawca*], erfordert das Seehandelsrecht aufgrund seiner fachlichen Spezialisierung und internationalen Dimension eine wesentlich differenziertere Herangehensweise.

Das deutsche *Handelsgesetzbuch* (HGB), insbesondere Buch V „Seehandel“, ist in sprachlicher Hinsicht besonders herausfordernd. Die Rechtsnormen sind durch hohe Nominalisierung, umfangreiche Komposita und lexikalisch kodierte Rollen gekennzeichnet, was eine starke Verdichtung der rechtlichen Bedeutung bewirkt. Eine wörtliche Übersetzung ins Polnische kann leicht zu Fehlinterpretationen der Rechtsfolgen oder zu einer Verzerrung der Normfunktion führen. Das polnische *Kodeks morski* (k.m.) bevorzugt dagegen beschreibende Konstruktionen, explizite Zeitangaben für die Rechtswirkungen und eine stärker verbal orientierte Normformulierung. Übersetzer müssen daher deutsche Komposita

‚entpacken‘ und funktionale Äquivalenz herstellen, anstatt auf rein lexikalische Übertragungen zu setzen. Diese methodischen und sprachlichen Hürden erklären, warum vergleichende Analysen des Seehandelsrechts bislang weniger verbreitet sind. Schon geringe Unterschiede in der Übersetzung können unmittelbare rechtliche und finanzielle Konsequenzen nach sich ziehen, etwa beim Risikoübergang, bei der Haftungsbegrenzung des Reeders oder bei den Rechten des Konnossementinhabers. Das Seehandelsrecht lässt sich demnach als besonders ‚fruchtbares Labor‘ für die Untersuchung funktionaler Äquivalenz in der juristischen Übersetzung betrachten. Die Analyse solcher Übersetzungen bedingt eine sorgfältige Berücksichtigung der normativen Funktion der Termini, ihrer rechtlichen Wirkungen sowie der internationalen Anwendbarkeit. Die gewonnenen Erkenntnisse sind nicht nur für das Seehandelsrecht relevant, sondern liefern wertvolle Impulse für die Übersetzung anderer transnationaler und handelsrechtlicher Rechtsbereiche.

2. Abriss des historischen Hintergrunds, Zielsetzung und Forschungsansatz

Die Entwicklung des deutschen Seehandelsrechts ist eng mit der langen Tradition des deutschen Handelsrechts verbunden. Bereits im 19. Jahrhundert wurden Grundlagen für die Regelung von Handelsschiffen, Frachtverträgen und Konnossementen im Rahmen des Handelsgesetzbuches gelegt. Mit der Verabschiedung des Handelsgesetzbuches 1897¹ wurden maritime Handelsbeziehungen systematisch kodifiziert und in Buch V (Seehandel) gesondert geregelt. Dieses Dokument vereinte bestehendes Gewohnheitsrecht, kommerzielle Praxis sowie internationale Standards und schuf eine einheitliche Rechtsgrundlage für die deutsche Seefahrt. Das polnische *Kodeks morski* hingegen wurde deutlich später verfasst, als Reaktion auf die zunehmende Bedeutung des Seehandels in der modernen Republik Polen. Die heutige Fassung stammt aus dem Jahr 2001 (Dz. U. 2001 Nr. 138, poz. 1545) und bildet eine eigenständige Kodifikation des Seehandelsrechts, die Elemente des Handelsrechts, des Zivilrechts und des öffentlichen Rechts integriert. Ziel war es, ein kohärentes System für die Regelung maritimer Rechtsverhältnisse zu schaffen, das internationalen Standards entspricht und gleichzeitig die spezifischen Bedürfnisse des polnischen Seerechts abbildet.

Für die vorliegende Studie bildet das Seehandelsrecht den zentralen Untersuchungsgegenstand. Analysiert werden insbesondere die gesetzestextlichen Normen des Handelsgesetzbuches (HGB), Buch V – Seehandel, und des polni-

¹ Handelsgesetzbuch (HGB) vom 10. Mai 1897, insbesondere Buch V – Seehandel. https://de.wikisource.org/wiki/Handelsgesetzbuch._Erstes_Buch._Handelsstand#.

schen *Kodeks morski*. Dieses Korpus dient als Grundlage für eine vergleichende Analyse ausgewählter Termini technici, Satzstrukturen sowie sprachlicher Konzepte im Seehandelsrecht und ermöglicht es, die Rolle der Sprache bei der Formulierung und Funktion von Normen in beiden Rechtsordnungen zu untersuchen. Die im Beitrag diskutierten Übersetzungsstrategien beziehen sich dabei nicht auf bereits vorhandene Übersetzungen, sondern beschreiben mögliche translatorische Verfahren, die aus der kontrastiven Analyse der genannten Rechtsnormen abgeleitet werden.

Der Beitrag verfolgt das Ziel, die sprachlichen Strukturen und funktionalen Mechanismen des Seehandelsrechts im deutsch-polnischen Vergleich zu beleuchten. Im Mittelpunkt steht die These, dass Sprache als unverzichtbares Instrument der Rechtssetzung fungiert: Gesetze existieren nur in sprachlicher Form, und auch die Rekonstruktion, Interpretation und Anwendung von Normen werden sprachlich vermittelt. Wie Sacco (1991: 17-18) hervorhebt, besteht ein Rechtssystem sowohl aus verbalen als auch aus außerlingualen Elementen, wobei letztere primär die Anwendung und Umsetzung von Recht, nicht jedoch dessen Formulierung betreffen (Beispiel: *per facta concludentia*). Selbst in Rechtsordnungen, in denen Gewohnheitsrecht maßgeblich ist, erfolgt die Weitergabe von Rechtswissen überwiegend über Sprache. Daraus folgt, dass Sprache nicht nur ein Medium der Formulierung, sondern auch der Vermittlung normativer Inhalte ist (Paluch/Owsiński 2020: 38). Diese Prämisse bildet die Basis für die vorliegende Untersuchung: Im Seehandelsrecht, sowohl im deutschen HGB als auch im polnischen *Kodeks morski*, ist die präzise sprachliche Formulierung von Normen entscheidend für deren funktionale Umsetzung und für die Übersetzungsstrategien zwischen beiden Rechtsordnungen. Die vorliegende Arbeit zielt ebenfalls darauf ab, diese Funktion im Bereich des Seehandelsrechts sichtbar zu machen und die Beziehungen zwischen sprachlicher Form und normativer Funktion in beiden Rechtsordnungen vergleichend zu analysieren.

3. Sprache als Normsetzungsmedium: Ein analytischer Einstieg in das Seehandelsrecht

Sprache bildet das Fundament menschlicher Kommunikation und ist zugleich Träger kultureller und sozialer Realitäten. In Platons Dialog *Kratylos* (424/423 v. Chr.-348/347 v. Chr.), der auch den deutschen Sprachwissenschaftler Bühler (1879-1963) inspirierte, wird Sprache als Werkzeug beschrieben, durch das ein Individuum einem anderen über Dinge berichtet. Aus dieser Perspektive erfüllt Sprache drei grundlegende Funktionen: 1) die referentielle Funktion, durch die die Realität abgebildet wird, 2) die expressive Funktion, durch die Gefühle und innere Zustände ausgedrückt werden, und 3) die appellative Funktion, die den Rezipienten zu Handlungen auffordert (Bühler 1965: 24-33; Trautmann-Voigt/

Voigt 2012: 43). Für das Rechtssystem sind insbesondere die referentielle und die appellative Funktion von wesentlicher Bedeutung: Sprache dient dem Gesetzgeber primär dazu, die von ihm geschaffene Rechtsrealität zu beschreiben (und damit die außerrechtlichen Realität zu regulieren) sowie normative Botschaften an die Adressaten der Rechtsnormen zu vermitteln. Lévi-Strauss (1908-2009) (2003: 22) hebt hervor, dass Sprache zugleich ein relevanter kultureller Faktor ist: Als erlernte Fähigkeit oder Gewohnheit wird sie vom Einzelnen übernommen, als Instrument ermöglicht sie die Erkenntnis der ihn umgebenden Kultur, und als Ausdruck geordneter Bedeutungsstrukturen dient sie der Analyse kultureller Artefakte, etwa von Kunst, Religion oder Recht. Die Sprache funktioniert zudem als Leitfaden durch die ‚soziale Wirklichkeit‘. Sie strukturiert in erheblichem Maße die kognitive Erfassung sozialer Prozesse und gesellschaftlicher Probleme. Menschen existieren weder ausschließlich in einer objektiven noch in einer rein sozialen Welt, sondern bewegen sich in einem durch sprachliche Konventionen strukturierten Raum, in dem die Realität durch die sprachlichen Gebräuche der jeweiligen Gemeinschaft geformt wird. Keine zwei Sprachen sind derart vergleichbar, dass sie dieselbe soziale Realität repräsentieren könnten. Die diversen Sprachgemeinschaften leben somit in separaten, sprachlich geprägten Welten, nicht jedoch in derselben Welt, lediglich unter unterschiedlichen Bezeichnungen (Sapir 2003: 81). Diese Überlegungen verdeutlichen, dass die Sprache nicht als bloßes Abbild von Gegenständen zu verstehen ist, sondern als Reflex des mentalen und kulturellen Weltbildes der Sprecher, was in engem Zusammenhang mit Humboldts (1767-1835) (1836: 41) Konzept des Sprachgeists steht, der die kollektive Weltsicht eines Volkes organisch mit der Sprache verknüpft. Ortega y Gasset (1883-1955) (2003: 169) betont, dass die Sprache eine charakteristische Form der Welt enthält und in ihr die Sichtweisen und die kulturellen Nuancen einer Gemeinschaft beobachtet werden können. Aus dieser Perspektive entsteht ein Wort nicht als direkte Abbildung eines Objekts, sondern als Spiegel des Bildes dieses Objekts im Bewusstsein des Sprachbenutzers (Casirer 2003: 68).

Vor diesem theoretischen Hintergrund wird im folgenden Kapitel der Fokus auf das Seehandelsrecht gelegt, ein Rechtsgebiet, in dem präzise sprachliche Formulierung und terminologische Genauigkeit aufgrund internationaler Transaktionen und fachsprachlicher Komplexität von extremer Relevanz sind.

3.1. Terminologische Analyse im Seehandelsrecht

Ein gewichtiges Problem bei der kontrastiven Analyse des deutschen Handelsgesetzbuches (HGB), Buch V – Seehandel, und des polnischen *Kodeks morski* besteht in den strukturellen Unterschieden zwischen der deutschen und der polnischen Fachterminologie, die sich besonders deutlich bei der Kodierung

rechtlicher Rollen, Institutionen und Rechtswirkungen in normativen Texten zeigen. Die deutsche Rechtssprache zeichnet sich durch eine hohe Produktivität von Nominalkomposita aus, d. h. durch Zusammensetzungen aus mehreren Substantiven oder substantivierten Verben, die in einem einzigen lexikalischen Ausdruck mehrere semantische Komponenten bündeln. Solche Komposita sind im HGB besonders häufig anzutreffen und tragen nicht nur lexikalische, sondern zugleich systematische Informationen über Rechtsverhältnisse und normative Funktionen. Im *Kodeks morski* werden solche Rollen und Funktionen hingegen meist durch beschreibende, syntaktisch transparente Phrasen wiedergegeben, welche die semantischen Komponenten explizit ausdrücken, anstatt sie in einem einzigen Lexem zusammenzufassen. Diese Konstruktionen sind funktional äquivalent, unterscheiden sich jedoch morphosyntaktisch deutlich vom deutschen Pendant.

Die Analyse der ausgewählten Beispiele aus dem Fünften Buch – Seehandel im HGB zeigt exemplarisch die hohe Produktivität von Nominalstrukturen im deutschen Seehandelsrecht. Auffällig ist, dass sowohl Nominalkomposita als auch nominalisierte syntaktische Konstruktionen zentrale normsetzende Funktionen erfüllen, indem sie in einem einzigen Ausdruck sowohl die Rechtsbeziehung als auch die normativen Pflichten kodieren. So kodiert beispielsweise der *Anspruch auf Ausstellung eines Konnossements* (§ 513 HGB) die Verpflichtung des Frachtführers, dem Absender auf dessen Verlangen ein Konnossement auszustellen, während das *Bord- und Übernahmekonnossement* (§ 514 HGB) unterschiedliche Momente der Dokumentenausstellung in einem einzigen Ausdruck zusammenfasst und damit die normativen Konsequenzen der jeweiligen Transportphase abbildet. *Inhalt des Konnossements* (§ 515 HGB) definiert komprimiert die erforderlichen Angaben des Frachtpapiers, wobei das Nominalkompositum durch die Präpositionalphrase den Bezug zum Dokument herstellt. Schließlich kodieren Ausdrücke wie *aus dem Konnossement Berechtigter* oder *legitimierter Besitzer des Konnossements* (§ 519 HGB) die legalen Rollen und Berechtigungen, die mit dem Konnossement verkoppelt sind, und zeigen, dass selbst nominalisierte Phrasen die systematische Normfunktion zum Vorschein bringen.

Die obigen Beispiele verdeutlichen, dass das deutsche Seehandelsrecht Nominalstrukturen nicht nur lexikalisch, sondern zugleich funktional einsetzt, um Rechte, Pflichten und normativ relevante Rollen zu kodieren. Für die Übersetzung ins Polnische bedeutet dies, dass die funktionale Entsprechung der deutschen Nominalstrukturen häufig analytische Konstruktionen erfordert, bei denen Präpositionen und substantivische Ergänzungen kombiniert werden, um die gleiche normative Funktion wiederzugeben. Während die deutschen Komposita mehrere semantische und normative Informationen in einem einzigen Ausdruck integrieren, muss im Polnischen diese Verdichtung aufgelöst werden, um die rechtliche Wirkung korrekt adäquat darzustellen. Dies ist besonders signifikant für Übersetzer, da ein zu wörtliches Übertragen der deutschen Nominalstruktur häufig zu Missverständnissen oder einer veränderten rechtlichen Qualifikation

Tabelle 1: Nominal geprägte Regelungsstrukturen im Seehandelsrecht des HGB und ihre funktionale Bedeutung

NOMINALSTRUKTUR / PHRASE	HGB	NORMATIVE FUNKTION
<i>Anspruch auf Ausstellung eines Konnossements</i>	§ 513 HGB	Kodierung der Verpflichtung des Frachtführers, das Konnossement auszustellen
<i>Bord- und Übernahmekonnossement</i>	§ 514 HGB	Kodierung der Art des Konnossements und des normativen Zeitpunkts der Ausstellung
<i>Inhalt des Konnossements</i>	§ 515 HGB	Festlegung der erforderlichen Inhalte des Konnossements
<i>aus dem Konnossement Berechtigter / legitimer Besitzer des Konnossements</i>	§ 519 HGB	Bestimmung der Berechtigten und deren Rechtsposition

führen würde. Die Kenntnis der systematischen Funktion dieser Nominalstrukturen ist daher unerlässlich, um die funktionale Äquivalenz zwischen HGB und dem *Kodeks morski* sicherzustellen.

Demgegenüber weist *Kodeks morski* eine deutlich analytischere Ausdrucksweise auf. Normative, im HGB nominal kondensierte Inhalte werden im Polnischen regelmäßig durch vollständige Satzstrukturen oder explizite verbale Prädikationen entfaltet. So entspricht dem deutschen *Anspruch auf Ausstellung eines Konnossements* die Formulierung, wonach *przewoźnik obowiązany jest wydać konosament* ‘der Frachtführer ist verpflichtet, den Konnossement auszustellen’ (Art. 129 k.m.). Dadurch wird die Verpflichtung sprachlich als Handlungspflicht des Normadressaten realisiert. Ein ähnlicher Unterschied zeigt sich bei der Differenzierung der Arten des Konnossements. Während das HGB mit den Komposita *Übernahmekonnossement* und *Bordkonnossement* (§ 514 HGB) diverse rechtlich relevante Zeitpunkte der Dokumentenausstellung terminologisch verdichtet, greift der polnische Gesetzgeber auf beschreibende Wendungen zurück, die den jeweiligen Sachverhalt explizit benennen (art. 130 k.m.). Auch die rechtlichen Wirkungen des Konnossements werden im deutschen Recht nominal abstrahiert (*Inhalt des Konnossements*, § 515 HGB), während der polnische Text dessen Beweis- und Legitimationsfunktion explizit ausformuliert (Art. 131 k.m.). Diese Unterschiede betreffen nicht lediglich stilistische Präferenzen, sondern spiegeln differente systematische Strategien der Normformulierung wider, die für die juristische Auslegung und für den Rechtsverkehr von hoher Relevanz sind.

Für die Übersetzung und terminologische Arbeit im deutsch-polnischen Seehandelsrecht ergeben sich aus diesen Beobachtungen substantielle Konsequenzen: Die nominalen Verdichtungen des deutschen Seehandelsrechts transportie-

Tabelle 2: Kontrastive Gegenüberstellung der normativen Funktionen des Konnossements im HGB und im polnischen Kodeks morski

NORMATIVE FUNKTION	HGB	KODEKS MORSKI
Anspruch auf Ausstellung eines Konnossements	<i>Anspruch auf Ausstellung eines Konnossements</i> (§ 513 HGB)	<i>obowiązek wydania konosamentu po przyjęciu ładunku</i> (Art. 129 k.m.)
Art des Konnossements	<i>Übernahmekonnossement / Bordkonnossement</i> (§ 514 HGB)	<i>konosament przyjęcia do załadowania / konosament przyjęcia ładunku na statek</i> (Art. 130 k.m.)
Rechtswirkungen des Konnossements	<i>Inhalt und Beweiswirkung des Konnossements</i> (§ 515 HGB)	<i>konosament jako dowód przyjęcia ładunku i dokument legitymujący</i> (Art. 131 k.m.)
Berechtigung zur Verfügung über die Ladung	<i>legitimierter Besitzer des Konnossements</i> (§ 519 HGB)	<i>legitymowany posiadacz konosamentu uprawniony do odbioru ładunku</i> (Art. 144 k.m.)

ren nicht nur lexikalische Information, sondern kodieren zugleich systematische Zuordnungen von Rechten, Pflichten und Rechtsfolgen. Eine formale, wortgetreue Transponierung solcher Nominalstrukturen ins Polnische würde regelmäßig zu terminologischen Verzerrungen oder zu nicht kodexkonformen Ausdrücken führen. Vielmehr erfordert eine adäquate Übersetzungsstrategie die funktionale Rekonstruktion der jeweiligen Norm: Der Übersetzer muss die im deutschen Nominalausdruck implizit enthaltenen Handlungselemente, normativen Rollen und zeitlichen Bezüge explizieren und sie in eine syntaktische Struktur überführen, die den polnischen legislativen Formulierungspraktiken entspricht. Dabei steht nicht die formale Äquivalenz einzelner Wörter im Vordergrund, sondern die funktionale Entsprechung der rechtlichen Wirkung innerhalb des jeweiligen Rechtssystems. Gerade im Seehandelsrecht, das durch internationale Sachverhalte, intensiven Dokumentenverkehr und hohe Anforderungen an Rechtssicherheit geprägt ist, wird deutlich, dass terminologische Entscheidungen unmittelbar auf die Verständlichkeit, Auslegbarkeit und praktische Anwendbarkeit von Normen wirken. Diese strukturelle Asymmetrie führt dazu, dass Übersetzungsentscheidungen regelmäßig nicht auf der Ebene formaler Äquivalenz getroffen werden können. Stattdessen ist eine funktionale Äquivalenz anzustreben, bei der nicht die äußere sprachliche Gestalt, sondern die rechtliche Funktion des Ausdrucks im jeweiligen Normgefüge maßgeblich ist. Die Übersetzung wird damit zu einem Akt der funktionalen Rekodierung innerhalb eines anderen sprachlich-rechtlichen Systems.

