

„Kwartalnik Filmowy” no. 127 (2024)
ISSN: 0452-9502 (Print) ISSN: 2719-2725 (Online)
<https://doi.org/10.36744/kf.3047>
© Authors; Creative Commons BY 4.0 License

Agata Hołobut

Jagiellonian University
<https://orcid.org/0000-0002-9445-9060>

Maciej Rapacz

AGH University of Kraków
<https://orcid.org/0009-0004-1665-9413>

Miłosz Stelmach

Jagiellonian University
<https://orcid.org/0000-0003-4004-7121>

Working Titles. Computational Analysis of Film Titling Practices: A Polish Case Study

Keywords:

film title;
metadata analysis;
title translation;
film history;
history of Polish
cinema

Abstract

This article traces historical trends in film titling practices in Poland, based on quantitative analysis of film metadata. The authors analyse a corpus of 2,519 films produced in Poland until 2023, looking at regularities in the naming of Polish original films: the average length of titles, title duplication, the most frequent content words typical of successive phases in the history of Polish cinematography, lexical diversity, and the presence of proper names and foreignisms. Then they replicate some of the tests on the corpus of foreign films translated into Polish, which includes 5,585 films distributed in Poland between 1945 and 1989 and 33,794 films released in Poland after 1989. The authors compare the peculiarities of the original titles and those translated into Polish in search of similarities and discrepancies, potentially symptomatic of *translationese*, or *tittlese*.

This publication was supported by a grant from the Faculty of Philology under the *Strategic Programme Excellence Initiative – Research University* at the Jagiellonian University in Kraków.

Titles have so far attracted relatively little attention from film scholars. Inherently verbal, ready to shapeshift, attached to productions by nobody-knows-whom (sometimes the screenwriter or director, on other occasions by the producers or marketing teams), they seem, at first glance, uncinematic in themselves. Most of the media coverage that they receive comes from viewers who discuss successful solutions and unsuccessful title translations. Most of the scholarly appreciation that they receive comes from linguists and marketing experts, who taxonomize them and explore their enormous power of persuasion.

However, as we wish to argue, film titles can be of true interest to film scholarship. They stage people's first encounter with a given production, even before its official release. They act as paratexts or *thresholds of interpretation*¹ that only some recipients decide to cross when viewing a given film. Others may be familiar with the title without entering the on-screen world to which the title is a passage. As such, titles are both co-dependent and autonomous, integral and external to the films they label. On the one hand, they are part of the cinematic experience, and incorporated into title sequences, they influence the entire work and its subcomponents.² On the other hand, they are set apart from the cinematic experience because they represent productions outside the movie theatre, securing their place in film history.³ That is why film-naming is of huge importance to the industry. According to Chung and Eoh's survey, filmmakers consider it as *one of the most important communication tools to explain their movies to movie viewers and pique their interest*; while also admitting that *movie titles significantly impact box-office performance*.⁴

Titles fulfil six major functions: (a) distinctive: helping differentiate the film from others; (b) metatextual: presenting themselves as valid film titles; (c) phatic: attracting the attention of potential viewers; (d) referential: offering information about the film in terms of plot and genre; (e) expressive: conveying the filmmakers' or production studio's stance towards their product and the world at large; and (f) appellative: sparking the curiosity of external observers, who might need additional encouragement to watch the movie.⁵ In other words, they label, individuate, position, advertise, and summarize productions. They provide context and create a mood for the cinematic experience. Their systematic study offers insights into the evolution of national film industries and international film distribution patterns by revealing dominant themes, genres, and aesthetic preferences typical of their respective eras. However, their investigation requires interdisciplinary instruments outside the usual film-scholarly toolbox.

In this paper, we borrow such instruments from the realms of data analysis and corpus linguistics to examine all the titles of feature films ever produced in Poland and juxtapose them with titles of foreign feature films that have been translated into Polish over time. We subject these titles to quantitative tests that may seem atypical of film studies, although the questions we pose have been inspired by film as well as literary and translation studies. We wish to find out: 1) whether the film titling practices in Poland have evolved over time with respect to the preferred title form (estimated in terms of length and syntactic structure) and content (estimated in terms of vocabulary choices); 2) whether they differ from film title translation practices; i.e., whether the original and translated titles exhibit discrepancies with regard to selected criteria, and if so, whether these dis-

crepancies are interpretable in terms of *titlese*, i.e., a form of *translatese* associated with translated titles; 3) whether film translation strategies have evolved over the years, especially considering the distribution of foreign films under original titles and original titles with structural modifications, including added or omitted elements, e.g., a subheading.

Thus, we hope to offer new insights into domestic and foreign film production and distribution practices in Poland. We intend to diachronically reconstruct film-naming trends, map out the themes most often touched upon by domestic and imported productions, and quantitatively confirm some of the intuitions expressed by film and media scholars on the basis of qualitative analyses of small title samples.

Below, we first delineate the research tradition in (film) title analysis and identify its most frequent concerns and findings. Then, we diachronically investigate regularities in the naming of Polish original films in terms of title length (character and word-count), part-of-speech makeup, and most frequent content words. Subsequently, we explore film title translation practices in Poland since 1945 based on similar criteria, supplementing the analysis with insights into dominant translation strategies, specifically focusing on the use of non-translation and non-translation with addition.

Titrology

The international circulation of film titles is governed by industry-specific norms,⁶ which are quite distinct from those that govern popular music titles (which remain almost exclusively untranslated) and literary works (which remain largely translated). Although not researched as thoroughly as the films they label, titles have attracted some academic attention from scholars of diverse disciplinary backgrounds. Their reflection continues the long tradition of *titrology* research, which has focused on the shape and function of literary paratexts.⁷

Among notable contributions from Polish film scholars, Marek Hendrykowski's early reflections are particularly relevant.⁸ The author discusses, among others, the divergent evolution of titling practices in different parts of the world, with the emergence of recurrent keywords (such as "love," "life," and "death") and the establishment of culture-specific titling conventions (Polish titles reportedly revealing strong affinities to literary models, American titles to advertising slogans and Italian titles to newspaper headlines). Titles are mentioned by film and media scholars interested in various paratexts, including trailers and title sequences.⁹

The most significant contributions to film title research have been made by linguists and translation scholars. Interested primarily in contrastive analyses, they conduct case studies of film title translation practices in particular regions, e.g., Iran,¹⁰ China,¹¹ Spain,¹² France, Romania,¹³ Germany, Croatia,¹⁴ or Poland.¹⁵ Their approaches range from detailed discussions of misconceived translation decisions¹⁶ to the investigation of dominant translation strategies based on title corpora of diverse temporal scopes, sizes, and language combinations. Among the examples of multilingual diachronic insights, one could mention the study of Academy Award-winning titles released over the last 90 years in Polish, German,

and Russian renditions¹⁷ or the investigation of James Bond film titles as rendered into French, German, and Polish.¹⁸

The typologies offered by scholars differ from publication to publication, but they mainly concern four broad categories of interpretive mechanisms that can be involved in the rendition of a source title into the target language. These, following Schubert's taxonomy,¹⁹ adopted for example by Gabrić et al.²⁰ and Fakharzadeh,²¹ include: (a) literal translation, which follows the source-language model; (b) free translation, which builds upon it; (c) transcription, which transfers the source-language material in its entirety (complete transcription); combines it with extra material, such as subheading (transcription with addition); shortens it (with omission) or replaces some of its elements (with substitution); (d) transcreation, which consists in devising a completely new title, unrelated to the original.²²

Most of the analyses conducted offer synchronic rather than diachronic perspectives, as scholars are often uncertain as to when exactly the particular titles were translated.²³ Relatively rare insights are provided by researchers interested in the institutional frameworks of translation, a notable exception being Ross's study into the influence of distributors and studios on the decision making.²⁴ All these linguistic discussions may be relevant for film studies because they draw attention to the cultural filters imposed on films in their international distribution. If we agree with Hendrykowski that the title is key to the interpretation of a movie, audiences around the world are invited to take different interpretation paths regarding the same production.

A different group of approaches to film titles comes from marketing and business scholars who apply quantitative methods to gauge the effectiveness of titling solutions. Among those, we should mention Bae and Kim's research into the correlation of title informativity (i.e., its capacity to explicate the genre and the storyline) and the box-office success;²⁵ Sood and Drèze's reflection on how numbering impacts the popularity of sequels and other sequential forms;²⁶ and Xiao et al.'s study of keywords and thematic words in American and Chinese film titles,²⁷ which confirms the universal use of some surefire favourites (such as "man," "movie," or "war") in the biggest-grossing movie titles across borders and the non-universality of others (for example, the second position of the keyword "love" on a Chinese list of biggest-grossing films, and its absence from the American top 50). Curiously, in 1975, the American scholar Ralph Dengler, S. J., used computational methods to trace the evolution of dominant themes in American movie titles based on over 7,598 examples spanning 70 years of film production.²⁸ He noticed diachronic changes in filmmakers' preoccupations, as well as the dominance of sex and danger themes, time references, and gender roles in film naming practices.

A particularly interesting contribution within the realm of market-focused quantitative studies is Chung and Eoh's approach to film titles as a variety of brand names whose success depends on information choice, expression strategies, and genre- and country-related factors. The scholars measured the impact of film-naming strategies on moviegoers' decisions and the box-office success of productions. As for informativity, they assumed that titles may impart information on film characters (e.g., their gender), subject matter, stories, or themes, as well as the positive/negative valence of film content.²⁹ They subsequently verified and

confirmed that story- and subject-matter indications influence moviegoers' decisions and so do movie valence signals in film titles. The authors proved that unlike in any other product, viewers are surprisingly attracted to negative words in film titles that reveal plot-informing conflicts.³⁰

With regard to the expressive aspect, Chung and Eoh assumed that filmmakers can opt for the use of distinctive words (non-words, proper nouns, and foreign words) or uncertainty-reducing familiar clichés ("love," "alien"), which help predict the story and genre; they can also regulate the length of the title. The authors confirmed that expression strategies (e.g., the use of unique distinctive expressions) can strengthen or weaken information choice effects and that the length of the title has a considerable impact. They found that *movie titles containing subject matter and negative content can be more attractive with unique expressions than popular expressions*.³¹ They also demonstrated that the effects of naming strategies differ depending on the genre and have the greatest influence on the success of science fiction and action films.

Finally, based on the data gathered in seven countries where English is the official language – the United States, the United Kingdom, New Zealand, Hong Kong, the Philippines, Singapore, and Malaysia – the researchers found that the preservation of the same title across these national markets has positive financial effects. They observed that *using the same movie title across countries (the standardization strategy) is better in international naming strategies than using different titles for different countries (the localization strategy)*.³² The above-mentioned contributions illustrate the methodologies and issues explored in titrological research, some of which are addressed in this study.

'Natively' Polish titles: Breaks and continuations

We shall start by characterizing the Polish film-naming practices over time with respect to selected formal properties (such as title length and syntactic structure) and content properties (most frequently recurring content words; proper names and foreignisms). Whenever possible, we will contrast our findings with patterns typical of Anglophone and Chinese movie-naming, as described in international scholarship.

