Marian Żmigrodzki
Uniwersytet Komisji Edukacji Narodowej, Kraków, Polska
ORCID 0000-0001-6783-0517

The techniques of introducing English acronyms in the Polish medical texts: a corpus-based study

Słowa kluczowe: akronim, zapożyczenie, język specjalistyczny, medycyna, język medycyny, kontakt językowy, badanie na korpusie

Keywords: acronym, borrowing, specialized language, medicine, medical language, language contact, corpus-based study

1. Introduction

The influence of English on other languages all around the world has continuously been on the rise throughout the last decades, and the Polish language has not been an exception to that process. The exposure to English has been propelled by the rapid expansion of the Internet and its accessibility to the average person. Direct and instant access to an almost infinite resource on every imaginable topic has pushed the already galloping globalization into overdrive.

Out of many different dimensions of this phenomenon, language contact and how the speakers incorporate foreign language words into their mother tongue have been of particular interest to linguistic researchers. English words are making their way into the language of different kinds of Polish speakers. The speakers of Polish are incorporating more and more English borrowings to supplant our native words; into both the colloquial language, as well as specialized languages connected to different professions.

In this paper, we analyse the appearance of English acronyms adopted into the language of medicine in Poland, on the basis of a corpus created out of 7 issues of a popular medical journal. We look at their usage, frequency, the methods of introducing them into the text by the authors, as well as the level of integration into the Polish language, and we attempt to find correlations between these factors. The aim is to establish the most common English medical acronyms in the chosen corpus and

determine which types of meanings are most likely to be communicated through them. This research is intended to serve as the foundation for further inquiry into medical acronyms, with the ultimate goal of creating practical tools for medical professionals, such as a dictionary of English acronyms.

2. Medical Polish as a specialized language

Languages used by practitioners of the same professions have been a subject of linguistic study both in Poland and worldwide for decades. Terms, such as *specialized language, language for specific purposes, professional language, sociolect, professiolect* (professional language), *technolect* and many others have been used throughout that time by different scholars, and their definitions remain unclear and differ between researchers. We shall now explore some of these definitions.

One of the early definitions (Szulc 1984: 106–107) epitomizes a specialized language as a particular form of a general language adjusted to more accurately describe a specific type of knowledge or technology. Wilkoń (1987: 101) describes *professiolect* as a type of language that combines informality with special vocabulary related mostly to field-specific terminology. Grabias (1997: 147), on the other hand, considers the language of a professional group to be a specific type of *sociolect*, which includes more vocabulary related to the profession than to other areas of life. The wider range of vocabulary is caused by the need to refer to different concepts and items found in said field. It is important to note that specialized languages are not independent languages, but instead, they are built with some other existing language as a foundation (Grucza 2002: 15). According to Grucza (2006: 34), specialized languages are only fully autonomous in their functional aspect, and partially autonomous in terms of specialized vocabulary, which is one of the most important parts that constitutes a specialized language.

The language used by the Polish practitioners of medicine can thus be considered a specialized language. It is abundant in special terminology exclusively needed to describe concepts connected to medicine, many of which are borrowed from other languages, mainly English. At the same time, however, it is built on the foundation of the Polish language and both its syntax as well as pronunciation should not feel alien to an average Polish speaker.

Medical language has been broadly discussed by researchers both internationally and domestically in Poland. Notable English-speaking works fully or partially dedicated to medical language include e.g. Warner (1976), Pilegaard (2000), and Fleischman (2003). In Poland, this issue has been tackled by linguists such as Górnicz (2003) and Grucza (2008), and more.

3. Indeterminacy of the terms acronym and initialism

Historically, the typology of abbreviations has been surrounded by considerable vagueness and inconsistency. Following Cannon (1989: 106), the way in which dictionaries have been explaining this term ever since the 15th century has greatly contributed to the confusion and fuzzy limits between different categories of abbreviations.

The term *acronym* itself in its current meaning was coined by Basil Davenport in his article *Initials into Words* in *American Notes and Queries* in 1943, as "[...] words made up of the initial letters or syllables of other words [...]". The definition was later expanded in 1950 by Funk & Wagnalls dictionary to also include using the initial syllable of a word, as well as agglutinating the initial and the final syllable of a word (Baum 1995: 103), which already muddied the waters by including different types of word shortening such as blending. Furthermore, the *Acronyms Dictionary: First Edition* (1960) contributed further to the problem by its inclusion of alphabetic designations, contractions, and initialisms. Jennifer Mosman, the author of the *Dictionary of Abbreviations, Acronyms and Initialisms* (1993), argued that abbreviations, acronyms, and initialisms are language units at the same level, and thus, acronyms are not considered a type of abbreviation (1993: 9).

Even in modern times, there does not seem to be a compromise on the finer points of the definition. Linguists offer differing accounts and viewpoints on the issue. According to Quirk and Greenbaum (1996: 449) *initialisms*, which they refer to as *alphabetisms*, are just *acronyms* that are pronounced as separate letters. Other scholars such as Bauer (1988: 39) and Booij (2005: 20) argue for the separation of *acronyms* and *initialisms* into completely independent categories.