Auf der Grundlage der kontrastiven Analyse der Terminologie des Handelsgesetzbuches (Buch V – Seehandel) und des polnischen *Kodeks morski* lassen

sich drei grundlegende translationsstrategische Verfahren rekonstruieren. Diese Verfahren orientieren sich zudem an den Grundsätzen der juristischen Fachübersetzung (Šarčević 1997; Sandrini 1996, 1999; Mattila 2016) und berücksichtigen die funktionale Äquivalenz zwischen Ausgangs- und Zielsprache. Die folgende Typologie hat heuristischen Charakter und dient der Systematisierung der im Material erkennbaren Übersetzungsmöglichkeiten:

- 1) strukturkonservierende Strategie, bei der versucht wird, eine nominale Struktur auch in der Zielsprache beizubehalten. Diese Strategie ist jedoch im deutsch-polnischen Seehandelsrecht nur eingeschränkt anwendbar, da das Polnische deutlich geringere Möglichkeiten zur Bildung komplexer Nominalkomposita bietet;
- 2) analytische Strategie, bei der ein deutsches nominales Verdichtungsprodukt durch eine beschreibende, syntagmatische Konstruktion im Polnischen wiedergegeben werden kann. Die Strategie wird auf der Grundlage eines Vergleichs von Termini aus dem Handelsgesetzbuch und dem polnischen *Kodeks morski* rekonstruiert und erweist sich als besonders funktional, da sie erlaubt, die normative Funktion eines Terminus transparent und systemkonform abzubilden, auch wenn dabei terminologische Kürze verloren geht;
- 3) funktional-adaptive Strategie, bei der weder eine formale noch eine semantisch unmittelbare Entsprechung angestrebt wird, sondern ein Ausdruck gewählt wird, der im Zielrechtssystem dieselbe rechtliche Wirkung entfaltet. Diese Strategie ist besonders relevant bei systemgebundenen Begriffen des Seehandelsrechts, deren Bedeutungsumfang erst aus dem Zusammenspiel mehrerer Normen erschlossen werden kann.

Die Ausarbeitung eines solchen Modells ermöglicht es, die beobachteten Unterschiede zwischen der deutschen und der polnischen Rechtssprache im Bereich des Seehandelsrechts systematisch zu ordnen und vergleichend zu beschreiben. Übersetzungsentscheidungen lassen sich in diesem Rahmen als differente Realisierungen derselben normativen Inhalte verstehen, die jeweils an die strukturellen und konzeptionellen Besonderheiten der beteiligten Rechtsordnungen angepasst sind. Besonders deutlich wird dies im Kontrast zwischen stark nominal geprägten Ausdrucksweisen des deutschen Seehandelsrechts und den häufiger analytisch oder paraphrasierend formulierten Lösungen im *Kodeks morski*. Das Modell dient damit in erster Linie der Beschreibung wiederkehrender sprachlicher Muster und ihrer funktionalen Entsprechungen im deutsch-polnischen Vergleich. Zugleich bildet es einen Übergang von der terminologischen Einzelbetrachtung hin zu einer umfassenderen Analyse der sprachlichen Gestaltung rechtlicher Normen im Seehandelsrecht beider Rechtsordnungen.

Der Tabelle 3 ist zu entnehmen, dass eine direkte Übernahme der deutschen Nominalstruktur ins Polnische nur selten eine adäquate Wiedergabe der normati-

Tabelle 3: Vergleich von Nominalstrukturen im deutschen HGB und ihren funktional-analytischen Entsprechungen im polnischen *Kodeks morski*

deutsche Nominalstruktur	polnische analytische Entsprechung	funktionale Transformation	Referenz
<i>Anspruch auf Ausstellung eines Konnossements</i>	<i>obowiązek wydania konosamentu po przyjęciu ładunku</i>	Nominalisierung → Verb + Präposition + Substantivische Ergänzung	§ 513 HGB ↔ Art. 129 k.m.
<i>Übernahmekonnossement / Bordkonnossement</i>	<i>konosament przyjęcia do załadowania / konosament przyjęcia ładunku na statek</i>	Nominalisierung → Präposition + Substantivische Ergänzung	§ 514 HGB ↔ Art. 130 k.m.
<i>Inhalt und Beweiskraft des Konnossements</i>	<i>konosament jako dowód przyjęcia ładunku i dokument legitymujący</i>	Nominalstruktur → sequenzielle Darstellung (Handlung + Rechtsfolge)	§ 515 HGB ↔ Art. 131 k.m.
<i>legitimierter Besitzer des Konnossements</i>	<i>legitymowany posiadacz konosamentu uprawniony do odbioru ładunku</i>	Nominalverdichtung → funktionale Beschreibung der Rolle und Rechte	§ 519 HGB ↔ Art. 144 k.m.

ven Funktion ermöglicht. Übersetzer: innen müssen daher die jeweilige Normfunktion präzise erkennen, einschließlich der Pflichten, Rechte, Bedingungen sowie des zeitlichen Bezugs. Die Analyse erfolgt nach einem kontrastiven Prinzip: Deutsche Nominalisierungen werden hinsichtlich ihrer funktionalen Entsprechungen im Polnischen betrachtet und systematisch unter die Lupe genommen, attributive Merkmale werden in Präpositionen und substantivische Ergänzungen aufgelöst, und die kompakte Darstellung der Norm wird als sequenzielle Abfolge von Handlung, Bedingung und Rechtsfolge rekonstruiert. Dieses Modell dient als Referenzrahmen zur Bestimmung funktionaler Äquivalenz im Seehandelsrecht, ohne normative Bewertungen zu implizieren.

4. Funktionale Äquivalenz und Übersetzungsrisiken im deutsch-polnischen Seehandelsrecht

Die in Kapitel 3 analysierten terminologischen und strukturellen Divergenzen zwischen dem Seehandelsrecht des HGB und dem *Kodeks morski* sind nicht lediglich sprachsystematisch zu erklären, sondern entfalten unmittelbare normative Relevanz für die juristische Übersetzung. Dies gilt insbesondere für Normbereiche, in denen rechtliche Wirkungen implizit durch lexikalische Verdichtung

kodiert werden, wie etwa bei der Ausstellung und Wirkung des Konnossements (§§ 513-519 HGB) oder bei der Haftungsbegrenzung des Reeders (§§ 611-617 HGB). Die Übersetzung solcher Normen entscheidet nicht nur über terminologische Angemessenheit, sondern über die korrekte Rekonstruktion rechtlicher Wirkungen im Zielrechtssystem.

Vor diesem Hintergrund ist die Übersetzung seehandelsrechtlicher Texte zwischen dem Deutschen und dem Polnischen als ein Vorgang funktionaler Rekodierung zu verstehen, bei dem die im Ausgangstext konkludent enthaltenen normativen Informationen identifiziert und in eine systemkonforme Zielstruktur überführt werden müssen. Die folgenden Unterkapitel analysieren zentrale Risikobereiche dieser Rekodierung und zeigen auf, warum formale Äquivalenz im deutsch-polnischen Seehandelsrecht regelmäßig hinter der funktionalen Äquivalenz zurücktreten muss.

4.1. Theoretische Grundlagen: Formale und funktionale Äquivalenz

Die Gegenüberstellung des deutschen und des polnischen Seehandelsrechts verdeutlicht, dass formale Äquivalenz – verstanden als lexikalische oder strukturelle Entsprechung (Nida 1964: 159) – nur eingeschränkt geeignet ist, normativ gleichwertige Übersetzungen zu gewährleisten. Zahlreiche zentrale Termini des HGB fungieren als Träger komplexer Rechtswirkungen, ohne diese explizit zu entfalten. So regelt § 513 Abs. 1 HGB den Anspruch des Abladers auf Ausstellung eines Konnossements, ohne die daraus folgende Legitimations- und Beweisfunktion des Dokuments ausdrücklich zu benennen; diese wird erst systematisch in den §§ 515 und 519 HGB deutlich. *Kodeks morski* verfolgt demgegenüber eine explizitierende Normtechnik: Die Verpflichtung zur Ausstellung des Konnossements wird ausdrücklich normiert (poln. *przewoźnik obowiazany jest wydać konosament* – Art. 129 § 1 k.m.), während dessen Beweisfunktion (Art. 131 § 1 k.m.) und Legitimationswirkung (Art. 131 § 2 k.m.) jeweils einer eigenständigen Regelung unterworfen werden. Eine formale Übersetzung deutscher Nominalstrukturen würde diese normative Entfaltung im Polnischen verkürzen und damit den funktionalen Gehalt der Norm verändern.

Funktionale Äquivalenz erfordert daher eine Übersetzungsstrategie, die sich nicht an der formalen Textfassung einzelner Termini orientiert, sondern an deren normativer Wirkung im jeweiligen Rechtssystem. Maßgeblich ist nicht, wie eine Rechtsfolge sprachlich kodiert ist, sondern welche rechtliche Funktion sie erfüllt.

4.2. Übersetzungsprinzipien und funktionale Rekodierung normativer Strukturen

Die Übersetzung seehandelsrechtlicher Normen zwischen dem Deutschen und dem Polnischen darf keineswegs als lineare Substitution sprachlicher Einheiten verstanden werden. Vielmehr handelt es sich um einen mehrstufigen Prozess, bei dem zunächst die normative Funktion einer Struktur im Ausgangssystem identifiziert und anschließend eine zielsprachliche Form gewählt wird, die diese Funktion im Zielsystem adäquat wiedergibt.

Besondere Aufmerksamkeit verdienen nominal verdichtete Normen des HGB, deren semantische Komponenten häufig in einem einzigen Lexem gebündelt sind. Um ihre rechtliche Wirkung im Polnischen zu reproduzieren, müssen diese Strukturen analytisch aufgeschlüsselt und in syntaktisch transparente, funktional äquivalente Formen übertragen werden.

Die funktionale Rekodierung dieser normativen Strukturen stellt sicher, dass die Übersetzung nicht nur sprachlich korrekt, sondern auch normativ wirksam ist. Sie bildet die Grundlage für die Sicherstellung von Rechtsklarheit und -wirkung, insbesondere in einem durch internationale Anwendbarkeit, Dokumentenverkehr und hohe Anforderungen an Rechtssicherheit geprägten Seehandelsrecht.

4.3. Praktische Risikobereiche im Seehandelsrecht

Aufbauend auf den im vorherigen Abschnitt erläuterten Übersetzungsprinzipien werden die zentralen praktischen Risikobereiche im deutsch-polnischen Seehandelsrecht analysiert. Im Fokus stehen normativ relevante Zeitpunkte, der Übergang der Gefahr sowie die rechtlichen Rollen der beteiligten Akteure, einschließlich Haftung und Haftungsbegrenzung. Hier ist die exakte Wiedergabe terminologischer Strukturen entscheidend für die korrekte Rechtswirkung im Zielsystem. Besonders problematisch sind nominal verdichtete Normen des HGB, etwa zu Wirkung des Konnossements (§§ 515, 519 HGB), Risikoübergang (§ 475 HGB) oder Haftungsbegrenzung (§§ 611 ff. HGB), die im Polnischen in analytische, syntaktisch transparente Strukturen überführt werden müssen, wie sie im *Kodeks morski* (Art. 129-131, Art. 137, Art. 207-211 k.m.) vorgesehen sind.

Die Analyse dieser praktischen Problemfelder macht deutlich, dass präzise Übersetzungsentscheidungen und funktional orientierte Strategien notwendig sind, um die Rechtswirkung der Ausgangsnormen zu wahren und die Anforderungen des international anwendbaren Seehandelsrechts zu erfüllen.

4.3.1. Normativ relevante Zeitpunkte und der Übergang der Gefahr

Ein besonders sensibler Bereich funktionaler Rekodierung betrifft normativ relevante Zeitpunkte. Das deutsche Seehandelsrecht verbindet den Übergang der Gefahr beim Seefrachtvertrag mit der Ablieferung des Gutes (§ 475 HGB), ohne diesen Zeitpunkt sprachlich weiter zu präzisieren. Die normative Wirkung – der Übergang des Risikos vom Verfrachter auf den Empfänger – ist lexikalisch impliziert und setzt eine systematische Auslegung voraus. Im *Kodeks morski* wird dieselbe Rechtsfolge explizit temporalisiert formuliert. Nach Art. 137 k.m. geht die Gefahr *z chwilą wydania ładunku* ‘mit der Ablieferung der Ladung’ über, womit der Zeitpunkt des Eintritts der Rechtswirkung eindeutig festgelegt wird. Diese explizite Zeitmarkierung ist integraler Bestandteil der Norm und kann in der Übersetzung nicht ausgelassen oder abgeschwächt werden. Eine wortgetreue Übertragung der deutschen Formulierung (*przy wydaniu*) würde den normativen Charakter des Zeitpunktes relativieren und könnte den Eindruck eines zeitlich ausgedehnten Vorgangs erwecken. Die funktional adäquate Übersetzung verlangt daher eine gezielte Explikation des Zeitpunktes, auch wenn diese Information im Ausgangstext nur implizit vorhanden ist. Übersetzung fungiert hier als normativer Präzisierungsmechanismus.

4.3.2. Rechtliche Rollen zwischen lexikalischer Verdichtung und normativer Beschreibung

Ein zentrales Übersetzungsrisiko im deutsch-polnischen Seehandelsrecht ergibt sich aus der unterschiedlichen Kodierung rechtlicher Rollen. Während das deutsche Recht solche Rollen häufig durch kompakte nominale Strukturen erfasst, werden sie im polnischen Recht systematisch als normativ definierte Berechtigungspositionen formuliert. Diese Divergenz wird besonders deutlich im Zusammenhang mit dem Konnossement. Das HGB vermeidet bewusst eine substantivische Rollenbezeichnung wie *Konnossementinhaber* und stellt die aus dem Konnossement folgenden Rechte ausdrücklich unter den Schutz der Legitimation durch den Besitz des Dokuments (§ 519 Abs. 1 HGB). Der *legitimierte Besitzer des Konnossements* ist berechtigt, die Ablieferung des Gutes zu verlangen (§ 519 Abs. 1 HGB). Die rechtliche Rolle wird somit nicht durch eine abstrakte Inhaberschaft, sondern durch eine qualifizierte Besitzposition definiert, die ihre normative Wirkung aus der Legitimationsfunktion des Konnossements als Traditions- und Wertpapier entfaltet.² Im *Kodeks morski* wird die Berechtigungsposition aus dem Konnossement nicht über eine Person oder

² Auffällig ist, dass das HGB bewusst auf eine abstrakte Rollenbezeichnung wie *Konnossementinhaber* verzichtet. Stattdessen knüpft § 519 Abs. 1 HGB die aus dem Konnossement folgenden Rechte ausdrücklich an die Legitimation durch den Besitz des Dokuments. Diese terminologische Zurückhaltung ist dogmatisch bedeutsam, da sie verhindert, dass die Berechtigungsposition vom Legitimationsmechanismus des Konnossements als Traditions-

den Besitzbezeichnung (poln. *posiadanie*) kodiert, sondern funktional durch das Dokument und die Indossamente definiert. Art. 131 § 2 k.m. regelt die Übertragung der Rechte aus dem Konnossement, wobei die maßgebliche Rolle durch die aus dem Dokument folgenden Anspruchsrechte bestimmt wird. Die polnische Regelung verzichtet damit bewusst auf eine terminologische Anknüpfung an den Besitz und unterstreicht die funktionale, normative Bestimmung der Berechtigungsposition. Für die Übersetzung ergibt sich daraus eine erhebliche Konsequenz: Eine formale Übertragung der deutschen Besitzterminologie in eine polnische Besitzbezeichnung (poln. *posiadacz konosamentu*) würde eine sachenrechtliche Konnotation aktivieren, die weder dem Regelungszweck des § 519 HGB noch der systematischen Auslegung des *Kodeks morski* entspricht. Die funktional adäquate Übersetzung erfordert vielmehr eine Rekonstruktion der normativen Rolle als Berechtigungsposition, unabhängig von der sprachlichen Kodierung im Ausgangstext. Diese Problematik ist nicht auf das Konnossement beschränkt, sondern betrifft auch weitere Akteursbezeichnungen des Seehandelsrechts. Termini wie *Verfrachter* (§ 476 HGB) oder *Reeder* (§ 484 HGB) sind im deutschen Recht systematisch eindeutig verankert, während ihre polnischen Entsprechungen regelmäßig nur im Kontext konkreter normativer Pflichten definiert werden, etwa durch die Stellung von *przewoźnik* (Art. 117 k.m.) oder des *armator* (Art. 73 k.m.). Auch hier kann eine rein lexikalische Entsprechung zu einer stillschweigenden Verschiebung der rechtlichen Qualifikation führen.

4.3.3. Haftung und Haftungsbegrenzung als terminologischer Hochrisikobereich

Besonders hohe Anforderungen an die funktionale Äquivalenz bestehen im Bereich der Haftung. Das deutsche Seehandelsrecht kennt detaillierte Regelungen zur Haftungsbegrenzung des Reeders (§§ 611-617 HGB), die sich auf den Umfang der Haftung beziehen, nicht jedoch auf deren rechtliche Natur. Der Terminus Haftungsbeschränkung bezeichnet somit keine beschränkte Haftungsform, sondern eine quantitative Begrenzung der Ersatzpflicht.

Das *Kodeks morski* übernimmt dieses Konzept systematisch, formuliert es jedoch terminologisch explizit als *ograniczenie odpowiedzialności* (Art. 207-210 k.m.). Eine Übersetzung als *ograniczona odpowiedzialność* wäre systemwidrig, da sie Assoziationen zum Gesellschaftsrecht und zur Haftung juristischer Personen hervorrufen würde. Gleiches gilt für den Haftungsbeschränkungs fonds (§ 612 HGB), der funktional dem *fundusz ograniczenia odpowiedzialności* nach Art. 211 k.m. entspricht.

und Wertpapier abgelöst wird. Die rechtliche Stellung wird somit nicht personalisiert, sondern funktional über die Dokumentenlegitimation definiert.

Diese Beispiele verdeutlichen, dass terminologische Ungenauigkeiten im Haftungsrecht nicht nur dogmatische Fehlinterpretationen, sondern auch erhebliche finanzielle Konsequenzen nach sich ziehen könnten. Die funktionale Rekodierung ist hier zwingende Voraussetzung für Rechtssicherheit.

5. Schlussfolgerungen und Forschungsperspektiven

Die vorliegende Analyse hat gezeigt, dass die Übersetzung des Seehandelsrechts zwischen Deutsch und Polnisch weit über die rein lexikalische Übertragung hinausgeht. Wie bereits dargelegt, ist die bewusste Vermeidung des Terminus *Konnossementinhaber* im Handelsgesetzbuch dogmatisch bedeutsam: Die normative Rolle des Besitzers eines Konnossements wird funktional über die Dokumentenlegitimation definiert, nicht über eine personalisierte Inhaberschaft. In der polnischen Regelung im *Kodeks morski* werden die Rechte aus dem Konnossement ebenfalls funktional über Besitz und Indossamente kodiert, ohne dass ein terminologisches Pendant wie *osoba uprawniona z konosamentu* verwendet wird. Dies illustriert die zentrale methodologische Lektion: Übersetzer:innen müssen die zugrunde liegende normative Struktur rekonstruieren, anstatt sich auf eine scheinbar Eins-zu-eins-Entsprechung lexikalischer Termini zu verlassen.

