1 Introduction

Let us picture history of literature as an unknown city one tries to explore. An individual may contemplate some details and visit places recommended by others. But one needs to climb higher to perceive regularities and connections between landmarks, to see hidden patterns of the city life. While climbing higher and higher we lose many details and subjective impression of particular buildings, yet, we get a fuller picture, and discover new relationships, invisible from the ground level. But you can never climb high enough to get a full picture. Or can you? This is precisely what this project aims to explore—the possibilities and limitations of a bird’s eye view on events, works, and processes in an important period of Polish literary history and comparing it with intersubjective claims by various critics and scholars.

Why the intersubjective view is not enough? All syntheses in literary history, as well as in other branches of humanistic inquiry are by necessity based on certain approximations due to obvious cognitive limitations of mortal, selective and forgetful human beings. How about expanding the capacities of individual researchers by providing them with the opportunity to read all the books at once, to learn all the facts (including the minor and seemingly unimportant ones) and combine everything into a meaningful theory? If we consider digital research tools not only as more effective typewriters, but—in McLuhanian terms—as extensions of our senses and capacities, we could treat such a ‘view from above’ as a great opportunity to generate new insights that could support our interpretative practices. Moreover, apart from cognitive limitations we may go beyond the set...
of works considered canonical and explore also those objects which remained unnoticed by contemporaries or next generations of critics. According to laws of ‘historical selection’, proposed by Robert Escarpit, 80% of literary production becomes forgotten within a year and 99% in twenty years. By looking ‘from above’ we may now account for those forgotten works and try to present a more accurate picture of literary processes during a given period.

Digital revolution in the humanities brought about not only the increased storage capacity and accessibility of literary resources but more importantly the possibility to reuse them for the large-scale quantitative research. In doing so we may take the advantage of existing resources that have been diligently collected and annotated in various institutions throughout the years with the intention to facilitate the research process: bibliographies, catalogues, chronicles, calendars, fact sheets, reports, etc. There is a growing body of new, computational methods and approaches to the study of literature, which have clearly marked the distinction between traditional, close-reading and detail-oriented approaches, on the one hand, and new, synthetic studies applying statistical methods to larger corpora, on the other. In order to highlight the qualitative and empirical angle of this approach, it is dubbed “distant reading”, “algorithmic criticism”, or “macroanalysis”. The main premise behind this research corresponds with the metaphor I used in the introduction: when looking ‘from above’ we may lose some granularity concerning individual works, but instead we are able to survey large-scale processes, sequences and cycles. Among these digital bird’s eye methods, we may distinguish between the scholarship analysing the features of actual texts and metadata-driven research.

The former approach concentrates on quantifying stylistic features of texts. For instance, stylometric research focuses on the most frequent words in a given corpus of texts. Such research may entail authorship attribution, genre analysis or large-scale analyses of national literary output. Stylometric analysis is also

---

combined with functional or grammatical linguistic categories for the study of prose\textsuperscript{9} or other literary genres\textsuperscript{10}.

The second strand of research employs metadata, i.e. data about texts like title, genre, author, publication date, etc. Although such research often draws from the book history, current scholarship proves its validity for literary studies. Franco Moretti visualized and interpreted genre life-cycle on the example of metadata of the 19\textsuperscript{th}-century British novels.\textsuperscript{11} Matthew Jockers used bibliographical data to explore history, topics and geographical distribution of Irish American fiction.\textsuperscript{12} Katherine Bode used various bibliographical datasets to “recalibrate” the history of Australian literature and challenge some established arguments including views on relationships between local and foreign publishing markets or gender distribution among authors and their coverage in press and scholarship.\textsuperscript{13} Through visualisation of data on poet-journal affiliation Long and So pictured the differences between literary networks in the U.S., Japan and China.\textsuperscript{14} It should also be noted that the division between linguistic and metadata approach is purely conceptual and they could be incorporated into a single research design, as in the case of Moretti’s “Reflections on 7000 titles.”\textsuperscript{15}

This paper explores a possibility of quantitative research into transitions in literary history on the example of Polish literary life 1989–2002. This period is by no means under-researched. On the contrary, the abundance of scholarly and critical writing makes it even more interesting from the data-driven research perspective, since it allows for validating critical claims against existing data.

