Methodological issues of the compilation of the Polish Academy of Sciences Great Dictionary of Polish

Piotr Żmigrodzki

The Institute of the Polish Language at the Polish Academy of Sciences
E-mail: piotr.zmigrodzki@ijp.pan.pl

Abstract

Polish Academy of Sciences Great Dictionary of Polish (pol. *Wielki słownik języka polskiego PAN*) is being created by a team of linguists, lexicographers and other specialists in The Institute of the Polish Language PAN. This process started in Kraków in 2006. The dictionary is published exclusively in the online format, free of charge. Before August 2018, 70,000entries, which include the most frequent words, idiomatic expressions and proper names of the Polish language, will have been finished. The paper provides a general description of the dictionary and focuses especially on following issues of its compilation: a) workflow and methods of organizing work, b) lexicographical system of the dictionary and some of its innovations, which are useful during the process of entry creation and the process of overseeing the entries by lexicographers, c) a new, introduced in 2015, graphical user interface and its features. The final part is devoted to a short presentation of the plans the team has for the future.

Keywords: Polish language, electronic lexicography, general dictionary of Polish, online dictionary

1   Introduction

Polish Academy of Sciences Great Dictionary of Polish (pol. *Wielki słownik języka polskiego PAN*, WSJP PAN, WSJP) has been created by a team of linguists, lexicographers and other specialists in The Institute of the Polish Language PAN in Kraków (including people from other scientific centers) since 2006. It is being published exclusively in the on-line format, free of charge, under the following URL: http://wsjp.pl). By August 2018, the second stage of the work on the dictionary will have come to an end (and at the same time, its second period of financing). After that stage, about 70,000 entries, including the most frequently used Polish words, as well as idiomatic expressions and proper names, will have been completed. The total size of the dictionary (not including inflectional paradigms, which are automatically imported, and which for verbs include several dozen forms for each word) corresponds to about 20 volumes of a printed traditional dictionary. The goal of this article is to concisely present the most important achievements of the lexicographical team, accomplished until this point during their work, as well as to point out the general plans for the future.

2   The General Characteristics of the Dictionary

When it comes to typology, WSJP PAN should no doubt be classified as a general dictionary. As mentioned above, it is an electronic dictionary "born digitally" (without a paper version). Apart from that, we could classify it in the following ways:

- A documentary dictionary, i.e. based on an authentic material base, texts in the Polish language. The sources will be discussed in more detail later.

- A dictionary of the contemporary Polish language in the broad meaning of the term – it has been decided, that it will include lexical items (meanings) recorded since 1945. This was substantiated by the fact that the previous great dictionary of the Polish language (Doroszewski, ed. 1958-1969) did include the materials from the post-World War II period, but in a very selective way. At the same time, the choice of sources as well as their description, especially the semantic one, were highly influenced by the socio-political circumstances in the country ruled by communists.

- A descriptive dictionary, not a normative one – the authors do not exclude linguistic realizations considered to be incorrect, or – for any reason – unworthy to be included in the dictionary. They only go as far as to include the information about the normative unacceptability of the listed facts, based on *Wielki słownik poprawnej polszczyzny PWN* (Markowski, ed., 2003) and stylistic qualification of the substandard units.

- An academic dictionary – in the sense that the authors aim to maximize the inclusion of the achievements of the 20th century linguistics, especially in the field of semantic, inflectional and syntactic description of the lexical items. At the same time, however, the description in the dictionary is intended to reach as many users of the Polish language as possible (more on this below).

3   Source Database, Resource of Entries, and the Scope of Information in the Dictionary

3.1   Source Database

During the planning stage of the dictionary, the National Corpus of Polish (Przepiórkowski et.al. ,2012) was intended to be its main (or even the sole) source database. The creation of this corpus was only slightly more advanced than the work on our dictionary. Therefore other corpora were also used: Polish Scientific Publishers PWN and Institute of Computer Science PAS corpus. An additional minor corpus was also created, solely for the purpose of complementing the dictionary. It included materials which could not make it into National Corpus of Polish and be made public for legal reasons. Unfortunately, the National Corpus of Polish project was concluded in 2011, and since that time the corpus has not been updated, which made it obligatory to search for other sources, especially ones that document the newest language phenomena. As a result, internet resources became more important. These were browsed with the use of the Monco tool (http://monco.frazeo.pl), as well as directly through an internet browser. At the same time, words collected in The Institute of the Polish Language PAN were used, both published and unpublished. The editors of the entries can also include materials which they find in printed publications they come across on their own, but this is only done in exceptional cases.