Darüber hinaus beweist die Analyse der Satzstrukturen und der sprachlichen Konzepte im Seehandelsrecht, dass deutsche kompakte Nominalstrukturen oft in polnischen normativ definierten Berechtigungspositionen aufgehen. Eine rein lexikalische Übersetzung kann hier ungewollt sachenrechtliche Konnotationen aktivieren, die weder dem Regelungszweck des HGB noch der Systematik von *Kodeks morski* entsprechen. Dies betrifft nicht nur das *Konnossement*, sondern auch weitere Rollenbezeichnungen wie *Verfrachter* (§ 476 HGB) oder *Reeder* (§ 484 HGB), deren polnische Entsprechungen regelmäßig im Kontext normativer Pflichten definiert sind (*przewoźnik*, Art. 117 k.m.; *armator*, Art. 73 k.m.).

Für die Rechtsübersetzungswissenschaft ergibt sich daraus ein doppeltes Erkenntnisinteresse: Einerseits ist eine dogmatisch fundierte Analyse der Ausgangs- und Zielrechtsordnungen unabdingbar, um funktional äquivalente Termini rekonstruieren zu können. Andererseits wird die Übersetzungspraxis selbst zu einem epistemischen Instrument, das die normative Logik und die semantische Struktur des Rechts sichtbar macht. Die enge Verzahnung von theoretischer Analyse, dogmatischer Sensibilisierung und praktischer Übersetzungspraxis stellt somit ein paradigmatisches Vorgehen dar, das auch über das Seehandelsrecht hinaus relevant ist.

Zukünftige Forschungen könnten auf der Grundlage dieser explorativen Studie systematisch untersuchen, wie sich unterschiedliche Kodierungen rechtlicher Rollen in weiteren Rechtsbereichen auf die Übersetzbarkeit auswirken. Besonders interessant wären Felder mit stark formalisierten Traditions- und Wertpa-

pierrechten, in denen die Verbindung zwischen Dokumentenlegitimation und normativer Berechtigungsposition zentrale Übersetzungsherausforderungen birgt. Durch einen vergleichenden Ansatz, der kontinentaleuropäische und Common-Law-Systeme einbezieht, ließen sich wiederkehrende Probleme und Muster im Umgang mit rechtlich kodifizierten Rollen identifizieren und in übergreifende Übersetzungsstrategien überführen.

Ein weiteres vielversprechendes Forschungsfeld liegt in der methodischen Entwicklung von Leitlinien zur Rekonstruktion normativer Rollen in Übersetzungen, die systematisch die semantische, syntaktische und dogmatische Dimension berücksichtigen und zugleich praktische Handlungsempfehlungen für Übersetzer:innen bieten. Empirische Ansätze, wie Corpus-Analysen von Originaltexten und Übersetzungen, Fallstudien zu konkreten Kontrakten oder Experteninterviews, würden die theoretischen Überlegungen validieren und konkrete Lösungen aufzeigen.

Darüber hinaus eröffnet die systematische Erforschung von Satzstrukturen, sprachlichen Konzepten und ihrer Interaktion mit normativen Funktionen die Möglichkeit, die Übersetzbarkeit komplexer Rechtstermini noch feiner zu erfassen. Die Ergebnisse könnten nicht nur die Ausbildung von Übersetzer:innen und juristischen Fachkräften bereichern, sondern auch die interdisziplinäre Zusammenarbeit zwischen Rechtsvergleichung, Linguistik und Übersetzungswissenschaft stärken.

Schließlich verdeutlicht die vorliegende Pilotstudie, dass eine enge Verzahnung von theoretischer Analyse, dogmatischer Sensibilisierung und praktischer Übersetzungspraxis neue Perspektiven für die Rechtsübersetzungswissenschaft erschafft. Sie zeigt, dass selbst begrenzte Fallstudien wertvolle Impulse für ein umfassenderes Forschungsprogramm liefern können, das theoretische Tiefe, normative Präzision und praktische Anwendbarkeit miteinander verbindet. Zukunftsweisend wäre die Fortsetzung solcher Untersuchungen in größerem Ausmaß, um robuste methodische Standards zu etablieren und die systematische Erforschung der Übersetzbarkeit normativer Rollen im internationalen Recht weiter voranzutreiben.

Quellen

Fünftes Buch Seehandel, Handelsgesetzbuch, <https://www.gesetze-im-internet.de/hgb/BJNR002190897.html#BJNR002190897BJNG035503377>.

Kodeks morski z dnia 18. września 2001 r., Dz. U. 2001 Nr 138, poz. 1545, <https://isap.sejm.gov.pl/isap.nsf/download.xsp/WDU20011381545/U/D20011545Lj.pdf>.

Literaturverzeichnis

- Bühler, K. 1965. *Sprachtheorie: Die Darstellungsfunktion der Sprache*. Jena: Gustav Fischer.
- Cassirer, E. 2003. Język i budowa świata przedmiotowego. In G. Godlewski (Hrsg.), *Antropologia słowa. Zagadnienia i wybór tekstów*, 68-76. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- European Commission Knowledge Centre for Global Food and Nutrition Security. 2025. Fisheries & Aquaculture and Food & Nutrition Security. https://knowledge4policy.ec.europa.eu/global-food-nutrition-security/fisheries-aquaculture-food-nutrition-security_en.
- FAO. 2022. The State of World Fisheries and Aquaculture 2022. Towards Blue Transformation. Rome: Food and Agriculture Organization of the United Nations. <https://openknowledge.fao.org/server/api/core/bitstreams/a2090042-8cda-4f35-9881-16f6302ce757/content>.
- Handelsgesetzbuch (HGB) vom 10. Mai 1897, insbesondere Buch V – Seehandel. https://de.wikisource.org/wiki/Handelsgesetzbuch._Erstes_Buch._Handelsstand#.
- Humboldt, W. von. 1836. *Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechts*. Berlin: Königl. Akademie der Wissenschaften.
- Lévi-Strauss, C. 2003. Kultura i język. In G. Godlewski (Hrsg.), *Antropologia słowa. Zagadnienia i wybór tekstów*, 21-24. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- Mattila, H.E.S. 2016. *Comparative Legal Linguistics. Language of Law, Latin and Modern Lingua Francas*. New York: Routledge.
- Nida, E.A. 1964. *Toward a Science of Translating. With Special Reference to Principles and Procedures Involved in Bible Translating*. Leiden: Brill.
- Ortega y Gasset, J. 2003. Mówienie jako zwyczaj społeczny. In G. Godlewski (Hrsg.), *Antropologia słowa. Zagadnienia i wybór tekstów*, 169-171. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- Owsiński, P.A., A. Paluch 2020. Zum sprachlichen Weltbild in ausgewählten Begriffen aus dem Bereich des Erbrechts in der deutschen und polnischen Sprache. *Colloquia Germanica Stetinensia* 29: 255-276.
- Owsiński, P.A., A. Paluch 2022a. (Un)Übersetzbarkeit der ausgewählten deutschen und polnischen Begriffe aus dem Bereich des Strafrechts unter Einbeziehung des sprachlichen Weltbildes. *Osteuropa Recht. Ukraine: Recht unter Kriegsbedingungen* 68 (4): 522-532.
- Owsiński, P.A., A. Paluch 2022b. Językowy obraz świata w tłumaczeniu wybranych pojęć z zakresu prawa karnego w języku polskim i niemieckim. In E. Kujawska-Lis, A. Krawczyk-Łaskarzewska und O. Letka-Spychała (Hrsg.), *Komunikacja międzykulturowa w świetle współczesnej translatologii* 9, 207-218. Olsztyn: Wydawnictwo Uniwersytetu Warmińsko-Mazurskiego.

- Paluch, A., P.A. Owsinski 2020. O (nie)poprawności terminologii prawa spadkowego w polskim i niemieckim języku prawnym na przykładzie wybranych pojęć. *Przegląd Prawniczy Allerhanda* 3 – nr 1 (5), poz. 4: 37-52.
- Sacco, R. 1995. Mute Law. *The American Journal of Comparative Law* 43: 455-467.
- Sandrini, P. 1996. *Terminologearbeit im Recht. Deskriptiver begriffsorientierter Ansatz vom Standpunkt des Übersetzers*. Vienna: Internationale Network für Terminology.
- Sandrini, P. 1999. Translation zwischen Kultur und Kommunikation: Der Sonderfall Recht. In P. Sandrini (Hrsg.), *Übersetzen von Rechtstexten. Fachkommunikation im Spannungsfeld zwischen Rechtsordnung und Sprache*, 9-43. Tübingen: Gunter Narr Verlag.
- Sapir, E. 2003. Język – przewodnik po kulturze. In G. Godlewski (Hrsg.), *Antropologia słowa. Zagadnienia i wybór tekstów*, 77-82. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- Šarčević, S. 1997. *New Approach to Legal Translation*. The Hague – London – Boston: Kluwer Law International.
- Trautmann-Voigt, S., B. Voigt 2012. *Grammatik der Körpersprache. Ein integratives Lehr- und Arbeitsbuch zum Embodiment*. Stuttgart: Schattauer.

RENÁTA GREGOVÁ

Pavol Jozef Šafárik University, Košice, Slovakia

Department of British and American Studies, Faculty of Arts,

ORCID: 0000-0003-4743-2559

renata.gregova@upjs.sk

RENÁTA PANOCOVÁ

Pavol Jozef Šafárik University, Košice, Slovakia

Department of British and American Studies, Faculty of Arts,

ORCID: 0000-0002-3235-2006

renata.panocova@upjs.sk

DO DISTINCTIVE FEATURES PLAY A ROLE IN SPEECH ERRORS? EVIDENCE FROM ENGLISH

This paper examines phonological speech errors through the lens of distinctive feature theory, addressing two central questions: whether distinctive features operate independently or interdependently, and how frequently speech errors involve single-feature consonant substitutions. The analysis demonstrates that feature dependency varies according to the theoretical framework applied. Drawing on 100 consonant speech errors extracted from *the Fromkin Speech Error Database*, nearly 50% involve phonemes differing by only one distinctive feature. These findings confirm the relevance of distinctive features in speech production research and establish a methodological foundation for future investigations extending the analysis to vowels and typologically diverse languages.

Keywords: speech error, phoneme, distinctive feature, English

1. Introduction

The notion of distinctive features of phonemes was introduced into phonological theory by Roman Jakobson as early as 1941 (see, e.g. Mannell 2008 for further details). Delimitation of phonemes as “bundles of simultaneously produced distinctive features” (Jakobson and Halle 1956: 59) indicates that not a phoneme but a distinctive feature is the crucial element in phonological ana-

lysis. In Jakobson's theory, features depend on the articulatory and acoustic properties of speech sounds (Jakobson and Halle 1956: 32). Subsequently, various distinctive feature theories were introduced by different authors emphasizing the importance of the phonetic nature of sounds in their phonemic delimitation (see, e.g. Hall 2007 for the summary of the issue). Later some linguists left out acoustic properties of sounds from the description of their distinctive features (e.g. Chomsky and Halle 1968), and some researchers criticize distinctive features completely (e.g. Wilson 1966).

Nevertheless, distinctive features of phonemes have gained their solid position in phonology and are considered central for specification of a phoneme and classes of phonemes, for the delimitation of rules of numerous sound changes in speech and for the description of some speech errors (Eulenberg and Farhad 2011).

Considering speech errors, much has been written about the possible application of distinctive features in the analysis of several phonetic/phonemic errors in children and/or adults (see, e.g. Pollack and Rees 1972 or Parker 1976 for further details). Diverse research results indicate not only the feasible role of features in speech errors mechanism (e.g. Eulenberg and Farhad 2011), but also possible limitations of such an approach (e.g. Pollack and Rees 1972). Victoria Fromkin – a leading personality in the field of psycholinguistic approach to speech errors – in the analysis and interpretation of the data from the corpus of more than 600 speech errors collected by herself (Fromkin 1973) indicates that “units smaller than the segments, i. e. properties or features of speech sounds, are ‘perceived’ and confused” (ibid. 223). Consider, for example, the pronunciation *mang the mail* instead of the intended *bang the nail* in (1):

(1) bang the nail /bæŋ ðə neɪl/ → mang the mail /mæŋ ðə meɪl/

The only difference between /b/ and /m/ in (1) is nasality: /b/ is [-nasal], /m/ is [+nasal], “all other features remaining the same” (ibid.: 264). In this case, the speech error is based on the substitution of a single feature (ibid.). Although in her analysis, there are more examples of this type, Fromkin states that “[i]t is certainly true that errors which involve a substitution of features are rare, compared to errors involving larger units.” (ibid.: 225) and indicates that “[...] we are unable to conclude that individual features are independently involved in all cases” (ibid.: 225). Taking the complexity of the relationship between phonetic properties of sounds and their distinctive features into account, the author suggests that certain distinctive features strongly depend on the existence of other characteristics of the sound. For example, “a simple switch of the feature [+coronal], [...], without other phonetic effects” (ibid.: 227) is not possible. Thus, despite several examples showing the change of one feature only, Fromkin seemed to be sceptical about the involvement of individual features in various

speech errors and preferred changes at the level of a segment (see Fromkin 1973 for more details).

Quotations above imply two reasons why it is not possible to admit that not only a phoneme but also a distinctive feature can be the object of misapplication in speech: (1) negligible occurrence of such cases and (2) mutual dependency of features.

Levelt (1993) in his popular step-by-step account of the entire process of speech production too casts doubt on the possibility to misuse a distinctive feature of a phoneme and not the whole phoneme while speaking stating that „[t]he only really clear cases of feature replacements, i.e., where new phonemes are formed, occur with negligible frequency; only a few cases have been reported in the literature” (ibid.: 341). However, a few pages later, when depicting segmental errors, the author gives evidence about “the feature-similarity constraint in experimentally elicited errors” (ibid.: 350) meaning that, as to segmental errors, the intended phoneme and the pronounced phoneme “tend to be similar in distinctive feature composition” (ibid.).

Consequently, thinking of speech errors and phonemes and their distinctive features, two questions arise:

(1) Are distinctive features of phonemes independent or is there a kind of mutual dependency between at least some of them as mentioned by Fromkin (1973)?

(2) Do feature-replacement based speech errors have so trivial frequency of occurrence that one can hardly accept the existence of speech errors resulting from the wrong activation of distinctive feature only as stated by Fromkin (1973) and Levelt (1993) two decades after?

We will try to find answers to those questions in the following lines of our paper. In section 2, we will concisely present the relevant facts from the original distinctive feature theory as introduced by Jakobson to answer the question (1). This is followed by a more modern version of this theory (Sabol 1989), which provides the distinctive feature delimitation of English consonant phonemes¹ required for our analysis. Section 3 briefly introduces speech errors. The results of our analysis can be found in 4.1. and part 5 offers concluding remarks.

2. Distinctive features – autonomous or dependent?

The answer to the question (1) depends on the theory chosen. In other words, the acceptance of the autonomous or dependent character of distinctive features is given by the extent to which the applied distinctive feature theory aligns with the original Jakobson’s view of features as autonomous units in (2):

¹ This is the first step of our research in which we concentrate on consonant speech errors only.

(2) Despite their multiform interdependence within the phoneme and within the entire phonemic pattern, the different distinctive features remain autonomous. Not only may any feature perform its distinctive function (/gip/ ≠ /gib/ ≠ /gid/), but the identification of a single feature regardless of the different phonemes in which it occurs is seen to play a significant part in language (Jakobson et al. 1963: 40-41).

It is obvious that in (2) Jakobson emphasises the autonomous nature of a distinctive feature and its autonomous ability to apply its distinctive power, contrary to some different distinctive feature theories which are based on mutual dependency of features (see, e.g. Sagey 1986 for different approach or Mannel 2008 for more details about various feature theories). Of course, it is true that some features are mutually dependent in a way that if a phoneme A has the feature X then it definitely has the feature Y. For example, to be sonorant means to be voiced. That is why the feature voicing is excluded from the specification of sonorants, because it is incorporated in the feature sonorant itself (Sabol 1989: 97). Therefore, the opposition based on its presence or absence (binary principle) is not applied in the given category of phonemes (see Sabol 1989 or section 2.2 below).

The subsection 2.1 offers the feature specification of English consonant phonemes in terms of original Jakobson's theory.

2.1. Jakobson's distinctive feature theory and specification of English consonant phonemes

Considering standard accent of British English, there are 24 consonant phonemes /p, b, m, w, f, v, ð, θ, t, d, n, l, s, z, ʃ, ʒ, ʒ, ʒ, dʒ, r, j, k, g, ŋ, h/ (see, e.g. Giegerich 1992; Roach 2000). As is well-known, Jakobson (1956, 1963) defines distinctive features in relation to production (articulation) and perception (acoustic and auditory aspect) of speech sounds. Distinctive features are either inherent, organized simultaneously, or prosodic features that operate on the axis of succession. The inherent distinctive features create twelve binary oppositions: vocalic/non-vocalic, consonantal/non-consonantal, interrupted/continuant, checked/unchecked, strident/mellow, voiced/unvoiced, compact/diffuse, grave/acute, flat/plain, sharp/plain, tense/lax, and nasal/oral.² These oppositions occur in a number of languages. However, no language has all features identified (Jakobson et al. 1963: 40). The authors emphasize that the occurrence of features is determined by implication, which can result in their universal validity or at least a high degree of their statistical probability. The phoneme pattern of

² Detailed depiction of the individual distinctive features can be found, for example, in Jakobson and Halle (1956) or in Jakobson et al. (1963).

standard British English consonants as specified by Jakobson et al. (1963) is captured in Table 1.

Looking at the inventory of the consonant phonemes in Table 1, we can see that eight distinctive features out of the set of twelve describe the consonant phonemes in British English (Received Pronunciation). In addition, Table 1 shows that two phonemes are omitted, /r/ and /w/. To complete the feature specification of English consonant phonemes in terms of distinctive features that are not hierarchically dependent, we had to rely on a feature theory that treats distinctive features as autonomous units and adheres as closely as possible to Jakobson's original framework. Although it does not represent the most recent theoretical model, Sabol's theory (1989) best meets these criteria and was therefore adopted for the purposes of the present study, as specified below.

Table 1. Phoneme specification of standard British English consonants by Jakobson's distinctive features (based on Jakobson et al. 1963: 43)³

	l	ŋ	ʃ	tʃ	k	ʒ	dʒ	g	m	f	p	v	b	n	s	θ	t	z	ð	d	h	j
vocalic/non-vocalic	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
consonantal/non-consonantal	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-
interrupted/continuant			+	-	-	+	-	-		+	-	+	-		+	+	-	+	+	-		
strident/mellow				+	+		+	-							+	-		+	-			
voiced/unvoiced	+	+	-	-	-	+	+	+	+	-	-	+	-	+	-	-	-	+	+	+	+	+
compact/diffuse		+	+	+	+	+	+		-	-	-	-	-	-	-	-	-	-	-			
grave/acute									+	+	+	+	+	-	-	-	-	-	-			
tense/lax			+	+	+	-	-	-		+	+	-	-		+	+	+	-	-	-	+	-
nasal/oral		+	-	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-			

³ Empty box means that the phoneme is not specified by a particular distinctive feature.

