

---

## **Tekstów świat. Przyczynek do makroanalitycznej monografii czasopisma literaturoznawczego<sup>1</sup>**

---

Maciej Maryl

---

INSTYTUT BADAŃ LITERACKICH PAN

*Pani Dalloway powiedziała, że sama kupi kwiaty.*

*Lucy ma wyliczoną każdą minutę. Zdejmie drzwi, przyjdą ludzie od Rumpelmayera.*

*A zresztą, pomyślała Klarysa Dalloway, co za ranek – jak gdyby ofiarowany w podarku  
dzieciom na plaży.*

W semestrze zimowym roku akademickiego 2001/2002, tj. w pierwszym semestrze mojego pierwszego roku akademickiego, miałem przyjemność analizować początek *Pani Dalloway* na zajęciach, które dla studentów kolegium MISH prowadził prof. dr hab. Ryszard Nycz. Zajęcia nosiły dość ogólny tytuł „Literatura Europejska XX wieku w perspektywie teoretyczno-kulturowej”, pod którym kryła się analiza przemian narracji w wielkich powieściach światowych. Nie sposób było wymagać od seminarzystów przeczytania w tak krótkim czasie całych książek takich autorów jak Proust, Musil, Dos Passos, Berent, Kafka czy Woolf. Zresztą nie chodziło tu o dyskusję o samych powieściach, tylko o przemiany w sposobie pisania. Nycz zastosował zatem ciekawy zabieg dydaktyczny – przygotował kserokopie z przepiszanymi pierwszymi akapitami tych powieści, które analizowaliśmy na seminarium zdanie po zdaniu. Próbką tekstu dawała wgląd w styl autorów, a kolejne przykłady – w charakterystykę powieści modernistycznej.

---

1 Praca naukowa finansowana w ramach programu Ministra Nauki i Szkolnictwa Wyższego pod nazwą „Narodowy Program Rozwoju Humanistyki” w latach 2015–2018. Projekt: „Polska Bibliografia Literacka – laboratorium wiedzy o współczesnej kulturze polskiej” NR 0061/NPRH3/H11/82/2014.

Nosiłem się wówczas z zamiarem napisania pracy zaliczeniowej poświęconej analizie nie dziesięciu, ale stu początków powieści, nic z tego jednak nie wyszło, ponieważ – o ile pamiętam – przerósł mnie ogrom spodziewanej pracy, czyli przepisywania (tak!) i ręcznej analizy tekstów. Dopiero kilka lat później poznałem odpowiednie narzędzia i metodologię. Brakowało mi też, trzeba przyznać, porządnego pytania badawczego, które pozwoliłoby uzasadnić dany dobór tekstów i kierunek ich analizy. W niniejszym tekście niejako wracam do tego zagadnienia, tym razem uzbrojony w teksty, dane oraz narzędzia i metody ich analizy. Choć nadal nie wiem, co mogłoby wyjść z analizy początków stu powieści (i dlaczego tylko sto, a nie sto tysięcy tekstów?), chciałbym zaproponować spojrzenie z lotu ptaka na 25 lat dwumiesięcznika „Teksty Drugie. Teoria literatury, krytyka, interpretacja”.

Przyjętą tu metodę nazywam za Matthew Jockersem „makroanalizą”<sup>2</sup>. Polega ona, w dużym uproszczeniu, nie tyle na badaniu poszczególnych tekstów, ile na operowaniu korpusami i zestawieniami danych w celu wykrycia procesów wyższego rzędu. Zestawienie strategii dydaktycznej Nycza z proponowanymi tu metodami makroanalitycznymi szybko ujawnia najważniejsze różnice między czytaniem uważnym (*close reading*) a czytaniem z dystansu (*distant reading*). Pierwszy termin odnosi się do tradycji hermeneutycznej i dekonstrukcji – z ogromu produkcji literackiej zostaje wybrany (nie wdajemy się tu w dyskusję, przez kogo) zestaw tekstów uznanych za istotne, wzorcowe lub symptomatyczne dla danego zjawiska historycznoliterackiego. Analiza tej próbki pozwala śledzić procesy historycznoliterackie na podstawie kanonu. Czytanie z dystansu zakłada wejście na poziom meta i obserwowanie zjawisk i procesów literackich na wielką skalę, bez wdawania się w takie szczegóły jak interpretacja pojedynczego tekstu. Weźmy analizy Franco Morettiego, który w *Wykresach, mapach, drzewach*<sup>3</sup> posługuje się metadanymi (tytuły powieści, kategoryzacja genologiczna), by pokazać złożone procesy historycznoliterackie, jak cykl życia gatunków powieściowych. Różnica zasadnicza polega zatem na tym, że makroanaliza nie posługuje się próbką, tylko możliwie jak największym zasobem danych. I takie to właśnie założenia przyświecają tej pracy.

2 Por. M. Jockers *Macroanalysis: Digital Methods and Literary History*, University of Illinois Press, Urbana 2013. Podobnym badaniom poświęcona jest część prac zebranych w numerze 2/2014 „Tekstów Drugich”.

3 F. Moretti *Wykresy, mapy, drzewa. Abstrakcyjne modele na potrzeby historii literatury*, przeł. T. Bilczewski, A. Kowalcze-Pawlik, Wydawnictwo UJ, Kraków 2016.

Wykorzystanie danych tekstowych i bibliograficznych do badania procesów życia literackiego na szeroką skalę ma stosunkowo krótką historię, wynikającą z braku dostępności materiału analitycznego. Postępująca digitalizacja i standaryzacja zasobów przyczyniła się do gwałtownego rozwoju takich projektów w ostatnich latach. Warto wspomnieć tu choćby prace Richarda Jeana So i Hoyta Longa z Chicago Text Lab<sup>4</sup>, którzy łączą socjologię relacyjną i teorię sieci z porównawczymi badaniami życia literackiego i powiązań między twórcami<sup>5</sup>. Andrew Goldstone i Ted Underwood skupiają się zaś na makroanalizie literaturoznawstwa amerykańskiego, wykorzystując modelowanie tematyczne<sup>6</sup> (*topic modeling*) w badaniu dwudziestowiecznych, amerykańskich czasopism naukowych.<sup>7</sup> Innym przykładem, najbliższym chyba zamierzeniom niniejszej pracy, jest projekt przygotowany na 40-lecie czasopisma „Signs”, poświęconego krytyce feministycznej<sup>8</sup>. Użytkownicy mogą przeglądać zawartość pisma, modele tematyczne i sieci współcytowań.

