



Językoznawstwo korpusowe. Fot. T. Piotrowski, 2015

Lem – nowator językowy?

Tadeusz Piotrowski / Katedra Leksykografii i Translatologii, Wyższa Szkoła Filologiczna, Wrocław

Czytelnicy utworów Stanisława Lema zauważają, że występują w nich nie tylko różnorodne style, ale także wiele zabawnych, pomysłowych innowacji językowych, takich jak nowe frazy, nowe słowa czy nowe znaczenia¹. Jeżeli przyjąć za Georges'em-Louisem Leclerkim Buffonem, że „*le style est l'homme même*” – styl to sam człowiek, styl odkrywa najważniejsze przymioty autora², badanie stylistyki tekstów Lema powinno, być może, dać nam wgląd w jego osobowość twórczą, a ponadto w epokę, w jakiej owe dzieła powstawały. Sama natura materiału pisarskiego nam to ułatwia, gdyż elementy dzieła literackiego – poszczególne słowa, frazy – są znaczące. Znaczące w dwojaki sposób. Z jednej strony, przejęte z mowy używanej przez autora i czytelników, skonstruowane z jednostek języka, nawiązują do zastanej konwencji społecznej, na mocy której ciągi głosek lub liter traktujemy jako nazwy pojęć, czynności *etc.* Całkowicie oryginalne twory językowe byłyby niezrozumiałe, ponieważ nie odwoływałyby się do wiedzy pisarza i odbiorcy³. Z drugiej – sam tekst wyposaża swe elementy w znaczenia obecne jedynie w danym dziele. Inaczej mówiąc, autor, używając zastanych środków, kreuje nowe elementy języka, umieszczając je w takim kontekście, aby można je było interpretować. Niekiedy bawi się konwencją, owymi zastanymi środkami, tworząc nowe jednostki, jakich w języku poza jego dorobkiem literackim nie znajdziemy – choć czasem sprawiają wrażenie, jakby tylko przez niedopatrzenie nie zaistniały. Czy tego rodzaju

¹ W artykule wykorzystałem fragmenty opisu projektu naukowego, o którym mowa w tekście. Współautorem opisu był dr hab. Łukasz Grabowski.

² G.-L. Leclerc de Buffon, *Discours sur le style. Discours prononcé à l'Académie Française par m. de Buffon le jour de sa réception le 25 aout 1753*, texte de l'édition J. P. Librairie, Ch. Poussielgue, Paris 1896; tekst na stronie: <http://pedagogie.ac-toulouse.fr/philosophie/textes/buffondiscourssurlestyle.htm> (data dostępu: 15 X 2015). Buffon sądził, że dzieło stanie się nieśmiertelne nie dzięki zaletom intelektualnym, ale estetycznym, ponieważ wiedza przemija, a dobry styl zostaje. „*Les ouvrages bien écrits seront les seuls qui passeront à la postérité: la quantité des connaissances, la singularité des faits, la nouveauté même des découvertes, ne sont pas de sûrs garants de l'immortalité: si les ouvrages*

qui les contiennent ne roulent que sur de petits objets, s'ils sont écrits sans goût, sans noblesse et sans génie, ils périront, parce que les connaissances, les faits et les découvertes s'enlèvent aisément, se transportent et gagnent même à être mises en œuvre par des mains plus habiles. Ces choses sont hors de l'homme, le style est l'homme même”.

³ Wierszyk Lema, a w *Cyberiadzie* Elekrybałta, zaczynający się słowami: „Apentula niewdziosek, te będą gruwaśne”, nie jest całkowicie niezrozumiałą, ponieważ do jego elementów niewątpliwie została zastosowana gramatyka polska. Można też ów tekst przełożyć na inne języki – zob. <http://www.wykop.pl/link/1620425/tlumacz-literatury-polskiej-jest-jak-indiana-jones> (data dostępu: 15 X 2015).

nowe jednostki mają szansę jednak przejść do tekstów innych użytkowników języka? Istotny pisarz potrafi ich przekonać, by przejęli nowe środki wyrazu, by zaczęli je wykorzystywać, nie tylko odnosząc się do jego utworów⁴. Te trzy aspekty języka pisarza (tzw. idiolektu⁵) omówię pokrótce w tym artykule. Razem z Łukaszem Grabowskim z Uniwersytetu Opolskiego mamy zamiar zająć się tym zagadnieniem dokładnie, badając teksty Stanisława Lema. Niniejszy artykuł jest rodzajem anonsu, zwiastunem planowanych studiów.