This exploration will be carried out on the basis of the data from \textit{Polish Literary Bibliography (PBL)}\textsuperscript{16}, a comprehensive database of Polish literary and cultural life. It indexes not only literary books but all instances of literary life and reception: reviews, journal articles, newspaper mentions, dramas, adaptations of literary works (stage, film, radio and television) and literary prizes. So, not only do we have access to rich data about literature but also to relationships between the objects (e.g. A is a review of B; C is an adaptation of D). The content is also annotated according its subject. This resource has been developed since 1954 and already covers Polish literary culture from 1944 to 2002. Unfortunately,

\textsuperscript{12}Jockers (Ann. 3), 35–62.
\textsuperscript{13}Katherine Bode, \textit{Reading by numbers. Recalibrating the literary field}, London 2014.
\textsuperscript{15}Moretti (Ann. 1), chapter “Style, Inc.: Reflections on Seven Thousand Titles (British Novels, 1740–1850),” 134–158.
\textsuperscript{16}https://pbl.ibl.waw.pl/ (retrieved on April 20, 2018).
most of the data are still available only in print, with over 1 million digital records covering only years 1988 to 2003. The recent development project\(^\text{17}\) was aimed at converting the Oracle database of *PBL* into a new, more accessible format with some tools for researchers. The project also entailed a prototype for retrodigitisation of older volumes but, given the vast scope of the task, it is hard to foresee, when the full database would be available.

Hence this paper consists of two interconnected goals: (1) to test selected qualitative hypotheses concerning Polish literature of the transition period (1989–2002); (2) to propose methodology for the data-driven study of literary history with the use of bibliographical datasets. Both goals will be supplemented with a reflection on how documentation methodologies and practices affect the quantitative analysis.

### 2 Literary transition as a challenge to literary theory

As Anna Nasiłowska observes the periodization of the 20th-century Polish literature is almost never based solely on artistic criteria and was heavily influenced by external—political and economic—events\(^\text{18}\). This is one of the reasons why she chooses 1976 as the starting point of her book on contemporary Polish literary history—the year of massive workers’ strikes which were violently broken by the authorities, what served as decisive impulse for the establishment of first organized oppositional bodies and, in turn, underground publishing. Nasiłowska uses the term “divided literature” to describe Polish literature between 1976 and 1989, what captures interesting dualism of literary life after the birth of *drugi obieg*, i.e. the second circulation, independent from the authorities and censorship. On the one hand there was official circulation, comprising of state-controlled publishing and dissemination institutions, and on the other a variety of *samizdat* books and journals, illegally printed and distributed by various dissident bodies. 1976 marks the beginning of organized underground printing in Poland, but Polish literature had been already published by émigré circles since the end of World War II, mostly in France and the UK but also in smaller centres in Germany, USA, Canada, and Sweden. The publications of those presses were targeted at local Polish communities, yet they were also smuggled to Poland and, after the establishment of illegal presses, reprinted (and *(vice versa)*—Polish *samizdat* was often republished abroad). And there was also the third circulation (*trzeci obieg*), i.e. counterculture publications, which were also published illegally but remained critical of political affinities and goals of the underground presses.

---


Defining literary transition poses a challenge to literary theory, because of the sheer fuzziness of the concept. First of all, it is difficult to set up a clear date of the political breakthrough—is it June 4th 1989, the day of the first partly free parliamentary election, or August 24th, when the first non-communist government was appointed, or, what is even more pertinent from the cultural point of view, April 11th 1990, the day when censorship was officially abolished? As Przemysław Czapliński notes, “Transition, whether we want it or not, lasts. It is not a rupture but a movement past a critical point [przesilenie], not an event but a process.”19 For this reason literary critics discuss their observations of literary change in a broader timeframe, which makes those elusive processes visible: Nasiłowska focuses on twenty years (1976–1996), she calls transitory period (okres przejściowy),20 Czapliński writes about the traces of the end in the literature of the 1980s and traces of the beginning in the 1990s,21 Janusz Sławiński, writes about disappearance of the centre (zanik centrali) in 1994, retrospectively naming the processes which gradually took place in the decade 1983–1993.22 There always has to be a point of reference. It should be noted here, although I am addressing this problem a bit later in this section, that a demarcation of the transition period ending point poses similar challenges.