Niniejsza praca – czerpiąc inspirację z powyższych przykładów – ma na celu wstępne zarysowanie monografii „Tekstów Drugich” przy zastosowaniu metod makroanalitycznych. Ze względu na ograniczone miejsce tekst ma charakter przyczynkarski, skupia się na możliwościach wykorzystania metadanych o tekstach kultury (czyli bibliografii oraz wyników statystycznych analiz samych tekstów). Próbuję w jednym tekście połączyć różnorodne, opisane wyżej podejścia – a zatem zebrać i zinterpretować dane literaturoznawcze i wykorzystać je do przedstawienia wstępnej monografii czasopisma. Celem dalekosiężnym jest refleksja nad dalszym wykorzystaniem takich metod do analizy procesów życia literackiego, a w szczególności przemian kierunków refleksji literaturoznawczej w Polsce. W niniejszym tekście zamierzam:

---

4 Zob. <https://lucian.uchicago.edu/blogs/literarynetworks/> (23.10.2016).

5 H. Long, R.J. So *Network Analysis and the Sociology of Modernism*, „boundary 2” 2013 nr 40 (2), s. 147–182; H. Long, R.J. So *Network Science and Literary History*, „Leonardo” 2013 nr 3 (46), s. 274–274.

6 Dobrym wprowadzeniem do tej problematyki jest artykuł M. Edera *Słowa znaczące, słowa kluczowe, słowozbiory – o statystycznych metodach wyszukiwania wyrazów istotnych*, „Przegląd Humanistyczny” 2016 nr 3.

7 A. Goldstone, T. Underwood, *What Can Topic Models of PMLA Teach Us About the History of Literary Scholarship?*, „Journal of Digital Humanities” 2012 No. 1; A. Goldstone, T. Underwood, *The Quiet Transformations of Literary Studies: What Thirteen Thousand Scholars Could Tell Us*, „New Literary History”, Summer 2014 No. 3, s. 359–384.

8 Zob. <http://signsat40.signsjournal.org/> (17.10.2016).

1) opisać statystycznie ćwierćwiecze „Tekstów Drugich”, wskazując pewne trendy w skali makro; 2) zbadać najważniejsze odniesienia autorów artykułów i na ich podstawie zbudować model grup intelektualnych w historii pisma; 3) zarysować perspektywy analizy tematycznej „Tekstów Drugich”.

## **Materiał**

Zebrań danych w dobie repozytoriów cyfrowych, bibliografii online i składu komputerowego wydaje się błahostką. W rzeczywistości nie jest już tak łatwo – pojawiają się pewne problemy z kompletnością zasobów i jakością danych. Samo przygotowanie i przetworzenie materiału zabrało prawdopodobnie więcej czasu niż ich analiza. W badaniu wykorzystano trzy źródła danych:

- 1 Dane bibliograficzne z bazy RCIN (Repozytorium Cyfrowe Instytutów Naukowych) za lata 1990–2012. W bazie są przechowywane archiwalne numery czasopisma wraz z następującymi informacjami: autor, współautorzy, tłumacz, tytuł, liczba stron, numer, rok, numery stron, bibliografia. Materiał pobrany z bazy został<sup>9</sup> odpowiednio przetworzony, by posłużyć do analiz i wizualizacji sieci odniesień oraz analizy długości tekstów, prezentowanych w dalszej części pracy.
- 2 Bibliografia zawartości „Tekstów Drugich” (BTD) dostępna na stronie czasopisma, uzupełniana na bieżąco. Zasoób powstał w 2012 roku na podstawie drukowanych (a następnie zeskanowanych i przetworzonych automatycznie) bibliografii zawartości pisma. Stąd też w bazie pojawia się trochę literówek, a starsze roczniki nie odnotowują tłumaczy. Dane z tego zasobu wykorzystano głównie do uzupełnienia metadanych z bazy RCIN oraz do prowadzenia statystyk dotyczących autorów.
- 3 Pliki pełnotekstowe „Tekstów Drugich”. Teksty za lata 1990–1998 pochodzą ze skanów z OCR, które przetworzyłem za pomocą programu Calibre i wstępnie wyczyściłem (np. z powtarzających się w nagłówkach tytułów lub nazw działów). Teksty za lata 1999–2014 zostały pozyskane z plików redakcyjnych (przesyłanych do składu) i przetworzone na format tekstowy. Zastosowana metoda siłą rzeczy dopuszcza pewien szum w danych (nieprawidłowe rozpoznania programu OCR dla starszych roczników, błędy wychwycone dopiero w korekcie szpaltowej), ale nie powinien on zbytnio

---

9 Dziękuję Marcinowi Helińskiemu z Poznańskiego Centrum Superkomputerowo Sieciowego za przygotowanie odpowiednich zrzutów z bazy RCIN.

rzutować na wyniki analizy. Wszystkie teksty zostały zlematyzowane i stworzono listy frekwencyjne lematów i cech gramatycznych za pomocą narzędzi lingwistycznych opracowanych przez konsorcjum CLARIN-PL<sup>10</sup>.

### **Ile tekstów w „Tekstach”**

Zacznijmy od podstawowych statystyk dotyczących ćwierćwiecza „Tekstów Drugich”. Na dobrą sprawę trudno jednoznacznie powiedzieć, ile tekstów opublikowano w tym okresie w czasopiśmie – zależy to od tego, co uznamy za tekst. W korpusie pełnotekstowym mamy 2533 pozycje, bez listów, sprowotowań, not o autorach i streszczeń. Bibliografia zawartości TD za ten okres odnotowuje zaś 2609 pozycji, zdaje się stanowić bardziej rzetelne źródło.