On the more practical side, popular dictionaries, such as *Merriam-Webster* or *Cambridge Dictionary*, tend to agree that *acronyms* are words formed from either the initial letters of each of the successive units of a compound term or initial syllables of said units, pronounced as a single word. *Initialisms*, also called *alphabetisms*, on the other hand, tend to be defined as items formed from initial letters or syllables, but which are pronounced as separate letters. As Crystal (2018: 120) notes, some linguists do not acknowledge this division and use the term "acronym" for both these linguistic units. As in this research, we are analysing a written corpus, the issue of pronunciation will not be at the forefront. In this paper, both *acronyms* and *initialisms* will be considered under the umbrella term *acronym*.

4. Acronyms as lexical borrowings

Language contact has been a subject of research by many linguistic scholars over the years including Haugen (1950), Weinreich (1953), and many more. Polish linguists adopted and conformed classifications created by their preceding notable names in the field. Following Witalisz (2016: 21–25, 37–39), acronyms belong to the category of lexical borrowings otherwise known as *loanwords*. Loanwords are a wide category of borrowings where both the meaning and the form of the foreign lexeme are borrowed, which causes these lexical items to be easily noticeable as foreign to the speakers of the language they get imported to. Some of these borrowings differ in form from their original English equivalents, because they get partially assimilated to the rules of Polish orthography. Witalisz (2016: 39) provides *TV*, *NATO*, *VAT*, *GPS*, *PIN*, *PIT*, *ABS*, *WWW*, *AIDS*, *LCD*, *PC* and more, as some examples of the frequently used borrowed English *acronyms* in the general Polish language.

Due to its international nature, and scientific basis, the field of medicine is prone to borrowing acronyms. Despite that, the issue has not seen much scientific scrutiny.

The most notable publication related to English medical acronyms intended for Polish medical professionals is the dictionary *Medical slang & acronyms* (2008) by Linda Perlińska and Janusz Krzyżowski.

In this paper, the emphasis is placed on the acronyms related to the general theme of medicine and its connected fields. We will take a closer look at the semantic groups they belong to and how this membership correlates to other aspects of the research, such as the frequency of these acronyms within the corpus and the methods authors used to introduce them to their reader.

5. The corpus and method

The corpus created specifically for this research was built from seven issues of a Polish journal on general medicine *Kurier Medyczny* [Eng. *Medical Courier*]. These issues were all released in 2021 and constitute a full annual run of the journal excluding issue 6, due to its unavailability on-line. Issues from year 2021 were chosen, because at the time of this paper's inception, this was the most recent complete year of *Kurier Medyczny* available on the website.

This journal in particular was chosen for the research because of the broad range of medical fields it discusses, as well as its accessibility to regular general practitioners of medicine. The intention was to analyse the language in which doctors communicate with other members of their professional group, rather than to scrutinize scholarly language used in academic papers connected to medicine.

The research was conducted through manual excerption of the tokens from the corpus, as there exist no tools capable of automatizing the process of finding English borrowings in Polish texts. The corpus consisted of 122 files in the .pdf format, each containing one article of varying lengths between 1 and 5 pages. Due to the formatting of said files, the articles were left in the files throughout the excerption process itself. They were only modified and worked on with the help of text and .pdf file editors later to acquire tertiary data such as the total number of words present within the corpus.

Due to the nature of the source material, certain pieces of text present in the journal were ignored for the analysis. The parts of texts that were ignored included advertisements, occasional bibliography under some articles, the text on graphs and pie charts, as well as any additional text boxes that were not related to the main body of the article. Certain parts of the text were also rewritten as quotations in a bigger font to presumably draw the readers' attention to them. The emboldened quotations also had to be excluded from the analysis.

The findings were then catalogued in separate text files, with one text file corresponding to one issue. Every instance of an English acronym was noted, along with the context it appeared in. Alongside the context, the parts where the acronym was explained to the readers were also separately added, if applicable. At the end of the process, one singular file was created which catalogued unique acronyms along with their total number of occurrences and select quotes. The number of occurrences was counted manually using the files previously created for each issue separately. The total number of all words within the corpus amounted to 226,340.

6. Findings and discussion

6.1. Presentation of the findings and the frequency of tokens within the corpus

A total of 326 unique English acronyms were found within the corpus. Including the repetitions, the total number of borrowed acronyms was 2724. As the corpus included 226,340 total words, foreign acronyms constituted 1.2% of all words, or, in other words, there was 1 English acronym per 83.1 words. The full table of acronyms in alphabetical order and their number of occurrences in the corpus can be found in Appendix 1. For clarity, the acronyms and their frequency are presented in five groups based on the amount of their repetitions in the corpus. All of the following findings and items of interests is further discussed in more detail in the section 6.2 of the paper.