Jeśli chodzi o pierwszy aspekt – wykorzystanie istniejących środków językowych – można postawić pytanie: jakich środków? Czy są to te same jednostki (częstki słowotwórcze, inaczej morfemy), których używano często i chętnie w języku współczesnym autorowi? Czy też są one inne? Po drugie: w jaki sposób Lem sugeruje czytelnikowi, co nowy element znaczy? Czy zasadniczym wskaźnikiem jest jej budowa słowotwórcza, czy też jej usytuowanie w tekście, odpowiednio zarysowany kontekst? I po trzecie: czy nowe jednostki zaproponowane przez Lema zostały zaakceptowane przez czytelników do tego stopnia, że weszły do ogólnego repertuaru użytkowników polszczyzny?

Odpowiedzi na takie pytania są ciekawe nie tylko z punktu widzenia analizy stylu Lema i stosowanych przez niego rozwiązań językowych. Dla lingwisty nowatorskie wykorzystanie elementów języka wskazuje na coś znacznie bardziej istotnego niż określenie, jaki był idiolekt jednej osoby czy jak ona umie się nim bawić. Teksty o innowacyjnej formie stanowią swoisty poligon doświadczalny, na którym testuje się możliwości języka, granice tzw. systemu językowego. Pisarz zwykle nie wychodzi poza zakres tych możliwości. Czytelnik stwierdzi niekiedy, że jakaś jednostka wygląda dziwnie, ale bardzo rzadko ma wrażenie, że jest sprzeczna z jego odczuciem językowym, inaczej mówiąc, że nie jest dopuszczalna na gruncie systemu. Dlatego pytanie o innowacje lingwistyczne Lema dotyczy w istocie także granic systemu polszczyzny z czasów Lema, przede wszystkim potencjału słowotwórczego.

Jak można wobec tego odpowiedzieć na postawione pytania? Co da się w takim razie orzec o języku Lema, a także o języku jego epoki? Odpowiedź musi być krótka: nie sposób stwierdzić nic pewnego, ponieważ nie ma wszechstronnych studiów czy to słownictwa Lema w szczególności, czy to jego tekstów w ogóle. Należałoby jednak tę generalizującą diagnozę rozwinąć, co uczynię w niniejszym artykule. Skonstatuję zwięźle: nie znam lingwistycznych analiz dzieł Lema, a zwłaszcza jego wielu stylistyk. Samej leksyce autora *Solaris* poświęcono natomiast stosunkowo wiele prac. Swoisty przegląd literatury znajduje się w książce Mileny Chomik i Moniki Krajewskiej⁶, rozprawie, zawierającej bardzo dobre studium problemów, jakie napotykają tłumacze literatury *science fiction*; dotyczy ona przekładów utworów Lema na język rosyjski oraz dzieł braci Strugackich na polski. Autorki stwierdzają, że neologizmom autora *Solaris* poświęcono ok. 50 prac, w znakomitej większości artykułów.

Po lekturze części z nich można ogólnie powiedzieć, że nie ma wśród nich takich, które by opisywały wszystkie neologizmy ze wszystkich utworów. Ma się też wrażenie, że w istniejących analizach wybrane neologizmy Lema są najczęściej dobierane tak, by służyły funkcji ilustracyjnej dla ogólnych tez teoretycznych. Owe tezy natomiast odnoszą się do funkcji neologizmów w literaturze *science fiction*, roli, jaką odgrywają w kreowaniu świata przedstawionego przyszłych, potencjalnie, czasów. To rola,

⁴ Od razu pospieszę z wyjaśnieniem, że w rzeczywistości język to żywioł rządzący się swoimi prawami, słabo poddający się więc życzeniom swoich użytkowników.

⁵ Uściślijmy, że fraza „język Lema” oznacza jego słownictwo i stylistykę. Nie sądzimy, by Lem posługiwał się innym językiem niż polski. Sam termin „język”

/ „język polski” / „język ogólny” odnosi się do słownictwa i stylistyki tekstów napisanych przez osoby mówiące po polsku.

⁶ M. Chomik, M. Krajewska, *Od nominacji do kreacji. Rzecz o przekładzie neologizmów science fiction*, Toruń 2011.

oczywiście, niebagatelna. Warto tu przytoczyć, za wymienionymi autorkami, słowa czy to Stanisława Barańczaka: „fantastyka naukowa nie może się bez neologizmu obejść; jest to środek stylistyczny [...], który dla tej dziedziny literatury jest koniecznością”⁷, czy to samego Lema:

Osoby naprzykrzające mi się mniej lub bardziej kąśliwymi uwagami, jakoby utrudniał zaznajamianie się z treściami wspomnień i podróży dzienników słowami-nowotworami, uprzejmie zachęcam do prostego eksperymentu, który wyjaśni im nieuchronność mego postępowania. Niechaj spróbuje taki krytyk opisać jeden dzień swego życia w dużej ziemskiej metropolii, posługując się wyłącznie słowami branymi z dowolnego słownika wydanego przed XVIII stuleciem. Temu, komu takiej próby nie chce się przeprowadzać, uprzejmie odradzam branie do ręki moich książek⁸.