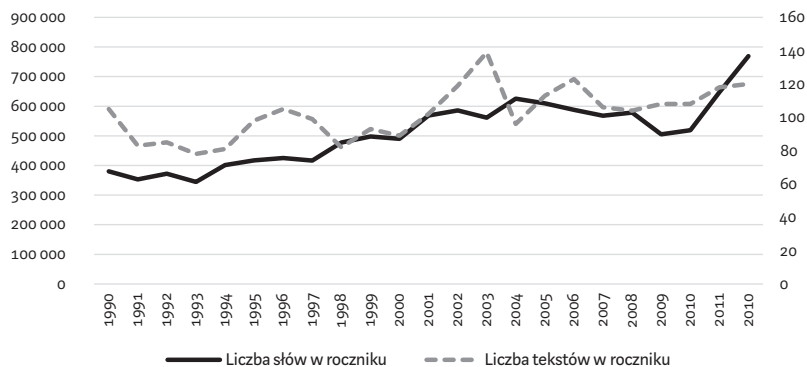
W omawianym okresie ukazało się 125 zeszytów pisma, w tym 23 numery podwójne i jeden potrójny, co daje łączną liczbę 150 numerów. Jest to średnio 21 tekstów na numer i 105 artykułów w roku. Objętość systematycznie rośnie – w pierwszym roku istnienia pisma ukazały się 83 artykuły, które łącznie zajęły 914 stron, a 22 lata później, w 2012 roku, mieliśmy już 111 artykułów na 1563 stronach, co daje wzrost o 34% i 71%. Czasopismo od początku stopniowo zwiększa objętość, co pokazuje Wykres nr 1 (opracowany na podstawie danych RCIN).

W ciągu ćwierćwiecza (wg BTd) pismo gościło na swoich łamach 1194 autorów, którzy łącznie opublikowali 2609 tekstów, co daje średnio 2,22 tekstu na autora. Jednakże, jeśli posłużyć się medianą (miarą odrzucającą wartości skrajne), otrzymamy liczbę 1. Oznacza to, że zdecydowana większość autorów (63,4%) opublikowała tylko jeden tekst, a łącznie 79,2% autorów opublikowało nie więcej niż dwa artykuły na łamach TD. 161 autorów (13,5%) opublikowało od 3 do 5 tekstów, 54 osoby (4,5%) od 6 do 10, ale już tylko 33 (2,8%) osoby opublikowały więcej niż 10 artykułów. Świadczy to o dużym przepływie autorów pisma i wąskiej grupie stałych twórców. Rzeczywiście – debiutanci stanowią istotną grupę w każdym numerze pisma. Wykres nr 2 przedstawia ten odsetek – dla lepszej czytelności wykresu skumulowano liczbę debiutantów z kolejnych numerów w danym roku.

---

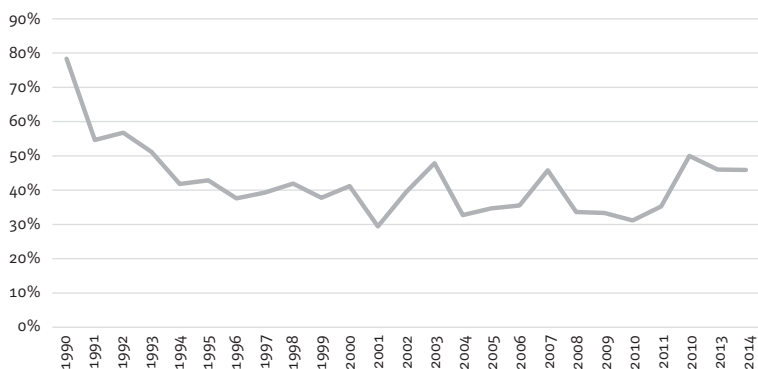
10 Dziękuję dr. inż. Tomaszowi Walkowiakowi i dr. inż. Maciejowi Piaseckiemu z Politechniki Wrocławskiej za przygotowanie plików i przeprowadzenie analiz.

## Wzrost objętości TD (1990–2012)



WYKRES NR 1. Wzrost objętości TD (1990–2012).

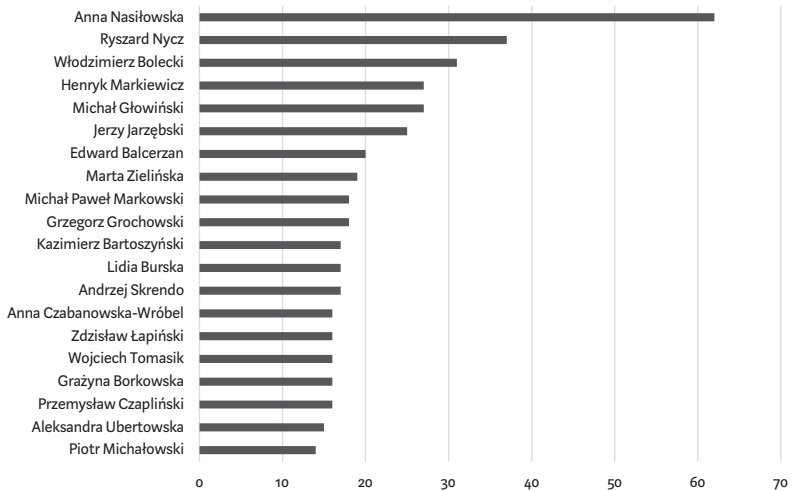
## Zainteresowanie najpopularniejszymi twórcami



WYKRES NR 2. Odsetek debiutantów w kolejnych numerach (wg BTĐ).

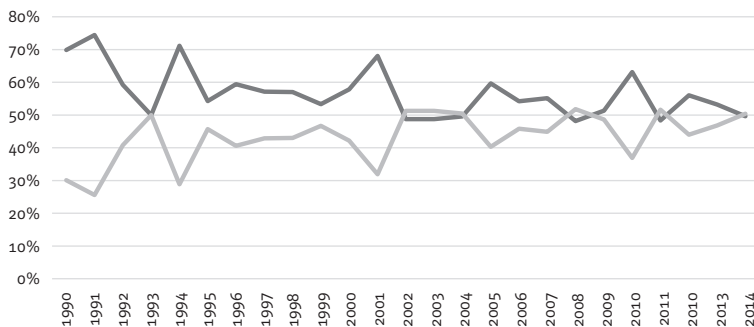
Niekwestionowaną liderką jest tu Anna Nasiłowska, zastępczyni redaktora naczelnego i częsta autorka wstępów, która opublikowała 62 artykuły, czyli niemal dwukrotnie więcej niż kolejny w rankingu Ryszard Nycz (Zob. Wykres nr 3).