Secondly, in thinking about the literary transition of 1989 there is an enthymematic assumption that political events bring about a cultural change. Many literary critics, as well as the audience anticipated a transformed, brand new literature.23 Those expectations are neatly captured as appetite for change in the title of Jerzy Jarzębski book (as well as in the book itself). Czapliński and Śliwiński note that the need for literary breakthrough was intensified by ongoing debates about transformations of the society and they conclude that “the question, whether a political change [przełom] triggers a literary one, is and is not unreasonable at the same time. It is unreasonable, because—luckily—literature is not an addition to political life. It is not, because literature is a part of reality, which underwent an enormous change.”24 So, perhaps, political change should be perceived as a sort of catalyst of cultural transition, because “it was not only a beginning of serious and extensive transformations of the infrastructure of the literary life, but it also … made the novelty legitimate, because it built up the expectation of novelty.”25

19 Przemysław Czapliński, Ruchome Marginesy: Szkice O Literaturze Lat 90, Kraków 2002, 6. All translations from the Polish sources are by the author.
20 Nasiłowska (Ann. 17).
21 Przemysław Czapliński, Ślady Przeleniu, Kraków 1997.
24 Czapliński/Śliwiński (Ann. 19), 213.
25 Ibid., 113.
In his book exclusively dedicated to the transition, Czapliński notes that literary breakthrough takes place in three spheres of poetics, ideas and institutions. It affects poetics, when certain genres and conventions lose their function of carriers of important content and new, different forms replace the accepted ones. The field of ideas undergoes change thanks to new modes of using literature. Finally, the institutions are transformed when the means of literary production and distribution are changed, together with relationships between the partners of literary communications: writers, patrons, publishers, educators. Given the vagueness of the plane of ideas (and, to some extent, the poetics), the change is mostly visible in the institutional field, Czapliński concludes. In this paper I try to map the evolution of literary institutions from the centralised model to the commercialised literary market of the late capitalism.

‘Mapping’ is not an accidental word here, since critics often used spatial metaphors to describe the processes of transition. The concept of disappearing centre was coined by Sławiński, an eminent Polish structuralist, in a tiny (3-pages long) and highly influential essay. Although he actually focused solely on poetry, his remarks were applied, with much success, to literary life in the broadest sense. Sławiński’s point of departure is a vision of literature in which individual works are parts of the greater Whole (większa Całość), in which authoritative (miarodajni) poets (i.e. those “read by everyone, whether they like them or not”) serve as orientation points for the assessment of other works. In other words, they constitute the canon, in relation to which the remainder of literature could be positioned. Sławiński observes the collapse of this system and concludes:

This Whole, which had supplied common frames of literary communication and had organised it around a certain centre, has been replaced by a polycentric system: multiplicity of—social and spiritual—micro worlds, in which poets and readers find each other. In each of them there are local tastes, local evaluation measures, local circulations, local relationships, authorities and hierarchies, local lucidities and unlucidities. One writes mostly for fellows … and expects new poems—from fellows.

This spatial metaphor is eventually elaborated by Bagłajewski, who claims that literature is in the state of dispersion (stan rozproszenia) by demonstrating how top-down mechanisms of literary prestige are being substituted by bottom-up canon formation: locally acclaimed writers, who write about their local

---

26 Ibid., 5.
27 Ibid.
28 Ibid.
29 Ibid.
30 Ibid., 5.
32 Ibid., 15.
33 Ibid., 16.
experience gain national recognition. ‘Locality’ in his understanding goes beyond geography, covering like-minded small initiatives in different parts of the country. What is important in this diagnosis is the feeling of temporariness: “Situation is still quite fluid. New initiatives appear all the time, transforming a—seemingly—ready cultural-institutional-literary system. We are in the state of dispersion.” This flexible relationship between the centre and peripheries will be later elaborated by Czapliński, who describes this process through a metaphor of moveable margins (ruchome marginesy). It is a vision of literary culture in which different phenomena coexist and either move to the centre or are pushed to the margins, depending on their importance for the greater community.

Those transformations are, if not welcome, at least accepted by the critics who—perhaps sharing the appetite for change—embrace this new configuration of cultural production, acknowledging its potential. Yet, quite quickly this optimism yields to disappointment, mostly due to the fact that decentralisation processes meet the market economy. Underfunded local initiatives find themselves on the margins, whereas popular and commercial literature occupies the centre. “Thanks to the domination of the market … the centre has become established and the margins fixed.”