3.2   Resource of Entries

The dictionary entries are not created alphabetically. The initial idea was to describe the most frequently used words in the Polish language. For this reason, a frequency list was taken from Institute of Computer Science PAS corpus. Fifteen thousand of the most frequently used words were chosen from the list to be described. The initial sorting according to thematic fields was conducted, and different groups of words were given to different editors to work on (at the same time, idioms and abbreviations connected to these words also were worked on). This task was finished by the end of 2012, after which the second stage of the work started. This involved describing synonymous or cognate lexemes to the aforementioned ones, as well as some of the newest words, which entered the language after 2000. This will also be the direction of further extensions of the dictionary in the future, because according to the initial plans of the authors it should include "pretty much all words used in the Polish

language". Among the currently completed entries, the biggest group (about 78%) are those describing single lexemes, which belong to traditionally perceived lexical categories such as nouns, verbs, adjectives, adverbs and numerals. About 18% of the entries are idioms (multi-word lexical units), and few percent are articles describing functional units (prepositions, conjunctions, etc.), proper names (mostly geographical), as well as abbreviations. It was also a norm in general Polish dictionaries of the 20th century, to include entries describing morphemes with high derivational productivity, as well as affixes such as *auto-, hiper-.* In WSJP we have decided to abandon this approach, because these are not lexical units of language, and only such elements should be included in a dictionary.

3.3   The Scope of Lexicographical Data in the Dictionary

A typical entry in WSJP includes the following elements:

- headword form (with variants)
- information about the pronunciation (only for the words with unpredictable pronunciation, especially fresh borrowings)
- chronology
- etymology
- description of meaning
- thematic classification
- hypernyms, synonyms and antonyms of the entry word in the specific meaning
- inflexion (especially the full paradigm of the words inflexion, its affiliation to a part of speech)
- syntactic requirements (especially for verbs)
- collocations and full sentence quotations
- abbreviations (if there are any)
- normative information (pertaining to some incorrect uses of the word)
- notes on usage (any other information pertaining to the usage of the word in texts).

These entries also include idioms, which use the particular word – the idioms themselves are described in separate entry articles with a different structure, though. The scope of lexicographical information in WSJP therefore largely overlaps with the scope of such information in other big Polish and European dictionaries (even though the syntactic requirements of different units or the semantic relations in the Polish lexicographical tradition occur less frequently). In the structure of the dictionary there are also elements of a less typical nature, taken from pedagogical (like the thematic classification of the meanings of each word or idiom) or scientific lexicography (chronology). Thematic classification is done according to the rules and pattern of the meaning division, created by Barbara Batko-Tokarz (2008). It is important not because of the description of a single lexeme, but because it allows the user of the dictionary, thanks to the usage of digital tools, to search for words connected to a given category. Etymological information, originally planned only for fresh borrowings to the Polish language, was extended to all one-word entries after receiving feedback from the readers of the dictionary. The specialists inserting the information are using the findings previously established in etymological dictionaries of the Polish language, but are also very often forced to conduct their own research and to create original etymologies (for details cf. Dębowiak et al., 2017). Chronology is typical for scientific dictionaries in Poland. Its goal is to provide the possibly earliest appearance of a given word in Polish texts. In contrast to, for example, the OED, it does not rely on quotes listed in the dictionary. Quotes are listed based on their illustrative usefulness, and also, as we mentioned earlier, WSJP quotes texts created after 1945. The aforementioned dates are therefore established, first and foremost, on the basis of texts which quote the old material, and also other collections of dated texts, including Google Books, which is slowly becoming a good source for establishing a chronology of function words since the beginning of the

19th century, up until the first half of 20th. Unfortunately, the chronology of words in WSJP does not include the division into meanings. Technical, financial and organizational limitations did not allow for this to happen, especially during the first stage of the project.