We based our analysis on the metadata provided by the *filmpolski.pl* website.³³ Managed and maintained by the Lodz Film School, it is a respected and reliable database containing information on all the professional (and many independent) movies ever produced in Poland, including casts and crew members, release dates, plot summaries, awards received, and, in some cases, additional details, such as technical specs and locations. Although the database covers various formats, such as short films and TV productions, we decided to limit ourselves only to feature films, which are the most recognized part of the film-making industry and are well represented in the other databases that we intended to consult for comparison.

We compiled a list of all Polish feature films matched with their production year, and subsequently manually verified it to exclude, among others, foreign

productions by Polish filmmakers (such as Andrzej Żuławski's 1989 film *Mes nuits sont plus belles que vos jours*, made in France) or international co-productions in which Poland was a minority co-producer (such as Ari Folman's 2013 work *The Congress*). Both groups are included in the database, but in our opinion, they do not reflect domestic film-naming practices. Since the data were consulted in May 2024, we excluded the ongoing year as well. After these operations, our dataset of Polish feature film titles included 2,519 items released between 1908 and 2023. Regarding the reliability of our dataset, some edge cases and minor errors may still be expected, but we believe that this list fairly accurately represents the entirety of Polish feature film production and that potential small errors do not distort the statistical image that we present.

For the tests discussed below, we also assumed a twofold historical periodization of our dataset. On a general level, we distinguished three time periods that correspond to the changing political system in Poland: 1) Until 1939 – this period covers the time when Poland was not an independent country (until 1918) and the Second Polish Republic (1918-1939).³⁴ Altogether, 369 titles belong to this category; 2) 1945-1989 – the period of the Polish People's Republic, a communist state where the film industry was nationalized, i.e., all films were produced, distributed, and exhibited through state-owned institutions and companies responsible directly before the government. The specific structure of the film industry changed many times during this period, but its general political and economic makeup remained stable. Altogether, 896 titles belong to this category; 3) 1990-2023 – the contemporary capitalist film industry dominated by private entities with limited support from state agencies and local governments. Altogether, 1,254 titles belong to this category.

For selected tests, we also applied a more granular periodization. Here, our defining criteria were mixed and included both internal transformations of the film industry (e.g., the introduction of sound to Polish cinema in 1930 and the foundation of the Polish Film Institute in 2005) and external factors connected to political shifts in the communist regime (such as the change of the First Secretary of the ruling Polish United Workers' Party). Our granular periodization covers the following segments: until 1929 (208 titles), 1930-1939 (161 titles), 1945-1956 (61 titles), 1957-1968 (249 titles), 1969-1980 (307 titles), 1981-1989 (279 titles), 1990-2005 (394 titles), and 2006-2023 (860 titles).

After creating the corpus and defining its periodization, we ran tests to reveal predominant patterns and historical changes in the linguistic makeup of titles given to Polish films. First, we examined the title length in terms of both the number of words and the number of characters. This parameter may seem overly simplistic at first glance, but it does reveal important information about Polish filmmakers' attitudes towards title memorability and informativity. Chung and Eoh argued that recall and recognition are connected to title length: *a longer name may have disadvantages against shorter names, as people have difficulty remembering longer movie titles. At the same time, longer titles are likely to deliver more information regarding the movie,*³⁵ thus piquing viewers' interest. The results of our analysis are depicted in Figure 1:

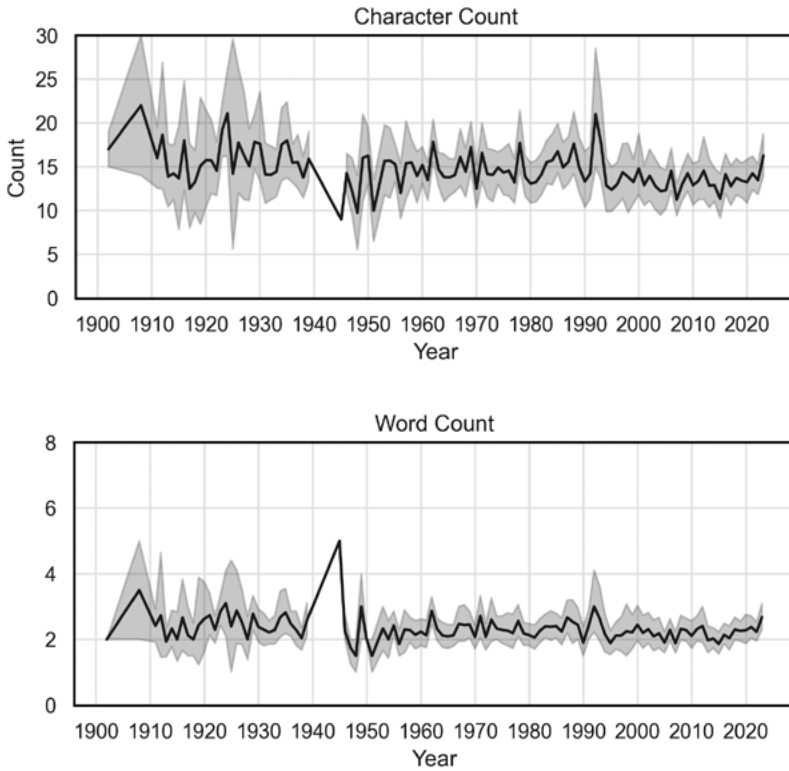


Fig. 1. Distribution of character and word count in Polish film titles across years, with a 95% confidence interval.

Although there are visible fluctuations in Figure 1, a subtle but clearly discernible tendency can be observed for the titles to become shorter. Considering the number of characters (including spaces), titles from the pre-war period, on average, consisted of 15.9 characters and were slightly longer than those from the communist period, which averaged 14.8 characters. The titles of modern movies made after 1989 are still shorter, with an average of 13.7 characters, a 14% decrease from the pre-war period. The same shortening tendency, although of a lesser magnitude, governs the number of words used in the titles. For our three periods, the average number of words are consecutively: 2.44 (until 1939), 2.31 (1945-1989), and 2.23 (after 1989), marking a less than 9% decrease throughout the entire history of Polish cinema. This indicates that the overall decrease in title length is due to the use of shorter rather than fewer words, as the trend line (linear regression) is slightly steeper for character count (Fig. 1b):

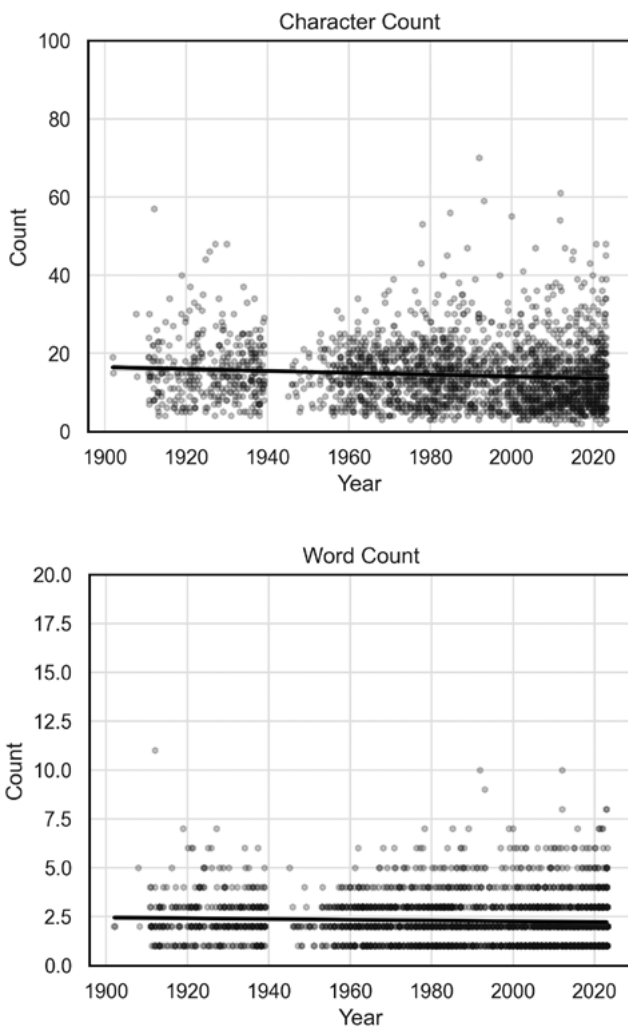


Fig. 1b. Regression plot – trend line of character and word count of Polish film titles.

Considering the influence of word length on reading speed and ease of comprehension,³⁶ these findings reveal a growing tendency to facilitate readability. By comparing the total number of words and characters in the corpus, we found an average of 6.5 characters per word in the pre-war period; 6.4 characters in People's Poland, and 6.14 characters after 1989 (spaces included). According to a recent estimation, an average word in written Polish is six characters long.³⁷ Polish titles have been composed of words oscillating around this value, with a tendency to decrease in word length. Hence, we conclude that trends in Polish film-titling are moving towards an increased ease of perception.

Also, we can see that in both metrics, the pre-war period reveals greater internal variation from year to year as well as within each year. The shaded area adjacent to the line represents the so-called confidence interval of 95%, which

means that if we were to take many random samples and calculate their means, 95% of those sample means would fall within this interval. For pre-war films, both the jolts of the line and the shaded areas are more pronounced, suggesting that titling practices in this period were more 'capricious' and only stabilized after the war.

The number of words used in Polish film titles averaged between 2.23 for most recent movies and 2.44 for the oldest ones, including all conjunctions, particles, prepositions, and other function words. This suggests that film titles tend to be short, perhaps surprisingly so. Indeed, we can notice that although some unusually lengthy titles (such as Rafał Wieczyński's record breaking 10-word and 60-character-long 1992 film *Naprawdę krótki film o miłości, zabijaniu i jeszcze jednym przykazaniu*) might stand out and distort our perception, the vast majority of film titles consist of just a few words, as presented in Figure 2:

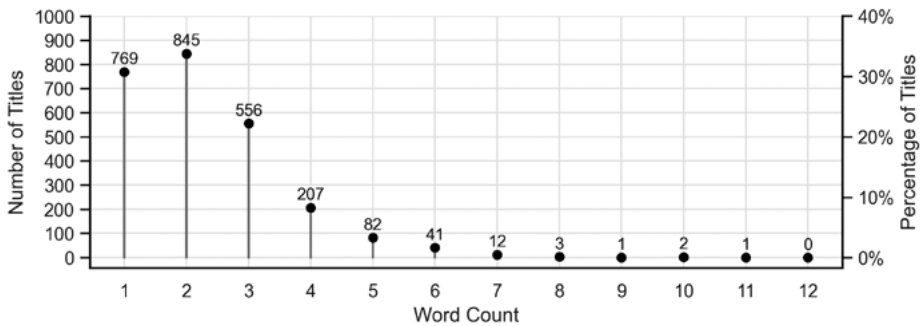


Fig. 2. Distribution of title lengths by word count.

The greatest share of titles (33.5%) consisted of only two words, followed closely by single-word titles (30.5%). This means that almost two-thirds of all Polish film titles contain only one or two words. If we go further down the graph and include slightly longer titles, we might conclude that more than 86% of all Polish movies have titles consisting of up to three words, and almost 95% have up to four words. Although the longer ones might catch attention, in fact, they are a marginal phenomenon.

