Justyna Jedlikowska
Uniwersytet Jana Kochanowskiego w Kielcach
https://doi.org/10.18778/8142-463-9.09

O kreatywnym wykorzystaniu narzędzi internetowych w badaniu tekstów

Kreatywność, *kreatywny* to leksemy, które dość niedawno zagościły w leksykonie Polaków, a już można uznać je za modne we współczesnym języku. W Narodowym Fotokorpusie Języka Polskiego odnajdujemy poświadczenie wystąpienia leksemu *kreatywny* datowane na 1954 rok, z kolei pierwsze użycie słowa *kreatywność* datowane jest w korpusie na 1971 rok, a wzrost frekwencji tego leksemu obserwowany jest po 1980 roku. Korzystając z Narodowego Korpusu Języka Polskiego, możemy wskazać na szereg połączeń z tymi leksemami, które świadczą o ich popularności i modzie na ich stosowanie. Są to np. *kreatywny człowiek*, *kreatywny aspekt*, *kreatywny potencjał*, *kreatywny charakter*, *kreatywny gatunek*, *czynnik kreatywny*, *kreatywna natura*, *kreatywny fotograf*, *dziecięca kreatywność*, *kreatywność wieku dziecięcego*, *kreatywność w tańcu*.

Czym jest ta *kreatywność*? Co to znaczy *być kreatywnym*? Z uwagi na większą frekwencję użycia leksemu *kreatywny* w związkach wyrazowych oraz fakt, że można odnaleźć poświadczenia użycia tego słowa już w latach 50. XX wieku, można uznać tę jednostkę leksykalną za pierwotną. W związku z tym przywołam tu tylko definicję tego przymiotnika. W *Słowniku języka polskiego* pod redakcją Stanisława Dubisza leksem *kreatywny* został zdefiniowany jako „zdolny do tworzenia czegoś nowego, oryginalnego; pomysłowy, twórczy;

mający na celu tworzenie, dający możliwość, podstawę do tworzenia czegoś nowego, oryginalnego" [*kreatywny* w: Dubisz 2003]. W tej definicji pojawia się wiele synonimów, a takie definiowanie słów spotykamy zwłaszcza wtedy, gdy trzeba wytłumaczyć znaczenie czegoś, co nie jest uchwytne. Ta nieuchwytność znaczenia mogła się przyczynić do wzrostu popularności tej jednostki, a tym samym wzrostu zainteresowania zjawiskami, osobami, procesami, których właściwości określane są wspomnianym przymiotnikiem.

Moda ta dotarła również do lingwistyki. We współczesnym polskim językoznawstwie coraz częściej mówi się o potrzebie badania kreatywności zjawisk językowych, o potrzebie kreatywności w prowadzeniu badań językoznawczych. W nurt ten wpisuje się niniejszy artykuł. Zostanie w nim bowiem omówione zagadnienie możliwości kreatywnego wykorzystania jednego z dostępnych w polskim internecie narzędzi możliwych do zastosowania w badaniach lingwistycznych – a w szczególności stylistyczno-składniowych – nad tekstami. Narzędziem, które zostanie tu omówione, jest dostępna za pośrednictwem internetu aplikacja Jasnopis. Powstała ona z myślą o tworzeniu tekstów przyjaznych dla czytelnika, lecz swoją funkcjonalnością powinna zainteresować wielu lingwistów, ponieważ może stanowić pomocnicze narzędzie w badaniach językoznawczych.

Ulotka informacyjna dołączona do opakowania farmaceutyku, instrukcja obsługi zmywarki do naczyń, umowa kredytu hipotecznego, ustawa, pismo z urzędu skarbowego – to tylko niektóre teksty określane przez wiele osób mianem trudnych i skomplikowanych. Nic w tym dziwnego – jako podstawowy cel czytania wskazuje się szybkie przyswojenie informacji celem ich dalszego wykorzystania. A szybkie przyswojenie informacji z tekstu jest możliwe wyłącznie wtedy, gdy czytelnik posiada ku temu stosowne kompetencje. Niestety, wiele tekstów funkcjonujących w przestrzeni publicznej jest formułowanych tak, że duże grono odbiorców, tj. ludzi nieposiadających wspomnianych kompetencji, ma problemy z ich poprawnym zrozumieniem.