2.2. Sabol's distinctive feature theory and specification of English consonant phonemes

Sabol's theory of distinctive features of phonemes is inspired by the aforementioned influential work of Jakobson and Halle (1956). This means that features are delimited in agreement with articulatory, acoustic and perceptual correlates of phonemes. Nevertheless, keeping the different functions of sounds in mind, Sabol – contrary to the original theory – delimits a different set of features for vowels and a different set for consonants. As to consonants, they too are divided into two categories depending on their acoustic properties: sonorants (their spectral structure is characterized by tone, similar to vowels) and obstruents (sounds whose acoustic structure is characterized by noise) (Sabol 1989: 73-74). The feature specification of English sonorants and obstruents on the basis of Sabol's theory can be found in Table 2 and Table 3, respectively.⁴

Table 2. Feature specification of English sonorants in terms of Sabol's theory

	l	r	w	j	m	n	ŋ
vocalic/non-vocalic	+	+	+	+	+	+	+
consonantal/non-consonantal	+	+	+	+	+	+	+
diffuse/non-diffuse	+	+	+	-	+	+	-
acute/non-acute	+	+	-	+	-	+	-
lateral/non-lateral	+	-	-	-	-	-	-
nasal/oral	-	-	-	-	+	+	+

Table 3. Feature specification of English obstruents in terms of Sabol's theory

	p	b	f	v	ð	θ	t	d	s	z	ʃ	ʒ	tʃ	dʒ	k	g	h
vocalic/non-vocalic	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
consonantal/non-consonantal	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
diffuse/non-diffuse	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-
acute/non-acute	-	-	-	-	-	-	+	+	+	+	+	+	+	+	-	-	-
occlusive/non-occlusive	+	+	-	-	-	-	+	+	-	-	-	-	+	+	+	+	-
sibilant/non-sibilant	-	-	-	-	-	-	-	-	+	+	+	+	+	+	-	-	-
voice/non-voice	-	+	-	+	+	-	-	+	-	+	-	+	-	+	-	+	-

⁴ Exhaustive delimitation of the individual distinctive features can be found in Sabol (1989).

Having specified all English consonant phonemes as bundles of simultaneously organized independent distinctive features, we can proceed to the core of our research – the analysis of speech errors.

3. Speech errors

Speech errors, the discrepancies between what we plan to say and what we say, also known as slips of the tongue, “reflect something important about how speech is planned and produced” (Cowles 2011: 162). Speech errors are not random and occur in systematic patterns and that is why psycholinguists use them to learn more about the speech production process (Traxler 2012: 43).

For the purposes of classifying speech errors, the present study adopts the framework proposed by Harley (2006: 739), who distinguishes two criteria. The first criterion concerns the linguistic unit affected by the error. As illustrated in (3), the unit involved may be a word (3a), a morpheme or phoneme (3b), or a distinctive feature (3c).

- (3) a. pass the **pepper** → pass the salt
- b. flock of **b**ats → **b**lock of **f**lats
- c. turn the knob /nɒp/ → /nɒb/ (ibid.: 739)

The type of error mechanism is the second criterion. As shown in (4), the principal error mechanisms include anticipation (4a), exchange (4b), and substitution (4c).

- (4) a. the **m**irst of May → the **f**irst of May
- b. Guess whose **m**ind came to **n**ame? → Guess whose **n**ame came to **m**ind?
- c. Get me a **f**ork → Get me a **k**nife (ibid.)

Additional mechanisms documented in the literature comprise perseveration, in which an earlier element is repeated in a later position, and blends, in which two competing forms merge into a single erroneous output (see, e.g. Fromkin 1973 for detailed account).

It is generally accepted that speech errors occur not only in the utterances of individuals with various brain damage or hearing impairment, but are also frequently observed in speech of speakers with normal hearing and without any brain injury. Their causes are varied and may include nervousness, tiredness, lack of concentration or, simply, discrepancies between speech and thinking (see, e.g. Harley 2006 for more details).

4. Research material

The material point of departure for our analysis was the corpus of 100 speech errors gained from the website Speech Error Database provided by Max Planck Institute for Psycholinguistics (https://www.mpi.nl/dbmpi/sedb/sperco_form4.pl). This database, officially called *The Fromkin Speech Error Database*, represents a significant resource for psycholinguistic research, comprising speech error data collected over several decades. The dataset was originally digitised at the University of California, Los Angeles (UCLA). Following Professor Fromkin's death in January 2000, the database underwent conversion to XML format as part of a collaborative preservation initiative sponsored by the Max Planck Society. The database encompasses nine corpora drawing on data from four languages – English, French, German, and Italian – and provides researchers with the ability to filter errors according to multiple linguistic parameters, including error type, process type, prosodic domain, syllabic position, and stress domain. This multi-dimensional filtering functionality makes the database a particularly versatile instrument for psycholinguistic inquiry (<https://www.mpi.nl/page/mpi-archives-databases>). For the purposes of the present study, the corpus UCLASEC.ENGLISH.FROMKIN was selected without sub-corpus specification, with the analysis restricted to phonological errors only.⁵ The search parameters applied are illustrated in Figure 1.

Table 1: Impoliteness type (Feng and Ren 2020)

⁵ It should be noted that in the present study, the speech error data drawn from the Fromkin database reflect productions by native speakers of English. English functions here as the speakers' first language (L1), which is a relevant factor to bear in mind when interpreting the results and considering their potential cross-linguistic implications.

4.1. Analysis

Set parameters provided us with more than 1000 various phonological speech errors. Keeping in mind the aim of our research, we concentrated on those phonological speech errors that are further classified as “*segment substitution*”. For example, error N. 328, target utterance: *and yeah it looks really good*, error utterance: *and yeah it **Rooks** looks really good* is accompanied with the information “target unit [l-], error unit [r-]”, so it has been included in our analysis since it is the error affecting one segment only. However, error N. 321, target utterance: *and now that's I'm finishing I want to thank all the speakers*, error utterance: *....I want to SPANK all the THINKers* with the additional information “target unit: #[T-] #[sp-], error unit: #[sp-] #[T-...-n-] “ was not the part of our analysis. Obviously, it is an error touching more than one segment.

Consequently, we have excerpted the sample of 100 speech errors based on the substitution of one segment and analysed them in terms of the distinctive features of consonant phonemes involved to find out the frequency of the occurrence of errors affecting distinctive features only.

4.2. Results

In our sample, 100 tokens of one-segment speech errors represent 64 distinct error types. Of these, 50 types (78%) involve speech errors in which the phonemes participating in the given phonological error mechanism differ in two or more distinctive features, as can be seen in (5)-(7) with target utterance on the left side and the error utterance on the right side.

(5) j – l (distinctive features diffuse, lateral)

he looks twenty years younger → *he looks twenty years lunger*
/ 'hi: 'ləks 'twenti 'jiəz 'jʌŋgə/ / 'hi: 'ləks 'twenti 'jiəz 'lʌŋgə/

(6) s – d (distinctive features occlusive, sibilant, voice)

to smoke such large cigars → *to smoke such large digarz*
/ tʊ 'sməʊk sʌtʃ 'lɑ:dʒ sɪ'gɑ:z/ / tʊ 'sməʊk sʌtʃ 'lɑ:dʒ dɪ'gɑ:z/

(7) dʒ – p/p – dʒ (distinctive features diffuse, acute, sibilant, voice)

jiggery pokery → *piggery jokery*
/dʒɪgəri 'pəʊkəri/ /pɪgəri 'dʒəʊkəri/

14 types (specified in Table 4), that is 22%, are based on the substitution of a single distinctive feature only. This substitution results in the replacement of one phoneme, possessing the given feature, with another phoneme characterised

by the absence of that distinctive feature, as illustrated by examples (8)-(12), or vice versa, as exemplified by (13) and (14).

(8) g – k (distinctive feature voice)

between the gloria and the credo → *between the cloria and the credo*
/bɪ'twi:n ðə 'glɔ:riə ən ðə 'kri:dəʊ/ → /bɪ'twi:n ðə 'klɔ:riə ən ðə 'kri:dəʊ/

(9) tʃ – ʃ (distinctive feature occlusive)

if there is a Swiss watch company → *if there is a Swiss wash company*
/ɪf ðə z ə swɪs 'wɒtʃ 'kʌmpəni/ → /ɪf ðə z ə swɪs 'wɒʃ 'kʌmpəni/

(10) l – r (distinctive feature lateral)

problems that are plaguing other schools → *problems that are*
praguig other schools
/'prɒbləmz ðæt ə 'pleɪɪŋ 'ʌðə sku:lz/ → /'prɒbləmz ðæt ə 'preɪɪŋ
'ʌðə sku:lz/

(11) p – k (distinctive feature diffuse)

beatrice potter country → *beatrice kotter country*
/'bi:tri:ks 'pɒtə 'kʌntri/ → /'bi:tri:ks 'kɒtə 'kʌntri/

(12) n – m (distinctive feature acute)

things that are seen as similar → *things that are seem as similar*
/'θɪŋz ðæt ə 'si:n əz 'sɪmələ/ → /'θɪŋz ðæt ə 'si:m əz 'sɪmələ/

(13) m – w/w – m (distinctive feature nasal)

that Edwards was mobbed → *that Edwards was wobbled*
/ðæt 'edwədz wəz mɒbd/ → /ðæt 'edwədz məz wɒbd/

(14) f – v (distinctive feature voice)

must be defined → *must be devined*
/məst bɪ dɪ'faɪnd/ → /məst bɪ dɪ'vaɪnd/

Those 14 types (22%) of segment substitutions between segments differing by only one distinctive feature (cf. Tables 2, 3, and 4) represent 49 individual instances in our sample, accounting for 49% of all cases.

Fromkin (1973) and Levelt (1993), when refuting the possibility that speech errors may affect only the distinctive features of phonemes rather than whole phonemes, argue, as quoted above, that cases of feature replacement are very infrequent. Nevertheless, as our analysis demonstrates, feature speech errors account for 49% of individual occurrences and 22% when error types are taken into account. Although the present sample is limited in size, these findings

Table 4. Types of distinctive feature substitutions

substituted segments	number of occurrences	distinctive feature differing them
r – l	11x	lateral
s – ʃ	6x	diffuse
p – k/k – p	5x	diffuse
t – p/p – t	5x	acute
f – p	4x	occlusive
m – w/ w – m	4x	nasal
f – v/v – f	3x	voice
n – m/m – n	3x	acute
z – d	2x	occlusive
t – d	2x	voice
k – g	1x	voice
ʧ – ʃ	1x	occlusive
b – d	1x	acute
r – j	1x	diffuse

provide preliminary evidence confirming the existence of speech errors that affect distinctive features only, suggesting that such errors, however infrequent, constitute a genuine phenomenon deserving of further investigation.

Figure 2 shows proportions of individual distinctive features in our sample of distinctive feature substitutions.

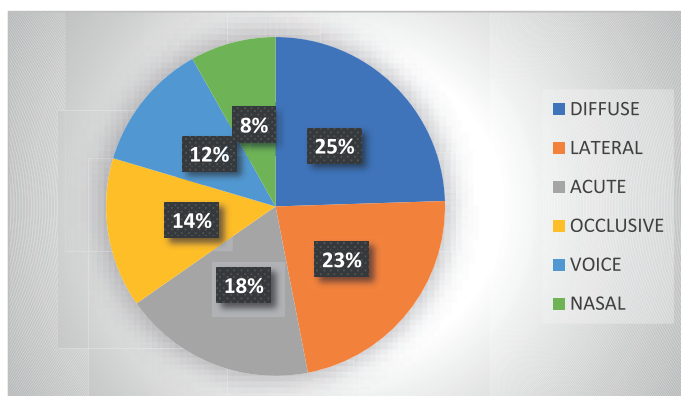


Figure 1. Specification of Parameters in Speech Error Database
(https://www.mpi.nl/dbmpi/sedb/sperco_form4.pl)

The results of segment substitutions between phonemes differing by only one distinctive feature, presented in Figure 2, indicate that the feature *diffuse* occurs most frequently in our sample. This is not in line with our initial expectation that the feature *voice* represents one of the fundamental distinctions between consonant sounds across languages. The distinctive feature *lateral* ranks as the second most frequent. In speech production, it is typically the feature lateral that is “wrongly activated” (using Levelt’s terminology). This feature differentiates exclusively between two phonemes, /r/ and /l/. It appears that the lateral – non-lateral opposition presents difficulties not only in English but, as our preliminary results based on the analysis of speech errors in Slovak suggest, in other languages as well. The remaining distinctive features – *acute*, *occlusive*, *voice*, and *nasal* – occur with gradually decreasing frequency. The size of our sample may be considered a limitation with respect to the interpretation of these results. Accordingly, these observations may be elaborated upon in greater detail in follow-up research.

5. Conclusions

With regard to speech errors and their possible distinctive-feature nature, the literature offers various standpoints. Some authors (e.g. Frisch and Wright 2002 or Eulenberg and Farhad 2011) acknowledge the role of distinctive features in speech errors. Another group of scholars (e.g. Fromkin 1973 or Levelt 1993) maintains that a speech error cannot be based on the change of a single feature only. Two arguments are put forward in support of this view: (1) the mutual dependency of distinctive features, and (2) the very rare occurrence of such changes. We demonstrate that, concerning “feature dependency”, this relies on the distinctive feature theory adopted. The original distinctive feature theory proposed by Jakobson (Jakobson and Halle 1956), as well as the theories of their followers (Sabol 1989), treats distinctive features as autonomous, independent elements participating in the formation of a phoneme. As for the second argument against feature-based speech errors – (2) the alleged rare occurrence – our preliminary results based on the analysis of 100 consonant speech errors have shown that nearly half of the speech errors in our sample are based solely on the substitution of a single distinctive feature. This finding provides empirical support to the existence of feature-based speech errors as a genuine (psycho)linguistic phenomenon. Our results are thus consistent with Frisch and Wright’s (2002: 140) observation that the misordered portions in speech errors correspond to abstract linguistic units such as onsets, codas, phonemes, segments, and features. This claim, in light of our data, we take as further confirmation that distinctive features may indeed function as independent units sensitive to error in speech production.

Finally, future research, as indicated above, will need to extend the sample of speech errors analysed, incorporate vowel changes into the analysis, and include a greater number of languages in order to obtain data of potential cross-linguistic value.

Acknowledgements

This work was supported by the Scientific Grant Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic and the Slovak Academy of Sciences under the research project VEGA 1/0226/24.

References

- Alderete, J. 2020. Speech errors and phonological patterns: Insights from psycholinguistic and linguistic theory. *Phonological Studies* 23: 77-91.
- Chomsky, N., and M. Halle 1968. *The Sound Pattern of English*. New York: Harper & Row.
- Cowles, H.W. 2011. *Psycholinguistics 101*. New York: Springer Publishing Company.
- Eulenberg, J., and A. Farhad 2011. *Descriptive phonetics. Distinctive features*. Retrieved from <https://www.slideserve.com/edita/csd-232-descriptive-phonetics-distinctive-features>.
- Frisch, S.A., and R. Wright 2002. The phonetics of phonological speech errors: An acoustic analysis of slips of the tongue. *Journal of Phonetics* 30: 139-162.
- Fromkin, V.A. (eds.) 1973. *Speech Errors as Linguistic Evidence*. The Hague, Paris: Mouton.
- Fromkin, V.A. 1973. Introduction. In V.A. Fromkin (ed.), *Speech Errors as Linguistic Evidence*, 11-45. The Hague, Paris: Mouton.
- Giegerich, H.J. 1992. *English Phonology. An Introduction*. Cambridge: Cambridge University Press.
- Hall, T.A. 2007. Segmental features. In P. de Lacy (ed.), *The Cambridge Handbook of Phonology*, 311-334. Cambridge: Cambridge University Press.
- Harley, T.A. 2006. Speech errors: Psycholinguistic approach. In K. Brown (ed.), *Encyclopaedia of Language and Linguistics* 2nd ed., Vol. 11, 739-745. Oxford: Elsevier <http://dx.doi.org/10.1016/b0-08-044854-2/00797-5>.
- Jakobson, R., and M. Halle 1956. *Fundamentals of Language*. The Hague: Mouton & Co's-Gravenhage.
- Jakobson, R., C. Gunnar, M. Fant and M. Halle 1963. *Preliminaries to Speech Analysis. The Distinctive Features and their Correlates*. 4th ed. Cambridge, Massachusetts: MIT Press.
- Levelt, W.J.M. 1993. *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.

- Mannell, R. 2008. *Distinctive Features*. Macquarie University.
- Parker, F. 1976. Distinctive features in speech pathology: Phonology or phonemics? *Journal of Speech and Hearing Disorders* 41(1): 23-39.
- Pollack, E., and N.S. Rees 1972. Disorders of articulation: Some clinical applications of distinctive feature theory. *Journal of Speech and Hearing Disorders* 37(4): 451-461.
- Roach, P. 2000. *English Phonetics and Phonology. A Practical Course*. 3rd ed. Cambridge: Cambridge University Press.
- Sabol, J. 1989. *Syntetická fonologická teória*. Bratislava: Jazykovedný ústav Ľudovíta Štúra SAV.
- Sagey, E.C. 1986. *The Representation of Features and Relations in Non-Linear Phonology*. Massachusetts Institute of Technology.
- Traxler, M.J. 2012. *Introduction to Psycholinguistics. Understanding Language Science*. Wiley-Blackwell.
- Wilson, R.D. 1966. A criticism of distinctive features. *Journal of Linguistics* 2(2): 195-206.

Online source

https://www.mpi.nl/dbmpi/sedb/sperco_form4.pl

TATIANA SZCZYGLÓWSKA

University of Bielsko-Biala, Poland

Institute of Modern Languages/Faculty of Humanities and Social Sciences

ORCID: 0000-0002-5051-4080

tszczyglowska@ubb.edu.pl

A CORPUS STYLISTIC APPROACH TO ERNEST HEMINGWAY'S FICTION: KEYWORDS AND RECURRENT PHRASES IN *THE SUN ALSO RISES*

Critics and readers alike revere Ernest Hemingway's novel *The Sun Also Rises* for its innovative writing style, characterised by omission and understatement, yet remaining semantically rich. However, the novel's keywords and recurrent phrases, and their respective meanings, have received little attention, probably because they are dispersed across the text. Consequently, it is difficult to analyse their narrative significance in detail within their immediate textual context. Corpus methods can address this issue by systematically identifying such repeated linguistic patterns in the text and revealing the hidden themes they convey. This study adopts this approach, using a corpus methodology to analyse the keywords and most frequent four-word n-grams and p-frames in Hemingway's debut novel, with the aim of demonstrating their contribution to the construction of its literary meanings.