Somewhere in the second half of the 1990s the trend is reversed, marking the beginning of a slow normalisation, understood as a process in which literature is marginalised by mass media, what in turn is treated as a natural order of things by the public. Critics start challenging the dominant discourse which establishes new, commercially-generated hierarchies. They call for a literature which would challenge the readers and face them with otherness. Those processes coincide with the return of the old masters, i.e. revaluation of the great writers who are again featured in mainstream and receive literary prizes. In the end, critics claim, we are witnessing the return of the centre (powrót centrali), but the centre has changed. It is not the official, state-controlled system, but the dominant media discourse and free market processes. Hence, the state of dispersion is considered a temporary stage between two centralized systems. As Czapliński claims, the

---

35 Ibid., 19.
36 Ibid., 17.
37 Czapliński (Ann. 18).
38 Ibid., 7.
40 Czapliński (Ann. 18), 21–36.
42 Przemysław Czapliński, Powrót centrali: Literatura w nowej rzeczywistości, Kraków 2007.
43 Dunin (Ann. 35).
44 Bagłajewski (Ann. 37).
only difference between those two being that the former was politically-driven and the latter is established on commercial principles.\(^{45}\)

Let us conclude with presenting the claims of the critics as set of hypotheses which we could try to verify on the basis of available data. Firstly, the institutions of literary life undergo dispersion in the early 1990s, followed by its subsequent re-centralisation according to the market principles. Secondly, literary reception follows the same pattern. Thirdly, literary reception in the ‘dispersed’ phase is debutocentric (i.e. focused on debutants), which changes in the re-centralisation phase with more attention being given to the ‘old masters,’ that is, eminent writers who earned their position before 1989. Finally, the processes of dispersion and re-centralisation also have a geographical dimension, which should be visible as the rise, and eventual decline, of smaller publishing centres.

\section{Operationalisation}

The first task is to translate elusive metaphors into variables or indicators. As Franco Moretti puts it in his pamphlet on character-space: “Operationalizing means building a bridge from concepts to measurement, and then to the world. In our case: from the concepts of literary theory, through some form of quantification, to literary texts.”\(^{46}\) In our case: from the concepts of literary history, through some form of quantification, to bibliographical data.

Apparently this task looks like an equation with two unknowns: on one hand there are metaphoric claims, on the other bibliographical material prepared by humans, full of possible errors and biases. However, in a spirit most recently advocated by Kathrine Bode, we should use the “bibliographical consciousness”\(^{47}\) and recognize the shortcomings of the data, what in turn will allow for a “data-rich literary history”, accounting for the phenomena which are graspable with this dataset (and such an account is given later in this section). But how should we approach the other unknown, namely claims made by literary critics of the transition period? Firstly, there is a question whether those generalisations are of a descriptive or rather prescriptive nature, as critics are \textit{nolens volens} actors of the literary life they investigate.

Moreover, one should not overlook the rhetoric glare of many metaphors coined by critics. Metaphors capture the ideas in an attractive manner but also sparkle the debate on their validity. One of the critics, Przemysław Czapliński, provides a perfect material for operationalisation, because he has coined and

\(^{45}\) Czapliński (Ann. 38).


defined many concepts concerning the period in question. Yet, his critical strategy was also criticised for this very reason: “He identifies, reads, recommends, diagnoses and describes trends, styles, poetics and circumstances but he does not show the path for literature because, despite his outstanding intellectual resourcefulness, he never gets attached to his ideas, proclaiming the new ones all the time. Fast enough he forgets about the previous ones.”

Unlike critics, who form their judgments on the basis of vast cultural material (although accessed unevenly and unsystematically), empirical approaches are faced with a scarcity and fragmentation of the data at hand. Unlike scholars in life sciences, who collect data relevant for their hypotheses, we need to tune our research scope to available material. Hence the operationalisation becomes a tricky activity which entails “building a bridge from concepts to measurement” with only that many bricks we can gather.

The most important thing to be noted about the PBL data is that it itself is a construct guided by certain methodological principles that has evolved overtime. Luckily the period I am working with would not witness major methodological changes, however there are important reservations to be made. First of all, although comprehensive and detailed, the PBL database does not contain every piece of literature published in Poland. For instance, it does not account for some regional literary journals. It may also favour certain eminent writers by collecting all their output, even of clearly non-literary nature. Secondly, literary documentation, unlike literary criticism, has to work with a sharp definition of literature. Hermeneuts can dedicate entire volumes to discussions of literariness, whereas literary documentation has to operate within well-defined borders. The consequences are grave—if something is considered literary it belongs in the database. Thirdly, only certain data fields (creator, reviewer, publisher, journal, particular work) have controlled vocabularies, and some entities have duplicate entries. I tried to minimise this influence in the data-cleaning process, however there may be human mistakes I overlooked. Moreover, an open ‘comment’ field for each data entry contains lots of unstructured information (e.g. if a critical work discusses two authors, the less featured one will be mentioned only in annotations, thus not available for quantitative analyses). Ongoing modernisation of the PBL will minimise those problems in the future.