How can we interpret these findings? First, the gradual decrease in title length may illustrate the slow transition from descriptive to seductive titles,³⁸ longer solutions enhancing informativity while shorter – recallability.³⁹ Second, the gradual decrease in average character count per title word may reflect a move towards readability. Third, it may also reflect the growing awareness of title design challenges; shorter names facilitating compact, efficient visual solutions. Fourth, the average word count of Polish film titles can act as an interesting counterpoint to the titling practices in English, Spanish, French, and German printed texts, where, according to Christiane Nord, *the optimum length of a book title (that is, a simple title or the main part of a title-subtitle combination) is between three and four words, with longer titles being made more memorable by using original structures, catch words, or a rhythmical pattern*. Consequently, a hypothesis may be proposed that film titling requires brevity by nature, but this should be verified using a multi-language corpus.

The second feature that we counted in an automated manner was all recurrent words in the dataset. Given the inflectional nature of the Polish language, which abounds in multiple grammatical forms of the same word, we performed automatic lemmatization of all titles in our dataset using the `pl_core_news_lg` pipeline in SpaCy – an open-source software library for natural language processing.⁴⁰ SpaCy is widely used for tasks such as named entity recognition, part-of-speech tagging, dependency parsing, and text classification. We subsequently excluded function words, such as particles, prepositions, conjunctions, or interjections, and focused on content words, which seem more informative and responsible for revealing the subject matter or theme, genre, and quite often the valency of the films.⁴¹ The list of most frequently recurring words for our dataset as well as specific time periods is presented in Table 1.

Until 1939		1945-1989		1990-2023		All-Time Top 10	
miłość	11	być	20	miłość	31	miłość	55
tajemnica	11	dzień	17	być	31	być	52
kobieta	9	miłość	13	dzień	20	dzień	42
życie	9	pan	13	życie	15	pan	33
człowiek	8	ziemia	12	2	15	życie	30
pan	8	ostatni	11	mieć	13	człowiek	28
córka	6	rok	10	kobieta	12	kobieta	27
serce	6	wielki	9	pan	12	ostatni	27
grzech	6	noc	9	ostatni	12	noc	22
dzień	5	człowiek	8	człowiek	12	mieć	21

Tab. 1. Top 10 most frequently used content words in Polish film titles across historical periods (English translations have been provided below in the discussion).

Some things have not changed; regardless of the period, Polish filmmakers eagerly use the word *miłość* [love] in their titles. It is not only the most popular word in the entire set but also in the pre-war period and contemporary cinema. It comes third in the communist period. The other recurring words are *człowiek* [human/man], *pan* [Mr], and *dzień* [day]. Interestingly, the word *być* [to be] is not that common before the war, but it dominates later periods, coming first in both the communist era and today (here tied with “love”).

Some historical peculiarities can also be noticed, namely words that appear only in certain periods with surprising regularity. A good case in point is the word *tajemnica* [mystery], which seems to be one of the key words for pre-war Polish cinema, recurring 11 times and almost disappearing from later periods (only 3 examples in the 1945-1989 set and 5 in the 1990-2023 set, which are both much bigger). Until 1939, *tajemnica* topped the list along with the all-time favourite *miłość*. This may indicate two of the most popular genres of the period: melodrama and mystery/suspense. The other two main currents in the Polish cinema of the time, i.e., patriotic historical films and comedies, evidently do not have sim-

ilar buzzwords and are thus not identifiable in the list of most frequent words. Half-jokingly, we might be surprised that no film titled *Tajemnica miłości* [*The Mystery of Love*] was ever released at the time, as this sounds like a surefire hit.

Another noteworthy change is the appearance of the number “2” in the most recent period, which clearly signals the post-communist shift towards commercial cinema, characterized globally by the focus on serialization and brand titles.⁴² Sequels appeared in Polish cinema earlier than 1990, but they were not that common and, most importantly, they rarely followed the pattern of numbering consecutive instalments, favouring original titles to announce consecutive parts (a good case in point is the 1967 cult classic comedy *Sami swoi*, followed in 1974 by *Nie ma mocnych* and in 1977 by *Kochaj albo rzuć*). It was only in the 1980s that the numbering of sequels, a practice popularized in Hollywood in the 1970s, was introduced into Polish productions with *Vabank II, czyli riposta* (1985), and *Galimatias, czyli kogel-mogel II* (1989). After 1989, the number “2” became the fifth most frequent word, appearing in 15 titles, and these estimations do not even include sequels marked by Latin numerals.

These findings would be more meaningful if compared to data available for other countries. Unfortunately, the only relevant studies we found used incompatible methods of data collection (i.e., sampling of the titles according to pre-defined criteria, such as box-office success) and keyword search (i.e., no lemmatization), thus yielding fragmentary results. Still, Dengler’s study of recurrent themes and key words in a random sample of 7,590 American movie titles produced between 1900 and 1968 demonstrates similar patterns: he reports the highest frequency scores for the words “man”/“men” (249 instances), “love” (185 instances), “woman”/“women” (151 instances), “lady”/“ladies” (104 instances), “girl” (139 instances, as opposed to “boy” with 52 instances), “night” (113 instances), and “life” (61; as opposed to “death” with 11 instances). Interestingly, no traces of “mystery” can be found in his database. Concerning more recent findings, the only relevant article we identified covered the presence of keywords in the titles of the top 50 most popular productions in the US and in China released in 2015-2019. This study reveals the dominance of military- and masculine-related keywords in both datasets (“man,” “men,” “captain,” “daddy,” “king” for the American titles; “man,” “Mr,” “men,” “dad,” “son,” “king” for the Chinese titles) and a surprising absence of any feminine references, which are very popular in Polish film titles (the only exception being the keyword “miss” in Chinese). The other interesting observation is the marked presence of negative-valence words in the American dataset, some of the high-ranking ones including “last,” “night,” “spider,” “war,” “death,” “black,” “dark,” “secret,” “evil,” “fallen,” and “furious.”⁴³

Automated tools of analysis can be assisted by human judgement, allowing us to identify broader tendencies with only minimal engagement of manual operations, such as identifying and tagging specific items in the dataset. At least a few properties of Polish film titles can be spotted and analysed on a statistical level using this method. Therefore, we went through the dataset and manually tagged: (a) all titles containing proper names and, separately (although sometimes both categories overlap), (b) all titles containing foreign expressions. Whenever in doubt, we tried to familiarize ourselves with the plot to judge whether a given

title contains a proper name or not (which is not always obvious, especially when it comes to character nicknames).

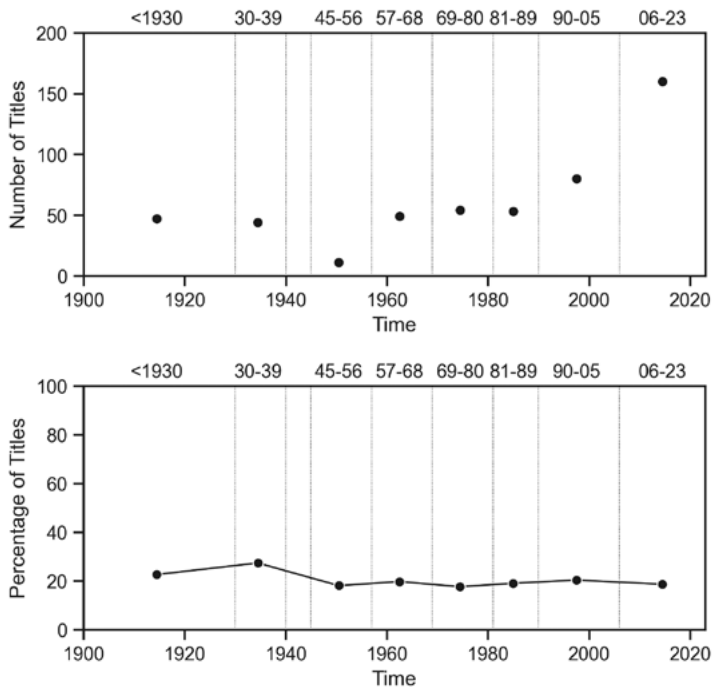


Fig. 3a and 3b. Number and percentage of proper names in Polish film titles across the analysed periods.

As Figure 3a demonstrates, the number of proper names (i.e., names of characters, places, historical events, organizations, etc.) used in Polish film titles has changed over time, ranging from just 11 in the first postwar decade to 160 in the last period (2006-2023). However, since these periods are not equal in terms of length or the volume of films produced, which varies from year to year, we also decided to check the percentage of titles that include proper names per time period, as presented in Figure 3b. Here, surprisingly, the numbers are very stable. Each period oscillates around 20%, with the only noticeable (but still not colossal) inflation in the 1930s reaching 27%. Polish filmmakers include a proper name in every fifth title regardless of the era in which they work, resulting in 19.8% instances for the entire corpus.

If we go down to the level of individual words (and not the entire titles), we will see that out of the 5,816 words forming all the titles in Polish cinema history, 607 are proper names (or parts thereof), constituting 10.4 % of the wordlist. As a point of reference, for example, as Friburger and Maurel demonstrated, in the large corpora of English and French newspaper articles, proper names also represent approximately 10% of the text.⁴⁴ This might suggest that although as much as one fifth of the titles include proper names, in fact, they probably do not appear unusually frequently – a somewhat surprising conclusion given the presence of

character names in many popular films, a tendency possibly more pronounced in mainstream Hollywood cinema.

In contrast to the stability of proper name usage over time, a notable upsurge in the frequency of the second category can be observed. It marks the use of foreign words by Polish filmmakers in recent decades as a title or part of it, no matter the origin of the word in question. A few examples appeared in the pre-war period (e.g., *Mascotte*, 1930; *Sabra*, 1933), but then this practice almost disappeared in the early communist period, only to reappear after the political thaw of the year 1956 (*Eroica*, 1957; *Rancho Texas*, 1958) and form a stable rising trend ever since, culminating in the 21st century. Thus, when counted and ordered diachronically, these instances represent a steep rising curve (Fig. 4a):

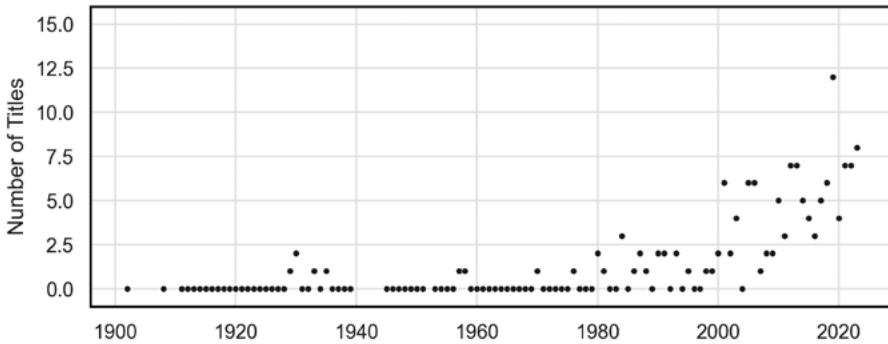


Fig. 4a. Number of film titles containing foreign words by year.