4   The Lexicographical Team and the Organization of Work

4.1   The Lexicographical Team

So far, over 80 people have been involved in the work on WSJP at some point. The amount of their input is of course different, and the team has been modified significantly during the 12 years in which the work on the project has been conducted. The core of the team consists of people employed indefinitely at The Institute of the Polish Language PAS, the lexicographical workshop, located in Kraków. Currently there are 17 such people. Previously, in the first period of the project, separate teams were organized in Warsaw (University of Warsaw), Katowice (University of Silesia) and Toruń (Nicolaus Copernicus University). There were also additional people working on fixed-term contracts (and these were the people whose jobs were affected the most by changes during the project's runtime). This was possible due to the dictionary being edited online, with the use of the Internet. Also, all the source materials are available either as corpora, or as digitalized texts, or are simply widely available in the web. The communication among the members of the project and the teams also took place mostly through the Internet, and only two times during the year were more general meetings of the teams organized. With time, the main work on the dictionary has moved to the team in Kraków, connected to the IJP PAN workshop. The teams in Katowice and Warsaw stopped working as separate parts of the undertaking. The only other group that is left is the team from Toruń, singled out due to the specific material they are tasked to work on (function words, i.e. prepositions, conjunctions, etc.). A group of computer specialists are cooperating with the team. They are responsible for creation and maintenance of the lexicographical system, creation and service of the user interface, processing of the inflectional data, and similar tasks. Finally, the administrative personnel of The Institute of the Polish Language PAS are responsible for financial backing of the project.

4.2   Organization of the Work

The work on WSJP is the very definition of a team effort. The rules governing the creation of the dictionary were created during long discussions held in the Workshop of the Great Polish Language Dictionary in The Institute of the Polish Language PAS, and every member of the team has also had an influence on the current state of these rules. It would however be hard to single out and credit any particular part to any particular team member. Due to the way the entries are created, there is never a single person responsible for any one of them. The hierarchy of the authors of the dictionary and the way of creating entry articles have already been described in more detail, e.g. in Żmigrodzki (2011, 2014), cf. also Fig.1. At this point, it should be remembered that apart from the editor, who prepares the initial version of the entry by inserting the data into the form of the dictionary database, the so-called specialists are also active. These are filling just one field in a given group of entries. This pertains to etymology, thematic classification and chronology. Each entry is also controlled by other editors, called supereditors, and before publication the entries are also read by a scientific editor of the dictionary, or another person authorized by one. Inflectional paradigms, prepared by the authors of *Grammatical Dictionary of Polish* (Saloni et al., 2007), are added to the entries. Select people can also fill other information in the entry article. Because of this, the final version of the entry can be a result of a combined effort of more than ten people. WSJP entries are therefore not credited to a specific person, which actually is not that peculiar, considering how such work is done in many dictionaries published around the world.
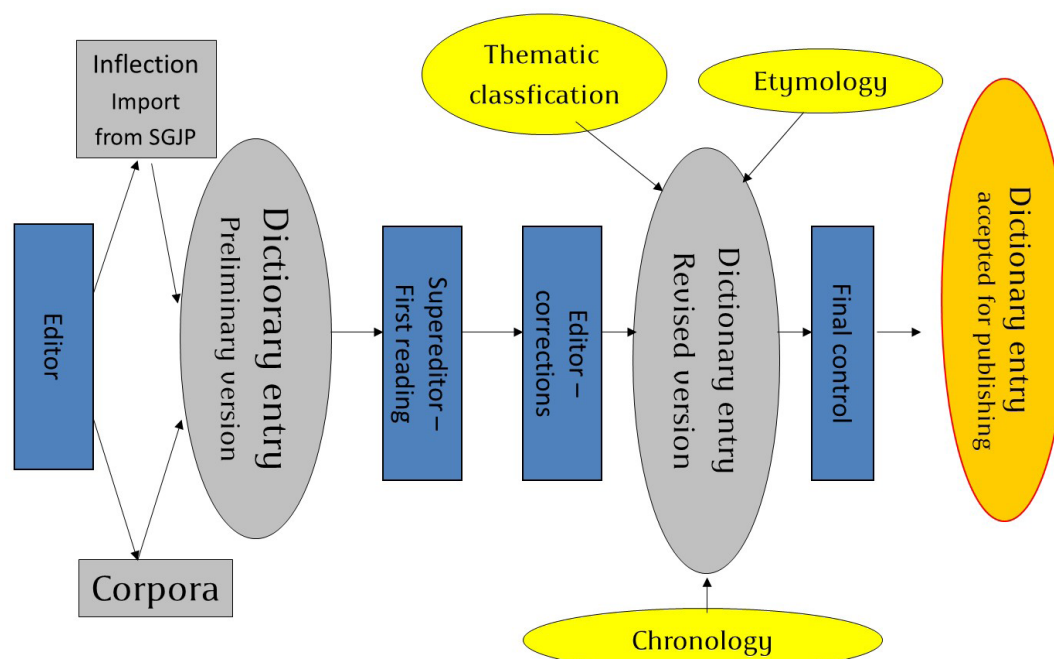
Figure 1: Lexicographical workflow in the WSJP team