Keywords: *The Sun Also Rises*, corpus stylistics, literary meaning, keywords, n-grams, p-frames

1. Introduction

Published in 1926, *The Sun Also Rises* (henceforth *TSAR*) is the debut novel of the Nobel Prize-winning 20th-century American novelist and short story writer, Ernest Hemingway. Initially deemed "skilfully written" but lacking "artistic significance" (Wagner-Martin 1990: 1), the novel soon gained recognition for its innovative writing style, characterised by lucidity, concision, understatement, and omission, yet infused with symbolism operating behind the curtain of stripped narrative (Svoboda 1983: 4; Wagner-Martin 1990: 2). While both the text and manuscript of the novel are now available to researchers, a substantial

portion of Ernest Hemingway's writing has yet to receive adequate scholarly attention (Rothman 1928: 336). One possible reason for this is that claims about style lack solid empirical support unless they are verified quantitatively (Leech and Short 2007: 57). This study therefore combines literary stylistics and corpus linguistics to analyse keywords and the most frequent four-word phrases in *TSAR*, attempting to identify their salient characteristics within their textual contexts and to reveal their respective interpretive weight, not readily accessible through close reading alone.

Previous research shows that the analysis of literary texts can be refined by examining their lexical and structural regularities. Biber (2011: 21) notes that researching literature through a corpus linguistics lens is not a new phenomenon, but that it has gained markedly in popularity in recent years and is now commonly conducted within the framework of corpus stylistics. Overviews of such analyses, as found in the works of Fischer-Starcke (2010), Mahlberg (2013) and Balossi (2014), demonstrate that corpus stylistic research in literature takes many forms, with scholars using corpus methods to analyse authorial style and authorship attribution, examine the style and discourse presentation of literary texts, and inform literary interpretation. While research aims are often pursued by focusing on frequency-based wordlists and other types of descriptive statistics, corpus methodologies for stylistic analysis also include "keyword analysis, identifying typical extended lexical phrases, and collocational analysis" (Biber 2011: 16). The latter allow to empirically identify linguistic patterns that are distinctive of the target text and divergent from those that are typical of the reference corpus. These patterns can then be interpreted for their stylistic and literary effects.

1.1. Keywords

The keywords of a text or corpus are items that occur statistically more frequently, in the case of so-called positive keywords, or less frequently, in the case of so-called negative keywords, in the text under analysis than in the reference corpus (Scott and Tribble 2006: 55). They are therefore characteristic of the text in question. Shepherd and Berber Sardinha (2013: 72) argue that keywords are crucial for interpreting literary meaning, since their unusual frequency in a text "is a useful indicator of style [...] as keywords may be indicative features of the study corpus, such as content lexical items, which in the case of narrative, may be plot explanatory". Distinctive keywords can serve specific literary purposes (Leech and Short 2007: 40), so analysing them can help to trace a text's semantic network and forms of expression. This has been demonstrated, for example, by Scott and Tribble (2006), who identify salient lexical features of *Romeo and Juliet* against Shakespeare's other plays; Culpeper (2009), who uses this method to show how character-specific keywords, such as Juliet's frequent

use of *if*, reflect underlying psychological and thematic concerns; and Fisher-Starcke (2010), who compares the keywords of *Northanger Abbey* with both Austen's other works and contemporary fiction to isolate text-specific lexis.

1.2. Extended lexical phrases

When extended lexical phrases are approached from a frequency-based perspective, they are often operationalised as n-grams – uninterrupted sequences of *n* word-forms (e.g. *and have a drink*) – and p-frames, which are n-grams with one variable lexical slot (e.g. *went out * the*) (Stubbs 2007: 166). Shepherd and Berber Sardinha (2013: 72, 84) argue that the extraction of n-grams from a corpus can reveal its phraseological or terminological characteristics, thereby uncovering either the structural patterns of a literary text or its thematic content. Additionally, focusing on the p-frames repository provides insight into the degree of phraseological fixedness and variability within the text. If such corpus-stylistic analyses of a literary text go beyond identifying its most frequent items and also consider their typical co-occurrence preferences, reflecting the collocational and colligational patterns exhibited by these items, the findings may shed light on the underlying structural and conceptual framework on which the text is based. Recurrent phraseological units in literary texts have been the subject of research in analyses such as Stubbs' (2005) study of Conrad's *Heart of Darkness*, Starcke's (2006) examination of Austen's *Persuasion*, Mahlberg's (2007, 2013) studies of Dicken's novels, and Fisher-Starcke's (2009, 2010) analyses of Jane Austen's novels.

1.3. Additional insights from corpus approaches to the study of literary texts

The applicability of corpus linguistics to the study of literary texts offers a clear advantage over more conventional methods of identifying, describing and interpreting linguistic – including stylistic – features, as it enables the extraction of actual patterns of language use. These patterns can then be analysed both quantitatively and qualitatively, in a manner that is more objective than traditional intuition-based approaches. This, in turn, allows for fresh insights into the unique textual features that are of literary significance but are difficult to detect through close reading alone. Some of these features result from the author's conscious choices, while others lie beyond their conscious control; together, however, they contribute to the distinctive linguistic imprint of an author's style. Its close analysis becomes particularly challenging in the case of the novel, which, as Short (1996: 255) observes, "is probably the most difficult genre to analyse stylistically", primarily due to its length.

Selected analyses of literary texts with the use of corpus methods have been mentioned in the previous subsections. A more comprehensive coverage of the

current state of corpus approaches to literature can be found, for instance, in Ho (2011: 8-11). Of particular relevance to the present study, which adopts the same methodological approach, is Fisher-Starcke's (2009) corpus-stylistic analysis of keywords and frequent phrases in Jane Austen's *Pride and Prejudice*. Her analysis reveals systematic linguistic patterns in the novel relating to family, emotion, communication, and social norms. These patterns illuminate literary meanings and narrative structures that cannot be easily discerned through traditional, intuitive literary criticism alone. As Shepherd and Berber Sardinha (2013: 85) suggest, "a lexical analysis of a text based on keyness, the extraction of concordancing lines for the keywords and a simple extraction of n-grams, may cast a light on the underlying (and often hidden) themes of a book".

Overall, it seems that an undeniable advantage of the frequency-based approach to literature is that it enables the identification and examination of salient lexis in a text in an objective and systematic way, within their linguistic context. The lexico-grammatical patterns in which these prominent words and phrases repeatedly occur provide information on the meaning relationships between the different aspects of the fictional world to which these lexical items refer. As Hunston and Francis (2000: 37) emphasise, the meaning of a word is determined by the words and structures with which it frequently co-occurs. Therefore, it can be assumed that the linguistic features revealed through corpus methods, owing to their recurrent nature, configure the thematic scaffolding of the narrative. Indeed, Shepherd and Berber Sardinha (2013: 71) note that the frequency of a lexical item is essential for understanding its meaning and for making wider assertions about the discourse. Thus, taking a corpus perspective to observe, describe and interpret statistically significant linguistic patterns in a literary text may help to uncover the subtle thematic threads present within it but operating below the narrative surface. This is particularly relevant in the case of *TSAR*'s famously sparse prose, characterised by Hemingway's signature 'iceberg' style of omission (Balassi 1990: 125) and "repetitive patterns" that "cohere throughout his text [...] producing meaning" (Elbaz 1984: 77).

1.4. Research questions

Taking the above into account, this study uses a corpus methodology to examine the literary text of *TSAR*, addressing the following research questions:

- 1) What literary meanings are encoded by the keywords and the linguistic patterns in which they occur throughout the novel?
- 2) What literary meanings are encoded by the novel's most frequent four-word phrases (i.e. 4-grams and 4-frames) and the linguistic patterns in which they occur?

2. The study

2.1. Materials

This study draws on Ernest Hemingway's novel *TSAR*, which forms the basis of the study, as well as two reference corpora: Hemingway's other six novels (the corpus *Hemingway*), and literary texts from roughly the same period (the corpus *ContempLit*), as summarised in Table 1. Standardised type/token ratios (i.e. STTR) are provided for comparing lexical richness across corpora, with *ContempLit* showing the lowest value. The *ContempLit* texts are listed in the Appendix.

TSAR is Hemingway's debut novel. Set in post-First World War Europe, it chronicles the lives of a group of British and American expatriates. The narrator and protagonist, Jake Barnes, is a war veteran and journalist who has been left impotent by his injuries. He is in love with Lady Brett Ashely, a woman he cannot truly have. Along with their hard-drinking, cynical friends, the couple moves between the nightlife of Paris and the bullfights of Pamplona in Spain. Jealousy, frustrated passions, and disillusionment with life after the war are prevalent themes. The novel portrays the vitality and the emptiness of the so-called 'Lost Generation' who are searching for meaning and identity in a world marked by trauma, loss, and disconnection.

The novel exemplifies modernist aesthetics, characterised by sparse prose, realism and dispassionate understatement. The narrative omits many details concerning the plot, setting and the characters' emotions, attributes and

Table 1. Corpora

Corpus	Content	Size (tokens)	STTR
<i>TSAR</i>	<i>The Sun Also Rises</i> (1926)	67,775	60.97
<i>Hemingway</i>	<i>A Farewell to Arms</i> (1929), <i>To Have and Have Not</i> (1937), <i>For Whom the Bell Tolls</i> (1940), <i>Across the River and into the Trees</i> (1950), <i>Islands in the Stream</i> (1970), ¹ <i>The Garden of Eden</i> (1986) ²	606,459	63.48
<i>ContempLit</i>	30 novels by various authors, published between 1911 and 1969 to be roughly contemporary with Hemingway's writing career	3,007,506	57.49

¹ The novel was written in 1950-1951 and published posthumously.

² The novel was written in 1946-1961 and published posthumously.

thoughts,³ encouraging readers to discover the meaning for themselves. It can also be described as a roman à clef, as Hemingway based its characters on real people from his social circle and mirrored events from his own life (Svoboda 1983: 9). *TSAR* was well received by readers and quickly became a commercial success, gradually establishing itself as Hemingway's most renowned book. Wagner-Martin (1990: 1) considers this unusual, since "most writer's first novels do not turn out to be their most important work". Using corpus-linguistic methods to delve deeper into the novel can reveal new literary meanings, particularly those that are difficult to detect with the naked eye. This is likely to be the case since, as Balassi (1990: 125) notes, "nowhere did Hemingway so thoroughly or so daringly" leave so much meaning below the surface, producing "a richly textured but elusive story that has been plagued by misreadings since its publication" (Balassi 1990: 125).

2.2. Method

This study examines the linguistic units of *TSAR*, which are identified on two hierarchical levels in language: lexis and phraseology. The lexical analysis was performed using *WordSmith Tools* v. 6. (Scott 2012) and involved examining keywords with a minimum frequency threshold of 25 that were calculated using the log-likelihood test with a probability value of 0.000001. Adopting the approach of Fisher-Starcke (2009), who studied the keywords and frequent phrases in Austen's *Pride and Prejudice*, *TSAR* was compared with two reference corpora: *Hemingway* and *ContempLit*. This revealed lexical features that are specific to the novel, as opposed to Hemingway's other books and literary works from the same period. As Baker (2004: 347) states, such quantitatively detected keywords "direct the researcher to important concepts in a text". The two lists of keywords were then compared to identify coincident semantic and functional sets, revealing the novel's "aboutness" and style (Scott and Tribble 2006: 59-60). This also helped to select the most relevant keywords for further analysis. Fisher-Starcke (2009: 501) explains that the relevance of keywords for analysis increases "when the same words are also identified by a second keyword analysis which uses another reference corpus".

The phraseological analysis was conducted using *kfNgram* (Fletcher 2012) to identify the novel's most frequent four-word sequences. This involved examining uninterrupted n-grams and their variants with one variable slot, known as p-frames. To yield a manageable number of items of each type for further

³ The writing technique of presenting only a small part of a story's deeper meaning, leaving the rest beneath the surface, is known as the "iceberg theory of omission" (Trodd 2007: 17). It was developed by Hemingway, who, as a journalist, opted for a minimalist and concise writing style due to the limited space available in newspapers.

analysis, the frequency threshold was set at six for n-grams and at four for p-frames. Four-word⁴ phrases were chosen because they are “frequent enough for good diversity” (Bestgen 2018: 206), exhibit greater phrasal, structural and functional distinctiveness than shorter units (Biber et al. 2021: 984; Goźdz-Roszkowski 2011: 110), and are particularly suitable for identifying collocational and colligational patterns (Fischer-Starcke 2010: 111), which typically span four words to the left and right of a node (Sinclair 1991). Examining the concordance lines of selected single- and multi-word units using *WordSmith Tools* provided additional insight into the important literary meanings expressed by keywords in the novel’s textual content and indicated by phrases in its textual structure.

3. Keywords

Keywords provide information about how unique the text or corpus is in terms of content, style, and structural features. Organising them into semantically and functionally related groups helps to reveal the propositional content of a text, and thus the literary meanings they encode. Examining keywords in their linguistic context further supports this intuitive process, as shown in concordance lines that highlight the items’ semantic and grammatical patterns of co-selection.

Comparing *TSAR* with *Hemingway* produced 100 keywords (85 positive) for the novel, whereas comparing *TSAR* with *ContempLit* produced 142 keywords (128 positive). Excluding proper names, the first comparison yields 71 keywords, while the second yields 120. Of the 191 positive keywords in total, 60 overlap, 60 result exclusively from comparing *TSAR* with *ContempLit*, and 11 result from comparing *TSAR* with *Hemingway*. When the raw frequencies of the keywords in both lists are compared, it emerges that the comparison with *ContempLit* generates only 1.57 times as many items as the comparison with *Hemingway* (27,245 vs. 17,256). These results can be compared with those of Fisher-Starcke (2009: 496-502), who examined the keywords in Austen’s *Pride and Prejudice*. When Austen’s other novels were used as a reference corpus, only 30 non-proper-name keywords were produced. However, when the book was compared to Austen’s contemporary literature, as many as 198 keywords were produced. In terms of raw frequency, these were nearly seven times more numerous than those resulting from the first comparison. It was concluded that *Pride and Prejudice* shares lexis and topics with Austen’s other novels, but differs significantly from its contemporary literature. In light of this, the findings of the present study suggest that *TSAR* and Hemingway’s other books are not particularly similar in terms of

⁴ Following Biber et al. (2021: 982), contracted forms were considered as single orthographic words.

topics and its respective lexis. While *TSAR* differs more sharply from its contemporary literature than from Hemingway's other works in terms of lexis, the difference is not that great.

The keywords revealed by the two comparisons are listed in Table 2. They are categorised into semantic and functional sets, which were developed intuitively based on their dominant meanings in the novel and verified through concordance-based analysis. For example, of the 28 occurrences of *dear*, the word is used as a term of address to refer to a character 24 times, and as an adjective only four times. Therefore, it was classified as a word referring to people. The middle column shows keywords that appear in both lists, sorted in decreasing order. The columns on the left and right contain keywords belonging to a given category that occur only in the list generated by comparing *TSAR* with *Hemingway* and *ContempLit*, respectively. Contractions were treated as single orthographic words, the role of which was determined by the initial element. Words marked with an asterisk appear twice in the Table because they are semantically related to different keyword categories. The numbers in brackets show how many items in a given keyword category were identified in each comparison, including overlapping items and items resulting from one comparison only.

Table 2. Keywords in *TSAR* identified by comparisons with *Hemingway* and *ContempLit*

<i>TSAR</i> vs. <i>Hemingway</i>	Overlapping keywords	<i>TSAR</i> vs. <i>ContempLit</i>
(1) <i>my</i>	personal pronouns (9) <i>I, I've, I'm, you're, he's, we, us, they're, it's</i>	(3) <i>you, me, it</i>
(1) <i>say</i>	forms of content verbs that describe the actions or events in the novel (8) <i>went, got, walked, started, came, paid, sat, stopped</i>	(13) <i>said, asked, go, get, put, come, sitting, looked, smiled, passed, know, took, eat*</i>
	referring to elements of the urban landscape (7) <i>café*, hotel, street, bus, square, taxi, town</i>	(4) <i>bar*, bridge, train, trees</i>
(3) <i>that's, around, what's</i>	functors (7) <i>a, the, out, down, up, on, all</i>	(7) <i>but, through, of, to, away, why, ahead</i>
(2) <i>lady, dear</i>	referring to (groups of) people (7) <i>crowd, count, bull-fighter*, friends, people, English, waiter*</i>	(1) <i>darling</i>
	related to bullfighting (6) <i>bull, bulls, bull-fight, bull-fighter*, ring, fiesta</i>	

Table 2. cont.

<i>TSAR</i> vs. <i>Hemingway</i>	Overlapping keywords	<i>TSAR</i> vs. <i>ContempLit</i>
(2) <i>didn't, doesn't</i>	inflectional forms of auxiliary verbs (4) <i>don't, was, were, isn't</i>	(3) <i>did, do, is</i>
	signalling intentions or desires (3) <i>let's, wanted, going</i>	(1) <i>want</i>
	related to visiting cafés and similar venues (3) <i>café*, table, tables, waiter*</i>	(8) <i>drink, wine, bottle, lunch, glasses, coffee, eat*, bar*</i>
	negatively connotated words that describe the characters' impressions of various aspects of the novel's plot (2) <i>damned, drunk</i>	(2) <i>hell, tight</i>
	interjections used to express greetings or emotions (2) <i>hello, oh</i>	
(2) <i>couldn't, wouldn't</i>	modal verbs (1) <i>can't</i>	
	positively connotated words that describe the characters' impressions of various aspects of the novel's plot (1) <i>nice</i>	(4) <i>good, fine, lovely, wonderful</i>
	intensifiers (1) <i>rather</i>	(1) <i>very</i>
	other (1) <i>hot</i>	(11) <i>white, thing, things, place, matter, open, Spanish, outside, yes, road, country</i>
	time references (0)	(4) <i>time, now, morning, years</i>

Given that “only words, and therefore topics, that are not dominant in the reference corpus are identified as key for the data analysed” (Fisher-Starcke 2009: 500), the following conclusions can be drawn about *TSAR* based on the data presented in Table 2:

- Nearly all of the keyword categories are distinctive for *TSAR* when compared with both *Hemingway* and *ContempLit*. This indicates that these topics and their respective vocabulary are particularly prevalent in the novel.

- The keywords include content-bearing terms (i.e. mainly lexical words such as those referring to people or to elements of the urban landscape) and style markers (i.e. mainly function words such as pronouns, functors or auxiliary verbs), offering insight into the novel's themes and Hemingway's linguistic choices.
- Time references are important in *TSAR* and common also in Hemingway's other novels, as indicated by their absence in the list of keywords resulting from a comparison of *TSAR* with *Hemingway*.
- The most prominent keyword categories reflecting the novel's important themes are those referring to people, bullfighting, the urban environment, and social spaces such as cafés.
- *TSAR* is strongly focused on human characters, as indicated by the occurrence of personal pronouns, words referring to people and denoting their actions (e.g. *paid, smiled, know*).
- The frequent occurrence of first and second person pronouns, along with verbs like *said* and *asked*, highlights extensive dialogues and the importance of communication.
- Keywords related to negation, intentions, and desires suggest a thematic emphasis on challenged assumptions or expectations, as well as on characters' motivations and aims.

3.1. Keywords related to bullfighting

The bullfighting keyword category was selected for further analysis because all of its terms appear as keywords in comparisons of *TSAR* with both reference corpora, confirming bullfighting as a significant theme of the novel. While not the most dominant category in terms of raw frequency, it has the highest overall keyness among content-bearing keyword categories in both comparisons. Table 3 shows the frequencies of occurrence and keyness values of the bullfighting keywords, ordered by statistical significance.