## 20 najczęściej publikujących autorów TD (1990–2014)



WYKRES NR 3. Dwadzieścioro najczęściej publikujących autorów TD (wg BTD).

## Odsetek mężczyzn i kobiet wśród autorów



WYKRES NR 4. Odsetek mężczyzn i kobiet wśród autorów TD (1990–2014) (wg BTD).

Rozkład płci wśród autorów jest w miarę równy, z małą przewagą mężczyzn (56,3%). Jak pokazuje Wykres nr 4, w pierwszej dekadzie istnienia pisma widać wyraźną przewagę mężczyzn (z zaskakującym rokiem 1993, w którym liczba autorów i autorek jest równa, zapewne za sprawą wydania potrójnego numeru feministycznego 4–5–6/1993). W ostatnich piętnastu latach różnica stabilizuje się z okresami przewagi autorek. Nie znaleziono statystycznie istotnych różnic między kobietami i mężczyznami wśród takich zmiennych jak: długość tekstu, jego lokalizacja w numerze (strona początkowa) czy liczba słów w tytule.

Powyższe zestawienia, choć składają się na charakterystykę czasopisma, nie pozwalają zbyt wiele o nim powiedzieć. Śledząc rankingi autorów dominujących w danych numerach czy rocznikach, moglibyśmy coś powiedzieć o przemianach aktywności, ale trudno wyciągać z tego jakieś wnioski merytoryczne. Żeby powiedzieć o „Tekstach Drugich” coś sensownego na podstawie naszych danych, potrzebujemy innych wskaźników – przydatne tu będą, po pierwsze, listy cytatów, po drugie, wzmianki osobowe i, wreszcie, słowa kluczowe.

### **Kto kogo i co z tego**

Heurystyki makroanalityczne wszyscy często stosujemy w odniesieniu do monografii naukowych, rozpoczynając ich lekturę od przejrzenia indeksu osobowego i rzeczowego. Obecność (lub brak) niektórych nazwisk i pojęć oraz częstotliwość ich stosowania często pozwala nam (czy zawsze słusznie?) wyrobić sobie wstępną opinię na temat dzieła. Teraz zrobimy coś podobnego, tylko na znacznie większym materiale – na wszystkich numerach „Tekstów Drugich” odnotowanych w bazie RCIN. Zajrzymy więc do indeksu księgi liczącej ponad 26 tysięcy stron.

Należy tu poczynić ważną uwagę metodologiczną. Cytowanie jakiegoś dzieła w artykule umieszczonym w bazie RCIN (na podstawie bibliografii tekstu mozolnie przygotowywanej przez doktorantów IBL PAN) oznacza, że w artykule odwołano się do tego utworu, ale nie znamy dokładnej liczby odniesień (tzn. jeżeli tekst jest recenzją danej książki, to będzie ona odnotowana wyłącznie jako jedno odwołanie w tym tekście, niezależnie od tego, ile razy została faktycznie przywołana). Dzięki temu przypadki skrajne (jak recenzja tekstu czy obfite cytowanie z jednej książki) pokażą, które z cytowanych postaci są ważne dla szerszego kręgu autorów.

Cytaty przetworzyłem automatycznie w taki sposób, że uwzględniają wyłącznie inicjał imienia i nazwisko autora (lub redaktora tomu, jeżeli odwołano



się do pracy zbiorowej, a nie do konkretnego tekstu w niej zawartego). Przedstawiane tu zestawienia należy traktować z pewną ostrożnością – za sprawą literówek, błędów w zapisie czy automatycznym przetwarzaniu – część wyników może być nieco zafałszowana. Biorąc jednakże pod uwagę skalę zestawienia, drobne błędy nie powinny wpływać na wnioski.

Po przetworzeniu cytatów udało się pozyskać do analizy odwołania z 1924 tekstów za lata 1990–2012. Mamy tu łącznie 28 613 odwołań (tj. liczbę zidentyfikowanych odniesień), z czego 10 191 unikatowych osób – autorów cytowanych tekstów. Ponad 2/3 cytowanych (67,89%, 6919) zostało przywołanych jednokrotnie i tylko 18,66% (1902) osób pojawia się częściej niż dwa razy. 385 autorów przywołano częściej niż dziesięciokrotnie (3,78%), a tylko 1% (103 osoby) uzyskał więcej niż 30 odwołań.

Rysunek nr 1 ukazuje zestawienie 100 najczęściej przywoływanych w cytatach osób (1990–2012). Rozmieszczenie jest przypadkowe, natomiast rozmiar nazwiska odpowiada liczbie odniesień (ze względu na wspomniane potencjalne błędy w zapisie ten typ prezentacji wydaje się bardziej odpowiedni niż rankingi ilościowe). Widzimy tu nazwiska twórców, badaczy i myślicieli polskich i zagranicznych, które stanowią najważniejszy punkt odniesienia dla autorów „Tekstów Drugich”. Pewne przemiany tej hierarchii widać, jeśli przyjrzeć się najczęściej przywoływanym nazwiskom w pięciu okresach, co pokazuje Tabela nr 1.

Widać, że część osób pozostaje stałym punktem odniesienia w artykułach publikowanych w „Tekstach Drugich”, pojawiając się we wszystkich (Barthes, Bolecki, Derrida, Głowiński, Gombrowicz, Nycz, Sławiński) lub prawie wszystkich (Markiewicz, Balcerzan, Janion, Ricœur) okresach wśród 30 najczęściej przywoływanych postaci. Ale znów, taki ranking nieco nam mówi o czasopiśmie, ale niewiele znaczy bez kontekstu, tzn. kto przywołuje i w jakim towarzystwie. Z pomocą mogą przyjść tutaj techniki wizualizacji sieci. Bardzo niewdzięczne do „statycznej” prezentacji w druku, ale pozwalające na wysnucie pewnych wniosków.

Każda metoda prezentacji wyników badań opiera się na pewnych założeniach epistemologicznych. W przypadku wizualizacji sieci takim założeniem jest traktowanie przedmiotu badań jako organizmu złożonego z powiązań między elementami, w których zarówno rozmiar, jak i położenie świadczą o sile związku między nimi. Do wizualizacji wykorzystałem program Gephi, tworząc sieć powiązań na podstawie ukierunkowanych relacji między nazwiskiem cytującej i cytowanej (przy czym wszyscy autorzy mogą występować w obydwu funkcjach). Do grupowania wykorzystałem algorytm Open Ord



RYS. NR 1. Sto najczęściej cytowanych osób (wg RCIN). Chmura słów wytworzona za pomocą Wordle.