5   Select Theoretical Aspects and Methodological Studies of the Dictionary

Since the 1980's, we have had increasing criticism of dictionaries in the Polish linguistic community. It has been claimed that their authors are not trying to include the newest achievements in the fields of semantics and syntax, which were heavily researched by the Polish scholars of the time. Because of that, the creators of WSJP decided to make it their goal to include these as much as possible. The fact is that the dictionary is based on the achievements of Polish linguistics of 1980's and 1990's, mostly in the structural paradigm. It develops the theoretical ideas contained in the academic *Grammar of the Contemporary Polish Language* (Grzegorczykowa et al., ed. 1984) and the works of such Polish scientists as Maciej Grochowski (esp. Grochowski 1993) and Andrzej Bogusławski (1997) (semantics), Mirosław Bańko and Zygmunt Saloni (inflection and syntax), Wieslaw Boryś (2005), Andrzej Bańkowski (2001) (etymology). The idea of the WSJP is to create a description of language phenomena based on the modern methodological foundations, which would at the same time be accessible to a wider audience. For this reason, theoretical achievements in linguistics cannot have been introduced directly, but have had to be simplified when it comes to the metalanguage and formal descriptive conventions.

6   The IT Aspects of the Work on the Dictionary

As mentioned above, the dictionary is being created and published exclusively in the digital format. Such was the aim of the team since the very beginning, even though in 2005, and even a couple years later, there were some doubts and suggestions to publish a paper version at the same time.[1]

---

1   In 2013 when the authors of the dictionary attempted to get a patronage of the Senate of Poland, doubts were also voiced by some senators, who were asking what would become of the dictionary in an event of a lack of electricity or a general downtime of the whole Internet.

However, the structure of the lexicographical description in WSJP was planned in such a way that converting it into a paper version would be impossible, even if financing was to be found. In this article we will not able to present an in-depth outlook at the technical side of the dictionary, because most members of the team are lexicographers who possess only the basic IT skills. The architecture of the system was created by programmers cooperating with The Institute of Polish Language PAS during the first decade of the 21st century, especially Mateusz Żółtak (currently University of Natural Resources and Life Sciences, Vienna)[2] Tomasz Żółtak and Paweł Fronczak. It is systematically modified and upgraded, currently with the last one as an overseer. The system is based on the idea of a database and PostgresSQL system. It involves three basic components: 1) the database of the dictionary, 2) the editing panel, i.e. internet form, used for introducing data into the database, and 3) the presentation panel, which is the representation of the lexicographical data available to the end-user as a website.