Szczególnym zainteresowaniem wśród badaczy, co zrozumiałe, cieszyły się groteskowe teksty Lema, napisane na wzór oświeceniowych przypowieści Woltera. Jest w nich neologizmów po prostu bardzo dużo, przez co zwracają one uwagę czytelnika. Sam Lem także często mówił o takich neologizmach, np. w wywiadzie-rzeczce przeprowadzonym przez Stanisława Beresia⁹. Powstały one przy użyciu różnych modeli. Z jednej strony, są tam nawiązania do słowotwórstwa niby staropolskiego, które znajdujemy i u Witolda Gombrowicza w *Trans-Atlantyku*, kiedy autor posługuje się nieproduktywnymi formantami; z drugiej – odwołania do słowotwórstwa naukowego. W rezultacie otrzymujemy dekolt swoiście Lemowy. Jedną ze wspomnianych badaczek, Krajewska, opublikowała w 2006 r. polsko-rosyjski słownik neologizmów zawartych w trzech utworach groteskowych Lema¹⁰. Jest to największy i najbardziej systematyczny zbiór innowacji leksykalnych tego autora; zawiera prawie 1500 jednostek. Wrócimy jeszcze do niego.

Jednak Lem nie tylko groteski pisał i nie tylko utwory fabularne. Stąd rola neologizmów w kreowaniu świata przedstawionego, choć ważna, nie stanowi ich roli jedynej. Każdy np. czytelnik traktatów teoretycznych tego autora ma świadomość ich przedziwnej aury stylistycznej. Ten styl można nazwać strasznym – jak to, zdaje się, uczynił Melchior Wańkowicz – w każdym razie nie jest on łatwy do przeniknięcia. Tworzy go nie tylko zawikłana struktura zdań, które są długie, bardzo rozbudowane, wiele w nich inwersji składniowych, ale i, w dużej mierze, właśnie kreacje słowne.

Posłużę się tu pewnym pośrednim przykładem. W *Głosie Pana*¹¹, formalnie sytuującym się pomiędzy dziełem fabularnym – powieścią, a dyskursywnym – traktatem filozoficznym, można znaleźć ok. 120 neologizmów, raczej nie notowanych w słownikach polskich¹². A oto część z tych wynalazków językowych, które – moim zdaniem – dość dobrze ilustrują ową niecodzienną stylistykę oraz umiejętność Lema:

akromegalicznie, akulturowy, algedoniczny, antykosmiczny, astrogeneza, astroinżynier, astroinżynieria, astroinżynierijny, atlantozaur, atomorodny, atomotwórczy, autodeskrypcja, autoewolucja, bezanonimowo, bezprecedentalny, bezprzydatność, bezwieczny, bezwina, biosympatia, ceremonializacja, ćwierćpolitykier, duumvirat, ergodyk,

⁷ S. Barańczak, *Elektrycerze i cyberchanioly*, „Nurt” 1972, nr 8, s. 14.

⁸ S. Lem, *Wizja lokalna*, wyd. 2, Kraków 1983, s. 309.

⁹ *Tako rzeczce... Lem. Ze Stanisławem Lemem rozmawia Stanisław Beres*, wyd. 2, popr., poszerz., Kraków 2002, s. 62.

¹⁰ M. Krajewska, *Polsko-rosyjski słownik Lemowych neologizmów*, Toruń 2006.

¹¹ S. Lem, *Głos Pana*, Warszawa 1968.

¹² Nie przeprowadziłem dokładnych studiów, stąd użycie słowa „raczej”.

fizjoanatomia, gimnazjalizm, hiperlaser, hipotetyzowanie, informacjonista, informacjoniści, infragrupa, kaźdopokoleniowy, kilotrump, komfortować, kosmogeneza, kosmogonista, kriotronowy, megatrump, megazwłoki, międzyludzie, mikrodetonacja, mikroeksplozja, mikrotona, mikrotonowy, odtłumaczyć, psychozoiczny, psychozoik, rejestrat, wysokomolekularny, życiosprawczość, życiosprawczy.