(lepiej pokazujący skupiska) oraz zastosowałem grupowanie algorytmem Louvain, pozwalającym na wyróżnianie grup na podstawie liczby wspólnych odwołań<sup>11</sup>. Wizualizacja to metoda dynamiczna, ukazująca wiele powiązań i najlepiej się chyba sprawdza jako narzędzie, jak to ma miejsce w przypadku, przywoływanego już wcześniej, projektu na czterdziestolecie czasopisma „Signs”<sup>12</sup>. Udało mi się wygenerować następujące grupy, które następnie zinterpretowałem.

Warto tu dodać, że postawa badacza analizującego sieci i przeprowadzającego grupowanie – w przeciwieństwie do urzędnika ministerialnego rozdzielającego dotację – opiera się na nieufności wobec algorytmu i stosowaniu kolejnych przybliżeń. Analizując grupy odniesień „Tekstów Drugich”,

- 11 Zob. V.D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre *Fast Unfolding of Communities in Large Networks*, „Journal of Statistical Mechanics: Theory and Experiment” 2008 No. 10, <http://arxiv.org/pdf/0803.0476v2.pdf> (27.09.2015).
- 12 Zwróćmy jednak uwagę, że sieć prezentowana w projekcie „Signs” jest dużo mniejsza (obejmuje tylko 233 źródła) i powstała na nieco innej zasadzie – wyłącznie na podstawie współwystępowania odnośień do tekstów w jednym artykule. Stąd też większa jej czytelność. Zob. <http://signsat40.signsjournal.org/cocitation/> (23.10.2016).

	1990–1994	1995–1999	2000–2004	2005–2009	2010–2012
1	W. Szyborska	Cz. Miłosz	Cz. Miłosz	R. Nycz	Cz. Miłosz
2	Cz. Miłosz	W. Gombrowicz	M. Głowiński	W. Gombrowicz	M. Foucault
3	M.R. Mayenowa	A. Mickiewicz	R. Nycz	J. Derrida	M.P. Markowski
4	M. Głowiński	M. Głowiński	J. Derrida	Cz. Miłosz	R. Barthes
5	J. Przyboś	Z. Nałkowska	Z. Herbert	M.P. Markowski	J. Derrida
6	W. Gombrowicz	J. Derrida	W. Gombrowicz	R. Barthes	B. Latour
7	J. Sławiński	M. Janion	F. Nietzsche	B. Latour	A. Wierzbicka
8	R. Barthes	R. Nycz	M.P. Markowski	M. Głowiński	P. Bourdieu
9	H. Markiewicz	J.N. Miller	R. Barthes	W. Bolecki	R. Nycz
10	S. Lem	J. Sławiński	J. Przyboś	J. Sławiński	S. Freud
11	R. Rorty	R. Barthes	J. Sławiński	F. Nietzsche	J. Sławiński
12	A. Wierzbicka	M. Białoszewski	W. Bolecki	M. Foucault	M. Heidegger
13	W. Odojewski	J. Błoński	J. Błoński	H. Markiewicz	S.J. Schmidt
14	M. Białoszewski	W. Bolecki	P. Ricœur	M. Heidegger	W. Gombrowicz
15	S. Barańczak	M. Foucault	H. Markiewicz	R. Shusterman	M. Janion
16	E. Balcerzan	P. Ricœur	M. Janion	M. Janion	M. Białoszewski
17	J. Ziomek	R. Rorty	U. Eco	P. Czapliński	W.W. Iwanow
18	R. Nycz	S. Barańczak	H. Grynberg	P. Ricœur	W. Benjamin
19	J. Błoński	E. Balcerzan	E. Balcerzan	A. Burzyńska	G. Agamben
20	M. Heidegger	H. Markiewicz	J. Jarzębski	J. Jarzębski	J. Baudrillard
21	A. Mickiewicz	J. Jarzębski	R. Rorty	E. Balcerzan	T.W. Adorno
22	J. Derrida	T. Różewicz	N. Frye	W. Benjamin	M. Głowiński
23	J. Kristeva	U. Eco	S. Barańczak	J.F. Lyotard	P. Ricœur
24	W. Bolecki	Z. Herbert	Z. Bauman	G. Didi-Huberman	T. Różewicz
25	K. Wyka	H.G. Gadamer	W. Benjamin	Z. Bauman	S. Fish
26	A. Zagajewski	L. Irigaray	K. Wyka	T. Różewicz	W. Bolecki
27	G. Herling-Grudziński	G. Ritz	P. de Man	W. Iser	E. Domańska
28	A. Wat	W. Szyborska	K. Irzykowski	Z. Herbert	G. Deleuze
29	R. Ingarden	K. Irzykowski	S. Fish	M. Czermińska	H.G. Gadamer
30	J. Culler	K. Bartoszyński	A. Sandauer	G. Herling-Grudziński	S. Brzozowski

TABELA NR 1. Najczęściej cytowane osoby z podziałem na okresy (wg RCIN).