W latach 2011–2012 w 24 krajach zostało przeprowadzone na grupie 166 tys. osób w wieku 16–65 lat Międzynarodowe Badanie

Kompetencji Osób Dorosłych (The Programme for the International Assessment of Adult Competencies, PIAAC). W ramach tego badania dokonano pomiaru trzech umiejętności, które można określić mianem niezbędnych do funkcjonowania we współczesnym świecie. Należą do nich: umiejętności wykorzystywania technologii informacyjno-komunikacyjnych, rozumowania matematycznego i rozumienia tekstu. Dla naszych rozważań istotne będzie zwrócenie uwagi na wnioski z badań nad trzecią z wymienionych umiejętności.

Czym jest kompetencja rozumienia tekstu? W ujęciu Organizacji Współpracy Gospodarczej i Rozwoju (OECD), która przeprowadziła to badanie, jest to „umiejętność rozumienia i oceny informacji zawartych w tekstach pisanych oraz wykorzystanie tych informacji do udziału w życiu społecznym, osiągania własnych celów oraz rozwoju własnej wiedzy i potencjału" [Rynko 2013: 24].

Autorzy tekstów uznawanych stereotypowo za trudne zwykle winą za trudności w odbiorze tekstów obarczają odbiorców. Zarzucają im brak doświadczenia czytelniczego, związane z tym trudności ze skupieniem uwagi na czytanym tekście, brak odpowiedniego wykształcenia, nieznajomość terminologii podstawowej dla dziedzin, z którymi przysłowiowy Kowalski ma stosunkowo mało do czynienia w życiu. Potwierdzają to rezultaty wspomnianego wyżej międzynarodowego badania, w którym Polacy „nie uplasowali się na czołowych lokatach". Wystarczy wspomnieć, że ponad 15% Polaków nie może wykazać się umiejętnościami z zakresu rozumienia tekstu. Z problemem tym borykają się szczególnie osoby starsze, choć spadek poziomu tych kompetencji zauważalny jest już po 30. roku życia. Przyczyn problemów z rozumieniem tekstu należy szukać również w braku odpowiedniego wykształcenia. Umiejętnością tą pochwalić się może, zgodnie z rezultatami badania, blisko 40% respondentów z wyższym wykształceniem i tylko 2% osób o najniższym wykształceniu [Rynko 2013: 5–7].

Czy wina leży tylko po stronie niekompetentnych czytelników? Członkowie zespołu projektowego, kórych dziełem jest aplikacja Jasnopis, w artykule naukowym *W poszukiwaniu metody automa-*

tycznego mierzenia zrozumiałości tekstów informacyjnych zauważyli, że w wielu przypadkach winnymi braku zrozumienia tekstu są jego twórcy. Odbiór tekstu utrudnia, ich zdaniem, sama jego struktura. Twórcy tekstów nader chętnie tworzą skomplikowane pod względem syntaktycznym wypowiedzenia, operują metaforami, stosują elementy leksykalne (np. specjalistyczną terminologię, archaizmy), których znaczenie dla przeciętnego czytelnika jest nieznane lub nie do końca zrozumiałe, co przyczynia się do tego, że ich wpływ na stopień trudności wydaje się niezaprzeczalny. Autorzy wspomnianego artykułu podkreślają konieczność zwrócenia uwagi na kwestię tworzenia tekstów z myślą o ich odbiorcach. Badacze stwierdzają, że w wielu krajach przywiązuje się wagę do tego, by „teksty przeznaczone dla przeciętnego odbiorcy (czasem nawet dla każdego odbiorcy, który umie czytać) były formułowane w sposób możliwie jasny (czytelny, przejrzysty)" [Gruszczyński i in. 2014].

Fakt, że na gruncie polskim dostrzeżona została potrzeba tworzenia tekstów przeznaczonych dla osób o przeciętnej kompetencji rozumienia tekstu przyczynił się do powstania aplikacji Jasnopis. Stworzyła ją grupa naukowców z Uniwersytetu SWPS i Polskiej Akademii Nauk pod kierownictwem prof. dr. hab. Włodzimierza Gruszczyńskiego, polskiego językoznawcy zajmującego się leksykografią, kulturą języka, formalnym opisem języka polskiego oraz historią polszczyzny. Narzędzie to powstało w ramach projektu finansowanego przez Narodowe Centrum Nauki.

Aplikacja ta pozwala na zmierzenie zrozumiałości tekstu na poziomie leksykalnym i syntaktycznym. Ma ona przede wszystkim zastosowanie praktyczne. Zgodnie z założeniami jej twórców pozwala ona na wygodne sprawdzanie stopnia trudności różnych tekstów polskich, które nie spełniają cech tekstów o charakterze artystycznym [Gruszczyński, Ogrodniczuk 2015: 5–9]. Adresatami tego narzędzia internetowego są jednak przede wszystkim autorzy tekstów użytkowych, które pojawiają się w przestrzeni publicznej. Do tekstów takich można zaliczyć m.in. pisma urzędowe, regulaminy, instrukcje, przeznaczone dla pacjentów ulotki dołączone opakowań środków farmakologicznych, podręczniki szkolne, artykuły infor-

macyjne przeznaczone do publikacji w tradycyjnych lub internetowych wydaniach gazet i czasopism.