Zwraca uwagę to, że większość tych neologizmów ma wyraźną strukturę słowotwórczą, często dwudzielną; osoba czytana bez trudności „odgadnie” ich semantykę ze znaczenia części, jakie się na nie składają. Oczywiście, jednostki te nie mają konwencjonalnego znaczenia, odbiorca może je dopiero wyposażyć w nie na podstawie ich struktury oraz umiejscowienia w kontekście, a ponadto swojej wiedzy. Mamy wśród tych neologizmów takie, w których wykorzystuje się tradycyjne środki słowotwórcze, jak w wyrazie „atomorodny”, opartym na wzorcu obecnym w formach: „jajorodny”, „korkorodny”. Wyraz ten jednak tak naprawdę jest metaforą, częśćka „-rodny” występuje bowiem zwykle z elementami o znaczeniu odnoszącym się do organizmów żywych. Są u Lema również neologizmy zbudowane z całkowicie obcych części, jak „astrogeneza”. Choć słowniki i encyklopedie ogólne nie notują tego ostatniego leksemu, sądząc z tekstów w internecie, chyba posługują się nim astrolodzy. Wyraz „informacjonista” przypomina słowa „informatyk” i „informatyka”. Lem nie mógł użyć tych terminów w r. 1968, kiedy powieść opublikowano, ponieważ drugi z nich został zaproponowany dopiero w tymże samym roku, przez Romualda Marczyńskiego¹³. U Lema termin „informacjonista” występuje zwykle w zestawieniu „matematyk-informacjonista”, zapewne po to, aby nasunąć czytelnikowi sposób interpretacji neologizmu, gdy jego sens okaże się mało oczywisty; przyrostek „-ista” jest bowiem wieloznaczny. Ostatni przykład pokazuje wyraźnie, jak Lem otwierał przed językiem polskim możliwość przyjęcia nowej jednostki. Przyjęcie to nie nastąpiło, ale jego propozycja pozostaje egzemplifikacją działania systemu. Każdy z takich neologizmów wydaje się wart krótkiej analizy na tle nie tylko elementów leksykalnych swej epoki, ale także środków słowotwórczych, jakich się w niej używało. Tego właśnie chcielibyśmy dokonać.

Dlatego – by powrócić do kwestii ogólnych – gdy badacz ogranicza się wyłącznie do neologizmów Lema z utworów groteskowych, żartobliwych, wówczas mocno zawęża pole swoich studiów, tracąc możliwość wyczerpującej analizy potencji słowotwórczej Lema. Traci także okazję do porównania środków językowych używanych w tekstach humorystycznych ze środkami umieszczonymi w tych poważnych. Te pierwsze miały nikłą szansę zaistnienia w języku, te drugie (jak „informacjonista”) – zdecydowanie większą¹⁴. Ponadto odnoszę wrażenie, że właśnie w dziełach, których aspekt językowy nie rzuca się w oczy tak bardzo jak grotesek, potencjał wyrazotwórczy Lema jako użytkownika polszczyzny jest „prawdziwszy”, właśnie w nich się odsłania. Jak wiadomo, autor *Solaris* wplatał neologizmy także do prywatnych listów. Dzięki temu, że obecnie ogłaszane są one drukiem, można wziąć pod uwagę również ich treść¹⁵. Niestety, do tej pory prób wydobycia wszystkich neologizmów ze wszystkich tekstów Lema nie było, mamy zamiar postarać się o to w projektowanych badaniach.

¹³ Zob. Romuald W. Marczyński (1922–2000), <http://apw.ee.pw.edu.pl/tresc/sylw/marczyński.htm> (data dostępu: 27 X 2015).

¹⁴ Za formą „informatyk” przemawia to, że jest dużo krótsza i bezpośrednio związana z nazwą dziedziny – „informatyka”. Za potencjalną formą „informacjonistyka” i derywatem „informacjonista” – to, że są one oparte na terminie „informacja”. „Informatyka” to tak naprawdę zapożyczenie z francuskiego.

¹⁵ Zob. W. Zemek, *Listy a literatura. Korespondencja Stanisława Lema z Michaelem Kandlem*, praca magisterska napisana pod kierunkiem prof. J. Jarczyńskiego, Uniwersytet Jagielloński 2009, <http://solaris.lem.pl/ksiazki/beletrystyka/pirx/109-teksty-o-lemie/prace-dyplomowe/527-listy-literatura-korespondencja-lem-kandel> (data dostępu: 15 X 2015).