With all those reservations in mind, we may treat the PBL database as a broad, representative sample of Polish literary life, which is carefully annotated and ready for the analysis. It is safe to assume that although the absolute results obtained on this data may not be comparable with other resources (e.g. actual number of books published or creators active in the given year), yet thanks to

consistent methodology it allows for relative comparisons overtime, i.e. looking at
certain trends in literary production.

The initial dataset was downloaded from the PBL database. The database is
relational, so the materials requested consisted of all books and articles published,
including secondary literature. Although PBL covers Polish literature from 1944,
the online database contains only the data for 1988–2003 with some earlier works,
if they were referenced in the modern material. At the time of the study year 2003
was still not finished and year 1988 was not complete, as it was used as a test
year for the setup of the first online database in 2000.50 Given those constraints
I limited the material to years 1989–2002. Furthermore, as the collection
methodology was changed after the year 1989 was entered into the database, most
notably through excluding some regional journals, I decided to exclude from the
sample all journals which appear only in 1989.

Limiting the transformation period to 1989–2002 is an arbitrary move, dictated
simply by the availability of the research data for comparative analysis. However,
basing on the critical assumptions discussed above, one should be able to trace
the processes of dispersion and subsequent re-centralisation in this period. Ideal
scenario would require a dataset spanning approximately five years more in both
directions. Hopefully, in the future data would allow for re-examination of the
findings presented below.

There are three major datasets I use in this study:

a) **Books**. Literary books either published in Poland (incl. translations of foreign
literatures), or by Polish authors abroad. Each entry includes information about
the author, publisher, place, and year of publication. In some cases book genre
was available or had been inferred from metadata or the title (e.g. play, children
literature, music, nonfiction, short stories, novel, reportage, art, poem, memoir).
b) **Articles**. Literary articles published in Polish literary journals, including
poems, fiction, letters, essays, nonfiction and adaptations. Metadata cover
information about the author, journal, year and the place of publication (in
some cases).
c) **Reception**. Various kinds of secondary literature, i.e. review, interview,
summary, letter, polemic, etc. Each piece of reception is connected to one
literary author (creator). Some reviewers were also authors but in this study I
omitted this relationship. Certain reception pieces noted in the bibliography are
about non-literary books and authors (e.g. review of a literary study).

For all datasets I prepared tables allowing for further computation. This entailed
data cleaning (esp. removing typos and mistakes) and enriching with additional
information (e.g. author’s gender, latitude and longitude of cities). Certain

---

50 Certain efforts are being made towards digitising printed volumes and hence expanding the
range of the online bibliography.
variables, have low reliability due to changing descriptive conventions and human errors, e.g. number of pages, or journal issue numbers.\textsuperscript{51} Altogether the dataset consists of 33,142 literary books and 138,925 literary articles by 21,865 authors. These were published by 8233 publishers and 958 journals. There are 93,874 reception texts, including pieces about 9286 authors and 22,335 books.

\section{Studies of dispersion}

The movement from dispersion towards centralisation and back, as described by critics, could be captured on many levels. The most obvious one, is simple bibliometric statistics, which could reveal the structure and evolution of literary life over the transformation period. I commence with some statistics of the literary production of that period, then combine those findings with the reception data, and subsequently, with comparison of reception of debutants and ‘old masters’. Next I analyse the dispersion on the example of geography and finally use network analysis for centrality measures.

\subsection{Literary production}

Let us start with sheer numbers concerning the literary production of that period. Dispersion, operationalised on the available dataset, would entail an increase in number of smaller publishers, journals or publishing centres disseminating literary content over time. Recentralisation would mean the reverse, i.e. smaller number of big players overtaking literary communication. Let us see how this look through some of the available measures.

To put the discussion in context we should start with a general statement that the period in question in Polish literature could be characterised in terms of growth. Figure 1 combines yearly numbers of publications (books and journal pieces), active literary journals (in this example only those which published literary works, not reception pieces), as well as count of publishers and authors (i.e. those who published at least one book or a journal piece). Trend lines show a steady growth of two first groups and quite a stable situation of the last one. This tendency is not so obvious if we take into consideration economic factors (diminishing of state-subsidized publishing) as well as the competition of other media in 1990s.