The analysis shows that keywords related to bullfighting are concentrated in the second half of the novel. As illustrated in Figure 1, which shows the dispersion plot for the six keywords listed in order of keyness, they become prominent from Chapter 10 onwards. This is where the word *bulls* first appears and becomes prevalent throughout the rest of the text, as do the other analysed terms. This concentration is preceded by an increasing frequency of the word *fiesta*, first introduced in Chapter 8, which marks the build-up to the climax of the novel: the Fiesta of San Fermin in Pamplona. The strong collocation of *fiesta* with *the* in L1 position (38 instances vs. 2 with *a*) indicates a reference to a specific event, underscoring its central role in the novel.

The concordance lines of *fiesta* with its most frequent collocates in L1 (*the*, 38 instances) and R1 (*was*, 8 instances) positions demonstrate a usage pattern that

Table 3. Keywords related to bullfighting

Keyword	Frequency in <i>TSAR</i>	Keyness	
		<i>TSAR</i> vs. <i>Hemingway</i>	<i>TSAR</i> vs. <i>ContempLit</i>
bulls	70	280.99	534.16
bull	105	258.65	463.47
fiesta	46	195.18	341.37
bull-fight	26	119.47	198.39
bull-fighter	26	119.47	189.88
ring	27	45.65	44.80
Total	300	1,019.41	1,772.07

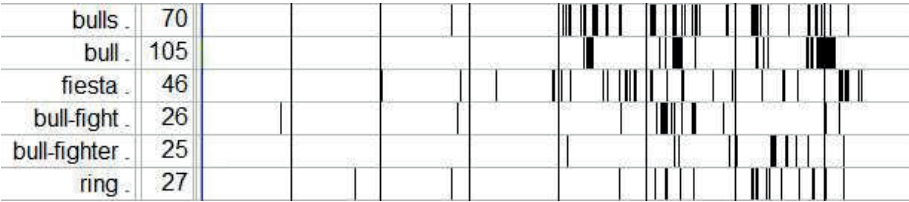


Figure 2. Proportions of distinctive feature substitutions in the sample

seems to align with a fundamental principle of Hemingway’s writing style: the iceberg theory. Examining the first six concordance lines in particular, it is evident that *fiesta* occurs in brief sentences that epitomise the writer’s sparse style, characterised by minimalism, simplicity and concision. As *fiesta* is directly linked to the novel’s climax, it is probably most susceptible to Hemingway’s avoidance of explicit explanations that would reveal its true symbolism. It is the readers who are expected to infer the significance of the fiesta for the story themselves.

t comes after a good bullfight. The fiesta was going on. The drums pounded and paper that stuck to the stone. The fiesta was over. I drank a coffee and after it stop. Outside in the square the fiesta was going on. It did not mean anythin " Cohn said. "Bring them here." The fiesta was really started. It kept up day an In the morning it was all over. The fiesta was finished. I woke about nine o'clock " It was beginning to get dark. The fiesta was going on. I began to feel drunk one of the beds in Cohn's room. The fiesta was going on outside in the night, but was the last day of the fiesta. The fiesta was solid and unbroken, but the motor

Figure 1. Keywords related to bullfighting in dispersion plot

The concordance analysis of *bull-fighter* reveals a usage pattern that is formed by the words found in its immediate co-text. In 14 instances, 11 of which are illustrated below, *bull-fighter* is directly preceded by *the* and most often followed by *chap* (6 instances)⁵ and *fellow* (5 instances) in the R1 position. This indicates that *bull-fighter* is often used attributively.⁶ While both right-hand collocates are informal and somewhat old-fashioned synonyms for *man*, *chap* sounds distinctly British, whereas *fellow*, which is used in both British and American English and carries prestige in professional or academic contexts, would more commonly be expressed as *guy*. Hemingway's alternation between *bull-fighter chap* and *bull-fighter fellow* appears to be a deliberate stylistic choice, intended to reflect the speech of British and American expatriates in the 1920s.

you know. She's gone off with the bull-fighter chap." "No." "Yes. She looked for now. All for forgiveness. And the bull-fighter chap hit him in the face again." e rest," Bill said. "It seems the bull-fighter chap was sitting on the floor. He . Said it would be wicked. So the bull-fighter chap sort of rather staggered over table. "Brett's gone off with the bull-fighter chap. They're on their honeymoon. e went in and found Brett and the bull-fighter chap in the bull-fighter's room, happened finally?" "It seems the bull-fighter fellow was sitting on the bed. He hey weren't having any of it. The bull-fighter fellow was rather good. He didn't nd wanted to shake hands with the bull-fighter fellow. He wanted to shake hands ohn. 'I'd be ashamed to.' "So the bull-fighter fellow hit him just as hard as he aned down to shake hands with the bull-fighter fellow. No hard feelings, you know

Figure 2. Example occurrences of *chap* and *fellow* in *TSAR*

Going beyond the immediate concordance lines of *chap* and *fellow*, it can be seen that British characters such as Mike Cambell or Lady Brett Ashley naturally say *chap*, while American characters such as Bill Gorton or Jake Barnes more often tend to say *fellow*. Evidence for this is provided in Figure 2. Interestingly, in Chapter 17, when Mike is recounting Robert Cohn's fight with Romero to Jake and Bill, who ask questions and encourage him to continue the story, he starts alternating between *chap* and *fellow*. One possible reason for this is that he is addressing his American friends. Alternatively, by using the casual, slightly humorous *chap*, he may be trying to reduce the aura of the bullfighter implied by the somewhat more serious and respectful *fellow*.

The novel contains 300 instances of keywords related to bullfighting, 214 of which are immediately preceded by *the*, most notably *ring* (88.88% of instances, i.e. 24 out of 27) and *fiesta* (82.6% of instances, i.e. 38 out of 46). While *the* is

⁵ In total, there are 7 instances of *chap* in the R1 position of *bull-fighter*, but in one instance, the keyword is preceded by *this*.

⁶ Of the 25 instances of *bull-fighter* in *TSAR*, 12 are attributive.

Mike speaking:

Chapter 15 - "I believe, you know, that she's falling in love with this bull-fighter chap," Mike said.

Chapter 17 - "Oh, go to hell, Cohn," Mike called from the table. "Brett's gone off with the bull-fighter chap. They're on their honeymoon."

Brett speaking:

Chapter 8 - "Chap bought it for me. Don't you like it?"

Brett and Mike speaking:

Chapter 9 - "By the way, where is Bill?" Brett asked.

"He's out at Chantilly dining with some people."

"He's a good chap."

"Splendid chap," said Mike. "He is, you know."

Bill speaking:

Chapter 8 - "You're right. Wasn't me at all. Was another fellow. Think we called him the local Harvard man. Remember him now. Studying music."

Chapter 17 - "They can't insult Mike," Bill said. "Mike is a swell fellow. They can't insult Mike. I won't stand it. Who cares if he is a damn bankrupt?" His voice broke.

Mike recounting the incident to Jake and Bill:

Chapter 17 - "It seems the bull-fighter chap was sitting on the floor. He was waiting to get strength enough to get up and hit Cohn again. Brett wasn't having any shaking hands, and Cohn was crying and telling her how much he loved her, and she was telling him not to be a ruddy ass. Then Cohn leaned down to shake hands with the bull-fighter fellow. No hard feelings, you know. All for forgiveness. And the bull-fighter chap hit him in the face again."

the most frequent word in English (Cappock and Beaver 2015: 378) and also the most frequent word in the novel, its presence in the L1 position in 71.33% of cases in this category appears deliberate. Among the other content-bearing categories, only the urban landscape category has a higher score, with its keywords collocating with *the* in the L1 position in 73.1% of instances (i.e. 424 out of 580). However, it is worth noting that this category consists of 11 keywords, almost twice as many as the bullfighting category. This repeated use of the definite article creates rhythm and precision, reflecting Hemingway's journalistic background, which is believed to have influenced his clear, "to the point writing" style (Keskesay 2021). When writing *the bull*, *the fiesta* or *the ring*, Hemingway invokes the archetypal elements of a culturally significant ritual, thereby grounding the reader in the shared reality of the characters who are direct spectators of the unfolding ritual: the fiesta.

The keywords belonging to the bullfighting category also reflect another defining feature of Hemingway's minimalist prose style: his avoidance of adjectives. He developed this technique while working as a cub reporter for the *Kansas*

City Star (Dewberry 1996: 19). Excluding 12 instances in which the six keywords are preceded by numerals acting as adjectives (e.g. *first, second, one*), only 24 other occurrences of the bullfighting keywords collocate with adjectives in the L1 and/or L2 positions. The concordance lines below provide examples of these rare cases.

en it yet. Tomorrow, if I get a **good bull**, I will try and show it to you." When h
 "I say," Mike said, "they were **fine bulls**, weren't they? Did you see their horn
 eder. So he had two small, **manageable bulls** without much horns, and when he felt th
 lls?" he asked. "Good. They were **nice bulls**." "They're all right"--Montoya shook h
 ps," Mike said. "It was a damned **fine fiesta**." "So long, Mike," Bill said. "I'll s
 hile Bill said: "Well, it was a **swell fiesta**," "Yes," I said, "something doing all
 , and went down-stairs. It was a **good bull-fight**. Bill and I were very excited abo
 . "He nearly killed the **poor, bloody bull-fighter**. Then Cohn wanted to take Brett
 after the cloth." "He's just a **young bull-fighter**. He hasn't learned it yet." "Bu

Interestingly, keywords relating to bullfighting mostly occur in narrative fragments; only 74 (24.66%) instances occur in dialogue in *TSAR*. One possible explanation for this is that Hemingway tied this motif to the narrative voice, rather than to the characters' superficial chatter, in order to emphasise the symbolic importance of bullfighting for the reader. After all, bullfighting is not simply a conversational topic, as such topics tend to lack depth. In line with Hemingway's austere writing style, the novel does not explicitly articulate the true significance of bullfighting, which must instead be decoded from Jake's narration.

3.2. Negative keywords

Analysing negative keywords "which occur significantly less often in the target corpus than in the reference one" (Evison 2010: 128) shows that, when compared with Hemingway's other novels, the top three most significantly underrepresented words in *TSAR* are *it, man* and *they*; and, when compared with his contemporary literature, *who, eyes* and *he*. Interestingly, comparing *TSAR* with each reference corpus yields items that explicitly refer to men: *he, his* and *man*. This is particularly revealing when *TSAR* is compared with Hemingway's other books: words such as *man, he, he's, him* and *his* are almost eleven times less frequent in *TSAR* than in his other works (1,977 vs. 21,047). Even considering that *he's* is among the keywords identified for *TSAR*, when both *Hemingway* and *ContempLit* are used as reference corpora, the novel still seems to focus less on a single male protagonist and more on a social group in which men are portrayed as either impotent (Jake Barnes) or weak (Robert Cohn), while women are depicted as sexually and socially dominant (Lady Brett Ashley).

4. Phrases

Phrases extracted from a text using corpus linguistic methods are word combinations of varying lengths that are identified based on their frequency. As Sinclair (2008: 408) puts it, they are the linguistic places “where structures are engineered to allow meanings to take shape”. When analysed, phrases provide insight into the grammatical, syntactic and lexical patterns of expression that dominate a text, shedding light on stylistic tendencies and recurring themes.

As outlined in Section 2.2., different frequency thresholds were assigned to the two phrase types: six for n-grams and four for p-frames. This resulted in the identification of 27 4-grams and 31 4-frames with 69 variants. It is worth noting that, when the frequency threshold for 4-frames was set above 4, there were too few recurrent units to obtain enough data for further analysis. When the minimum occurrence was set to 5, only three 4-frames were identified, and when it was increased to 7, the *kJNgram* tool returned no results. A similar pattern emerged for 4-grams: 17 units were obtained with the frequency threshold set at 7, but only 8 units were identified when it was increased to 8. For comparison, Fischer-Starcke (2009: 509) set the frequency threshold to 8 for 4-grams 17 for 4-frames in her analysis of frequent phrases in Austen's *Pride and Prejudice*, identifying 21 and 24 units respectively; she explained that this was “a restriction that has been imposed to generate a manageable amount of data”.

In this context, it is also worth comparing the lexical diversity of Hemingway's *TSAR* and Austen's *Pride and Prejudice*, which can be done by calculating the STTR. The value calculated for *TSAR* is 60.97, which is higher than the value of 57.49 calculated for Hemingway's contemporary literature, and much higher than the value of 42.25 calculated for *Pride and Prejudice*.⁷ For comparison, the STTR value for James Joyce's 1922 novel *Ulysses* is 50.91. Green (2017: 294, 296) considers this to be a high value, indicating “Joyce's lexical complexity”.

The above findings lead to the conclusion that Hemingway's *TSAR* is characterised by creative linguistic production. This is reflected in his unique word usage, as indicated by the relatively high STTR, and the low degree of phraseological recurrence reported above. The overall scarcity of recurring multi-word units and the relatively low number of variants for the 4-frames suggest that Hemingway does not rely on conventionalised, prefabricated sequences, preferring individualised, context-dependent phrasing instead. This aligns with the concise, minimalist style often attributed to the writer, who constructs meaning through expressive precision and resists formulaic repetition.

⁷ The STTR for *Pride and Prejudice* was calculated using the Project Gutenberg version of the novel, which contains 6,690 types and 124,757 tokens. This version was downloaded from: <https://www.gutenberg.org/cache/epub/1342/pg1342.txt> (accessed October 2025).

The phrases generated for the novel are listed below, alongside their raw frequencies (RF): Table 4 shows the 4-grams and Table 5 shows the 4-frames. For each 4-frame, Table 5 also lists the number of variants (V), the variant/p-frame ratio (VPR),⁸ and the predictability score (PM). V refers to the number of distinct lexical items that appear in the open slot of a p-frame, reflecting its overall productivity. The VPR, introduced by Römer (2010: 105) to express “the relation of different words that fill the blank (*) slot in a p-frame to the number of p-frame tokens”, shows internal variability in p-frames. The PM, as defined by Gray and Biber (2013: 114), represents “the percentage of all occurrences of a frame in which the most frequent filler occurred”, thereby indicating how predictable the filler word in the open slot is.

Table 4. 4-grams generated for *TSAR*

4-gram	RF	4-gram	RF	4-gram	RF
in front of the	19	and there was a	7	at the edge of	6
the edge of the	15	doesn't make any difference	7	at the far end	6
a hell of a	14	i don't want to	7	don't be an ass	6
do you want to	11	i went to the	7	i did not want	6
came out of the	9	it was a good	7	i would like to	6
went over to the	9	the other side of	7	on the other side	6
did not want to	8	we went into the	7	the centre of the	6
stopped in front of	8	you want me to	7	the door of the	6
all right i said	7	and have a drink	6	the top of the	6

The study identified 58 continuous and discontinuous sequences as *TSAR*'s most frequent phrases. 27 follow the ABCD pattern, 13 follow the *BCD pattern, 9 follow the ABC* pattern, 5 follow the A*CD pattern, and 4 follow the AB*D pattern. Among discontinuous phrases, initial- or final-slot patterns occur far more often than medial ones (22 vs. 9). This indicates that the novel's most frequent 4-frames resemble conversation, in which most frames “represent continuous sequences of fixed elements, with a preceding or following variable slot” (Biber 2009: 294).

In terms of their structural characteristics, the phrases presented in Tables 4 and 5 can be categorised into the three grammatical types proposed by Fuster-

⁸ The VPR is also known as the type-token ratio, which Gray and Biber (2013: 113) define as a measure of “internal fixedness” that “provides a way of measuring the degree to which the slot is relatively fixed or variable”.

Table 5. 4-frames generated for *TSAR*

4-frame	RF	V	VPR	PM	4-frame	RF	V	VPR	PM
the * of the	46	7	0.15	32.60	* one of the	9	2	0.22	55.55
in * of the	23	2	0.08	82.60	* on the bed	9	2	0.22	55.55
* out of the	17	3	0.17	52.94	* a good time	9	2	0.22	55.55
do you want *	15	2	0.13	73.33	* to the hotel	9	2	0.22	55.55
* the edge of	15	3	0.20	40.00	back to the *	9	2	0.22	55.55
* over to the	13	2	0.15	69.23	a bottle of *	9	2	0.22	55.55
went * to the	13	2	0.15	69.23	* came out of	8	2	0.25	50.00
* went into the	11	2	0.18	63.63	went out * the	8	2	0.25	50.00
* went to the	11	2	0.18	63.63	what do you *	8	2	0.25	50.00
i went to *	11	2	0.18	63.63	what's the mat- ter *	8	2	0.25	50.00
i * to the	11	2	0.18	63.63	in the * and	8	2	0.25	50.00
* did not want	11	2	0.18	54.54	and went * the	8	2	0.25	50.00
i did not *	10	2	0.20	60.00	in the square *	8	2	0.25	50.00
* top of the	10	2	0.20	60.00	the * it was	8	2	0.25	50.00
in the morning *	9	2	0.22	55.55	* looked at me	8	2	0.25	50.00
on the * of	9	2	0.22	55.55					

Márquez (2014: 96). Specifically, “NP-based types are noun phrases with or without a post-modifying fragment, e.g. *the * of the, the rear of the, [...]* The PP-based type [...] phrases [...] start with a preposition followed by a noun phrase with or without a post-modifying fragment, e.g. *in the * of, at the time of, [...]* the VP-based types are any phrases which contain a verb, e.g. *will be * for, we reserve the right*”. Table 6 summarises the structural types of the novel’s most frequent phrases.

Most *TSAR* phrases are VP-based, a pattern typical of conversation (Gray and Biber 2013; Biber et al. 2021), reinforced by nine subject-pronoun + VP structures (e.g. *I went to *, we went into*). In the 4-frames, slot-fillers are almost evenly split between function words (52%, 36 items, mainly prepositions and personal pronouns) and content words (47.8%, 33 items, mainly nouns and lexical verbs), further reflecting a conversational style (Biber 2009). The prevalence of nouns and prepositions among the slot-fillers indicates the importance of defining the people, objects, places and concepts that make up the novel’s fictional world, and of relating these elements to one another in terms of space, time, or logic. The strong presence of personal pronouns further emphasises the

Table 6. Structural categories of the most frequent phrases in *TSAR*

Type	4-grams (%)	4-frames (%)
NP-based	22.2	19.3
PP-based	14.8	32.3
VP-based	63.0	45.2
Other	0	3.2

central role of characters in *TSAR*, and of lexical verbs, reflects the need to bring the fictional world to life by generating energy and rhythm. Interestingly, adjectives were not used as slot-fillers and are rare, even in the fixed slots of the novel's recurring phrases. This aligns with Hemingway's austere writing style.

The 4-frames have 69 variants in total, giving an average of 2.22 per unit. Only three of the 4-frames have more than two variants, with *the * of the* being the most diverse unit with 7 different fillers. The low internal variability of the 4-frames is reflected in the VPR values, all of which are less than 0.30, indicating that the units are relatively fixed.⁹ The 4-frames also demonstrate relatively high predictability: 1 item is classified as highly predictable, 19 as predictable, 9 as almost predictable, and 2 as unpredictable; there are no highly unpredictable items.¹⁰ Based on the preceding discussion, it can be concluded that the range of recurring phrases in *TSAR*, particularly the 4-frames, is limited yet lexically and structurally stable. Hemingway tends to reproduce his few preferred multiword patterns with minimal modifications, producing strong phraseological cohesion in the novel. He uses this economical and precise phraseological core to develop the main themes of *TSAR* in a stylistically consistent and rhythmically regular way.