Trzeba powiedzieć, że wbrew pozorom niełatwo określić, co jest neologizmem. Zmiany językowe zwykle nie mają charakteru rewolucyjnego, innowację lingwistyczną – jak zapożyczenie, neologizm – przyswajają się stopniowo, a to trwa lata, jeśli nie dziesiątki lat. Oczywiście, wyrazy modne mogą pojawić się szybko, ale równie szybko przemijają, nie pozostawiając po sobie śladu. W tekstach napisanych po polsku będzie sporo takich elementów, jakich badacz nie potraktuje jako jednostek języka polskiego. Oprócz błędów i pomyłek, zdarzających się nieraz, wystąpią również zapożyczenia, cytaty, makaronizmy. Które z nich traktować jako polskie wyrazy? Jako typowy przykład wśród utworów Lema może służyć *Szpital Przemienienia*, obejmujący długie passusy po niemiecku. W tym przypadku decyzja jest akurat prosta. Pojawiają się jednak i takie twory, co do których badacz będzie się wahał, czy zaliczyć je do repertuaru jednostek polskich, ponieważ są bardzo rzadkie. Stąd też najczęściej językoznawcy odwołują się do istniejących rejestrów, czyli do słowników, przerzucając odpowiedzialność za decyzje na słownikarza. Jeżeli danego wyrazu nie ma w słowniku, to znaczy, że jest neologizmem.

Największym zbiorem współczesnego słownictwa polskiego jest *Słownik gramatyczny języka polskiego*, zawierający ok. 334 733 haseł (technicznie: leksemów)¹⁶. Będziemy się do niego często odwoływali. Nawet największe zbiory leksyki polskiej nie są jednak wyczerpujące – choćby właśnie z tego względu, że nie wszyscy słownikarze zgodzą się co do tego, czy jakiś wyraz jest polski, czy nie; poza tym tak naprawdę do większości istniejących wyrazów można dodać przedrostki lub przyrostki, liczba potencjalnych elementów leksykalnych jest – jak niektórzy twierdzą – nieskończona. Trzeba więc badać teksty. Obecny rozwój językoznawstwa w Polsce pozwala na takie badania, nie tylko dlatego, że mamy dostęp do ogromnych zasobów tekstów zamieszczonych na stronach internetowych (ich użycie jest jednak mało efektywne, ponieważ należałoby zweryfikować ich zasoby¹⁷), ale także dlatego, że dla potrzeb nauki powstały wyspecjalizowane bazy tekstów w postaci cyfrowej, zwane korpusami. Największy korpus polszczyzny liczy sobie ponad 1,5 mld wyrazów. To Narodowy Korpus Języka Polskiego, w skrócie NKJP, dostępny pod adresem: www.nkjp.pl¹⁸.

Aby zilustrować wskazane problemy, posłużę się przykładami. W słowniku neologizmów Lema autorstwa Krajewskiej¹⁹ pojawia się hasło „defenestracja”, a w nim samym mamy inny neologizm – „flagellacja”²⁰. Oba wyrazy zwróciły moją uwagę, gdyż na podstawie intuicji sądziłem, że występują one w tekstach polskich. Z pewnością słowo „defenestracja” nie jest nowe, choć, co prawda, w dużym słowniku języka polskiego pod redakcją Doroszewskiego nie odnotowano go²¹; można je jednak znaleźć nie tylko w encyklopediach polskich, ale także w ogólnych słownikach (np. tych opublikowanych przez PWN w latach 70. XX w.²²) oraz w przedwojennych słownikach wyrazów obcych. Natomiast w słownikach polskich brak hasła „flagellacja”, choć występuje „flagellant”; osoba umiejąca używać tego ostatniego

¹⁶ Z. Saloni [et al.], *Słownik gramatyczny języka polskiego*, Warszawa 2007 (wyd. 2: Warszawa 2012). Najobszerniejsza wersja słownika jest dostępna w internecie pod adresem: <http://sgjp.pl> (data dostępu: 15 X 2015).

¹⁷ Bardzo istotne jest to, że jeżeli jakaś wyszukiwarka internetowa podaje liczbę wystąpień pewnej formy, to nie tylko nie wiadomo, co dokładnie podaje (faktyczne wystąpienia jednostki czy liczbę stron internetowych, na których się ona pojawia), lecz również ta liczba niczego nie mówi, ponieważ brak punktu odniesienia – a takowy byśmy mieli, gdybyśmy wiedzieli, ile jest wszystkich jednostek leksykalnych w Internecie. Ale nie wiemy.

¹⁸ Korpus jest opisany w publikacjach dostępnych na wskazanej stronie, np. w: *Narodowy Korpus Języka Polskiego*, red. A. Przepiórkowski [et al.], Warszawa 2012. Analizą tekstów cyfrowych przy użyciu narzędzi korpusowych, wyko-

rzystując statystykę, zajmuje się dział językoznawstwa komputerowego zwany językoznawstwem korpusowym. On także został opisany w ww. publikacjach.