In order to add some granularity to those frequencies I divided publishers and journals into big and small. If critical assumptions were correct we should see lots of small, dispersed presses in the early 1990s and subsequent market

\textsuperscript{51}I am very thankful to Cezary Rosiński from the \textit{PBL} team, who extracted the data and prepared a list of debutants, referred to in the next section.
re-concentration towards the turn of the decades. In order to differentiate the publishers and journals I took their annual publication numbers and applied the 90th percentile as the threshold for qualifying a journal or publisher as big or small. In other words, big players were 10% of the most prolific journals and publishers each year. It should be noted that the quality of being “big” or “small” is always relative to a single year and the publishing numbers of other entities; a publisher or journal with the same number of publications over time may be qualified as a big one in one year, and as a small one in the other, depending on the outcomes of other publishers.

Area graphs in Fig. 2 show interesting discrepancy between the actual number of journals and publishers and their average output, i.e. average number of publications by active entities in all four categories. The production seems to be dominated by large players, which is especially true for big journals after 1995 and big publishers since around 1998. Interestingly, this growth of big publishing coincides with visible drop of the number of active players in the early 2000s.

It should be noted that the overrepresentation of small literary players may be partially caused by the data collection methodology, which would not register non-literary books. For instance, if a big publisher in the field of history releases one volume pertinent for literary studies (e.g. collection of letters by a literary figure), only this book will be counted and the publisher would appear in PBL as a small one.

![Evolution of the literary field, 1989-2002](image-url)
4.2 Critical reception

Sheer production statistics could be misleading, if we fail to take into consideration the reception practices. If we want to analyse the tendency for centralising the literary taste, we need to compare the number of books with its reception, namely how many times certain book is reviewed, discussed, or mentioned. Data in Fig. 2 seem to at least partially corroborate the claims of literary critics about eventual recentralization of literary life by a few bigger players. This trend is even more visible if we take into consideration the critical attention. Figure 3 plots the number of reviews of books by big and small publishers.

The critical attention devoted to books mirrors the processes described by the critics. On the centralised market of the 1989 Poland there was a clear domination of books by big publishers, which attracted most of the critical attention. Yet, they are swiftly outpaced by small publishers after 1991. This trend quickly reversed around 1995 and within two years big publishers regained the dominance in the struggle for critical attention. This trend is perhaps even more strikingly visible when compared with the actual publishing output.

In Fig. 4 the books by big (blue) and small (red) publishers are contrasted with the ratio of critical attention (green line represents the number of book reviews...
of publications from small publishers divided by reviews of big publishers’ output; score above 1 means higher interest in small publishers, whereas value below 1 indicates that critical attention shifts towards big players). It is worth noting that 1989 was the only year when big publishers published more than the small ones (55%). In the entire period small publishers had published on average 57% of all books (median = 58%), and in 2001 and 2002 the record numbers of 58% and 62% of all books published respectively. This pattern is compatible with the critical narrative about the recentralisation and the emergence of mainstream in the late 1990s.

In order to better understand those relationships we apply some measurements from economy. In this case a particularly useful is Gini coefficient, usually used to measure inequality in societies. We may consider literary life as a system in which attention is given to authors in three ways: by publishing their books, shorter pieces in journals, and by writing about their work or themselves. Measuring the distribution of this attention allows for judging the level of equality of writers according to those variables and for making further inferences about the

---

52 I am grateful to Mark Algee-Hewitt and Ryan Heuser for pointing me in that direction.
Operationalising the Change

processes of literary life. Data for computation was prepared for three variables (books published, journal pieces published, personal reception pieces about the author) for every year between 1989–2002. Authors not present in a given year were omitted. Gini coefficient was computed in ineq package for R.\(^{53}\) The results are presented in Fig. 5.