The first group consists of the two most frequent acronyms: *COVID-19* (294) and *SARS-CoV-2* (143). These two have been singled out due to their overwhelming numerical advantage over all the other items of interest in the text. With 294 occurrences, *COVID-19* makes up for just above 10% of all English acronyms in the corpus, *SARS-CoV-2* followed closely with its 143 occurrences, or about 5% of the total number of acronyms.

Following the first group, a number of frequently occurring acronyms with the number of repetitions between 69 and 118 was sectioned out, these will be henceforth referred to as Group 2. This group includes the following acronyms: *HPV* (118), *SM* (108), *LDL* (91), *HCV* (84), *CMV* (75), *mRNA* (70), *PARP* (69). With 615 combined repetitions across all of the members of Group 2, these acronyms constitute about 23% of all English acronyms that exist within the corpus.

Going into the third group, the number of repetitions drops significantly. For this batch of acronyms, the decision was made to include acronyms that have less occurrences than 69, but more than 20. Group 3 thus includes the following: *Allo-HSCT* (47), *BRCA* (45), *NAFLD* (44), *SMA* (41), *NOAC* (35), *SIBO* (35), *ACE-2* (35), *HIV* (34), *USG* (34), *GERD* (33), *SPMS* (33), *IBS* (31), *HAE* (30), *NEN* (28), *ESC* (25), *HDL* (23), *NET* (22), *PCSK9* (20), *WHO* (20), *AF* (20). All these acronyms combined total up to a very slightly higher count than Group 2, which is 635, and constitute about 23% of all token occurrences within the corpus.

The two final groups include the majority of the unique acronyms but have the lowest count of repetitions per acronym. All 157 of the acronyms with more than 1 but less than 20 repetitions were considered as part of Group 4. With a total of 898 occurrences, this group constitutes almost 33% of all acronyms within the text, while at the same time constituting almost half of all unique acronyms in the corpus.

The acronyms that only occurred once, or in other words had just one repetition, make up the final group, i.e. Group 5. This group consists of 140 unique acronyms, which is almost 43% of all unique acronyms, and only about 5% of total repetitions across the whole corpus. The specific acronyms for the two final groups will not be listed here for the sake of conciseness. Please refer to Appendix 1 for a complete list.

The table below includes a breakdown of all the aforementioned groups and their share of the total count of the borrowed acronyms with.

Table 1. Acronym statistics.

Acronym group	Combined repetitions	Percent of all unique acronyms (326)	Percent of all acronyms incl. repetition (2724)
COVID-19 (294) and SARS-CoV-2 (143)	437	0.07%	16.04%
HPV (118), SM (108), LDL (91), HCV (84), CMV (75), mRNA (70), PARP (69)	615	2.15%	22.58%
Allo-HSCT (47), BRCA (45), NAFLD (44), SMA (41), NOAC (35), SIBO (35), ACE-2 (35), HIV (34), USG (34), GERD (33), SPMS (33), IBS (31), HAE (30), NEN (28), ESC (25), HDL (23), NET (22), PCSK9 (20), WHO (20), AF (20)	635	6.13%	23.31%
Acronyms below 20 but above 1 repetition – 157 in total	898	48.16%	32.97%
Acronyms that only appear one time – 140 in total	140	42.94%	5.14%

6.2. Semantic analysis of the found acronyms

Due to the nature of the corpus and its medical theme, almost all of the acronyms were at least loosely related to medicine. In this section, a more in-depth consideration of their meanings and their relation to their frequency will be explored.

The two most frequently used acronyms which were considered as the only ones making up Group 1, were *COVID-19* and *SARS-CoV-2*. The ubiquity of these two items of interest should be immediately understandable considering the historical context, as during the year 2021, the whole world was still at the height of the COVID-19 pandemic. The newly formed acronyms COVID-19 and SARS-CoV-2, which stand for "coronavirus disease 2019" and "severe acute respiratory syndrome coronavirus 2", respectively, were on everyone's tongue. COVID-19 had profound implications on every field of medicine and affected treatments of patients suffering from a wide range of different diseases, even seemingly unrelated to body organs most vulnerable to COVID-19 itself.

Group 2, the second numerous group frequency-wise, consisted of the following acronyms: *HPV* (118), *SM* (108), *LDL* (91), *HCV* (84), *CMV* (75), *mRNA* (70), *PARP* (69). Similarly to the first group, their high count within the corpus can be attributed to the fact that they all signify concepts that appear in many areas of medicine or are considered to be topical issues that had multiple articles devoted to them within the analysed journal issues.