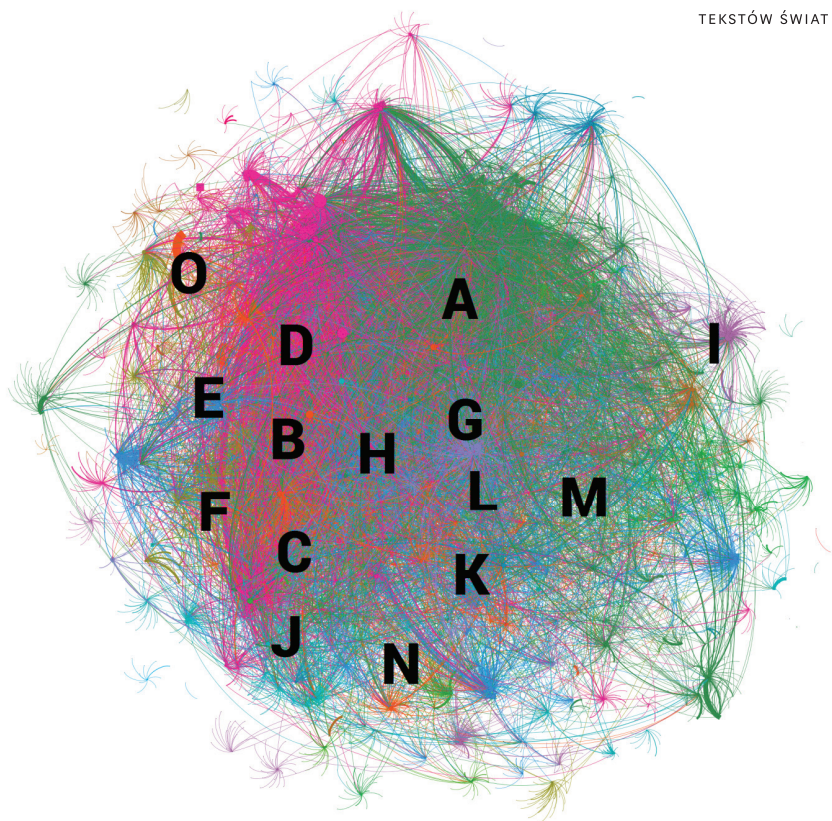
Klaster	Liczba osób	% (N=10543)	SUMA
A	2510	23,81%	Postrukturalizm, filozofia i literatura
B	1284	12,18%	Teoria literatury i kultury (strukturalizm, hermeneutyka, socjologia literatury)
C	928	8,80%	Historia literatury polskiej XX wieku I
D	912	8,65%	Historia literatury polskiej XX wieku II
E	748	7,09%	Modernizm, kultura i sztuka
F	683	6,48%	Romantyzm
G	535	5,07%	Zagłada i świadectwo
H	439	4,16%	Krytyka literacka, komparatystyka
I	423	4,01%	Neurosemiotyka i darwinizm
J	404	3,83%	Myśl feministyczna
K	387	3,67%	Poetyka teoretyczna i wersologia
L	387	3,67%	Postkolonializm
M	347	3,29%	Staropolska
N	293	2,78%	Formalizm i strukturalizm praski
O	220	2,09%	Darwinizm
Reszta	43	0,41%	[12 sporadycznych skupisk rzadkich autorów]

TABELA NR 2. Grupy autorów i cytowanych osób na łamach TD (1990–2012) (wg RCiN).

przyglądałem się grupowaniu, które dawało od kilkunastu grup po 44. Analiza wykazała, że wariant z 27 grupami, prezentowany w Tabeli nr 2 (z pominięciem 12 najmniejszych grup), daje najbardziej sensowne wyniki. Zauważmy, że kolejne obliczenia mogą umieścić autorów w nieco innej konfiguracji, a zatem niezwykle istotna jest właśnie procedura interpretacyjna, towarzysząca badaniu ilościowemu.

Zanim przejdę do omówienia wyników, muszę podkreślić, że nie wolno traktować tych grup rozłącznie – granice między nimi są płynne, co najlepiej widać w przypadku prezentacji całej sieci, na której – jak już wspominałem – istotne jest sąsiedztwo w przestrzeni. Ta sieć – składająca się z 10 543 węzłów (punktów) i 20 952 krawędzi (powiązań między nimi), to w istocie pajęczyna relacji między osobami pojawiającymi się na kartach „Tekstów Drugich” (Rys. nr 2).

Dla każdej grupy starałem się wyróżnić najważniejsze punkty odniesienia – cytowanych (choć na cytujących też patrzyłem) badaczy i twórców.



RYS. NR 2. Sieć powiązań między autorami i osobami przez nich cytowanymi w artykułach TD 1990–2012 (wg RCIN). Wizualizacja za pomocą Gephi.

Dość interesujące, że dani twórcy funkcjonują częściej w konkretnych kontekstach teoretycznych niż inni, ale to już temat na kolejną pracę. Ciekawe też, że w największej grupie (A) pisarze prawie się nie pojawiają, wyjąwszy takich twórców jak Celan, Eliot, Kafka, Homer czy Ważyk, przywoływanych – jak się zdaje – raczej w kontekście treści intelektualnych niż analizy twórczości. Możemy też na wstępie wyróżnić cztery największe grupy, których skład i zakres najbardziej zmieniały się przy kolejnych przybliżeniach i modyfikacjach algorytmu – powiązania są tu tak złożone, że możliwe jest wiele konfiguracji. Grupy mniejsze, o lepiej zdefiniowanej tematyce (dzielącej wyraźnie odrębny od pozostałych zestaw odniesień), pojawiały się w zbliżonej postaci

w różnych wariantach grupowania. Omówię te grupy skrótowo, przy czym warto zerkać co chwila na sieć (Rys. nr 2), która pokazuje relacje topograficzne między zbiorami, istotne dla interpretacji.

Dwie największe grupy łączę raczej z teorią niż z interpretacją. Grupa A zbiera badaczy związanych z teorią poststrukturalistyczną i literaturoznawstwem inspirowanym filozofią – kluczowe nazwiska (przedstawiane wedle liczby odniesień do nich) to Nycz, Derrida, Barthes, Markowski, Ricœur, Rorty, Latour, Bauman, Fish, Lyotard, de Man, Burzyńska, Domańska, Nietzsche, Benjamin, Adorno, Lacan, Agamben, Bal, Lejeune, Bielik-Robson, Dybel, Zeidler-Janiszewska itd. Jest to oczywiście grupa rozległa, rozciągająca się w prawej górnej części wykresu.

Druga grupa (B), nieco mniejsza liczebnie, wiąże się mocniej z teorią literatury, poetyką historyczną i socjologią literatury. Znajdziemy tu takie nazwiska jak Głowiński, Sławiński, Markiewicz, Eco, Irzykowski, Gadamer, Ziomek, Bachtin, Ingarden, Culler, Bartoszyński, Skrendo, Genette, Frye, Łebkowska, Kuźma, Okopień-Sławińska, Danek, Hopfinger.