The dominant grammatical and functional patterns that can be identified in the novel's 4-grams and 4-frames are listed in Tables 7 and 8, respectively. Some phrases belong to multiple categories, which were developed intuitively through concordance analysis.

As the tables show, selected keywords recur within the phrases, notably the lexical words *went* (9 tokens) and *want* (7 tokens), and personal pronouns, especially *I* (8 tokens) and *you* (4 tokens). Incorporating the novel's most distinctive words into its recurrent phrases creates a coherent stylistic effect to which the reader is repeatedly exposed. This reinforces the novel's main themes in the reader's mind, eliminating the need for explicit explanations.

⁹ Values approaching 1 indicate variable slots, while values approaching 0 indicate more fixed patterns, "with fewer attested fillers" (Gray and Biber 2013: 113-114).

¹⁰ Garner (2016: 40) proposed the following categories of p-frames based on their predictability score: highly predictable with PM>075%, predictable with PM=50-75%, unpredictable with PM=25-50%, and highly unpredictable with PM<25%.

Table 7. Patterns of 4-grams in *TSAR*

Patterns of 4-grams	Realizations
Fourteen 4-grams specify location and/or movement in space	e.g. <i>at the far end, went over to the, the top of the</i>
Ten 4-grams include personal pronouns, five of which are <i>I</i>	<i>I, me, you, we, it</i> (impersonal)
Ten 4-grams contain past forms	<i>came, went, did, stopped, said, was</i>
Six 4-grams express volition, three of which negate it	<i>do you want to, did not want to, I don't want to, you want me to, I did not want to, I would like to</i>
Six 4-grams signal interactional discourse	<i>do you want to, all right I said, don't be an ass, doesn't make any difference, you want me to, and have a drink</i>
Five 4-grams contain the grammatical negation <i>not</i> or <i>n't</i>	<i>did not want to, doesn't make any difference, I don't want to, don't be an ass, I did not want</i>
Five 4-grams express evaluation or emotion	<i>a hell of a, all right I said, doesn't make any difference, it was a good, don't be an ass</i>
Three 4-grams describe familiar social spaces, moments and objects	<i>and have a drink, the door of the, the centre of the</i>

Table 8. Patterns of 4-p-frames in *TSAR*

Patterns of 4-frames	Realizations
Twenty-one 4-frames specify location, movement in space and/or temporal relationships	e.g. <i>went * to the, back to the *, in the morning *</i>
Eleven 4-frames contain past forms	<i>went, did, came, was, looked</i>
Seven 4-frames describe familiar social spaces, moments and objects	<i>* on the bed, * to the hotel, in the square *, a bottle of *, the * it was, in the morning *, * a good time</i>
Six 4-frames include personal pronouns, three of which are <i>I</i>	<i>I, me, you, it</i> (impersonal)
Four 4-frames express volition, two of which negate it	<i>do you want *, * did not want, what do you *, I did not *</i>
Three 4-frames signal interactional discourse	<i>do you want *, what do you *, what's the matter *</i>
Two 4-frames contain the grammatical negation <i>not</i>	<i>* did not want, I did not *</i>
One 4-frame expresses evaluation	<i>* a good time</i>

The tables also reveal overlaps between the dominant 4-gram and 4-frame patterns, particularly in phrases expressing spatial movement and/or location. Alongside the temporal 4-frame *in the morning* *, they build the narrative's dynamics, focusing it on the characters' restless routines. The frequent use of personal pronouns, particularly *I* and *you*, which is also implicitly present in phrases such as *and have a drink* and *don't be an ass*, foreground spoken language. While *I* emphasises the subjective experiences of the protagonists over external narration, *you* further reinforces the novel's interpersonal orientation and conversational rhythm. This spoken-style narration is marked by interpersonal friction, as evidenced by evaluative and affect-laden expressions such as *don't be an ass*, *doesn't make any difference* or *a hell of a*. Clusters of volitional and negative forms reflect the characters' disillusionment and emotional frustration, stemming from their inability to fulfil desires in the post-war reality. The negative attitude permeating the narrative is also indicated by the frequent presence of the grammatical negation *not* in the phrases, while the frequent occurrence of past-tense verbs reinforces the novel's focus on reporting lived experience rather than introspection.

4.1. A case study of selected phrases

To uncover additional literary meanings, selected frequent phrases and their co-texts were analysed to identify lexical and grammatical patterns in their usage. The phrases examined – *do you want to*, *you want me to*, *I don't want to*, *I would like to* and *I did not want* – the first and last of which are realisations of one or more frequent 4-frames, express volition and interpersonal negotiation, often in direct speech. This reinforces *TSAR*'s spoken-style narration. In two of the phrases, volition is modified by the negative particle *not*, highlighting the characters' inner conflict between desire and aversion. As *do you want to*, *I would like to* and *I did not want* have not been identified by Stubbs and Barth (2003: 73) as frequent in their fiction corpus, they may reflect Hemingway's idiolect rather than general literary usage.

The phrase *do you want to* (11 occurrences) mainly collocates with activity and communication verbs, typically referring to practical actions such as eating, going, or doing. This semantic preference is evident in the concordance lines below, in which both verb types are particularly frequent to the right of the phrase but also to the left. The phrase is a second-person interrogative and is often embedded in dialogue questions, accompanied by forms of conversational scaffolding, such as *Cohn asked*, *said Montoya*, or *I asked*. This suggests that the characters use the phrase to negotiate everyday activities and logistics, deflecting attention from the emotional emptiness they experience as members of the Lost Generation.

ave some supper?" Cohn **asked**. "Do you want to **eat**?" "Yes. Why not? I'm hungry." ard and **looked** me in the eye. "Do you want to **know** something, Jake?" "Yes." "I a tablet up. Like at Lourdes. Do you want to **hear**, Robert? I'll **tell** you. It's bout the car?" Bill **asked** me. "Do you want to keep it on?" "It doesn't make any oya. "**Come** on," **said** Montoya. "Do you want to **meet** Pedro Romero?" "Fine," **said** e. She's well **taken care of**." "Do you want to ?" "Would I **ask** you if I didn't w tside on the pavement I **said**: "Do you want to **go** to the fight?" "Sure," **said** Bi this pub in Saint Jean." "**What** do you want to **do** about the car?" Bill **asked** me. Boulevard du Port Royal. "**What** do you want to **do**?" I **asked**. "**Go** up to the café **do**?" I **asked**. Brett **turned** her he **leaned** his head around. "**Where** do you want to **go**?" I **asked**. Brett **turned** her he s looking over my desk. "**Where** do you want to **eat**?" "**How** about Wetzels? They've

The phrase *you want me to* (7 occurrences) shows similar collocational patterns, but it co-occurs more strongly with mental concepts such as *think*, *know* and *have no idea*. Evidence for this can be seen in the concordance lines below. Compared with *do you want to*, this collocational pattern is stronger, with five instances. The phrase signals interpersonal pressure between the characters, as the structure 'want someone to [verb]' involves an asymmetry of will, whereby one person expects another to act.

u've no idea, Jake." "Anything you want me to **do**?" "No, just **go** to the fight wi hen. **Think** of that." "Anything you want me to **think** about it?" "Don't be an ass r **felt** such a bitch." "**What** do you want me to **do**?" "**Come** on," Brett **said**. "Let' u'd **know** these things. **What** do you want me to **do**? **Come** over here and **tell** you e "Darling," she **said**. Then: "Do you want me to **send** him away?" "No. He's nice." k like that, darling." "**How** do you want me to **look**?" "Oh, don't be a fool. I'm let's **eat**," **said** Bill. "Unless you want me to **tell** you some more travel stories

The remaining three phrases also collocate with activity, communication, and mental concepts, as can be seen in the concordance lines below. *I would like to* (6 occurrences) is particularly associated with mental verbs and conveys a desire that is either softened or hypothetical. Jake Barnes, the novel's main protagonist, uses it three times, twice introspectively to report his thoughts. *I did not want* (6 occurrences) appears exclusively in Jake's reflective narration, signalling his inability to act on desire. The fact that these two phrases are mainly used by only one speaker suggests that they are significant in characterising that person as someone driven by inner conflict between tentative desire and reluctance. In contrast, *I don't want to* (7 occurrences) only appears in dialogue across different characters, foregrounding negation, which signals unwillingness to articulate desires.

The above findings provide interesting insights into the literary meanings of *TSAR*. As activity and communication verbs are among the novel's keywords,

out to the park?" Brett **asked**. "I don't want to **go** up yet. I fancy he's sleeping **shouted** from the wine-barrels. "I don't want to **eat** up your meal," I **said** when s
l you. I **support** that boy, but I don't want to have him around." "He is rather
never **liked** children much, but I don't want to **think** I'll never have them. I al
aw me **notice** it. "I **say**, Jake, I don't want to **ruin** your rugs. Can't you **give** a
aid. "I don't **know**," she **said**. "I don't want to **go** through that hell again." "We
d **come** back later," Bill **said**. "I don't want to **sit** here on a night like this."

a yellow, sun-baked color, and I did not want to **leave** the café. But we had to
o loosen. I was very drunk and I did not want to **shut** my eyes because the room
ething else. I **told** the barman I did not want anything to **drink** and **went** out th
e. It was really only two, but I did not want to **explain** after I had made the m
ng with my face away from her. I did not want to **see** her. "**Sent** him for champag
"**Go** on. **Go** on up and **see** him." I did not want to **climb** another flight of stairs

Is there an Englishwoman here? I would like to **see** this English lady." "Muy bue
s to **see** it," the critic **said**. "I would like to **be** convinced." "He doesn't **like**
han your life." What a **speech**! I would like to have it illuminated to hang in t
ve the bottles a twirl. "Still I would like to **hear** you talk some time." "Isn't
other high headland. I **thought** I would like to **swim** across the bay but I was af
I might **pray** for and I **thought** I would like to have some money, so I **prayed** tha

their collocation with frequent phrases confirms the importance of these themes in the text. Conversely, mental concepts are rare among the keywords, with the only examples being the verbs *know* and *want*, which are generally common in fiction and conversation (Biber et al. 2021: 367, 373-374, 378).

However, the analysed phrases themselves include mental verbs that describe emotional or attitudinal states, such as *want* and *like*. This points to an important thematic and stylistic pattern. While mental states and emotions do not emerge as a dominant theme of the novel from the list of keywords, they are granted this status implicitly by their collocation with some of the text's frequent phrases, as well as by being part of singular phrases. This is particularly evident with the verb *want*, which conveys the notions of longing, need, or lack. Its repeated use in the novel, particularly in conjunction with personal pronouns, suggests that the reality depicted in *TSAR* centres on desire rather than fulfilment. The characters seem to be constantly reaching for something that either eludes them or that they cannot realise due to restraints, as indicated by *I don't want to* and *I did not want*. The fact that some of the phrases containing *want* occur in dialogue highlights its role as a social verb, frequently involved in interpersonal interactions in the novel. The meanings and usage patterns of *want* in *TSAR* thus seem to reflect the characters' social and psychological dislocation after the war.

5. Conclusion

This study analysed the keywords and frequent 4-word phrases in *TSAR* to reveal their subtle literary meanings. The findings show that the novel differs from Hemingway's other works and his contemporary literature through its distinctive vocabulary and themes.

The keywords indicate that the novel centres on characters, their conversations, routine actions and movement through familiar urban spaces. This corroborates Elbaz's (1984: 80) assertion that "Hemingway's characters" in *TSAR* "develop functionally to the spaces they inhabit and visit, and the actions they undertake in these spaces". The frequent use of *not*-negation and negatively connotated lexis points to interpersonal tension and emotional strain among the characters, alongside a reluctance to express their desires. The presence of fewer male-referencing keywords than in other Hemingway novels suggests a shift away from a single dominant male protagonist, despite the fact that the narrator remains male. In this narrative world, keywords related to bullfighting form an important motif, gaining prominence through their concentration later in the novel, their rhythmic collocation with the definite article, their placement in narration rather than dialogue, and their appearance in concise, adjective-free sentences. These factors all serve to heighten the symbolic status of bullfighting in the novel.

Analysis of the novel's most frequent phrases corroborates these observations. The recurrent phrases form a small, lexically stable repertoire built largely on verbs and personal pronouns. This highlights the focus on characters, their spoken interactions and their concrete, everyday experiences manifested in past-tense reporting. The lexical items within the phrases and their nearest collocates reinforce these themes. This limited phraseological core shows little variation, reflecting Hemingway's linguistic minimalism and creating stylistic and thematic consistency, and a documentary rhythm throughout the novel. At the same time, the high STTR value indicates a diverse vocabulary, suggesting precise and context-sensitive lexical choices.

Overall, the analysis shows that, despite his stylistic minimalism, Hemingway's use of language in *TSAR* generates statistically significant patterns that enrich existing interpretations of his first novel. Future research could compare the novel with its translations to determine whether Hemingway's limited phraseological repertoire and distinctive stylistic features, such as sparse adjectives and stable 4-frames, survive translation. As the novel's phraseology indexes spatial movement, interpersonal negotiation, volition, and negativity, it would also be interesting to map how these patterns cluster around particular narrative fragments (e.g. before and after the fiesta) or characters, and to examine how they shape the novel's meanings. Finally, cross-novel comparisons could further identify which of *TSAR*'s keyword and phraseological features reappear elsewhere in Hemingway's works.

References

- Baker, P. 2004. Querying keywords: Questions in difference, frequency, and sense in keyword analysis. *Journal of English Linguistics* 32(4): 346-359.
- Balassi, W. 1990. Hemingway's Greatest Iceberg: The Composition of *The Sun Also Rises*. In J. Barbour and T. Quirk (eds.), *Writing the American Classics*, 125-155. Chapel Hill: North Carolina UP.
- Balossi, G. 2014. *A Corpus Linguistic Approach to Literary Language and Characterization: Virginia Woolf's 'The Waves'*. Amsterdam/Philadelphia: John Benjamins.
- Bestgen, Y. 2018. Evaluating the frequency threshold for selecting lexical bundles by means of an extension of the Fisher's exact test. *Corpora* 13(2): 205-228.
- Biber, D. 2009. A corpus-driven approach to formulaic language in English: multi-word patterns in speech and writing. *International Journal of Corpus Linguistics* 14(3): 275-311.
- Biber, D. 2011. Corpus linguistics and the study of literature Back to the future? *Scientific Study of Literature* 1(1): 15-23.
- Biber, D., and S. Johansson, G. Leech, S. Conrad, E. Finegan 2021. *Grammar of Spoken and Written English*. Amsterdam/Philadelphia: John Benjamins.
- Cappock, E., and D. Beaver 2015. Definiteness and determinacy. *Linguist and Philosophy* 38: 377-435.
- Culpeper, J. 2009. Keyness: Words, parts-of-speech and semantic categories in the character-talk of Shakespeare's Romeo and Juliet. *International Journal of Corpus Linguistics* 14: 29-59.
- Dewberry, E. 1996. Hemingway's journalism and the realist dilemma. In S. Donaldson (ed.), *The Cambridge Companion to Hemingway*, 16-35. Cambridge: Cambridge University Press.
- Elbaz, R. 1984. The mechanics of repetition in the discourse of Ernest Hemingway. *Zagadnienia Rodzajów Literackich* 27(1): 77-90.
- Evison, J. 2010. What are the basics of analysing a corpus? In A. O'Keeffe and M. McCarthy (eds.), *The Routledge Handbook of Corpus Linguistics*, 122-135. London: Routledge.
- Fisher-Starcke, B. 2009. Keywords and frequent phrases of Jane Austen's *Pride and Prejudice*: A corpus-stylistic analysis. *International Journal of Corpus Linguistics* 14(4): 492-523.
- Fischer-Starcke, B. 2010. *Corpus Linguistics in Literary Analysis: Jane Austen and her Contemporaries*. London: Continuum.
- Fletcher, W. H. 2012. *KfNgram* (Software). Annapolis: USNA. Available at: <http://www.kwicfinder.com/kfNgram/kfNgramHelp.html> (accessed 12 September 2025).
- Fuster-Márquez, M. 2014. Lexical bundles and phrase frames in the language of hotel websites. *English Text Construction* 7(1): 84-121.
- Garner, J. 2016. A phrase-frame approach to investigating phraseology in learner writing across proficiency levels. *International Journal of Learner Corpus Research* 2(1): 31-67.

- Goźdz-Roszkowski, S. 2011. *Patterns of Linguistic Variation in American Legal English: A Corpus-Based Study*. Frankfurt am Main: Peter Lang.
- Gray, B., D. Biber 2013. Lexical frames in academic prose and conversation. *International Journal of Corpus Linguistics* 18(1): 109-135.
- Green, C. 2017. Introducing the *Corpus of the Canon of Western Literature*: A corpus for culturomics and stylistics. *Language and Literature: International Journal of Stylistics* 26(4): 282-299.
- Ho, Y. 2011. *Corpus Stylistics in Principles and Practice: A Stylistic Exploration of John Fowles' 'The Magus'*. London: Continuum.
- Hunston, S., and G. Francis 2000. *Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English*. Amsterdam/Philadelphia: John Benjamins.
- Keskesay (2021, April 24). *Ernest Hemingway, journalism in fiction*. Medium. <https://medium.com/@keskesay productions/ernest-hemingway-journalism-in-fiction-3492d3f7a1aa>.
- Leech, G., and M. Short 2007. *Style in Fiction: A Linguistic Introduction to English Fictional Prose*. Harlow: Pearson Education.
- Mahlberg, M. 2007. Clusters, key clusters and local textual functions in Dickens. *Corpora* 2: 1-31.
- Mahlberg, M. 2013. *Corpus Stylistics and Dickens's Fiction*. New York and London: Routledge.
- Rothman, N.L. (1928, April 1). Hemingway Whistles in the Dark. *The Dial* 84: 336-338. https://upload.wikimedia.org/wikipedia/commons/e/e3/The_Dial_%28Volume_84%29.pdf.
- Römer, U. 2010. Establishing the phraseological profile of a text type. The construction of meaning in academic book reviews. *English Text Construction* 3(1): 95-119.
- Scott, M. 2012. *WordSmith Tools (version 6.0)*. Stroud: Lexical Analysis Software.
- Scott, M., and C. Tribble 2006. *Textual Patterns: Key Words and Corpus Analysis in Language Education*. Amsterdam: John Benjamins.
- Shepherd, T.M.G., and T. Berber Sardinha 2013. A rough guide to doing corpus stylistics. *Matraga* 20(32): 66-89.
- Short, M. 1996. *Exploring the Language of Poems, Plays and Prose*. London: Longman.
- Sinclair, J. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J.M. 2008. The phrase, the whole phrase, and nothing but the phrase. In S. Granger and F. Meunier (eds.), *Phraseology: An Interdisciplinary Perspective*, 407-410. Amsterdam: John Benjamins.
- Starcke, B. 2006. The phraseology of Jane Austen's *Persuasion*: Phraseological units as carriers of meaning. *ICAME Journal* 30: 87-104.
- Stubbs, M. 2005. Conrad in the computer: examples of quantitative stylistic methods. *Language and Literature* 14(1): 5-24.