This rise in the latter part of the graph is spectacular, but it must be noted that since the datasets for later periods contain more movies, the likelihood of foreign language use increases. Therefore, we also checked the percentage of titles including foreign-language borrowings for specific periods:

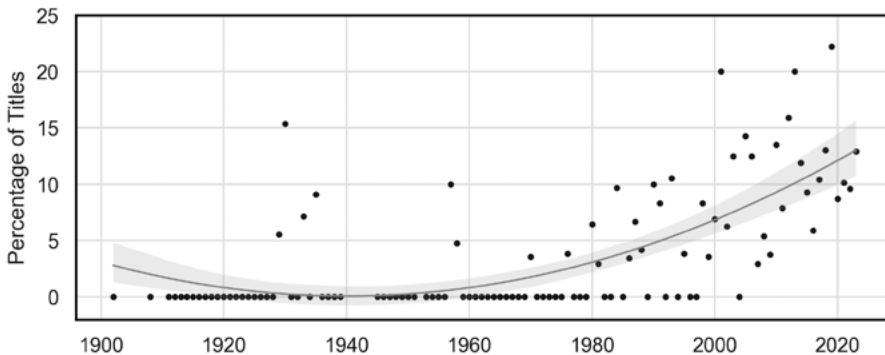


Fig. 4b. Percentage of film titles with foreign language borrowings across years, with a second-order (quadratic) trend line and a 95% confidence interval.

This does not nullify the general trend described above but rather flattens the curve, demonstrating that the contemporary rise of foreign-language-induced titles is a noticeable and stable phenomenon, although not as marked as the ab-

solute values would make us believe. Apart from the number of borrowings, we also tried to trace their sources to determine which foreign languages were used in the Polish titles (Fig. 5):

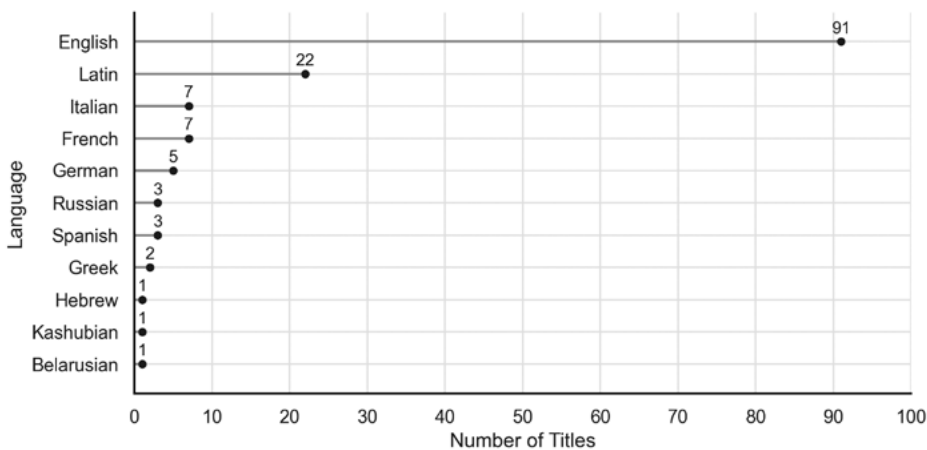


Fig. 5. Distribution of source languages for borrowings used in the Polish titles.

Undoubtedly, the dominant donor language is English, with most of the identified examples containing an Anglicism. Here included is the curious practice of unconventionally ‘Polonizing’ the spelling of foreign words, mostly English, likely to be familiar to the Polish audience. Started in the 1980s with Juliusz Machulski’s films *Vabank* (1981, adapted from the French *va banque*) and *King-sajz* (1987, adapted from the English “king size”), it recurs regularly ever since, as exemplified by *Ajlawju* (1999, from “I love you”), *Lejdis* (2008, from “ladies”), *Hardkor Disko* (2014, from “hardcore disco”), and *Autsajder* (2018, from “outsider”). This tendency is usually associated with the filmmakers’ humorous or ironic intentions: the encounter with misspelt foreignisms produces an amusing incongruity that is resolved by the audience. The percentage of titles containing Polonized borrowings is presented in Figure 6, indicating that although marginal and relatively new, it is a rising phenomenon in Polish title-giving:

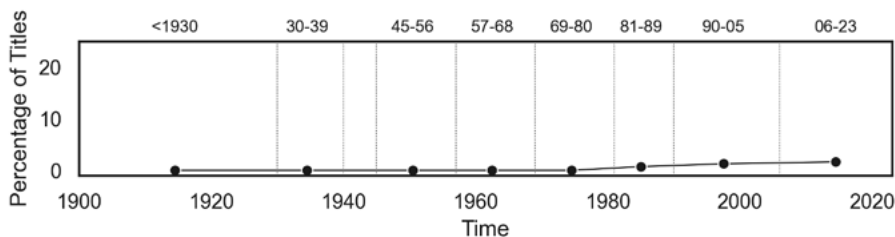


Fig. 6. Percentage of titles containing unconventionally Polonized loanwords in the overall film production across the analysed periods.

A few other peculiarities can be observed. Surprisingly, many Latin references can be found in Polish titles. These include mostly well-known phrases (*Persona non grata*, 2005; *Memento mori*, 2019), words that are either used in Polish

(*Constans*, 1980) or at least identifiable because of their similarity to their Polish counterparts (*Eucalyptus*, 2001; *Jasminum*, 2006; *Erratum*, 2010), but sometimes they are less intuitively understandable (as in the case of Artur Tomczak's little-known 2006 film *Walczak retro agit*). Another interesting influence is the Italian language, which is sometimes used to evoke musical connotations (as in *Eroica*, 1957; *Con amore*, 1976; *Belcanto*, 2010), but not exclusively (*In flagranti*, 1991; *Italiani*, 2011). Some borrowings in Polish titles were most likely intended to evoke a general feeling of exoticism without referring to a specific donor language or culture (such as the 1924 *Atakualpa* or 1979 *Kung-fu*).

Hybrid titles, using a mix of Polish and a foreign language, following the main trend, appeared a few times in the 1920s and 30s, and then only in the 1980s and later, with films such as *Smażalnia story* (1984), *Szkoła kochanków, albo cosi fan tutte* (1987). Since then, they have been in rare but regular use, with examples such as *Body/Ciało* (2015) and *Sala samobójców. Hejter* (2020; here also the Polonized spelling of "hater") or *Doppelgänger. Sobowtór* (2023). Their presence in Polish cinema is illustrated in Figure 7:

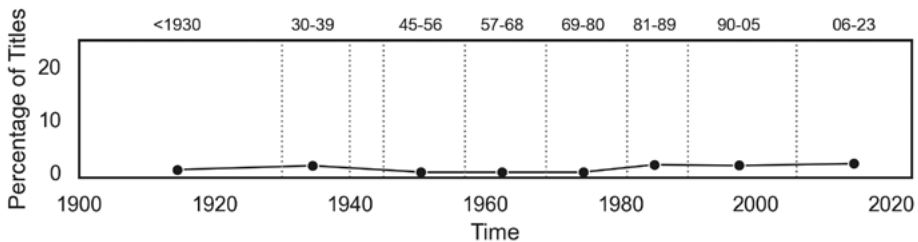


Fig. 7. Percentage of hybrid titles in Polish cinema across the analysed periods.

These findings suggest a growing interest in foreign-sounding titles among Polish filmmakers, possibly suggesting their international aspirations. According to Chung and Eoh, such naming strategies (specifically, the introduction of proper nouns, non-words, and foreign words) contribute to the distinctiveness of titles. The scholars point out that *distinctive words in a name are likely to attract consumers' attention and enhance their memorability*; hence, some filmmakers may prefer to use them in their movie-naming practices.⁴⁵ This is apparently achieved at the risk of higher uncertainty levels and lower understandability. The growing exoticization of Polish titles, alongside the aforementioned word shortening, demonstrates the filmmakers' move towards fast recall and serves as another proof of their preference for attractiveness over informativity.

The SpaCy model, which we used previously to lemmatize the titles in our dataset, also includes a part-of-speech tagging module, which allowed us to perform morphological analysis of the titles in our corpus. Because it distinguishes 18 different parts of speech, many of them appearing only marginally or not being relevant for our purposes (such as conjunctions etc.), we selected 5 dominant parts of speech and relegated the rest into the category "Other," as shown in Figure 8:

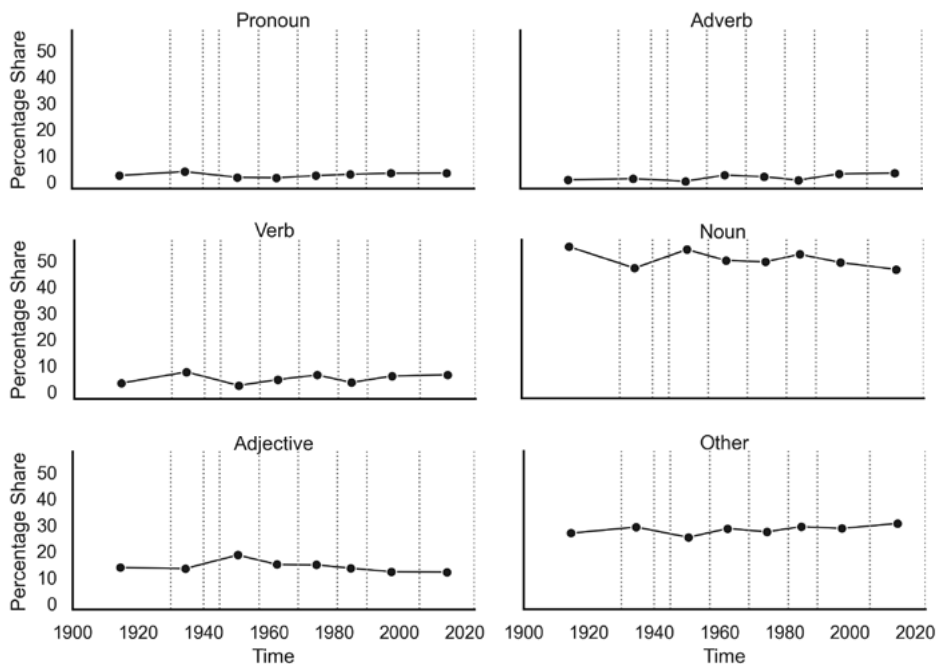


Fig. 8. Percentage share of the dominant parts of speech in the Polish original film title corpus across the analysed periods.

The most striking feature is the domination of nouns, which make up about half of all words used in Polish film titles, followed by adjectives as the distant second (around 15%), and a group of parts of speech with a visibly smaller share: verbs, pronouns, and adverbs. Although the share of all parts of speech in the entire corpus remains quite stable, with fluctuations limited to just a few percentage points between the periods, one trend is visible: a slow but steady decline in the usage of nouns, which is not compensated by an increase in any particular part of speech. Since titles can function and evolve relatively independently of the general use of language, be it literary or vernacular, additional studies are required to determine whether these changing patterns in our dataset reflect broader tendencies in the Polish language as such or in global titling practices.