Dwie kolejne grupy (C i D) są zakresowo niezwykle podobne i dotyczą szeroko pojętej historii literatury polskiej. Różnicę między nimi można wytłumaczyć wyłącznie za pomocą topografii sieci – teksty z grupy C rzadziej wchodzi w relację z kontekstem teoretycznym grupy A. W grupie C znajdziemy takich autorów jak Białoszewski, Zagajewski, Czechowicz, Karpowicz, Lechoń, Ficowski czy badaczy jak Balcerzan, Balbus. W grupie D znaleźli się Miłosz, Gombrowicz, Schulz, Witkiewicz, Tuwim, Dąbrowska, Whitman i badacze: Błoński, Bolecki, Jarzębski, Stala, Zaleski, Fiut, Chwin, a także spora grupa polonistów zagranicznych.

Grupa E jest dość niejednoznaczna i ogólnie została sklasyfikowana jako prace odnoszące się do modernizmu, kultury i sztuki – centralnymi postaciami są w niej Maria Janion i Kazimierz Wyka, a także Popiel, Brzozowski, Berent, Eliade, Zawadzki, Hegel, Czermińska. Wydaje się, że są to sieci blisko związane z grupami B i D, choć operują innymi powiązaniami.

Kolejne grupy są dużo lepiej sprofilowane. Grupa F odnosi się do romantyzmu (Mickiewicz, Ritz, Kleiner, Górski, Stefanowska, Witkowska, Nawarecki, Sokołowski, Piwińska, Kuziak) i szerzej pojmowanego wieku XIX (Chachulska, Orzeszkowa). Pojawiają się w tej grupie także Iwaszkiewicz, Leśmian, Sandauer oraz badacze i badaczki z kręgu krytyki feministycznej i *queer* (Brach-Czaina, Kofta, Filipiak, Śmieja), co świadczy o dużym przenikaniu się wspólnych odniesień między tymi obszarami badawczymi. Wydaje się, że podobnie należy interpretować grupę H, którą roboczo można określić mianem

krytyki literackiej a którą chyba należy rozumieć wyłącznie w kontekście przecięcia z pozostałymi grupami (zwróćmy uwagę na jej centralną pozycję w sieci). Znajdziemy tu takich twórców jak Lem czy Peiper i badaczy: Czaplinski, Komendant, Jameson, Sontag, Nowacki, Burek, Kaniewska, Rutkowski.

Grupa J poświęcona jest w największym stopniu krytyce feministycznej: Nasiłowska, Irigaray, Cixous, Showalter, Bakke, Bator, Rudaś-Grodzka, Smoleń, choć pojawiają się w niej również Mann i Lipski. Grupa G łączy nazwiska autorów prac poświęconych Zagładzie, Żydom i problematyce świadectwa (Wat, Grynberg, Tokarska-Bakir, Panas, LaCapra, Zimand, Leociak, Ringelblum). Wyraźnie rysuje się też klaster (K) prac z zakresu poetyki teoretycznej, językoznawstwa i semiotyki (Mayenowa, Łotman, Witosz, Dobrzyńska, Tabakowska, Wysłouch, Bartmiński, Kloch, Kaźmierczak, Szary-Matywiecka). Mocno powiązane wewnętrznie są grupy L – postkolonialna (Thompson, Hutcheon, Burszta, Gosk, Skórczewski, Spivak, ale i Wańkowicz), N – obejmująca formalizm rosyjski i strukturalizm praski (Jakobson, Szklowski, Propp, Ulicka, Brzostowska-Tereszkiewicz, Macura, Balík) oraz teksty o literaturze, dawnej zebrane w grupie M (Biblia, Kochanowski, Hernas, Brückner, Curtius, Mrowcewicz, Nowicka-Jeżowa, Stępień, Karpiński, Śnieżko). Dwie ostatnie grupy są bardzo idiosynkratyczne i powstały za sprawą odniesień egzotycznych (tj. bardzo nietypowych w skali pozostałych tekstów), chodzi tu głównie o teksty z zakresu neurosemiotyki (I) i teorii ewolucji (O).

Powyższy podział musi budzić pewne wątpliwości dotyczące przypisania konkretnych postaci do grup (a wymieniam tylko osoby najważniejsze, cieszące się największą liczbą odniesień). Wydaje się jednak, że pozwala to mówić o formacjach intelektualno-problemowych w obrębie czasopisma. Oczywiście twórcy i badacze nie są jednowymiarowi, ale być może także kierują do „Tekstów Drugich” tylko określone swoje teksty, pasujące ich zdaniem do profilu czasopisma. Widzimy jednak, że „profil pisma” to koncept niejednoznaczny, na który składają się lepiej lub gorzej zdefiniowane obszary odniesień.

Niezwykle interesującym wątkiem, wykraczającym poza ramy tego tekstu, ale wymagającym pogłębionych badań (także w perspektywie diachronicznej), jest powiązanie określonych pisarzy i poetów z danym kontekstem teoretycznym. Warto by zbadać, przez jaki pryzmat są zazwyczaj czytani i z czego to może wynikać.



### O czym piszą „Teksty”

Ostatnią z proponowanych tu analiz jest analiza tematyczna. Dzięki zlematyzowaniu tekstów (czyli sprowadzeniu wyrazów do form podstawowych) możemy sporządzić listy frekwencyjne i przyjrzeć się popularnym słowom. Rysunek nr 3 przedstawia chmurę trzystu najpopularniejszych słów pojawiających się w „Tekstach Drugich” (wykluczono słowa nienaznaczone semantycznie, jak spójniki itp.). Wielkość słowa odpowiada jego frekwencji.

Widać, że największą popularnością cieszą się takie terminy jak: tekst, literatura, świat, człowiek, język, literacki, życie, autor, słowo, sposób, historia, kultura, czas, wszyscy, książka. Trudno się dziwić, że to akurat te słowa – choć oczywiście ich wysoka frekwencja z pewnością pozwoliłaby odróżnić ten korpus od zbiorów tekstów użytkowych lub poświęconych innej dyscyplinie. Czy pozwala nam to jednak wysnuć jakieś wnioski o samej problematyce poruszanej w „Tekstach Drugich”? W pewnym sensie tak – pozwala tropić pojawianie się określonych problemów i odwołań na przestrzeni lat.<sup>13</sup>

Możemy np. przyjrzeć się historii zainteresowania danymi twórcami. Na Wykresie nr 5 widać zestawienie dla twórców, których wzmiankowano przynajmniej tysiąc razy w ćwierćwieczu (Miłosz, Gombrowicz, Herbert, Różewicz, Białoszewski).