¹⁹ M. Krajewska, *op. cit.*

²⁰ Taki układ wskazuje, że słownik jako dzieło leksykograficzne nie został dobrze wykonany. Każdy neologizm powinien mieć własne hasło.

²¹ *Słownik języka polskiego*, red. W. Doroszewski, t. 1–10, Warszawa 1958–1969; wersja *online*: <http://sjp.pwn.pl/doroszewski>. Słownik zawiera ok. 130 000 haseł.

²² Dla większej precyzji – aby nie uwzględniać efemeryd – słowniki zwykle zaczynają rejestrować wyrazy dopiero wtedy, kiedy od pewnego czasu używa się ich w tekstach. Stąd słowniki z lat 70. XX w. opisywały leksykę z poprzedniej dekady.

leksemu raczej nie będzie miała problemu z utworzeniem derywatu czynności. Faktycznie, w słowniku z 1939 r.²³ znaleźć można hasło „flagellancja”. Forma „flagellacja” pojawia się również w NKJP, a ponadto – dość często – w tekstach internetowych. Trudno więc uznać oba wyrazy za neologizmy²⁴. Trzeba tu więc podkreślić, że nie wystarcza intuicja nawet doświadczonego leksykografa – a za takiego się uważam – aby zakwalifikować jakąś formę jako innowację.

Zgodnie z zasadami językoznawstwa korpusowego, w proponowanym projekcie neologizmy zostaną wyodrębnione automatycznie, z większości dzieł drukowanych Lema, a także, w pewnym wyborze, z dzieł niedrukowanych (korespondencji). Ze względu na to, że do przetwarzania komputerowego niezbędne są teksty w postaci elektronicznej, skorzystamy z tego, że większość znanych utworów Lema – zarówno literackich, jak i eseistycznych – wydana została w takiej formie. Istnieją 44 ebooki z prozą tego autora²⁵. Na podstawie wstępnych badań można ostrożnie szacować, że ten korpus tekstów (bez korespondencji) liczy ok. 2–2,5 mln wyrazów. Te same badania wykazały, że komercyjne ebooki bez większego problemu poddają się dalszemu przetwarzaniu. Korpus danych uzupełniony będzie o listy Lema²⁶. Ich treści nie są opublikowane w formie cyfrowej, zwrócimy się do wydawców o dostarczenie ich kopii elektronicznych lub sami opracujemy przynajmniej ich część, skanując edycje książkowe i podając powstałe w ten sposób obrazy obróbce, aby mieć do dyspozycji zapis tekstu. Dzieła Lema przekonwertujemy (do formatu XML) oraz opatrzymy tzw. metadanymi – informacjami o typie tekstu, roku wydania *etc.* Nie będziemy dążyć do ustalenia jak najlepszej postaci tych dzieł, tzw. formy kanonicznej. Piszę o tym choćby dlatego, że niektóre utwory, zwłaszcza teoretyczne, mają wiele wersji. Przygotowanie kanonicznej to zupełnie inne zadanie; my będziemy traktować treści ebooków jako obowiązujące, zgodne z wolą spadkobierców.

Na podstawie tych tekstów automatycznie opracowane będą listy wyrazów, jakie się w nich znajdują, odrębnie dla każdego utworu, a następnie scalone dla wszystkich utworów i ich grup (jak traktaty, eseje, powieści, listy). W wyniku kolejnych procesów abstrahowania zostaną wykonane spisy leksemów (wyrazów sprowadzonych do postaci tzw. podstawowej, hasłowej²⁷). Z tych list przez porównanie ich zawartości ze spisami wyrazów umieszczonych w słownikach oraz używanych w tekstach polskich wyodrębnimy neologizmy; jak widzieliśmy, nie są to dokładnie te same zbiory wyrazów. Posłużą nam do tego narzędzia komputerowe opracowane dla celów konsorcjum CLARIN-PL. Celem jego działań jest wykreowanie ogólnoeuropejskiej infrastruktury naukowej, która ma umożliwić badaczom z dziedziny nauk humanistycznych i społecznych przetwarzanie bardzo dużych zbiorów tekstów²⁸. Konsorcjum obejmuje takie instytucje jak Grupa Technologii Językowych G4.19 Politechniki Wrocławskiej, IPI PAN w Warszawie, Uniwersytet Łódzki. Dla potrzeb obróbki zasobów w ramach CLARIN-PL wykonano indeksy wyrazów, jakie znajdują się w tekstach polskich; ich zawartość przewyższa liczbowo to, co znamy ze słowników tradycyjnych.