The major theme in this group are the names of viruses and diseases: *HPV* (human papillomavirus), *SM* (sclerosis multiplex), *HCV* (hepatitis C virus), and *CMV* (cytomegalovirus). These acronyms are either exclusively used for their related meanings in the Polish language (*CMV*), or seem to be strongly preferred over their Polish equivalents, which are as follows: *HPV* – wirus brodawczaka ludzkiego, *SM* – stwardnienie rozsiane, *HCV* – wirus zapalenia wątroby typu C. It is worth noting that there are no

Polish acronyms for these, and instead, the English acronyms are often used together with the Polish names. The remaining two acronyms from this group are *mRNA* (messenger ribonuclease inhibitor) and *PARP* (poly-ADP ribose polymerase). The former is closely related to the COVID-19 pandemic and the vaccines against the SARS-CoV-2 virus, while the latter is usually used in conjunction with the word "inhibitor" to describe a modern anti-cancer drug. Neither of these two have any Polish equivalents.

The third group singled out in the previous section consists of the following 20 acronyms: *Allo-HSCT* (47), *BRCA* (45), *NAFLD* (44), *SMA* (41), *NOAC* (35), *SIBO* (35), *ACE-2* (35), *HIV* (34), *USG* (34), *GERD* (33), *SPMS* (33), *IBS* (31), *HAE* (30), *NEN* (28), *ESC* (25), *HDL* (23), *NET* (22), *PCSK9* (20), *WHO* (20), *AF* (20). The acronyms in this group still seem to mostly describe popular illnesses and viruses. *NAFLD* (non-alcoholic fatty liver disease), *SMA* (spinal muscular atrophy), *SIBO* (small intestinal bacterial overgrowth), *GERD* (gastroesophageal reflux disease), *SPMS* (secondary-progressive multiple sclerosis), *IBS* (irritable bowel syndrome), *HAE* (hereditary angioedema), *NEN* (neuroendocrine neoplasms), *NET* (neuroendocrine tumor), *AF* (atrial fibrillation) are all names of diseases, syndromes or organ failures of different kinds. All of these have completely overtaken their original counterparts, in rare cases where they even existed in the first place, such as with *SMA* – rdzeniowy zanik mięśni, *IBS* – zespół jelita wrażliwego, *AF* – migotanie przedsionków. The rest never had a Polish equivalent, or had one solely created to be able to explain the meaning of the acronym to the non-English speaking Polish practitioners of medicine and are not used within the corpus. The remainder of the acronyms above 20 repetitions is where we see increased variety as far as the semantic field of the acronyms is concerned. Among these, we can observe names of chemical compounds and minute elements of the human body, such as *BRCA* (breast cancer gene), *ACE-2* (angiotensin-converting enzyme 2), *NOAC* (non-vitamin K antagonist oral anticoagulants), *HIV* (human immunodeficiency virus), *PCSK9* (proprotein convertase subtilisin/kexin type 9), and *HDL* (high-density lipoprotein), which is often used in conjunction with cholesterol to denote its type. There are also names of medical procedures, such as *Allo-HSCT* (allogeneic hematopoietic stem cell transplantation), or *USG* (ultrasonography). Closer to the bottom of this group there are also two names of international medical organizations, namely *ESC* (European Society of Cardiology), and *WHO* (World Health Organization). Similarly to the previous group, most acronyms here do not have a Polish equivalent at all. The exceptions to that are *WHO*, which sometimes appears as "Światowa Organizacja Zdrowia", though not in a shortened form, and *USG*, which arguably may even be treated as a Polish acronym because of how old and well-established the word *ultrasonography* is in the Polish language.

Groups 4 and 5 are much more extensive than the previous ones and include 297 unique acronyms. Within all these acronyms, we can enumerate several distinct semantic groups, however, we will only take a closer look at several notable examples for each of them, due to their sheer number.

The most prominent group of meanings found within the acronyms with a low number of occurrences are the names of medication, hormones, chemical compounds and the like. These often tend to have very complex names, which is why they are

shortened and presented in the form of an acronym. These terms are also never translated into Polish, which is a sign that they function in the Polish language of medicine in their original form. Some notable examples include *TG* (triglyceride), *RNA* (ribonucleic acid), *LAL* (lysosomal acid lipase) and *SSRI* (selective serotonin reuptake inhibitor).

Some other meanings conveyed through acronyms in Group 4 and Group 5 include names of medical procedures, like *CPT* (current procedural terminology), *ABPM* (ambulatory blood pressure monitoring) or *HYVET-COG* (hypertension in the very elderly trial cognitive function assessment). Names of illnesses and medical conditions can also still be encountered, *HBV* (hepatitis B virus), *NASH* (non-alcoholic steatohepatitis) and *LUTS* (lower urinary tract symptoms) being some of them.

More scarcely populated groups of acronyms include names of lesser-known medical organizations, such as *ASCO* (American Society of Clinical Oncology), or *ESMO* (European Society for Medical Oncology); miscellaneous acronyms related to the human body such as, for example, *BMI* (body mass index), or *ARR* (absolute risk reduction), and even terms used to refer to different types of patients, such as *HFA* (high functioning alcoholics).