6.1   WSJP Editing Panel

The editing panel is the basic element of the IT system, from the point of view of its creators. It allows creation of new entries, filling them with data, introducing corrections and remarks directed at editors, by proof-readers or the leader of the project. Apart from that, it contains various functionalities that allow effective management of the entries and the whole dictionary. These include:

- A search engine for the editing panel, which includes such criteria as searching by author, by status (degree of completion), by filled or unfilled fields, by type of entry (connected to the type of the object it describes, e.g. single lexemes, idioms, abbreviations etc.) – with the possibility of combining the criteria and sorting the results based on alphabetical order, editors, degree of completion, date of last change. The list of results includes direct links to specific entries, which allows going directly into editing mode.
- A system of managing editors (available to the main editor and select other individuals), which allows the creation of editor accounts, giving access levels, and setting permissions to modify different groups of entries or the extent of this modification.
- A system of reporting about the work of specific editors, which shows how many entries (and which particular ones) a given editor created (or modified to a satisfactory extent) during a particular time period (which can be changed freely) – this serves to monitor the tasks assigned to particular people, and also for the purposes of calculating wages for the members of the project, as well as for statistical purposes (Fig. 2). It is possible to choose (by ticking the appropriate box) one or multiple editors, whose initials are listed in the upper part of the screen, or even all of them, as well as the type of the report (entry creation, control, thematic classification, etymology, chronology).
- A possibility to track changes in each entry, which makes it possible to establish which editor modified a particular entry and to what extent, including the possibility of displaying the old version of the entry.
- An advanced search engine for the editing panel, released in 2018. Thus allows free searching of any string of characters in any field of the database and any entry, searching for entries labeled with a specific tag, entries coming from a specific language, searching for entries with empty fields of a specific type, etc. The aim of implementing this kind of mechanic is to simplify the final corrections of the dictionary, which requires unification of the descriptions between entries, deletion of inconsistencies in the descriptions and other errors, which were not spotted before (even though each entry is controlled twice).
- A tool to import inflectional paradigms and to transform them.

---

2   He is the author of the IT groundwork of another dictionary being created in IJP PAN, namely, the Electronic Dictionary of 17th and 18th Century Polish (https://sxvii.pl).

Figure 2: Starting screen of the report about the work of the editors page.

- A preview of the current state of the entry in the presentation panel (see below).
- Dictionary autocorrect tools. These allow users to search for errors and (formal) inconsistencies in the entries, as well as to simplify correcting these errors without having to open particular entries. These include:
    - correction of the semantic relations – this tool produces a list of pairs of entries, where in entry A entry B is listed as its synonym or antonym, but in entry B there is no reference back to entry A (as we know, synonymy and antonymy are reciprocal relations), and allows users to correct this without the need to open the entries;
    - correction of the aspectual oppositions – working similarly to the above;
    - searching for entries without imported inflectional paradigms or with imported paradigms, which are not included in the description of a given entry article;
    - searching for entries describing idioms, which are not linked to proper entries about the single lexemes, with the possibility of automatically collecting them;
    - searching for sequences, inserted by mistake as collocations and idioms in the same entry, with the possibility of easily removing them;
    - correction of the spelling, conversion of the "computer"-style inverted commas, which are not normally used in the Polish language, reduction of excess spaces, etc.

6.2   The Presentation Panel of the Dictionary and Its Website

The presentation panel is how we refer to the user interface of the dictionary. Its main functions have already been discussed in Żmigrodzki (2015). The changes which have been introduced since then mainly pertain to the graphical interface of the website, designed from scratch by professional computer graphic designers, and to some optimization of the advanced search engine (though its efficiency will require some improvements as the number of entries grows). Currently, there are the following possible ways of searching in the dictionary:

- Alphabetically – after choosing a letter from the ones listed at the bottom of the screen (Fig. 3) the entries will be displayed in the form of a scrollable list.
- Simple search, typing the form in the window situated in the middle of the initial screen (as a result, all the entries containing the form are shown).

- Advanced search (according to criteria such as: part of speech, qualifier, etymology, thematic category); also searching for specific words, forms and clusters of letters in the specific parts of the entry article.