Jak działa aplikacja Jasnopis? Pod tą nazwą kryją się dwa programy komputerowe. Pierwszy przeznaczony jest do statystycznej analizy trudności „gotowego" polskojęzycznego tekstu. Teksty przeznaczone do analizy mogą być albo wprowadzane bezpośrednio za pomocą narzędzia wklejania, albo poprzez wskazanie drogi dojścia do pliku tekstowego znajdującego się na nośniku pamięci (np. dysku twardym komputera), albo przez podanie adresu URL. Drugi program to wtyczka (*plug-in*), dzięki której można korzystać z Jasnopisu w trakcie edycji tekstu w popularnych edytorach tekstu, np. MS Word, LibreOffice. Przydatność tej wtyczki nie jest obiektem naszych zainteresowań.

Dla omówienia przydatności narzędzia internetowego Jasnopis w badaniach lingwistycznych konieczne jest wyszczególnienie typów informacji o tekście, jakie uzyskuje się w wyniku analizy tekstu. W aplikacji wyliczane są następujące statystyki tekstu: ocena trudności tekstu w siedmiostopniowej skali, wskaźniki mglistości, indeks Pisarka, średnia długość zdania, średnia długość akapitu, procent słów trudnych (tj. słów, które mają cztery lub więcej sylab), procent rzeczowników, procent rzeczowników trudnych, procent czasowników, procent czasowników trudnych, procent przymiotników, procent przymiotników trudnych, stosunek rzeczowników do czasowników.

Badając przydatność tego narzędzia internetowego, sprawdziłam możliwości aplikacji dostępnej *online*, w której można zbadać tzw. gotowe teksty. Za jej pomocą dokonałam analizy dwudziestu tekstów należących do trzech realnych makrodyskursów, wydzielanych w ramach funkcjonalnego pragmatyzmu: dyskursu profesjonalnego, codziennego i społeczno-etycznego [Leszczak 2010: 37]. W związku z tym, że twórcy aplikacji zaznaczają, iż jest ona przeznaczona do analizy tekstów o charakterze użytkowym i nie powinna być stosowana w analizie tekstów dyskursu artystycznego, uznałam, że nie stanowi narzędzia przydatnego w badaniach nad tekstami pozostałych dwóch wirtualnych makrodyskursów

(tj. naukowego i filozoficznego). Teksty reprezentujące te dyskursy nie są z założenia przeznaczone dla odbiorców o przeciętnej kompetencji rozumienia tekstu. Do tekstów, które posłużyły mi w analizie przydatności aplikacji, należały m.in.: tekst obowiązującej konstytucji, tekst decyzji o przyznaniu stypendium doktoranckiego, wybrane posty użytkowników Facebooka, teksty z bloga poświęconego kwestiom zdrowego stylu życia, teksty informacyjne z portalu lokalnej gazety, tekst kazania, teksty z podręczników szkolnych do biologii, chemii i wiedzy o społeczeństwie. Zadbałam o to, by były one wprowadzane do aplikacji na różne sposoby.

Na podstawie dokonanej analizy tego – oczywiście niewielkiego – korpusu zauważyłam, że aplikacja najlepiej sobie „radzi" z opracowaniem statystycznym tekstu wprowadzonego bezpośrednio oraz poprzez wskazanie drogi dojścia do pliku z tekstem (w formacie .doc i .docx). Problematyczne dla programu okazało się przeanalizowanie tekstu wprowadzonego poprzez wskazanie dojścia do pliku z tekstem w formacie pdf lub adresu URL. W pierwszym przypadku program najczęściej interpretował koniec wersu jako koniec zdania, a w drugim pojawiły się problemy z wczytaniem strony internetowej przeznaczonej do analizy. Oczywiście nie kwestionuję tu dokonań twórców aplikacji. Moje wnioski wysnuwam na podstawie analizy skromnego pod względem ilościowym korpusu, w którym mogły się znaleźć teksty niestarannie przygotowane pod względem edytorskim lub takie, podczas zapisu których popełnione (często nieumyślnie) zostały błędy natury informatycznej, które wpłynęły bezpośrednio na jakość plików z tekstem. Korzystając z omawianej aplikacji należy zatem zadbać o właściwe przygotowanie tekstu przeznaczonego do analizy oraz zweryfikować uzyskane wyniki poprzez skontrolowanie w oknie z rezultatami analizy tego, czy np. podział tekstu na zdania jest zgodny z oryginałem.