²³ Trzaski, Everta i Michalskiego encyklopedyczny słownik wyrazów obcych. *Pochodzenie wyrazów, wymowa, objaśnienia pojęć, skróty, przysłowia, cytaty*, red. S. Lam, Warszawa 1939. Trzeba powiedzieć, że słownik ten nie jest zbyt wiarygodny, zawiera bardzo dużo wątpliwych haseł.

²⁴ Pewną wskazówkę dla autorki słownika (M. Krajewska, *op. cit.*) mogło stanowić to, że ekwiwalenty rosyjskie obu wyrazów mają taką samą postać jak polska oraz że notują je słowniki rosyjskie. To oznacza, że są to zaświadczane formy.

²⁵ Zob. „*Dzieła Stanisława Lema*”, <http://www.publio.pl/dzieła-stanisława-lema,s3153.html> (data dostępu: 15 X 2015).

²⁶ Np. S. Lem, S. Mrozek, *Listy. 1956–1978*, Kraków 2011; S. Lem, *Stawa i Fortuna. Listy do Michaela Kandla 1972–1987*, Kraków 2013.

²⁷ Trzeba tu zauważyć, że jeżeli jakiś neologizm występuje jedynie w przypadku zależnym, np. w dopełniaczu, mimo uwzględnienia kontekstu mogą być pewne problemy z ustaleniem formy podstawowej. Dlatego neologizmy będą opracowywane na poziomie typów, czyli klas form znajdujących się w tekstach.

²⁸ Więcej informacji oraz dostęp do gotowych narzędzi można uzyskać na stronie internetowej: <http://clarin-pl.eu/pl/strona-glowna> (data dostępu: 15 X 2015).

Po wyodrębnieniu neologizmów dokona się ich klasyfikacji ze względu na ich występowanie w typie tekstu, jego chronologię, a także ich funkcjonowanie w tekście, poprzez analizę kontekstów. Konteksty w językoznawstwie korpusowym można badać również półautomatycznie, wyliczając za pomocą mierników statystycznych współwystępowanie wyrazów w tekstach (to tzw. kolokacje), a ponadto rozpatrując zbitki wyrazów, w jakich dany wyraz się pojawia. Przyjęte metody w gruncie rzeczy pozwolą na wykonanie ogólnego studium słownictwa Lema oraz jego stylu – dzięki opracowaniu podstawowych indeksów stylistycznych i wskaźników stylometrycznych (np. długość jednostek tekstu, takich jak wyrazy, akapity, zdania; gęstość i bogactwo leksykalne oraz wzajemne podobieństwo rozmaitych dzieł Lema powstałych w różnych odstępach chronologicznych). Tym samym, oprócz wszechstronnej analizy neologizmów, dokonana zostanie analiza statystyczno-stylometryczna twórczości Lema, również w ujęciu diachronicznym. Wreszcie porównamy język Lema z językiem tekstów jego czasu.

Badania pozwolą zweryfikować wiele hipotez cząstkowych, np. w odniesieniu do neologizmów: na ile są one podobne do wyrazów zapożyczonych, czy to długością, czy to strukturą morfologiczną, czy to upodobaniami kolokacyjnymi (intuicyjnie wydaje się, że powinny być bardzo podobne). Każda z tych czynności można wykonać automatycznie lub półautomatycznie. Oczywiście, surowe wyniki narzędzi informatycznych muszą zostać zinterpretowane, co zajmie sporo czasu. Wszystkie efekty studiów będą udostępnione na stronach internetowych. Nie zamierzamy wydać naukowego słownika języka Lema, choć w Polsce jest długa tradycja publikowania takich dzieł. Językoznawcom wystarczą dokładne indeksy, aby mogli badać rozmaite aspekty języka Lema. Niewykluczone, że warto byłoby za to przygotować do druku słownik lub indeks neologizmów Lema; edycja tego rodzaju służyłaby dalszym studiom ekwiwalentów idiolektu autora *Solaris* z tłumaczeń obcych, podobnym do tych wykonanych przez Krajewską i Chomik. Taki wielojęzyczny materiał może być bardzo ciekawy dla teoretyków przekładów.