Import importance: Trends in translated titles

To expand the scope of this analysis, we examined and compared film-naming trends observed in original Polish productions with those guiding foreign films distributed in Poland. We used two sources to obtain information on the Polish translations of foreign film titles. We referenced the database Oglądane w PRL (henceforth abbreviated as PRL), hosted by the Department of Film and Audiovisual Media at the University of Lodz,⁴⁶ which contains well-researched and highly reliable information on all foreign full-length feature films distributed theatrically in Poland from 1945 to 1989: a total of 5,585 titles. We have portioned

their distribution into the previously used historical periods: 1) 1945-1956: 835 titles; 2) 1957-1968: 1,909 titles; 3) 1969-1980: 1,911 titles; 4) 1981-1989: 930 titles.

Unfortunately, no similar compilation exists for the pre-war period, and it would be very demanding to create one, with only partial historical data and studies available.⁴⁷ Therefore, we excluded this period from our quantitative analysis. Similarly, no analogous database has been created for the post-communist period, and an additional difficulty in such an enterprise would be posed by the proliferation of film distribution channels. Initially, cinemas had been the main means of film dissemination; since the 1960s, it was gradually supplemented by television, and in the declining communist era, by VHS cassettes. After 1989, the range of distribution channels has steadily expanded; more and more terrestrial television channels were supplemented by cable and satellite TV; videotapes initially experienced a boom, but were later replaced by digital media, including DVDs and Blu-ray discs, and finally by internet distribution in various streaming models as well as pirate content circulation. As a result, even if we had managed to compile the cinema repertoire for that period, it would no longer serve as the main point of reference for film-titling practices in Poland.

Therefore, we drew on a much larger data resource in the form of the Internet Movie Database (IMDb),⁴⁸ which includes the "Also known as" section, or the AKAs, i.e., localized title variants for specific distribution regions/languages, including Poland. Therefore, the second corpus that we compiled, sourced from IMDb datasets,⁴⁹ were films made after 1989 that had Polish AKAs. We only considered productions with at least 100 votes in the IMDb, to filter out films with a documented Polish AKA that did not have a sizeable circulation. This is the largest set, containing 33,794 titles, of which 10,020 were released between 1990 and 2005 and 23,774 between 2006 and 2023.⁵⁰ Juxtaposed with the Polish films already analysed, a graphical representation of the size of these sets and their distribution is shown in Figure 9:

Corpus	FP	208	161		61	249	307	279	394	860
	IMDB								10020	23774
	PRL				835	1909	1911	930		
		<1930	30-39	40-44	45-56	57-68	69-80	81-89	90-05	06-23
		Period								

Fig. 9. Number of films in each film title corpus per period.

Why do we need them? There are two possible uses for a corpus containing localized Polish variants of foreign movies. First, they may serve as a point of reference to check whether the translated titles resemble their original Polish counterparts in any quantifiable way, especially considering the multiplicity of potential translation strategies available to title localizers. Among these strategies, scholars distinguish source-oriented solutions (such as Schubert's transcription or literal translation), which transfer or mimic foreign models and may suggest the recipients' implied fascination with (or forcible exposition to) exoticism, and target-oriented solutions, adaptative to Polish domestic norms, which involve free translation or transcreation. A marked discrepancy between the formal features of domestic and translated titles may indeed point to the presence of *titlese*, a neolo-

gism coined by analogy to *dubbese* or *translatese* to describe the translation-specific language of film titles that reveals foreign interferences unusual in the domestic context. At the metatextual level,⁵¹ such titles manifest themselves as representations of foreign film titles, which sets them apart from domestic productions. At the referential level, they point to different themes and subject matters explored in foreign productions. Second, since both the PRL and IMDb datasets contain the original titles of those films, we can include a new field of interlingual comparative analysis inspired by translation studies and examine the relationship of source-language titles to their Polish renditions.

For the former task, we replicated tests that did not require manual tagging and selection, which was unfeasible due to the volume of the data. With regard to the character and word count of titles translated into Polish, Figure 10 shows the results obtained:

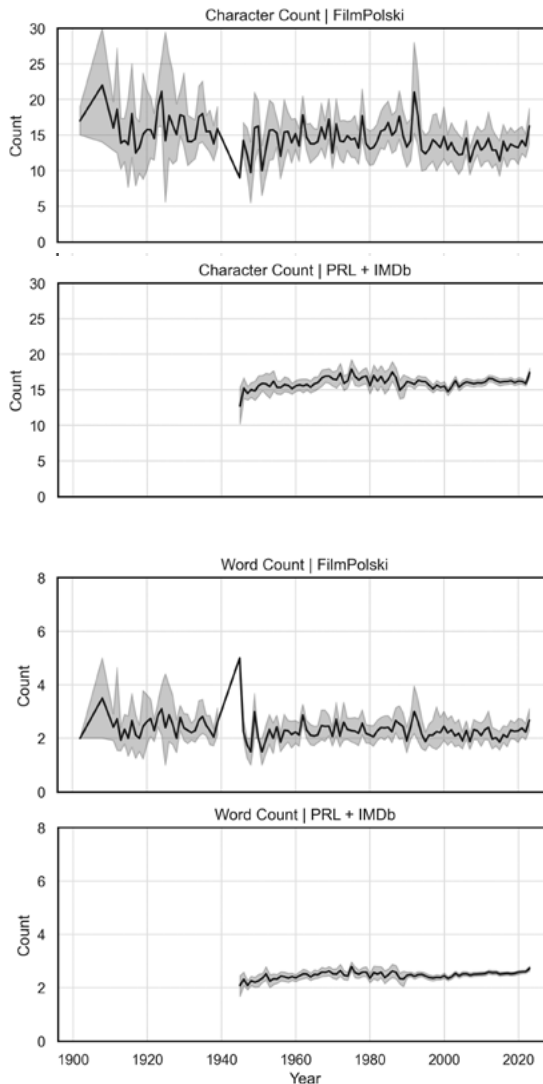


Fig. 10. Original Polish and translated title length over time in terms of character and word count.

A relative similarity of length trends in original Polish titles and foreign title translations can be noticed, although translated titles tend to be slightly longer in terms of characters used. Polish titles of foreign films distributed during the communist period averaged 16.1 characters (spaces included), while modern ones averaged 16.0, both being longer than their native Polish counterparts by 9% (PRL) and 17% (IMDb). They also used more words on average – 2.47 for PRL (against 2.31 for *filmpolski.pl*) and 2.52 (against 2.23). Comparing the total number of words and characters in the corpora, we find an average of 6.5 characters per word in People’s Poland and 6.34 characters per word in recent times, both testifying to the possibility of slightly lower readability of translated titles compared to domestic inventions.

Concerning the most frequently used content words, Table 2 lists the top 10 for communist times and the post-communist era:

1945-1989				1990-2023			
Filmpolski.pl (Orig.)		PRL (Transl.)		Filmpolski.pl (Orig.)		IMDb (Transl.)	
być	20	miłość	132	miłość	31	miłość	532
dzień	17	być	100	być	31	być	503
miłość	13	człowiek	91	dzień	20	życie	494
pan	13	życie	78	życie	15	2	443
ziemia	12	noc	70	2	15	dzień	368
ostatni	11	dzień	69	mieć	13	ostatni	353
rok	10	ostatni	58	kobieta	12	świat	278
wielki	9	dziewczyna	57	pan	12	śmierć	267
noc	9	kobieta	55	ostatni	12	noc	259
człowiek	8	śmierć	51	człowiek	12	człowiek	259

Tab. 2. Comparison of the 10 content words most frequently used in title translations and original Polish titles.

The numbers differ by order of magnitude due to the incomparable sizes of the datasets. What is highly interesting, however, is how similar these lists of the most frequently used content words are. Translations of foreign film titles in People’s Poland included as many as 6 of 10 content words most frequently used by Polish filmmakers during the period. Similarly, the most popular content words in the IMDb set of translations coincide in 7 out of 10 cases with the Polish original titles created after 1989. Furthermore, the entire top 5 are the same, only differing in order once. Interestingly, *śmierć* [death] appears to only top the lists of translated titles, reflecting the proclivity towards negative valence in American movie-naming patterns. The similarity between the PRL and IMDb sets is also strong: as many as 8 of the top 10 content words. After 1989, *dziewczyna* [girl] and *kobieta* [woman] were replaced by *świat* [world] and the numeral “2.”

This suggests that both translated and indigenous Polish titles are characterized by a limited set of similar expressions, although the Polish ones stand out a little against the foreign ones. This may testify not only to the shared referentiality of both title groups (the imported and domestic films may have touched upon similar themes and storylines), but also the appellative power of the keywords mentioned. One may presume that allusions to love and finality ‘sell’ films in Poland as they do in other parts of the world.⁵²

The issue of vocabulary variation can also be verified using computational methods. This is done using a measure called the type-token ratio (TTR), which checks the range of different words used in a set by dividing the number of unique words by all words in the set. If the set contains only unique words, the ratio is 1; however, the more words are repeated, the more it leans towards 0, representing a decreasing fraction. Smaller sets promote greater lexical diversity and offer fewer opportunities for repetition; thus, a direct comparison of all indigenous Polish titles with those from IMDb would not make sense because the latter set is several times larger. For this reason, we decided to sample these larger sets; an algorithm took a random sample of 128 words (to match the smallest set, i.e., the number of words making up the titles of 61 Polish films from 1945-1956), repeated this sampling procedure 1,000 times, and created an average of these results, which should reflect the set in question in a representative way. We measured the lexical diversity for particular periods and sets, as shown in Figure 11:

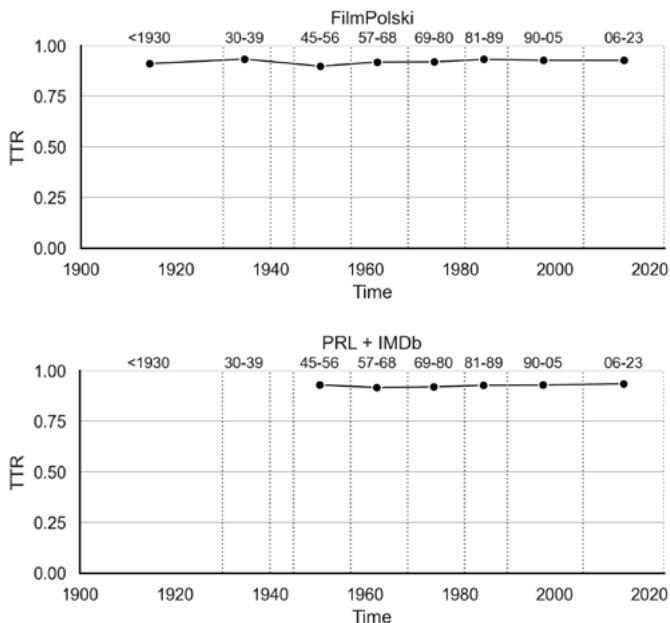


Fig. 11. Sampled lexical diversity values for original Polish and translated title corpora across the analysed periods.