The default layout of the entry is structuralized, with tabs (Fig. 4), in which the entry article is shown in a format similar to a traditional paper catalogue, and particular parts of the entry are shown or hidden by choosing the appropriate tabs. In the case of entries with more meanings, finding the proper one is made easier by a menu. This consists of semantic tags, which guide the user towards the meaning they are interested in. The second option is a consolidated layout (Fig. 5), which displays the whole content of the entry article in linear fashion. This format allows users to easily print or save the article as a PDF file, if need be.
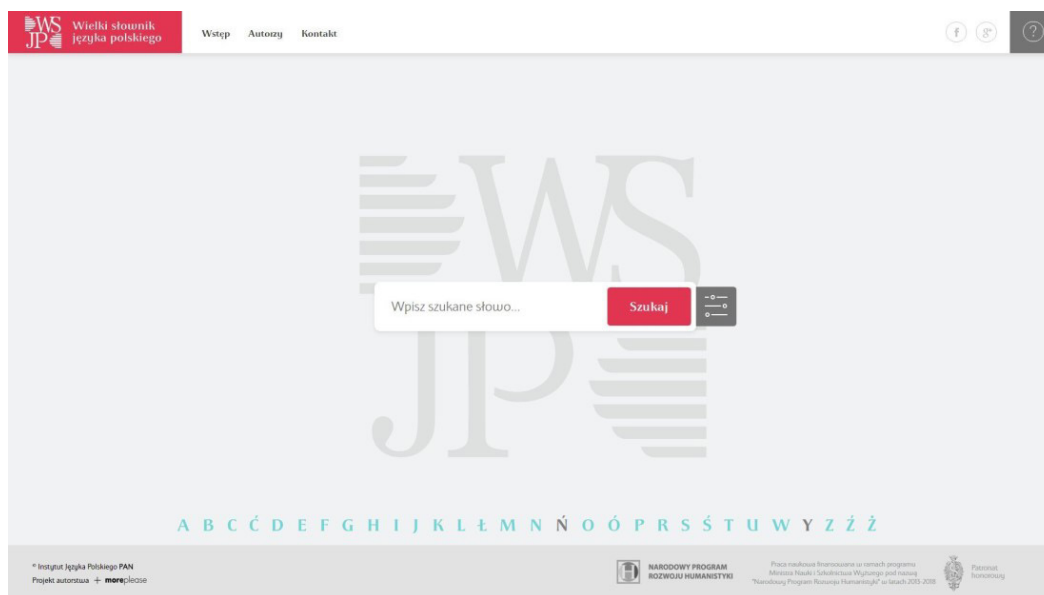

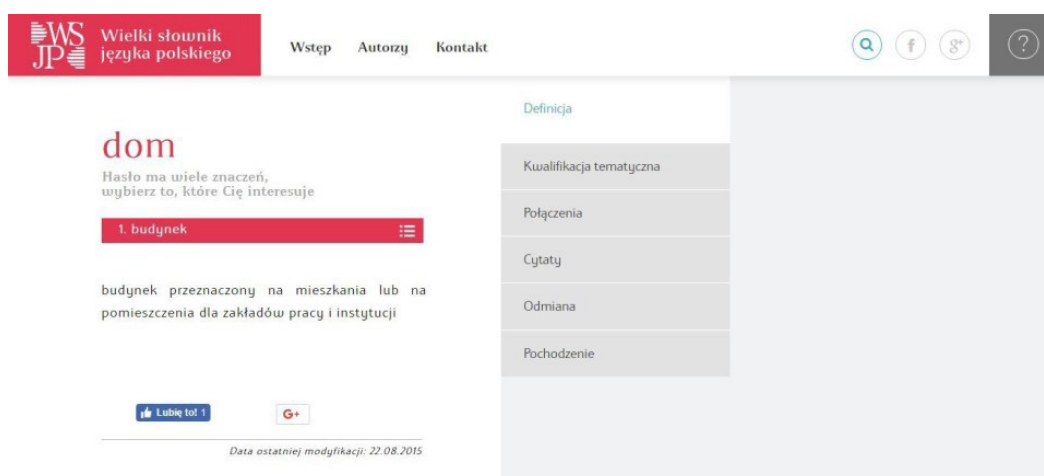Figure 3: The homepage of the dictionary with the search box.