Lastly, the non-medical acronyms, which were nonetheless used to complement a medical description in the corpus, must be mentioned. Among these, we can find examples such as *HTP* (heated tobacco product), *3D* (three-dimensional), *2D* (two-dimensional) or *XRES* (extreme resolution).

6.3. The techniques of introducing English acronyms in Polish medical texts

Alongside the acronyms themselves, the context in which they appear, and the techniques writers use to explain the meanings of these acronyms to Polish readers were also analysed. It must be noted that in the vast majority of cases, not every single instance of an acronym is introduced together with a full explanation of the words behind it, but rather it is done together with the first mention of said acronym and its subsequent repetitions receive no such treatment. As a result of this analysis, multiple techniques that were found can be enumerated.

One common way to introduce acronyms is by providing the readers with the Polish translation of all the words that the acronym combines and putting the original English acronym in brackets. In these cases, the readers do not learn what the original acronym stands for at all, only its Polish rendition. It can therefore be assumed, that the writer does not perceive the original acronym as important. This can be observed in the following excerpts from the corpus:

[1] W ubiegłym roku podczas dwóch kongresów medycznych – Amerykańskiego Towarzystwa Onkologii Klinicznej (**ASCO**) (*Kurier Medyczny*, Issue 1, 2021) [Within the last year, during the two medical summits of the American Society of Clinical Oncology (ASCO)]

[2] Zgodnie z aktualnymi wytycznymi oznaczenie wskaźnika aldosteronowo-reninowego (**ARR**) należy do badań uzupełniających […] (*Kurier Medyczny*, Issue 5, 2021)

[In accordance with the current guidelines, the designation of the aldosterone-renin ratio (ARR) is a complementary procedure [...]]

[3] Dobra informacja jest taka, że duża część z nich została objęta leczeniem antyretrowirusowym (***ARV***) – ta grupa zwiększyła się z 5,6 tys. do ponad 11 tys. (Kurier Medyczny, Issue 4, 2021) [The good news is that a huge part of these patients received antiretroviral therapy (ARV) – this group increased from 5,6 to 11 thousand.]

[4] W kwietniu tego roku szef chińskiego Centrum Kontroli i Prewencji Chorób (***CDC***) stwierdził publicznie [...] (*Kurier Medyczny*, Issue 3, 2021) [In April, earlier this year, the chief of the Chinese Center for Disease Control and Prevention (CDC) stated publicly [...]]

[5] Zgodnie z zaleceniami panelu ekspertów Europejskiego Towarzystwa Miażdżycowego (***EAS***) [...] (*Kurier Medyczny*, Issue 1, 2021) [According to the guidelines created by the experts of European Atherosclerosis Society (EAS) [...]]

The approach is sometimes modified to also include the original English words that the acronym stands for, on top of the Polish ones, such as in examples (6–11):

[6] Znaczenie ma zarówno roczny wskaźnik rzutów (annualized relapse rate – ***ARR***), jak i kryteria rozpoznania SM. (*Kurier Medyczny*, Issue 1, 2021) [The important factors are both the annualize relapse rate (annualized relapse rate – ARR), as well as the criteria of diagnosing SM.]

[7] [...] oraz sieci centrów lipidowych Europejskiego Towarzystwa Miażdżycowego (European Atherosclerosis Society – ***EAS***). (*Kurier Medyczny*, Issue 3, 2021) [[...] as well as the network of centres of the European Atherosclerosis Society (European Atherosclerosis Society – EAS).]

[8] [...] oraz certyfikatem Europejskiego Towarzystwa Badań nad Otyłością (The European Association for the Study of Obesity – ***EASO***). (*Kurier Medyczny*, Issue 1, 2021) [[...] as well as the certificate of The European Association for the Study of Obesity (The European Association for the Study of Obesity – EASO).]

[9] Dotychczas zdefiniowano ok. 8 tys. chorób rzadkich, a jedną z nich jest dziedziczny obrzęk naczynioruchowy (hereditary angioedema – ***HAE***). (*Kurier Medyczny*, Issue 2, 2021) [So far, about 8 thousand rare diseases have been defined, one of them being the hereditary angioedema (hereditary angioedema – HAE).]

[10] Osoby te określa się jako otyłe metabolicznie z prawidłową masą ciała (metabolically obese normal weight – ***MONW***). (*Kurier Medyczny*, Issue 3, 2021) [These people are referred to as metabolically obese normal weight (metabolically obese normal weight – MONW).]

[11] Aż 37–38 proc. pacjentów z cukrzycą typu 2 ma już niealkoholowe stłuszczeniowe zapalenie wątroby (non-alcoholic steatohepatitis – *NASH*) [...] (*Kurier Medyczny*, Issue 2, 2021) [As many as 37–38 percent of patients with type 2 diabetes has non-alcoholic steatohepatitis (non-alcoholic steatohepatitis – NASH)]

In some cases, particularly with organization names, the acronym is only explained in English, such as in examples (12–16):

[12] Dla przykładu American Academy of Sleep Medicine (*AASM*) rekomenduje stosowanie tylko dwóch BZD [...] (*Kurier Medyczny*, Issue 7, 2021) [For example, American Academy of Sleep Medicine (AASM) recommends using only two BZD [...]]