Vocabulary variation in the original and translated film titles does not change much over time (apart from a slight decrease for Polish films made between 1945 and 1956), and there is virtually no difference in terms of vocabulary richness between these two categories. We also replicated the automated detection of parts of speech using SpaCy. It yielded practically the same results for both datasets, as demonstrated in Figure 12a, which focuses on the PRL and IMDb datasets, and Figure 12b, which juxtaposes the results with those previously presented for *film Polski.pl*. This may suggest more general syntactic patterns in Polish titling practices, dominated by nominal phrases.⁵³

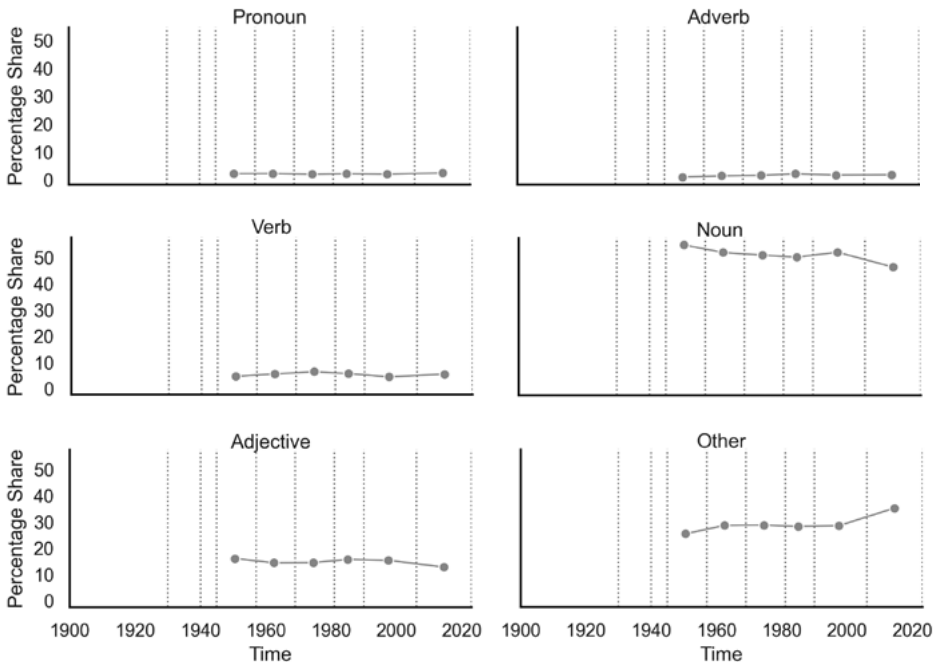


Fig. 12a. Percentage share of the dominant parts of speech in the translated film title corpus per period.

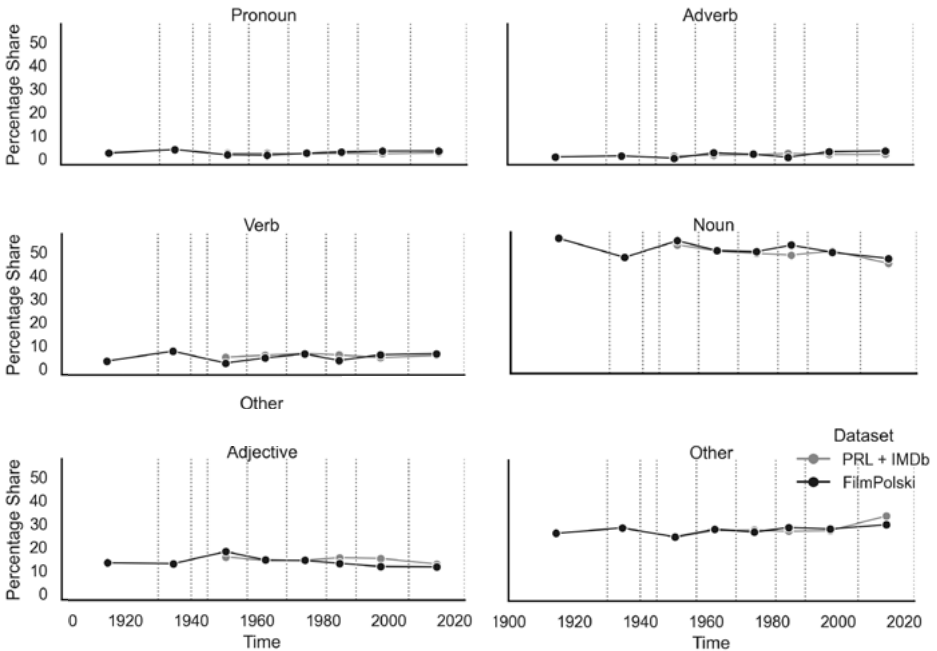


Fig. 12b. Percentage share of the dominant parts of speech in the translated and original film title corpus per period.

Overall, the original and translated titles in our dataset appear to closely resemble each other in terms of character and word count, keywords, and preference for nominalization. This may suggest the presence of film-naming norms in Poland, especially in terms of preferred syntactic forms, that run across the divides of internal vs. external production. What remains to be explored is the relationship between translations and their foreign models. This, however, requires adopting a translational perspective.

Title translation trends from an interlingual perspective

Translation scholars have distinguished multiple taxonomies of title translation strategies and adaptation procedures. Therefore, we conducted a comparative analysis of Polish translations of foreign film titles (the AKAs) and their multilingual originals, employing a relatively easy-to-automate method. We decided to verify the share of foreign films distributed in Poland under their original, unmodified title – a tendency that has been perceived as growing by multiple scholars⁵⁴ and confirmed as a global trend in our previous study.⁵⁵ We also wanted to verify which films reuse the original titles with some modification, such as an added subheading in Polish (a common practice illustrated for example by *Babe: Świnka z klasą* /1995/ or *Legion samobójców: The Suicide Squad* /2021/).

The taxonomy we used was designed to facilitate processing large title corpora: we assigned each Polish AKA in both PRL and IMDb datasets into one of three categories, for which we selected descriptive labels, thus replicating the above-mentioned study we conducted on international film titles: 1) Unchanged – the Polish title is identical to the original title, i.e., the foreign-language title is used in its unmodified form, a strategy classified by Schubert as “complete transcription”; 2) Hybrid – the Polish title transfers the original with modification, such as an added Polish subheading or partial removal of original components. This category subsumes Schubert’s “transcription with subtraction” and “transcription with addition”;⁵⁶ 3) Other – in which the Polish title belongs to neither of the above groups and involves other forms of target-oriented translation, such as transcreation, literal, or free translation.

For our trend analysis, we used a rolling mean with a window of size 5, so every point in our diagrams represents not only the corresponding year on the axis but also the preceding and succeeding 2. The minimum number of observations in the window required to have a value was 5. Figure 13 demonstrates the title translation trends in Poland since 1945:

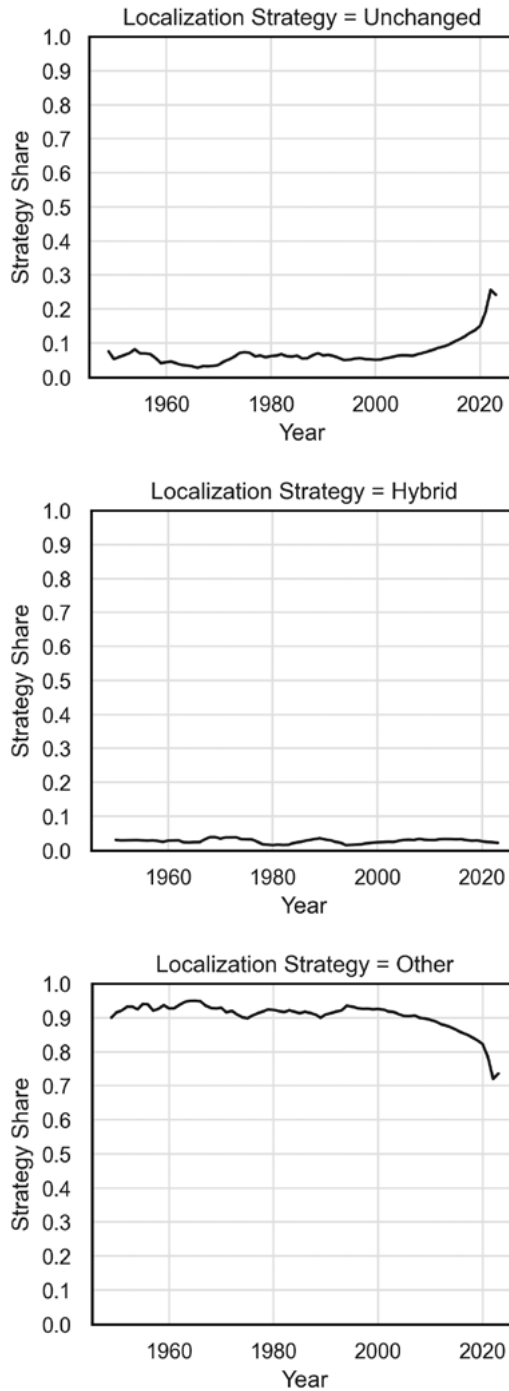


Fig. 13. Title translation trends in Poland over time divided into three broad categories concerning the reuse of source-language materials.

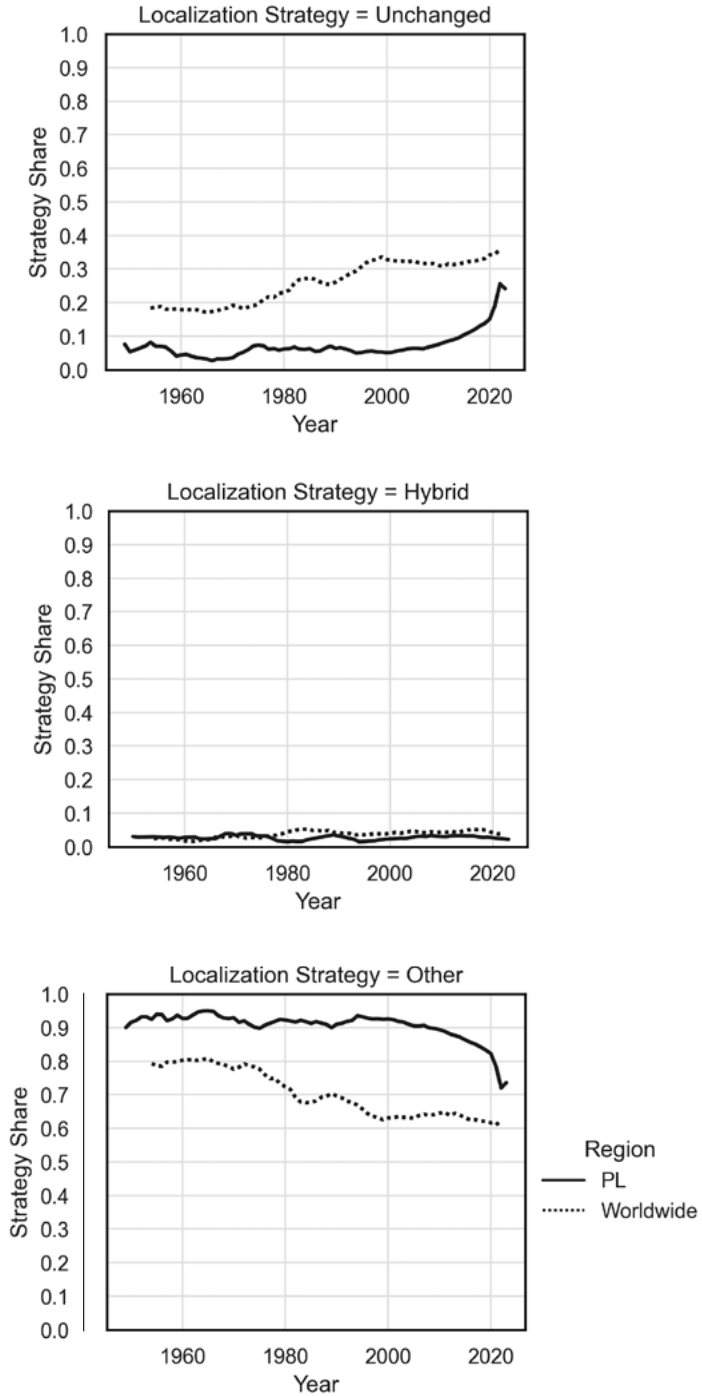


Fig. 14. Title translation trends in Poland over time against global trends.