Problemy te nie przekreślają możliwości wykorzystywania aplikacji Jasnopis w badaniach lingwistycznych. To właśnie te statystyki tekstu mogą być najbardziej przydatne dla lingwisty. Automatyczne wyliczenie statystycznych właściwości tekstu pozwala przede wszystkim na zaoszczędzenie czasu badacza oraz umożliwia pozy-

skanie potrzebnych wyników przy minimalnym nakładzie pracy. Badania statystyczne nad tekstem (korpusem tekstów) są bowiem żmudne i czasochłonne, wymagają dokonania precyzyjnej, starannej i skrupulatnej analizy materiału językowego oraz wykonywania szeregu obliczeń i zestawień. Dane statystyczne uzyskane dzięki analizie wykonanej przy pomocy aplikacji pozwalają na wskazanie głównych właściwości opracowywanego materiału językowego. Przyspiesza to z całą pewnością przebieg procesu badawczego, pozwala bowiem na wstępie wskazać te elementy materiału językowego, które powinny zostać zbadane szczególnie skrupulatnie.

Czy jednak możliwe jest ograniczenie wkładu pracy językoznawcy do tego stopnia, by mógł on uzyskać potrzebne statystyki w ciągu kilku minut? Nie da się jednoznacznie odpowiedzieć na to pytanie. Wszystko zależy od tego, do czego lingwiście potrzebne są owe statystyki. Monograficzne opracowanie korpusu tekstów, w którym powinny być wskazane szczegółowe i precyzyjne statystyki, nie stanowi publikacji, w której mogłyby być zastosowane statystyki uzyskane za pomocą aplikacji takiej jak Jasnopis. Jest to bowiem typ tekstu, w którym może być tolerowany niewielki margines błędu. Z powodzeniem jednak można stosować rezultaty otrzymane dzięki tej aplikacji w tekstach naukowych, popularnonaukowych lub dydaktycznych, w których można przyjąć za wiarygodne statystyki o większym marginesie błędu. Są to przede wszystkim teksty, w których wskazanie rezultatów o charakterze statystycznym ma pomóc w pokazaniu głównych tendencji w ramach badanego zjawiska, we wskazaniu właściwości typowych dla danego typu tekstu itp. Dlaczego nie można bezkrytycznie ufać danym uzyskanym dzięki analizie tekstu za pomocą tego programu? Moja ostrożność jest tu związana m.in. z wątpliwościami dotyczącymi automatycznego przyporządkowania leksemów do poszczególnych części mowy. Na podstawie mojej wiedzy o specyfice tego typu programów komputerowych przypuszczam, że przyporządkowanie to odbywa się na drodze formalnego przypasowania do grupy reprezentującej zbieżne cechy gramatyczne. Najprawdopodobniej może być nie brany w tym momencie pod uwagę aspekt semantyczny leksemów. Można

wskazać tu np. taki wyraz jak *chory*, który da się przyporządkować zarówno do klasy przymiotników, jak i rzeczowników.

Czy zatem Jasnopis stanowi narzędzie przydatne w badaniach stylistyczno-składniowych nad tekstami? Czy można je zastosować w badaniach nad szczególnie specyficznymi tekstami, takimi jak np. tekst konstytucji? Badania stylistyczno-składniowe to typ badań nad stylem tekstu, w których badacz koncentruje swoją uwagę na badaniu jego składni [Ruszkowski 2004]. W mojej ocenie omawiane narzędzie może być przydatne w badaniach nad szeregiem typów tekstów, w tym również tekstów specjalistycznych. Rozważmy to na przykładzie wspomnianego tekstu konstytucji. Mówiąc o wykorzystaniu omawianej aplikacji w badaniach nad konstytucją, nie mam, rzecz jasna, na myśli badania jej zrozumiałości w siedmiostopniowej skali trudności tekstu, indeksu mglistości (FOG) ani indeksu Pisarka. Konstytucja to tekst dyskursu społeczno-etycznego [Leszczak 2010: 88] szczególnie trudny w interpretacji. Konstytucja jest aktem prawnym o najwyższej mocy prawnej w systemie źródeł prawa danego państwa [*konstytucja* w: Wikipedia 2018]. Ten akt prawny „stanowi podstawę całego porządku prawnego, z którego wywodzą się wszystkie inne normy prawne" [Malinowska 2012: 171]. Nie dziwi zatem fakt, że jej interpretacja i orzekanie o zgodności przepisów prawnych z jej treścią wymaga ogromnej wiedzy natury prawnej, o czym może świadczyć konieczność obecności szczególnego organu władzy sądowniczej – Trybunału Konstytucyjnego, czyli instytucji powołanej w celu kontrolowania konstytucyjności prawa. Ocenianie za pomocą wskaźników trudności tekstu dokumentu prawnego szczególnie skomplikowanego na poziomie semantycznym byłoby wręcz kuriozalne, a trudność tego typu tekstu jest w pełni zrozumiała. Oczywiście wspomniane wyżej wskaźniki możemy z powodzeniem wykorzystać w tekście naukowym, w którym zestawione byłyby teksty różnego typu pod względem poziomu trudności. Jednak i w tych tekstach wysoki poziom trudności dokumentu prawnego nie powinien być oceniany negatywnie.