- Stubbs, M. 2007. Quantitative data on multi-word sequences in English: the case of the word *world*. In M. Hoey and M. Mahlberg, M. Stubbs, W. Teubert (eds.), *Text, Discourse and Corpora*, 163-190. London: Continuum.
- Stubbs, M., and I. Barth 2003. Using recurrent phrases as text-type discriminators: a quantitative method and some findings. *Functions of Language* 10(1): 65-108.
- Svoboda, F.J. 1983. *Hemingway & The Sun Also Rises: The Crafting of a Style*. Lawrence, Kansas: University Press of Kansas.
- Trodd, Z. 2007. Hemingway's camera eye: The problem of language and an interwar politics of form. *The Hemingway Review* 26(2): 7-21.
- Wagner-Martin, L. 1990. Introduction. In L. Wagner-Martin (ed.), *New Essays on Sun Also Rises*, 1-18. New York: Cambridge University Press.

Appendix. The literary texts included in the corpus *ContempLit*

Jack London, *Adventure* (1911)
 Edith Wharton, *The Age of Innocence* (1920)
 Willa Cather, *A Lost Lady* (1923)
 E.M. Forster, *A Passage to India* (1924)
 F. Scott Fitzgerald, *The Great Gatsby* (1925)
 Theodore Dreiser, *An American Tragedy* (1925)
 Thornton Wilder, *The Bridge of San Luis Rey* (1927)
 William Faulkner, *As I Lay Dying* (1930)
 John Fante, *The Road to Los Angeles* (1933)
 Gertrude Stein, *The Autobiography of Alice B. Toklas* (1933)
 John Steinbeck, *Of Mice and Men* (1937)
 Margaret Mitchell, *Gone with the Wind* (1939)
 Carson McCullers, *The Heart is a Lonely Hunter* (1940)
 Sinclair Lewis, *Kingsblood Royal* (1947)
 Ralph Ellison, *Invisible Man* (1947)
 J.D. Salinger, *The Catcher in the Rye* (1951)
 James Baldwin, *Go Tell It on the Mountain* (1953)
 Ray Bradbury, *Fahrenheit 451* (1953)
 Jack Kerouac, *On the Road* (1957)
 Harper Lee, *To Kill a Mockingbird* (1960)
 John Updike, *Rabbit Run* (1960)
 Joseph Heller, *Catch-22* (1961)
 Ken Kesey, *One Flew Over the Cuckoo's Nest* (1962)
 Sylvia Plath, *The Bell Jar* (1963)
 Saul Bellow, *Hertzog* (1964)
 Cormac McCarthy, *The Orchard Keeper* (1965)
 Anne Tyler, *Tin Can Tree* (1965)
 James Salter, *A Sport and a Pastime* (1967)
 Philip Roth, *Portnoy's Complaint* (1969)
 Kurt Vonnegut, *Slaughterhouse-Five* (1969)

MARZENA BĘDKOWSKA-OBLĄK

Schlesische Universität in Kattowitz, Polen

Institut für Sprachwissenschaft

ORCID: 0000-0002-6758-7796

marzena.bedkowska-oblak@us.edu.pl

Rezensionsbeitrag

INTERDISZIPLINÄRE ZUGÄNGE ZU DER MODERNEN FORSCHUNG DER SPRACHKOMMUNIKATION

IZABELA KUJAWA UND ROMAN OPIŁOWSKI (HRSG.). 2026.
LINGUISTISCHE GRENZÜBERGÄNGE. EINBLICKE
IN DIE SPRACHFACETTEN UND DISKURSWELTEN DER MODERNEN
KOMMUNIKATION. STUTTGART: FRANZ STEINER VERLAG. 207 S. PRINT

Die zeitgenössische Forschung zur Sprachkommunikation erstreckt sich heute weit über die Grenzen der Kommunikationswissenschaft hinaus und integriert zunehmend andere linguistische Fachgebiete. Das Forschungsinteresse liegt nicht nur in der Untersuchung des Kommunikationsprozesses, seiner Organisation oder Realisierung. In der Analyse rücken die beteiligten Akteure, der Kontext sowie die Informationsmedien verstärkt in den Fokus. Das zentrale Anliegen des vorliegenden Sammelbandes liegt darin, die neuesten Forschungsergebnisse aus drei grundlegenden Bereichen der modernen Kommunikationsforschung vorzustellen: der Diskursanalyse, der Medienlinguistik und den methodisch-didaktischen Aspekten des berufs- und fachbezogenen Fremdsprachenunterrichts.

Im ersten Teil des Bandes gehen die Autoren auf das Charakteristische und Spezifische der Diskursanalyse und Medienkommunikation ein und behandeln solche Aspekte wie u.a. mediale Darstellung deutsch-polnischer Beziehungen, die multimodale Kommunikation in sozialen Netzwerken, Migrationsdiskurse, die Thematisierung der COVID-19-Pandemie sowie die Analyse der Topoi in der Wahlkampfkommunikation.

Die Erörterungen des Beitrages von Waldemar Czachur werden der Problematik kollektiver Identität in sozialen Interaktionen aus diskurslinguistischer Perspektive gewidmet. Die kollektive Identität ist für den Autor ein sprachliches, kommunikatives und diskursives Phänomen, das als ein System von Normen und Regeln der sprachbezogenen Identifikation und zugleich als ein dynamischer Prozess der sprachbezogenen Bestätigung sowie Modifikation der Identifikation (Positionierung) von Mitgliedern einer bestimmten Gemeinschaft in einer bestimmten soziokulturellen Realität zu betrachten ist. Czachur nutzt für seine Untersuchung die framesemantische und multimodale Analyse, um im Hinblick auf das Konzept des „Doing Identity“ die Rolle der Sprache als Instrument und Medium zur Konstruktion von Identität hervorzuheben. Zur Visualisierung werden Fallstudien zur polnischen und deutschen Identität angeführt, anhand derer der Autor die diskursive Gegenüberstellung von „Uns“ und „Anderen“ treffend illustriert. Die Gegenüberstellung fungiert hier als Forschungsinstrument, das es ermöglicht zu verstehen, wie Identität durch das Aufzeigen von Unterschieden aufgebaut wird. Dank der Framesemantik wird deutlich, wie Sprache und multimodale Bilder unsere Identität und unsere Wahrnehmung der Welt mitgestalten.

Beata Mikołajczyk untersucht wiederum in ihrem sowohl diskurs- als auch medienlinguistisch ausgerichteten Aufsatz die Funktion des Topos der historischen Analogie als Ausdruck eines Vertrauensverlustes gegenüber den Regierenden innerhalb des Corona-Anti-Impf-Diskurses. Diese universell anwendbaren und vertrauten Topoi – wie etwa der Topos des totalitären Staates, der Freiheit und Selbstbestimmung, der Krise oder der Geldgier – fungieren als persuasive Instrumente, mit denen staatliches Handeln rechtlich sowie moralisch dekonstruiert und folglich negativ bewertet wird. Wie Mikołajczyk stimmig darlegt, wird der Analogie-Topos zu einem hocheffektiven Überzeugungswerkzeug instrumentalisiert, das insbesondere im deutsch- und polnischsprachigen Corona-Anti-Impf-Diskurs eine signifikante persuasive Wirkung entfaltet. Dieser Topos besitzt ein manipulatives Potenzial, welches selbst dann erfolgreich sein kann, wenn die zugrunde liegende Analogie sachlich nicht fundiert ist. Diese persuasive Kraft ermöglicht es Diskursakteur*innen, gezielt negative Assoziationen zu evozieren, welche die öffentliche Wahrnehmung beeinflussen und zur Konstruktion von Feindbildern beitragen. Das der Untersuchung zugrunde liegende Korpus umfasst 2500 deutsch- und polnischsprachige Einträge auf der Plattform X. Analysiert werden Texte, in denen historische Ereignisse mit aktuellen Vorkommnissen im Kontext der Corona-Krise verglichen werden. Ziel dieser Vergleiche ist es, mittels gezielter sprachlicher und visueller Elemente kollektive Ängste sowie historische Traumata zu reaktivieren und dadurch die argumentative Aussagekraft der jeweiligen Stellungnahmen zu stützen.

Der Beitrag von Michał Smułczyński thematisiert hingegen die zunehmende Bedeutung der Kommunikation in sozialen Medien. Da sich durch diese Medien ein kommunikativer Wandel im Internet vollzogen hat, bei dem One-to-Many-

Monologe durch sozial-mediale Dialoge (Many-to-Many) erweitert wurden (vgl. Rybszleger 2014), werden Plattformen wie X (ehemals Twitter) verstärkt von prominenten polnischen und deutschen Politikern genutzt, um Anhänger und Wähler zu gewinnen. Im Fokus der eingehenden Analyse stehen insbesondere die Beiträge rechtspopulistischer Politiker. Es werden jeweils 1000 der neuesten X-Einträge kontrastiv untersucht, um zu ermitteln, welche Argumentationen im Rahmen der rechtspopulistischen Narration verwendet und wie diese sprachlich konstruiert werden. Im Fokus seiner Untersuchung stehen die Stimmungskodierung (u.a. *positive, negative, neutrale* Stimmung) und Knotenkodierung (Vorkommen solcher Themen, Motive, Konzepte, wie z. B. *Kampf, Flüchtlinge, kraj, bezpieczeństwo*).

Gegenstand der Ausführungen von Izabela Kujawa ist dagegen eine linguistische Analyse der rhetorischen Strategien auf Wahlplakaten rechtspopulistischer Parteien im Kontext der Europawahl 2024, der polnischen Parlamentswahl 2023 sowie der deutschen Kommunalwahlen 2024. Diese Strategien der Rhetorik dienen, wie die Autorin bemerkt, primär der Unterstützung von Politikern, die sich als Verteidiger nationaler Werte und Identität sehen. Sie richten sich gezielt an eine Zielgruppe, die eine Entfremdung von etablierten Parteien oder internationalen Organisationen erfährt. Analog zu Mikołajczyk rückt Kujawa das Topos-Konzept in das Zentrum ihrer Untersuchung. Dabei werden Topoi als argumentative Muster begriffen, die von populistischen Parteien in Deutschland und Polen funktionalisiert werden. Izabela Kujawa arbeitet in ihrer Analyse sowohl koinzidierende Argumentationsmotive heraus – etwa den Schutz der nationalen Souveränität und des kulturellen Erbes – als auch erhebliche Divergenzen in den rhetorischen Mitteln. Während in den polnischen Wahlplakaten eine stärkere religiöse Fundierung festzustellen ist, dominiert im deutschen Diskurs das Thema der Remigration.

Der erste Teil des Sammelbandes wird mit dem Aufsatz von Roman Opiłowski abgeschlossen, in dem er seine Aufmerksamkeit der Intermedialität widmet. Diese wird als eine Verbindung von Medien verstanden, durch welche die zwischenmenschliche Kommunikation effizienter gestaltet wird (vgl. Schneider/Stöckl 2011 et al.). Unter Medien werden hier wiederum materielle und technische Träger von Texten verstanden. Der Autor behandelt mediale Texte unter anderem aus dem Fernsehen, der Presse, dem Internet sowie dem urbanen Raum („der Stadt“) und kontrastiert diese aus einer intermedialen Perspektive. Ziel ist es, die Besonderheiten kommunikativer Praktiken – wie Kommentieren, Zitieren, Erklären oder Unterhalten – in Abhängigkeit von der jeweiligen Medienverknüpfung zu ermitteln. Intermedialität gilt dabei als Methode der Verknüpfung und Analyse, die sich in der Regel auf zwei oder mehrere Medien bezieht. Der Schwerpunkt liegt dabei weniger auf der Textsorte an sich, sondern vielmehr auf einzelnen kommunikativen Aspekten. Wie Opiłowski aufzeigt, lassen sich

infolge dieser intermedialen Kontrastierung innovative und diverse Praktiken erkennen, die erst an der Schnittstelle der Medien zum Vorschein kommen.

Der zweite Teil des Bandes ist dem Aspekt der Sprachkommunikation in Unterricht und Beruf gewidmet. In ihrem einleitenden Beitrag widmet sich Beata Grzeszczakowska-Pawlikowska der Entwicklung rhetorischer Redekompetenz, die als komplexe Fertigkeit verstanden wird und besonders in der akademischen sowie beruflichen Ausbildung von Belang ist. Ein sicheres Auftreten und die Präsentation vor Publikum sollten gezielt in das Lehrprogramm aufgenommen, vermittelt und trainiert werden. Die Autorin arbeitet hierfür praktische didaktisch-methodische Hilfsmittel und Instrumente aus, die den Prozess der Kompetenzentwicklung unterstützen können. Dabei wird insbesondere auf psycholinguistische Aspekte wie Selbstbeobachtung, Selbstevaluation und Reflexion hingewiesen.

Den Schwerpunkt der Erörterungen im Aufsatz von Jacek Makowski bilden hingegen die didaktisch-methodischen Aspekte der fremdsprachlichen Fachkommunikation sowie des Berufsbezuges innerhalb von Geschäftsprozessen. Wie der Autor treffend feststellt, stellt die Komplexität des Forschungsfeldes Fach- und Berufskommunikation sowohl eine Herausforderung als auch einen attraktiven Forschungsgegenstand dar. Im Beitrag werden erste Ergebnisse eines Forschungsprojektes berichtet, in dessen Rahmen der Bedarf an sprach- und fachbezogenen Berufskompetenzen in der Personalwirtschaft untersucht wurde. Wie sich gezeigt hat, gehört Polen zu den Ländern, in denen der Sektor der modernen Unternehmensdienstleistungen starke Entwicklungstendenzen aufweist. Im Projekt werden jene Sprachkenntnisse und fachsprachlichen Kompetenzen ermittelt, die in der akademischen Ausbildung zu implementieren sind. Daraufhin unternimmt der Autor den Versuch zu erforschen, inwiefern die geforderten Sprachkenntnisse und Kompetenzen in das für diese Zwecke ausgewählte DaF-Lehrwerk *Schritte international neu 5/B1.1* eingebunden sind. Abschließend werden die spezifischen Teilprozesse der Kompetenz im Bereich der Personalbeschaffung analysiert und anhand ausgewählter Sprachbeispiele aus dem Untersuchungskorpus verdeutlicht.

Der letzte Teil des Bandes widmet sich der Problematik urbaner Minitexte aus lexikalischer, interkultureller und multimodaler Sicht. Camilla Badstübner-Kizik thematisiert in diesem Zusammenhang die Merkmale kleiner Texte – wie Größe, Komplexität, sprachliche Gestaltung und Funktionalität – und weist auf spezifische Techniken des Spiels mit linguistischen Irritationen hin. Dazu zählen unter anderem Mehrdeutigkeit, Hybridisierung sowie Auslassungen, die sie anhand zahlreicher aktueller Beispiele (erhoben 2024) aus dem polnisch- und deutschsprachigen urbanen Raum verdeutlicht. Wie die Autorin betont, ist der Charakter dieser Texte als ‚kleine Texte‘ im öffentlichen Raum ausschlaggebend: Sie treten in großer Zahl auf und müssen sich auf beschränkter Fläche gegenüber konkurrierenden Textangeboten durchsetzen. Sprachliche und visuelle Irritatio-

nen dienen hierbei als notwendiges Mittel, um Aufmerksamkeit zu erregen und den Rezeptionsprozess überhaupt erst auszulösen. Badstübner-Kizik führt viele interessante Beispiele für diverse Formen sprachlicher Ironisierung, wie z. B. sprachliche Mehrdeutigkeit, scherzhaft-ironische Vertauschung oder Hybridbildungen, mit denen das zu erwartende Effekt der Wahrnehmung und Aufmerksamkeitslenkung eingetreten ist.

Magdalena Makowska widmet die Ausführungen ihres Beitrags indessen dem Einsatz von Minitexten in einem spezifischen Diskursbereich: der Politik. Sie rezipiert politische Protestkontexte in der urbanen Öffentlichkeit Polens (beispielhaft anhand zweier Demonstrationen in Warschau im Jahr 2023) vor dem Hintergrund der Linguistic-Landscape-Forschung. Wie die Autorin darlegt, fungiert die städtische Umgebung als Kommunikationsraum; die politischen Transparente mit ihren unter anderem kommunikativ-informativen, expressiven, kulturell-kreativen sowie erzieherischen Funktionen tragen zur Transformation der urbanen Umgebung in eine Protestplattform bei. Zwar sind Transparente keine prototypischen Minitexte, wie die Autorin anmerkt, dennoch gelten sie als multimodale Kommunikate in knapper, semiotisch differenzierter Form, mit denen facettenreiche kommunikative Aufgaben bewältigt werden. In ihrer detaillierten Analyse verfolgt Makowska zwei Ziele: Sie verdeutlicht, dass selbst diese Textsorte zu den mobilen urbanen Kleintexten gehört, und erörtert die sprachlichen sowie visuellen Mittel, mit denen Proteste in Form kleiner Botschaften visualisiert und verbalisiert werden.

Im Aufsatz von Katarzyna Siewiert-Kowalska finden wir wiederum eine Vertiefung der bisherigen Darlegungen des Bandes über Plakate, die der kommunikativen Handlung des Protestierens dienen. Es wird eine Analyse der Sprachmittel durchgeführt, die zum Ziel hat, die sprachlichen Strategien zur Kodierung von Emotionen im Protestplakat zu erörtern. Das Protestplakat ist eine besondere Form des Ausdrucks von Emotionen; es handelt sich um einen multimodalen Text, in dem diverse Zeichen das Gesamtgefüge des Inhalts bilden. Die Kodierung von Emotionen in diesen Texten und ihre Interpretation werden hier durch den soziokulturellen, diskursiven und medialen Kontext determiniert. Den Textkorpus des Beitrags bilden die Plakate von Bauernprotesten in Deutschland. Wie die Analyse zeigt, werden in den Bauernprotesten vorwiegend negative Emotionen wie Unzufriedenheit, Ärger, Ungerechtigkeit und Verlustängste thematisiert. Diese werden sprachlich u. a. auf der syntaktischen Ebene vorwiegend durch Exklamativsätze, lexikalisch hingegen durch negativ konnotierte Verben, Phraseologismen sowie prosodische Elemente ausgedrückt.

Der Sammelband liefert aufschlussreiche Ergebnisse zur modernen Sprachkommunikationsforschung mit Fokus auf kontextuelle und instrumentelle Aspekte, wie etwa Diskursbereiche, multimodale Träger von Kommunikaten sowie deren Einbeziehung in die didaktischen Prozesse. Er ist all jenen zu empfehlen, die ihr Wissen über multiperspektivische Betrachtungsweisen der Sprachkom-

munikation unter besonderer Berücksichtigung von Diskursen und Medien vertiefen möchten. Das Werk kann zudem einen Anstoß für weitere theoretisch und empirisch ausgerichtete Forschungen sowie neue Anregungen geben.

Bibliographie

- Rybszleger, P. 2014. Soziale Medien als multimodale Communities of Practice. In G. Antos, R. Opilowski, J. Jarosz (Hg.), *Sprache und Bild im massenmedialen Text. Formen, Funktionen und Perspektiven im deutschen und polnischen Kommunikationsraum*, 287-298. Wrocław / Dresden: Atut / Neisse.
- Schneider, J.G., und H. Stöckl 2011. Medientheorien und Multimodalität: Zur Einführung. In J.G. Schneider, H. Stöckl (Hg.), *Medientheorien und Multimodalität. Ein TV-Werbespot – Sieben methodische Beschreibungsansätze*, 10-44. Köln: Herbert von Halem.