Figure 4: Entry *dom* (*house*) in the structuralized layout (1st subentry 'building').

Figure 5: Entry *dom (house)* in the consolidated layout (fragment).

Apart from the entries, the reader of the dictionary is also able to view the materials, which in classic lexicography would be categorized as *front matters*. So we have a tab called "Introduction", the contents of which are akin to the introductions of multiple-volume paper dictionaries. It contains a bibliography of the works about the dictionary together with hyperlinks to their electronic versions, if they are available, as well as a subpage with short information about the authors of the dictionary and a contact form for the readers.

7   The Internet Life of the Dictionary

The issues of presence and availability of WSJP on the Internet were discussed in a few papers delivered at scientific conferences both in Poland and other European countries, as well as in some published texts, like the most recent Kozioł-Chrzanowska (2017). Not to reiterate the information available in these sources, I will just state that the interest in the dictionary is steadily rising, together with the increase of available entries. The overall number of sessions (according to Google Analytics) in 2017 amounted to 1,066,693, out of which 44.88% came from smartphones[3], 52.22% from a personal computer or laptop, and 2.91% from tablets (cf. Biesaga 2018). The majority of the connections were tracked to an Internet search engine (about 85%), about 11% to manually inputting the dictionary's address in the address bar of the browser, and about 3% to links from other sources, which includes social network sites, which constitute about 1% on their own. Apart from the webpage of the dictionary, there is also an official profile on Facebook, as well as Google Plus. It should be noted though that

---

3   The dictionary webpage is responsive, meaning it adjusts the display to the characteristics of the device which is used to browse the dictionary.

the makers of the dictionary are currently focused mostly on the lexicographical work, while neglecting promotional activities. So adverts for the dictionary on other sites were not purchased, efforts at positioning the dictionary in search engines were also not made. There is definitely much to be done in this regard in the future. Because the dictionary is of rather a scientific nature, and not social one, we are not aiming at an intensive communication with the readers, even though we do share a contact form with them. The interest in this form of communication is not very great, and thus the form is not used by many people, only a few per month on average. The feedback we get from this is, however, mostly valuable, usually pertaining to errors found in the published entries, or to suggestions about including currently missing entries. From time to time we also get requests for clarification on the proper usage of some words and expressions, but the WSJP team is not providing this kind of advice. In such cases we direct these queries to appropriate people or institutions, which can provide this kind of expertise.

8   The Future of the Dictionary

A couple words should be said about the dictionary team's plans for the future. The most important thing is further extension of the number of entries. At the end of 2017, it was assumed that within the next five years (until 2023) it would be possible to have around 150,000 entries, which would already be quite a serious number (the aforementioned dictionary by Doroszewski had about 125,000 entries in 11 printed volumes), while the biggest general dictionary of Polish, the so-called Warsaw dictionary (Karłowicz et al. 1900-1927), has almost 280,000., but with a very poor microstructure and scope of information.

Some technological improvements are also planned:

• Automation of the allocation of entries and opening them. So far allocations were done manually. The editor had to create them in the editing panel, download the inflectional paradigm, and start describing them, after having received the list of entries. It was established that an automatic method of creating the initial versions of entries, including inflectional information, readily available to the user, which the editor could start working on at a convenient time, would be desirable (they would be unavailable to the end-user until published).
• Adjusting the dictionary to the new receiving devices, which might be popularized in the future, which possibly includes a mobile application (that would, however, require the involvement of a specialized IT company).
• Integration with other sources of lexicographical information. The main idea here was to add links to Słowosieć (the Polish version of WordNet) to the dictionary entries, as well as other dictionaries created and published in The Institute of the Polish Language PAN.
• Extension of the internet analytics and promotion of the dictionary, to make it the basic source of knowledge of the Polish language to as many users as possible, not only to the scientific community.