[13] Sądzę, że procedura rejestrowania przez European Medicines Agency (*EMA*) wkrótce się rozpocznie. (*Kurier Medyczny*, Issue 2, 2021) [I believe that the procedurę of registration through European Medicines Agency (EMA) will soon commence.]

[14] European Alliance for Personalised Medicine (*EAPM*) to sojusz na rzecz medycyny spersonalizowanej. (*Kurier Medyczny*, Issue 1, 2021) [European Alliance for Personalised Medicine (EAPM) is an Alliance for personalized medicine.]

[15] Zgodnie z kryteriami diagnostycznymi International Diabetes Federation (*IDF*) otyłość trzewną u dorosłych Europejczyków rozpoznaje się przy obwodzie talii od 94 cm u mężczyzn i od 80 cm u kobiet. (*Kurier Medyczny*, Issue 2, 2021) [According to the diagnostic criteria of International Diabetes Federation (IDF), visceral obesity in adult Europeans is diagnosed at 94cm abdominal circumference in men, and at 80cm in women.]

[16] [...] jak i w ramach kontynuacji terapii, co znalazło odzwierciedlenie m.in. w aktualnych wytycznych American College of Chest Physicians (*ACCP*). (*Kurier Medyczny*, Issue 8, 2021) [[...] as well as part of the continuation of therapy, which is reflected within the recent guidelines of American College of Chest Physicians (ACCP).]

Finally, cases where no information at all is given about the words that the acronym is comprised of need to be mentioned. This way of introducing acronyms, or rather complete lack thereof is employed for acronyms that the writers assume are known to the readers and, which are either well-established in the Polish medical language, or are simply ubiquitously known around the world. The two most frequent acronyms in the corpus, *COVID-19* and *SARS-CoV-2* are the best examples of this phenomenon, as the words behind these two acronyms were not even once spelled out for the readers. Outside of these two special cases, no direct correlation between the frequency of acronyms and the ways of introducing them was found. In contrast to *COVID-19* and *SARS-CoV-2*, there were acronyms within the groups with high counts of occurrences which had in-depth explanation attached to them many times, such as for example *HPV* or *SM*. On the other hand, there were also cases of acronyms

that were pretty rare within the corpus but, nevertheless, did not receive any special introduction at all, such as for example *NMDA*, *OAS* or *WBC*. Some other examples of acronyms which never received any kind of explanation include *CAT*, *PET*, *HDL*, *LDL*, *HIV*, *HBV*, *ICD-10*, *IgA*, *pH*.

7. Conclusions

The impact of English borrowings in the Polish language is not exclusive to informal registers of language, nor the younger speakers. In this paper, we have analysed the language used in a Polish medical journal *Kurier Medyczny*. We have looked at a full list of borrowed acronyms found in the corpus, the meanings behind the words that the acronyms stand for, the techniques writers used to introduce and explain them to their readers, as well as at different figures related to the frequency of the acronyms in the analysed language sample.

We have successfully managed to establish a connection between the frequency of occurrences of the acronyms and their meanings. The acronyms that stand for the names of the most well-established and known diseases, viruses and drugs are among the ones with the most repetitions within the corpus. On the other hand, the ones that refer to chemical compounds and lesser-known illnesses, as well as highly complex professional terminology turn out to be the ones repeated less often. Being able to pinpoint acronyms with certain meanings as potentially more frequent in medical texts allows us to make broader general predictions about other acronyms in the language of medical professionals as a whole in the future.

No direct link between the frequency of occurrences of a particular acronym and the technique of introducing said acronyms was found. The analysis of the most frequent group of acronyms in the corpus, which consisted of only two acronyms connected to the COVID-19 pandemic, which was the topical issue at the time the texts in the corpus were being released, initially seemed to confirm it. However, further analysis of other items of interest showed that there is no correlation between the types of introduction and the repetition count. Furthermore, many acronyms seem to be introduced through multiple different approaches in different parts of the corpus (cf. section 6.3). Outside of the most obviously well-established acronyms, within the corpus, the technique of introducing them seemed to rely solely on the preference of the writer.

Further research is needed for the linguists to be able to more closely assess the impact of English borrowings in the Polish medicine, which was preliminarily explored in this limited corpus-based study. By borrowing acronyms from English and using them as the main way to refer to diseases and illnesses instead of using original Polish terms, the Polish doctors make significant changes to both our language and the field of medicine as a whole. Thus, it is in the best interest of not only linguists and medical professionals but also patients and the general public to understand this role and the impact it may have on all of our lives.

References

Bauer L., 1998, *Is there a class of neoclassical compounds, and if so is it productive?*, "Linguistics" 36, no. 3.