W podobny sposób możemy śledzić np. zainteresowanie określoną problematyką. Dobrym przykładem są tu badania nad Zagładą Żydów. Problematyka w zasadzie nieporuszana w pierwszej dekadzie istnienia pisma, zostaje następnie nierozzerwalnie spleciona ze wszelkimi odniesieniami do Żydów.

Podobnie można obserwować recepcję teorii postkolonialnej na łamach „Tekstów Drugich” – pojawienie się problematyki i chwile większego nią zainteresowania.

Powyższe zestawienia pokazują także narzędziowy charakter tych rozpoznawień – statyczne wykresy nie oddają złożoności problemu. Badacze powinni mieć do dyspozycji interfejs, który pozwoli wizualizować takie zestawienia i analizować ich występowanie w poszczególnych kontekstach. Trudno też w większości przypadków mierzyć popularność wyłącznie za pomocą pojedynczych słów kluczowych – należy przyjrzeć się ich współwystępowaniu na przestrzeni wszystkich numerów.

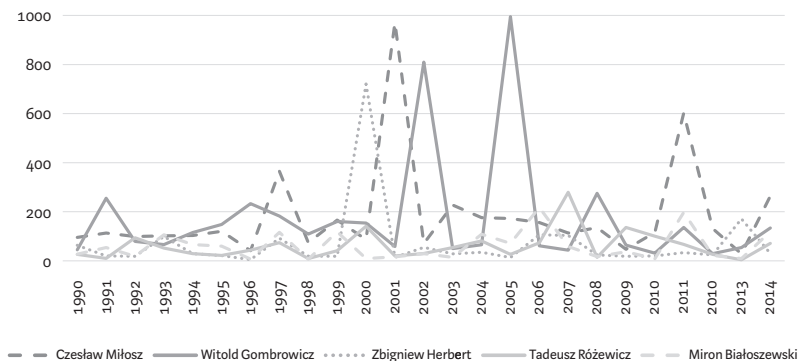
13 Pełne opisanie tego zagadnienia wymagałoby modelowania tematycznego, jak w przypadku badań Goldstone’a i Underwooda, które przytoczyłem na wstępie tego artykułu. Ze względów objętościowych ograniczam się jedynie do zestawień frekwencyjnych, a modelowaniu tematycznemu planuję poświęcić osobny artykuł.





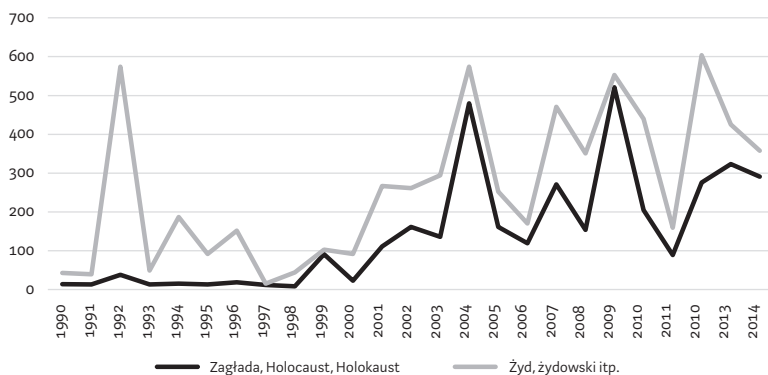
RYS. NR 3. Trzysta najczęściej pojawiających się słów w TD (1990–2014). Korpus przetworzony przez CLARIN-PL.

## Zainteresowanie najpopularniejszymi twórcami



WYKRES NR 5. Przekrój diachroniczny zainteresowania najpopularniejszymi pisarzami w TD (1990–2014).

## Badania nad kulturą żydowską i Holocaustem

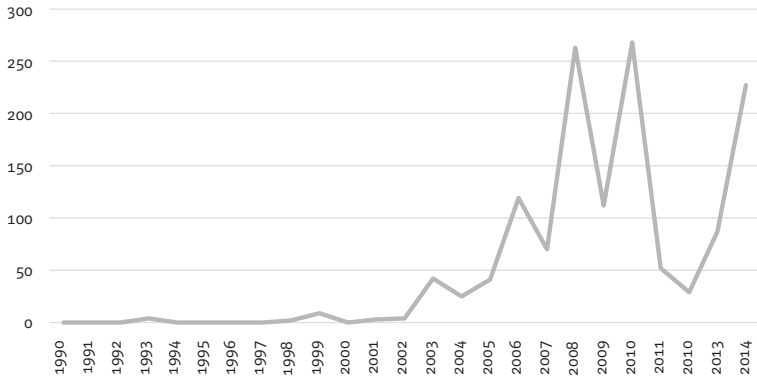


WYKRES NR 6. Odniesienia do słów związanych z Holocaustem i Żydami.

### A to dopiero początek

Celem ukrytym, który przyświecał tej pracy, była próba opisanego przemian intelektualnych w polskim literaturoznawstwie ostatnich lat. Niemniej jednak analiza cytowań pozwoliła wskazać pewne grupy intelektualne w obrębie pisma, które z pewnością – w jakimś stopniu – mogą pokrywać się z kierunkami zainteresowań polskich badań literackich ostatniego ćwierćwiecza.

## Badania postkolonialne



WYKRES NR 7. Recepcja teorii postkolonialnej na łamach TD (1990–2015).

Wydaje się jednak, że takie zadanie wymaga szerszej bazy materiałowej (m.in. uwzględnienia zasobów innych najważniejszych pism literaturoznawczych). Ciekawym eksperymentem byłaby także analiza porównawcza z innym czasopiśmem polskim lub przekrojowe zestawienie popularnych teorii badań literackich w różnych krajach i środowiskach. Prezentowane tu rozważania to zatem jedynie skromna przymiarka do kolejnych, szerszej zakrojonych badań wykorzystujących dane o życiu literackim. W tym krótkim tekście chciałem zaprezentować i poddać pod dyskusję różnorodne instrumentarium, które będzie rozwijane w kolejnych badaniach. Dopracowanie pytań badawczych i metod zdobywania odpowiedzi pozwoli w pełni wykorzystać tak obszerne zasoby danych, jak choćby Polska Bibliografia Literacka, która – dzięki unikalnym zapisom szczegółowym – może stanowić zaplecze do badań nad polską kulturą współczesną.