Unfortunately, the sudden change of the method of financing Polish scientific projects, which was introduced at the beginning of 2018 (and further reforms are planned) created a situation in which the main goal is currently to find funds to sustain the dictionary team, so that the lexicographical work can be continued, even in the most basic format. Because of that, the aforementioned new ideas have had to be put on hold. We can only hope that the determination of the scientific editor and the team working on the dictionary, as well as maybe the voice of its users and the Polish linguistic community, will sway the authorities towards providing a proper financial base for the further development of WSJP PAN.

References

Bańkowski, A. (2000): *Etymologiczny słownik języka polskiego*. Warszawa: Wydawnictwo Naukowe PWN.

Batko-Tokarz, B. (2008). Tematyczny podział słownictwa w Wielkim słowniku języka polskiego. In P. Żmigrodzki, R. Przybylska (eds.) *Nowe studia leksykograficzne*, vol. 2, Kraków: Lexis pp. 31–48.

Biesaga, M. (2018). Wielki słownik języka polskiego PAN w Internecie. In P. Żmigrodzki, M. Bańko, B. Batko-Tokarz, M. Biesaga, J. Bobrowski, A. Czelakowska, M. Grochowski, R. Przybylska, J. Waniakowa, K. Węgrzynek (eds.) *Wielki słownik języka polskiego PAN. Geneza, koncepcja, zasady opracowania*. Kraków: Instytut Języka Polskiego PAN (in print).

Bogusławski, A. (1988). *Język w słowniku. Desiderata semantyczne do wielkiego słownika polszczyzny*. Wrocław: Zakład Narodowy im. Ossolińskich.

Boryś, W. (2005). *Słownik etymologiczny języka polskiego*. Kraków: Wydawnictwo Literackie.

Dębowiak, P., Ostrowski, B., Waniakowa, J. (2017).   Etymology in the Polish Academy of Sciences Great Dictionary of Polish. In *Lexikos*,. 27, pp. 597–608.

Doroszewski, W., (ed.) (1958-1969). *Słownik języka polskiego PAN*, vol. 1-11. Warszawa: Państwowe Wydawnictwo Naukowe.

Grochowski, M. (1993): *Konwencje semantyczne a definiowanie wyrażeń językowych*. Warszawa: Zakład Semiotyki Logicznej Uniwersytetu Warszawskiego, Polskie Towarzystwo Semiotyczne.

Grzegorczykowa, R., et. al., ed. (1984). *Gramatyka współczesnego języka polskiego. Morfologia*. Warszawa: Państwowe Wydawnctwo Naukowe.

Karłowicz, J., Krynski, A., Niedźwiedzki, W. (eds.) (1900-1927). *Słownik języka polskiego*, vol. 1-8. Warszawa: Kasa im. J. Mianowskiego.

Kozioł-Chrzanowska, E. (2017). What Do Users of General Electronic Monolingual Dictionaries Search for? The Most Popular Entries in the Polish Academy of Sciences Great Dictionary of Polish. In I. Kosem, J.Kallas, C. Tiberius , S. Krek, M. Jakubíček, V. Baisa (eds.) Electronic lexicography in the 21st century: Lexicography from scratch, Proceedings of eLex 2017 Conference. Brno: Lexical Computing CZ s.r.o., pp. 202–220.

Markowski, A. (ed.) (2003). *Wielki słownik poprawnej polszczyzny PWN*. Warszawa: Wydawnictwo Naukowe PWN.

Przepiórkowski A., Bańko, M., Górski, R., Lewandowska-Tomaszczyk, B. (eds.) (2012). *Narodowy Korpus Języka Polskiego*. Warszawa: Wydawnictwo Naukowe PWN.

Saloni, Z., Gruszczynski, W., Woliński R., Wołosz R. (2007). Grammatical Dictionary of Polish. Presentation by the Authors. In: *Studies in Polish Linguistics*, 4, pp. 5-26.

Żmigrodzki, P. (2011). Polish Academy of Sciences Great Dictionary of Polish. History, presence, prospects In *Studies in Polish Linguistics*, 6, pp. 7-26.

Żmigrodzki, P. (2014). Polish Academy of Sciences Great Dictionary of Polish [Wielki słownik języka polskiego PAN]. in: *Slovenščina 2.0*, 2 (2), pp. 37–52.

Acknowledgements