In the earliest pre-war part of the diagram, no translation tendencies are visible because no data were available. After 1945, what can be noticed is the stable but relatively unremarkable presence of hybrid titles that utilized the source-language models with some modifications and a marked growth of the tendency to distribute foreign films in Poland under their original names in the second decade of the 21st century. Around the year 2020, the share of untranslated titles amounted to almost one-third of all foreign productions. Still, a lion's share of foreign films is distributed in Poland with translated titles that involve other strategies, which we have not yet managed to nuance and quantify.

We can contrast Polish trends with those observed in our previous study on global film distribution patterns, based on the subsets of films included in the official selection of the Cannes International Film Festival and an annual sample of the most popular productions in multilingual translation. We observed tendencies similar to those demonstrated in Figure 13, such as the niche status of hybrids and a steady increase in non-translated titles at the expense of others, except that in our country, this process was slower for most of the period under study and then accelerated greatly towards the end of the second decade, as if catching up with the rest of the world. This suggests that systemic factors, shaped by the political and economic situation, played a major role in the evolution of translation strategies in Poland. Under communism, Poles relied heavily on Polonized film titles, and the global trend of an ever-increasing presence of original titles in international film distribution has bypassed Poland by a wide margin. However, since the 21st century, Polish film distributors have started to catch up, with foreign transplants rapidly approaching the global average.

Conclusions

Our preliminary analysis of Polish film-naming practices has confirmed certain intuitions and raised additional questions. Based on the criteria adopted, we learned that the titles of original Polish films have not evolved much over time in terms of length, part-of-speech composition, vocabulary richness, and most frequently used content words. We also found that in some respects, such as the presence of love themes or the use of foreign vocabulary, there was greater affinity between the pre-war and post-communist periods than between either of them and the Polish People's Republic. This suggests that perhaps it is the nonpublic, commercial nature of the film industry that conditions certain developments in film titling. Yet other explanations are also possible: perhaps, for political reasons, the communist state suppressed certain broader cultural phenomena unrelated to the commercial orientation of the film industry that would otherwise have emerged.

It is also interesting to observe how, in some respects, the 1980s resemble and foreshadow the 1990s and contemporary times more closely than the preceding decades. Until now, the communist era has been mostly treated as a heterogeneous, although distinct, period in the history of Polish cinema, with a cut-off date of 1989 and subsequent changes in the organization of film production. Meanwhile, some recent publications indicate the need to re-evaluate

this historical assumption, showing that, in fact, it was the 1980s and the introduction of martial law in Poland (1981-1983) that marked the breakthrough in Polish filmmaking and the passage towards commercial cinema, entertainment, and genre films under the umbrella of the communist state – more than the politically symbolic date of 1989.⁵⁷ Restricted to film titles, our data also points to this interpretation, for example, in the use of numerals and foreign-sounding titles, often Polonized or hybridized.

We also found that domestic and translated titles resemble each other quite closely by the same criteria, although the discovery of potential discrepancies may require the application of additional measures, which are beyond the scope of this paper. We confirmed that Polish viewers are currently exposed to an increasing number of foreign film titles retained in their untranslated form. This may indicate the growing influence of major distribution companies that are closely affiliated with large production studios and shape the global tactics of film title translation.⁵⁸

As it is preliminary, this study also deserves further refinement and elaboration on our part. Our future research avenues include: 1) a more in-depth analysis of semantic fields evoked by Polish and translated titles; 2) a more in-depth syntactic analysis; 3) replication of tests on larger corpora, perhaps involving other languages; 4) effective automation, e.g., by means of machine learning, of the tasks currently conducted manually, such as keyword delimitation, proper name tagging, and foreign language recognition; 5) refinement of the adopted translation strategy taxonomy and effective automation of strategy recognition; 6) comparative studies, allowing us to juxtapose our data and findings with other areas, such as literary or musical titles, and regularities or changes observed in other text corpora.

Although our study is clearly Polonocentric, we hope it can serve as a blueprint for exploring other national film industries. We tried to address one of the most pressing needs of the computational humanities now – to produce hard data in the yet uncharted territories, in this case of Polish film titles, providing a foundation for possible future interpretations, comparisons, and further studies by other scholars, helping to advance the field and its methods, and consequently creating opportunities for the enrichment of the traditionally delineated realm of film studies.

¹ G. Genette, *Paratexts: Thresholds of Interpretation*, trans. M. Lewin, Cambridge University Press, Cambridge 1997.

² M. Hendrykowski, *Slowo w filmie*, PWN, Warszawa 1982, p. 158.

³ Ibidem.

⁴ J. Chung, J. Eoh, "Naming Strategies as a Tool for Communication: Application to Movie Titles", *International Journal of Advertising* 2019, vol. 38, no. 8, p. 1140.

⁵ C. Nord, "Text-Functions in Translation: Titles and Headings as a Case in Point",

Target 1995, vol. 7, no. 2, pp. 261-84.

⁶ J. Ross, "How Global Conglomerates Influence Translation Practice: Film Title Translation in Turkey", in: *Translation and Global Spaces of Power*, eds. S. Baumgarten, J. Cornella-Detrell, Multilingual Matters, Bristol 2018, pp. 194-210.

⁷ L. Hoek, *La marque du titre: Dispositifs sémiotiques d'une pratique textuelle*, De Gruyter Mouton, Berlin – New York 1981; A. Rothe, *Der Doppeltitel: Zu Form und Geschichte einer literarischen Konvention*, Verlag

- der Akademie der Wissenschaften und der Literatur, Mainz 1970; idem, *Der literarische Titel. Funktion, Formen, Geschichte*, Klostermann, Frankfurt am Main 1986; G. Genette, op. cit.
- ⁸ M. Hendrykowski, op. cit.
- ⁹ A. Böhnke, *Paratexte des Films*, Transcript Verlag, Bielefeld 2007; J. Gray, *Show Sold Separately: Promos, Spoilers, and Other Media Paratexts*, New York University Press, New York – London 2010; J.-C. Horak, "Film Titles: Theory and Practice", in: *Saul Bass: Anatomy of Film Design*, University Press of Kentucky, Lexington 2014, pp. 71-128; G. Stanitzek, "Texts and Paratexts in Media", *Critical Enquiry* 2005, vol. 32, no. 1, pp. 27-42; idem, "Reading the Title Sequence", trans. N. Aplevich, *Cinema Journal* 2009, vol. 48, no. 4, pp. 44-58; W. Straw, "Letters of Introduction: Film Credits and Cityscapes", *Design and Culture* 2010, vol. 2, no. 2, pp. 155-166.
- ¹⁰ M. Fakharzadeh, "A Sociological Approach to Official and Non-Official Audiovisual Translators' Practice in Iran: The Case of Movie Title Translation", *Journal of Intercultural Communication Research* 2022, vol. 51, no. 4, pp. 430-449.
- ¹¹ L. Yin, "On the Translation of English Movie Titles", *Asian Social Science* 2009, vol. 5, no. 3, pp. 171-173.
- ¹² M. S. Peña-Cervel, "Motivating Film Title Translation: A Cognitive Analysis", *Círculo de Lingüística Aplicada a la Comunicación* 2016, vol. 63, pp. 301-339; J. Santaemilia, B. Soler-Pardo, "Translating Film Titles: Quentin Tarantino, on Difference and Globalisation", *Babel* 2014, vol. 60, no. 2, pp. 193-215.
- ¹³ L. Pungă, G. Badea, "Lost in Translation: On Some English Movie Titles and Posters and Their French and Romanian Equivalents", *Skase* 2018, vol. 11, no. 1, pp. 55-68.
- ¹⁴ P. Gabrić, I. Brajković, L. Licchetta, J. Bezuh, D. Kelčec Ključarić, "A Comparative and Diachronic Analysis of Film Title Translations and Appellative Effect Transfer into Croatian and German", *Hieronymus: Journal of Translation Studies and Terminology* 2017, vol. 4, no. 1, pp. 1-27.
- ¹⁵ C. Galilej, "Z problematyki tłumaczeń anglojęzycznych tytułów filmowych na język polski", *Roczniki Humanistyczne* 2012, vol. 8, pp. 15-29; M. Nowak, "Tłumacz jako autor tytułów filmów i seriali", *Białostockie Archiwum Językowe* 2012, vol. 17, pp. 163-180; A. Naruszewicz-Duchlińska, "Onimiczna kreatywność translatorska w tytułach filmów i seriali", *Prace Językoznawcze* 2016, vol. 18, no. 3, pp. 147-158; J. Suwaj, "Problemy kulturowe w tłumaczeniu anglojęzycznych tytułów filmów na język polski", *Rocznik Przekładoznawczy* 2018, no. 13, pp. 225-234; J. Suwaj, "Przekład tytułów filmów anglojęzycznych i ich funkcjonowanie na rynku polskim", *Między Oryginałem a Przekładem* 2014, vol. 20, no. 1, pp. 189-201.
- ¹⁶ K. Wook-Dong, "Lost in Translation", *Babel* 2017, vol. 63, no. 5, pp. 729-745.
- ¹⁷ M. Mocarz-Kleindienst, "Tytuły filmów 'oskarowych' w wielojęzycznym przekładzie", *Acta Polono-Ruthenica* 2018, vol. 23, no. 2, pp. 75-84.
- ¹⁸ G. Gwóźdź, P. Mamet, E. Pawlikowska-Asendrych, B. Rusek, A. Skwara, M. Gola-Brydnyk, "Tłumaczenia tytułów filmów z Jamesem Bondem na język francuski, niemiecki i polski", in: *Przeźrenie przekładu*, eds. O. Lubocha-Kruglik, O. Małysa, Wydawnictwo Uniwersytetu Śląskiego, Katowice 2016, pp. 73-88.
- ¹⁹ C. Schubert, "Die Appellwirkung englischer Filmtitel und ihrer deutschen Neutitel: Techniken interkulturellen Transfers", *Arbeiten aus Anglistik und Amerikanistik* 2004, vol. 29, no. 2, pp. 239-259.
- ²⁰ P. Gabrić et al., op. cit.
- ²¹ M. Fakharzadeh, op. cit..
- ²² P. Gabrić et al., op. cit.
- ²³ M. Mocarz-Kleindienst, op. cit.
- ²⁴ J. Ross, op. cit.
- ²⁵ G. Bae, H. Kim, "The Impact of Movie Titles on Box Office Success", *Journal of Business Research* 2019, vol. 103, pp. 100-109.
- ²⁶ S. Sood, X. Drèze, "Brand Extensions of Experiential Goods: Movie Sequel Evaluations", *Journal of Consumer Research* 2006, vol. 33, no. 3, pp. 352-360.
- ²⁷ X. Xiao, Y. Cheng, J.-M. Kim, "Movie Title Keywords: A Text Mining and Exploratory Factor Analysis of Popular Movies in the United States and China", *Journal of Risk and Financial Management* 2021, vol. 14, no. 2, p. 68.
- ²⁸ R. Dengler, S. J., "The Language of Film Titles", *Journal of Communication* 1975, vol. 25, pp. 51-60.
- ²⁹ J. Chung, J. Eoh, "Naming Strategies as a Tool for Communication: Application to Movie Titles", *International Journal of Advertising* 2019, vol. 38, no. 8, pp. 1139-1152.
- ³⁰ Ibidem, p. 1149.
- ³¹ Ibidem, p. 1148.
- ³² Ibidem, p. 1149.