Baum S.V., 1962, *The acronym, pure and impure*, "American Speech" 37.

Booij G.E., 2005, *The grammar of words: An introduction to linguistic morphology*, Oxford.

Cannon G., 1989, *Abbreviations and acronyms in English word-formation*, "American Speech" 64, no. 2.

Crystal D., 2018, *The Cambridge Encyclopedia of the English Language*, Bangor.

Davenport B., 1943, *Initials into words*, "American Notes and Queries" 2.

Fleischman S., 2003, *Language and medicine*, [in:] *The handbook of discourse analysis*, edited by D. Tannen, Oxford, pp. 470–502.

Grabias S., 1997, *Język w zachowaniach społecznych*, Lublin.

Grucza F., 2002, *Języki specjalistyczne – indykatory i/lub determinanty rozwoju cywilizacyjnego*, [in:] *Problemy technolingwistyki*, edited by J. Lewandowski, Warszawa, pp. 9–26.

Grucza S., 2006, *Komunikacja specjalistyczna a idiokontekst specjalistyczny i konstytuacja specjalistyczna*, [in:] *Glottodydaktyka i jej konteksty interkulturowe, Studia Rossica XVIII*, edited by A. Wołodźko-Butkiewicz, W. Zmarzer, Warszawa, pp. 209–223.

Grucza S., 2008, *Lingwistyka języków specjalistycznych. Języki. Kultury. Teksty. Wiedza*, Warszawa.

Górnicz M., 2003, *Terminologia tekstów specjalistycznych*, [in:] *Lingwistyczna identyfikacja tekstów specjalistycznych*, edited by B.Z. Kielar, S. Grucza, Warszawa, pp. 106–117.

Haugen E., 1950, *The analysis of linguistic borrowing*, "Language" 26, pp. 210–231.

Mosman J., 1993, *Dictionary of abbreviations acronyms and initialisms*, Gale.

Pilegaard M., 2000, *Introduction*, "HERMES. Journal of Language and Communication Studies" 13 (25), pp. 7–9.

Perlińska L., Krzyżowski J., 2008, *Medical slang & acronyms*, Warszawa.

Quirk R., Greenbaum S., Leech G., Svartvik J., 1985, *A comprehensive grammar of the English language*, London – New York.

Szulc A., 1984, *Podręczny słownik językoznawstwa stosowanego*, Warszawa.

Warner R., 1976, *The relationship between language and disease concepts*, "International Journal of Psychiatry in Medicine" 7, pp. 57–68.

Weinreich U., 1953, *Languages in contact*, The Hague.

Wilkoń A., 1987, *Typologia odmian językowych współczesnej polszczyzny*, Katowice.

Witalisz A., 2016, *Przewodnik po anglicyzmach w języku polskim*, Kraków.

The techniques of introducing English acronyms in the Polish medical texts: a corpus-based study

Abstract

The paper outlines the research on the usage of English borrowings, specifically acronyms, in the Polish medical language. The raw data is presented alongside a discussion about the findings and their implications. The research was conducted manually, with the assistance of basic electronic tools, on a personally prepared corpus consisting of an annual run of *Kurier*

Medyczny, a journal that covers general medical topics. The analysis included the techniques of introducing the borrowed acronyms to the readers, their general meanings, as well as data related to their frequency within the corpus.

Metody przedstawiania angielskich akronimów czytelnikom w polskich tekstach medycznych: badanie na korpusie

Streszczenie

W niniejszym artykule zbadano użycie angielskich zapożyczeń – akronimów – w polskim języku medycznym. Przedstawiono zarówno dane, jak i wyniki badań oraz ich implikacje. Badanie zostało wykonane ręcznie, za pomocą podstawowych narzędzi elektronicznych. W tym celu stworzono specjalny korpus składający się z wydanych w 2021 r. numerów „Kuriera Medycznego", czasopisma opisującego tematy z różnych dziedzin medycyny. Prezentowana analiza zawiera sposoby przedstawiania zapożyczonych akronimów, ich rozwinięć i znaczeń, a także dane dotyczące ich częstotliwości w korpusie.

Appendix 1. Full list of acronyms

Acronym	Count	Acronym	Count	Acronym	Count
2D	1	3D	1	17p	6
AAD	1	AASM	2	ABPM	7
ACC	3	ACCP	2	ACE	3
ACE-2	34	ACEI	2	ACG	2
ADP	3	AF	20	AHA	1
AHI	1	AIT	2	AKI	5
AIAT	3	ALK	1	Allo-HSCT	47
ALT	8	AMA	1	ANDA	1
APC	2	API	3	APT	1
ARB	2	ARNI	9	ARR	7
ARR (2)	8	ARV	15	ASCO	8
ASCOT-LLA	1	ASH	1	AspAT	3
AST	7	AT1	1	ATP	3
BALB/c	2	BB	2	BCL2	1
BMI	15	BPH	2	BR2	3
BRCA	45	BRCA1/2	11	BSA	1
BSH	1	BTK	1	BTS	1
BZD	17	C1-INH	18	CA	1
CAD	7	CAP	1	CAT	1
CD4(+)	5	CD8(+)	7	CDC	5
CESD	1	CFTR	4	CGM	3
CHA2DS2-VASc	1	ChAD	1	CIS	1