Dla badań językoznawczych (w tym stylistyczno-składniowych) nad tekstem konstytucji szczególnie cenne mogą być statystyki ta-

kie, jak średnia długość zdania, średnia długość akapitu, procent słów trudnych oraz określona w procentach liczba leksemów należących do poszczególnych, podstawowych części mowy. Dane te są często niezbędne przy określaniu właściwości stylistycznych tekstów. Informacje o nich, uzyskane za pomocą aplikacji, z uwagi na nieznany – choć zapewne niewielki – margines błędu powinny być wykorzystywane jedynie w publikacjach naukowych, w których nie jest konieczne podawanie niezwykle precyzyjnych danych. Mogą być zatem przydatne w przypadku prac naukowych, w których konstytucja byłaby zestawiana z innymi tekstami dyskursu prawnego, makrodyskursu społeczno-etycznego lub innych dyskursów. W pracach poświęconych badaniu wyłącznie tekstu konstytucji informacje uzyskane dzięki analizie w aplikacji nie są wystarczające i nie budzą zaufania.

Podsumowując zamieszczone rozważania, można stwierdzić, że aplikacja Jasnopis to narzędzie pomocne szczególnie dla osób tworzących teksty o charakterze użytkowym, których odbiorcą docelowym ma być osoba o przeciętnych kompetencjach rozumienia tekstu. Może być ona również z powodzeniem wykorzystywana w badaniach lingwistycznych, w których zakłada się, że hipotezy powinny być weryfikowane za pomocą danych liczbowych. Z uwagi na możliwość występowania ewentualnych błędów pojawiających się w toku automatycznej analizy statystycznej tekstu, rezultaty analizy tekstu wykonanej przy użyciu omówionego w artykule narzędzia powinny być wykorzystywane tylko w tekstach naukowych, w których podanie ścisłych danych statystycznych lub liczbowych nie jest obligatoryjne.

Bibliografia

Gruszczyński W., Ogrodniczuk M. [2015], *Jasnopis, czyli mierzenie zrozumiałości tekstów użytkowych*, Oficyna Wydawnicza ASPRA-JR, Warszawa.

Gruszczyński W., Broda B., Nitoń B., Ogrodniczuk M. [2014], *W poszukiwaniu metody automatycznego mierzenia zrozumiałości tekstów informacyjnych*, http://nlp.ipipan.waw.pl/Bib/gru:etal:15.pdf (dostęp: 10.04.2018).

Konstytucja, [w:] *Wikipedia. Wolna encyklopedia* [2018], https://pl.wikipedia.org/wiki/Konstytucja (dostęp: 7.03.2018).

Kreatywny, [w:] Dubisz S. (red.) [2003], *Uniwersalny słownik języka polskiego*, t. 2, Warszawa, s. 295.

Leszczak O. [2010], *Lingwosemiotyka kultury. Funkcjonalno-pragmatyczna teoria dyskursu*, Wydawnictwo Adam Marszałek, Toruń.

Malinowska E. [2012], *Konstytucja jako gatunek tekstu prawnego*, Wydawnictwo Uniwersytetu Opolskiego, Opole.

Ruszkowski M. [2004], *Statystyka w badaniach stylistyczno-składniowych*, Wydawnictwo Akademii Świętokrzyskiej, Kielce.

Rynko M. (red.) [2013], *Umiejętności Polaków – wyniki Międzynarodowego Badania Kompetencji Osób Dorosłych (PIAAC)*, http://eduentuzjasci.pl/images/stories/publikacje/ibe-raport-PIAAC-2013.pdf (dostęp: 10.04.2018).