- ³³ Internetowa Baza Filmu Polskiego, film Polski.pl, <https://www.filmpolski.pl/fp/index.php> (accessed: 10.06.2024).
- ³⁴ The Second World War (1940-1944) has not been included in any category because no new Polish feature films were made at that time.
- ³⁵ J. Chung, J. Eoh, op. cit., p. 1143.
- ³⁶ T. Moździerz, "Proporcja liczby wyrazów faktycznych i przeciętnych tekstu w kontekście siedmiostopniowej skali trudności jaśnopis.pl", *Acta Universitatis Lodzianensis* 2021, vol. 28, p. 326.
- ³⁷ Idem, "Długość przeciętnego polskiego wyrazu w tekstach pisanych w świetle analizy korpusowej", *Acta Universitatis Lodzianensis* 2020, vol. 27, pp. 177-192.
- ³⁸ M. Hendrykowski, op. cit., p. 158.
- ³⁹ J. Chung, J. Eoh, op. cit., p. 1143.
- ⁴⁰ SpaCy, https://spacy.io/models/pl/#pl_core_news_lg (accessed 01.06.2024).
- ⁴¹ J. Chung, J. Eoh, op. cit.
- ⁴² S. Sood, X. Drèze, op. cit.
- ⁴³ X. Xiao, Y. Cheng, J.-M. Kim, op. cit., pp. 67-68.
- ⁴⁴ N. Friburger, D. Maurel, "Textual Similarity Based on Proper Names", *Proceedings of the Workshop Mathematical/Formal Methods in Information Retrieval*, ACM SIGIR 2002, pp. 155-167.
- ⁴⁵ J. Chung, J. Eoh, op. cit., p. 1142.
- ⁴⁶ G. Balski, K. Klejsa, *Watched in PL: Database for Polish Film History*, www.ogladanewprl.uni.lodz.pl (accessed: 10.06.2024).
- ⁴⁷ An excellent example of which, although limited both temporally and geographically and thus insufficient for our purposes, is W. Świdziński, *Co było grane? Film zagraniczny w Polsce w latach 1918-1929 na przykładzie Warszawy*, IS PAN, Warszawa 2015.
- ⁴⁸ IMDb, <https://imdb.com> (accessed: 15.05.2024).
- ⁴⁹ IMDb, <https://developer.imdb.com/non-commercial-datasets/> (accessed: 15.05.2024).
- ⁵⁰ Information courtesy of IMDb (<https://imdb.com>). Used with permission.
- ⁵¹ C. Nord, op. cit.
- ⁵² X. Xiao, Y. Cheng, J.-M. Kim, op. cit., p. 68.
- ⁵³ C. Galilej, op. cit., pp. 15-29.
- ⁵⁴ E.g., J. Santaemilia, B. Soler-Pardo, op. cit.; Gabrić et al., op. cit.; C. Galilej, op. cit.
- ⁵⁵ A. Hołobut, M. Rapacz, M. Stelmach, "Translating Film Titles: A Qualitative and Quantitative Approach", in: *The Palgrave Handbook of Multilingualism and Language Varieties on Screen*, eds. I. Ranzato, P. Zabalbeascoa, Palgrave Macmillan 2024 (forthcoming).
- ⁵⁶ C. Schubert, op. cit.
- ⁵⁷ See: *Kino w cieniu kryzysu. Studia i szkice o polskiej kinematografii pierwszej połowy lat 80.*, eds. P. Zwierzchowski, P. Kurpiewski, Bydgoszcz 2022; M. Piepiórka, "Jałowa ziemia. Jak kapitalizm wykończył polskie kino gatunkowe", *Ekrany* 2024, no. 70, pp. 29-35.
- ⁵⁸ J. Ross, op. cit.

Agata Hołobut

Linguist and translation scholar. She works as an Associate Professor at the Institute of English Studies, Jagiellonian University in Kraków, Poland. Her main areas of interest include literary and screen translation, visual arts, humour studies, cognitive semiotics, and digital humanities. She has published papers on literary, audiovisual, and intersemiotic translation, as well as screen adaptation and film remake; she has also co-authored, together with Monika Woźniak and Jan Rybicki, the monograph *Historia na ekranie: Gatunek filmowy a przekład audiowizualny [History on Screen: Film Genre in Audiovisual Translation]* (2017).

Maciej Rapacz

Computer scientist, software and data engineer, and a PhD candidate at the AGH University of Kraków. His interests include the application of computer science to solving problems in the humanities, with a particular focus on machine translation and natural language processing methods for Ancient Greek.

Miłosz Stelmach

Research and Teaching Assistant in film studies at the Jagiellonian University in Kraków. Editor-in-chief of the film magazine *Ekrany*. His main research interests are cinematic modernism, cinema of Eastern Europe, and transformations of contemporary arthouse cinema.

Bibliography

- Bae, G., Kim, H.** (2019). The Impact of Movie Titles on Box Office Success. *Journal of Business Research*, (103), pp. 100–109. <https://doi.org/10.1016/j.jbusres.2019.06.023>
- Chung, J., Eoh, J.** (2019). Naming Strategies as a Tool for Communication: Application to Movie Titles. *International Journal of Advertising*, 38 (8), pp. 1139–1152. <https://doi.org/10.1080/02650487.2019.1593719>
- Dengler, R., S. J.** (1975). The Language of Film Titles. *Journal of Communication*, 25, pp. 51–60.
- Fakharzadeh, M.** (2022). A Sociological Approach to Official and Non-Official Audio-visual Translators' Practice in Iran: The Case of Movie Title Translation. *Journal of Intercultural Communication Research*, 51, (4), pp. 430–449. <https://doi.org/10.1080/17475759.2021.1966075>
- Friburger, N., Maurel, D.** (2002). Textual Similarity Based on Proper Names. In: *Proceedings of the Workshop Mathematical/Formal Methods in Information Retrieval*, ACM SIGIR 2002, pp. 155–167.
- Gabrić, P., Brajković, I., Licchetta, L., Bezuh, J., Kelčec Ključarić, D.** (2017). A Comparative and Diachronic Analysis of Film Title Translations and Appellative Effect Transfer into Croatian and German. *Hieronymus: Journal of Translation Studies and Terminology*, 4 (1), pp. 1–27.
- Genette, G.** (1997). *Paratexts: Thresholds of Interpretation* (trans. J. E. Lewin). Cambridge: Cambridge University Press.
- Hendrykowski, M.** (1982). *Słowo w filmie*. Warszawa: PWN.
- Moździerz, T.** (2020). Długość przeciętnego polskiego wyrazu w tekstach pisanych w świetle analizy korpusowej. *Acta Universitatis Lodziensis*, (27), pp. 177–192.
- Nord, C.** (1995). Text-Functions in Translation: Titles and Headings as a Case in Point. *Target*, 7 (2), pp. 261–84.
- Ross, J.** (2018). How Global Conglomerates Influence Translation Practice: Film Title Translation in Turkey. In: Baumgarten, S., J. Cornella-Detrell (eds.), *Translation and Global Spaces of Power* (pp. 194–210). Bristol, Blue Ridge Summit: Multilingual Matters. <https://doi.org/10.21832/9781788921824-015>

- Schubert, C.** (2004). Die Appellwirkung englischer Filmtitel und ihrer deutschen Neutitel: Techniken interkulturellen Transfers. *Aaa-Arbeiten aus Anglistik und Amerikanistik*, (29), pp. 239-259. <https://api.semanticscholar.org/CorpusID:151513454>
- Sood, S., Drèze, X.** (2006). Brand Extensions of Experiential Goods: Movie Sequel Evaluations. *Journal of Consumer Research*, 33 (3), pp. 352-360. <https://doi.org/10.1086/508520>
- Xiao, X., Cheng, Y., Kim, J.** (2021). Movie Title Keywords: A Text Mining and Exploratory Factor Analysis of Popular Movies in the United States and China. *Journal of Risk and Financial Management*, 14 (2), p. 68. <https://doi.org/10.3390/jrfm14020068>

Słowa kluczowe:

tytuł filmowy;
analiza metadanych;
tłumaczenie tytułu;
historia kina;
kino polskie

Abstrakt

Agata Hołobut, Maciej Rapacz, Miłosz Stelmach

Tytuły robocze. Analiza obliczeniowa praktyk w nazewnictwie filmów na polskim przykładzie

W artykule autorzy przyglądają się historycznym trendom w praktykach tworzenia tytułów filmowych w Polsce, opierając się na badaniach ilościowych nad metadanymi filmowymi. W poszukiwaniu prawidłowości w rodzimym nazewnictwie analizują korpus 2519 filmów wyprodukowanych w Polsce do 2023 r.: badają średnią długość tytułów, dublowanie się ich, najczęstsze słowa znaczące występujące w tytulaturze typowej dla poszczególnych etapów historii polskiej kinematografii, bogactwo leksykalne i obecność nazw własnych oraz zapożyczeń w tytułach polskich produkcji. Następnie przenoszą niektóre z powyższych badań na korpus filmów obcych, których tytuły przełożono na język polski. Uwzględniają w nim 5585 filmów dystrybuowanych w Polsce w latach 1945-1989 oraz 33794 filmy, które ukazały się w Polsce po roku 1989. Autorzy porównują specyfikę tytułów oryginalnych i tłumaczonych, szukając śladów językowej interferencji. Badają również, jakie strategie przekładu dominują w Polsce od 1945 r., skupiając się na udziale tytułów w wersji oryginalnej oraz zmodyfikowanych przez dodatek lub opuszczenie wybranego elementu.