Acronym	Count	Acronym	Count	Acronym	Count
CK	1	CMS	1	CMV	75
COPD	1	COVID-19	294	CPT	6
CRP	6	CS	3	Long COVID	2
Post-COVID	11	DAPA-HF	1	DH-CCB	2
DLQI	1	DMT	3	DNA	18
DPP-4	2	DSM-5	2	DvD	1
E	1	EAACI	1	EAN	1
EAP	3	EAPM	3	EAS	3
EASL	1	EASO	1	EBM	2
ECCO	1	ECIL	1	ECOG	1
ECTRIMS	1	EFLM	1	EFSUMB	3
eGFR	12	EHRA	3	EMA	19
EMGS	1	ESC	25	ESH	5
ESMO	15	EuroA	4	EWG	1
FAB	1	FDA	14	FESC	2
FEV1	8	FFR	4	FH	6
FIB-4	1	FIH	1	FODMAP	1
FTU	1	FXR	2	GABA	5
GEP-NEN	1	GERD	33	GFR	5
GGTP	1	GLP	1	GLP-1	9
GMP	1	GOG213	1	GVHD	2
H	1	HA	3	HAE	30
HbA	1	HBV	6	HCC	1
HCTZ	17	HCV	84	HCV-RNA	11
HDL	23	HEK293T	1	HELP	1
HF	1	HFA	8	HFpEF	6
HFrEF	13	HGSOC	1	HIT	1
HIV	34	HOMA-IR	1	HPV	118
HR	3	HRD	4	HTP	9
HUS	1	HYVET-COG	6	IBD	1
IBS	31	IBS-D	4	IBS-M	2
ICD-10	3	ICD-9	1	ICH	1
ICON7	1	ICS	3	IDF	1
IFN-β	7	IFSO	2	IgA	4
IgE	3	IgHV	4	IHME	1
IL-2	1	IL-6	1	ILD	18
ILEP	1	IMDC	1	iPAAC	4
IPF	3	IPSS	2	ISH	1

Acronym	Count	Acronym	Count	Acronym	Count
IT	3	LAL	14	LAL-D	14
LDH	1	LDL	91	LDL-C	5
LES	1	LGA	2	LIPA	1
LMWH	10	LNP	3	LNP-mRNA	1
LOLA	6	LTFU	1	LTV	1
Lung-RADS	1	LUTS	16	MAP	2
MAPK	1	MDA	1	MERS	1
MHC	2	MHRA	3	MMSE	1
MONW	1	MPP	2	MRA	2
mRNA	70	MSKCC	2	MSM	1
MSMUnit	1	NAFLD	44	NAS	1
NASH	18	NCCN	2	NCT	1
NEC	2	NEN	28	NESARC	1
NET	22	NHC	5	NICE	1
NMDA	1	NMO	1	NNT	1
NOAC	35	NRS	1	NYHA	1
OAB-V8	2	OAS	1	OCD	4
OCDR	2	OECD	1	OS	2
OTC	1	OXR	1	P407	3
PARP	69	PASI	2	PBL	6
PCI	2	PCOS	2	PCR	2
PCSK9	20	PD-L1	4	PEG	1
PET	5	PET-CT	2	PFS	1
pH	5	PHQ-9	2	P-IBS	2
P-IBS-D	1	PIMS	1	Pol-SCORE	4
PP	1	PPAR	1	PPMS	3
PVd	3	pVNT	1	QALY	1
RAA	13	RAASi	1	RBRVS	1
Rd	1	REM	6	RNA	13
RRMS	9	RTG	11	RT-PCR	3
RX	1	SAFT	1	SAMSON	1
SARS	2	SARS-CoV-2	143	SCFA	1
SCORE	2	SCORE2	2	scRNA-seq	1
sGC	1	SGLT2	11	SHEP	1
SIBO	35	SLIT	4	SM	108
SMA	41	SMN	2	SMN2	3
SNRI	2	SOLO1	1	SOLO-1	1
SPMS	33	SSc	19	SSRI	15

Acronym	Count	Acronym	Count	Acronym	Count
T	1	TEM	1	TG	10
Th1	2	Th2	2	TK-FFR	17
TKI	1	TLR7	1	TLR8	1
TMPRSS2	5	TMPRSS4	1	TNF-a	1
TP53	5	TSH	2	tTG	1
UDCA	10	ULN	1	UNAIDS	1
USG	34	VEGFR	1	VITT	1
VKA	4	VLDL	1	VOC	1
VSV	1	VTD	1	VTE	9
WAO	1	WA-OAC	1	WBC	1
WHO	20	XRES	1		