Tak, w ogromnym skrócie, będzie wyglądała praca nad analizą jednostek leksykalnych dzieł Stanisława Lema. Pozostaje jeszcze napisać kilka słów na temat badań nad występowaniem tychże jednostek w innych tekstach. Intuicyjnie rzecz sądząc, nowe twory Lema nie wywarły wpływu na słownictwo języka polskiego. Choć pojawiają się poza dziełami autora *Solaris*, w większości używane są jako aluzja do jego prozy. Pobieżne sprawdzenie frekwencji najbardziej bodaj znanego wyrazu tego rodzaju, „sepulka”, zdaje się potwierdzać te intuicje. Występuje on w NKJP 26 razy na 1,5 mld słów, jest to więc wyraz rzadki; najczęściej używa się go jako nazwy serwera jednej uczelni. Na stronach internetowych także można go znaleźć, nierzadko w nazwach użytkowników forów. Takie badania należy jednak przeprowadzić systematyczniej. Dlatego w NKJP oraz w tekstach zamieszczonych w sieci będziemy wyszukiwać konteksty, w jakich występują wszystkie neologizmy Lema. Po wykonanej analizie, którą także da się zautomatyzować, sprawdzimy, czy te innowacje są wykorzystywane poza jego utworami. Dopiero po wykonanych badaniach będzie można spróbować zinterpretować ich wyniki, aby stwierdzić, na ile język Lema jest podobny do innych tekstów jego czasu czy był oryginalny, a jeśli tak, to co o tym decydowało, i jakie elementy tego idiolektu są wykorzystywane przez innych użytkowników polszczyzny. Niebagatelną zaś zaletą korpusowych metod analizy tekstów jest to, że nie wiadomo, jakie wyniki one dadzą.

Słowa kluczowe / Keywords

językoznawstwo korpusowe, idiolekt, słownictwo, stylistyka, neologizm, analiza statystyczna / corpus linguistics, idiolect, vocabulary, stylistics, neologism, statistical analysis

Bibliografia / References

- Barańczak Stanisław, *Elektrycerze i cyberchanioty*, „Nurt” 1972, nr 8.
- Chomik Milena, Krajewska Monika, *Od nominacji do kreacji. Rzecz o przekładzie neologizmów science fiction*, Toruń 2011.
- Krajewska Monika, *Polsko-rosyjski słownik Lemowych neologizmów*, Toruń 2006.
- „Dzieła Stanisława Lema”, <http://www.publio.pl/dzieła-stanisława-lema,s3153.html> (data dostępu: 15 X 2015).
- Lem Stanisław, *Głos Pana*, Warszawa 1968.
- Lem Stanisław, *Stawa i Fortuna. Listy do Michaela Kandla 1972–1987*, Kraków 2013.
- Lem Stanisław, Mrożek Sławomir, *Listy. 1956–1978*, Kraków 2011.
- Lem Stanisław, *Wizja lokalna*, Kraków 1983.
- Narodowy Korpus Języka Polskiego, <http://nkjp.pl> (data dostępu: 15 X 2015).
- Narodowy Korpus Języka Polskiego, red. A. Przepiórkowski [et al.], Warszawa 2012.
- Saloni Zygmunt [et al.], *Słownik gramatyczny języka polskiego*, Warszawa 2007 (wyd. 2: Warszawa 2012); wersja online pod adresem: <http://sgjp.pl> (data dostępu: 15 X 2015).
- Zemek Wojciech, *Listy a literatura. Korespondencja Stanisława Lema z Michaelem Kandlem*, praca magisterska napisana pod kierunkiem prof. J. Jarzębskiego, Uniwersytet Jagielloński 2009, <http://solaris.lem.pl/ksiazki/beletrystyka/pirx/109-teksty-o-lemie/prace-dyplomowe/527-listy-literatura-korespondencja-lem-kandel> (data dostępu: 15 X 2015).

prof. dr hab. Tadeusz Piotrowski (t.piotrowski@wsf.edu.pl)

Kierownik Katedry Leksykografii i Translatologii w Wyższej Szkole Filologicznej we Wrocławiu. Zajmuje się słownictwem języka polskiego i angielskiego, a zwłaszcza sposobami jego opisu w słownikach, jak również teorią i praktyką tłumaczeń.

Tadeusz Piotrowski (Department of Lexicography and Translatology, Philological School of Higher Education in Wrocław) / Lem – language innovator?

The paper describes briefly the present state of research into Stanisław Lem’s language, especially into his vocabulary, in particular his linguistic innovations (called neologisms), and outlines the main principles of a study into Lem’s language that the author plans (with Łukasz Grabowski). It is argued that so far it has been predominantly the humorous stories by Lem, such as *Cyberiada* (*The Cyberiad*) that were studied, and that a future study should take into account all neologisms by Lem from all his works. The planned study will do that, using the methods of a branch of computer linguistics called corpus linguistics. Most texts by Lem are available as ebooks now, and the authors will use them as their corpus, adding texts from his correspondence. All the texts will be analysed by statistical methods, and a list of neologisms will be produced automatically. Lem’s texts, his vocabulary and neologisms will be studied in comparison with the texts from his period.