Gini coefficient takes a value between 0 and 1 and indicates the distance between actual distribution (as represented by Lorenz curve) and the perfectly equal one. Hence, the value 0 indicates perfect equality (all authors publish the same number of books), and 1 the opposite, absolute inequality (one author publishes all the books). First, it should be noted that it is hardly surprising that literary culture seems to be unequal, yet what strikes in Fig. 5 is the discrepancy between the distribution of the books published and other variables. How to interpret it? If we look at the World Bank statistics of economical distribution in various countries, we will notice that even most unequal societies rarely score higher than 0.5. Through this comparison we may conclude that distribution

of critical attention and journal pieces is highly unequal, whereas the authors publishing books are represented more equally. Moreover, the publishing data show a slightly downward trend, whereas the inequality of critical attention is on the rise and seems to follow the pattern visible in previous figures (sharp rise in the second half of the 1990s and a drop in the last year). This is by far the only variable in this figure, which follows the claims of the critic and seems to point out to some re-centralisation processes in the field of secondary literature. Let us examine it more closely, through examination of another metaphor describing the transformation—the return of the ‘old masters’.

4.3 Debutants and masters

One of the critics claimed that in the first part of the 1990s we witnessed debutocentrism in Polish literature, i.e. concentration on debutants, which eventually shifted to the interest in the “old masters”, i.e. eminent writers, who earned their position before 1989. In order to assess this claim the notions of debutants an masters were operationalised. Debutants are those authors, who have not published anything before 1989, whereas the masters where selected qualitatively as those authors, who were highly popular before 1989 and featured

Fig. 5 Gini coefficient of literary attention
Operationalising the Change

in criticism (e.g. Czesław Miłosz, Wisława Szymborska, Tadeusz Różewicz, Zbigniew Herbert, Stanisław Lem). The third, largest group of other writers (including historical figures) was left out. This selection led to establishing a two groups consisting of 11 masters and 3907 debutants. The reception of both groups is plotted in Fig. 6. The y axis represents the percentage of critical texts dedicated to authors from both groups, relative to all articles about authors published in a given year. Yet, despite similarities of scores, we must not forget that the groups compared are extremely different in size.

We may start with the observation that trends in both cases point to increasing interest. Furthermore, although the steep climb of writings dedicated to debutants in the early 1990s may be an indicator of debutocentrism, we hardly see a decline in the 2nd part of the decade. Yet, this could be a side effect of operationalisation, which defines debutants as all authors, who published for the first time in 1989 or later. Hence the number of debutants increases overtime, what may cause the gradual rise of critical attention. Yet, if we concentrated on the trends in the first half of the 1990s we can actually detect a decline in interest in acclaimed writers, who are overtaken by debutants by 1995. Suddenly, this trend almost reverses between 1996 and 1998 and afterwards.

A closer scrutiny of the data allows to link those irregular spikes on the masters’ line with particular events such as awarding Wisława Szymborska
with the Nobel prize in 1996, deaths of important writers—Zbigniew Herbert in 1998 and Gustaw Herling Grudziński in 2000, as well as with the 90th birthday of another Polish Nobel-prize winner, Czesław Miłosz in 2001. In all cases the abundance of writing about those authors would dramatically increase the overall reception of masters. And such events are not only limited to the living authors. For instance, the 200th birthday of Adam Mickiewicz in 1998, an eminent poet of Polish romanticism, generated over one thousand articles about him. It seems that the Nobel prize for Szymborska may have been a turning point in (or at least an influence on) Polish literary history in that respect that it renewed the interest in established writers. Yet, it should be highlighted that this pattern of interest seems to be incidental and related solely to big events regarding particular writers. Szymborska is a good example here—in the Nobel year there were 306 critical pieces about her, almost twice as much as during the preceding seven years combined (127). A year later, there were 180 pieces about her and in 1999 she was featured in roughly the same number of articles as in the year before winning the prize (18). This interest seems to be quite short-lived and seasonal, yet it may explain inequality patterns in critical attention of the late 1990s, when individuals were subjects of disproportionately high attention.

4.4 Geography

Processes of dispersion have also geographical dimension. If the critics were right, we should witness the rise and subsequent decline of smaller publishing centres throughout the transformation period. In order to map this process I compiled the information about the locations of publishers and journals. I excluded publishing locations outside Poland in order to minimise the number of accidental locations. For each of the 607 Polish cities I computed the number of books and articles published there each year. In order to locate larger centres I used the same procedure as in the case of publishers and journals, setting a high publishing level above the 90th percentile of all the city publications for a given year. It is important to add that levels were computed separately for books and journal pieces, hence, for instance, in terms of book publishing a given city could be considered large centre, and be simultaneously classified as small in terms of the journal pieces published (and vice versa). The classification of a given city may also differ overtime.

Figure 7 contains two representations of these data. The graph above shows the number of works published in cities, where blue colour signifies the high level of literary production in the city, and red represents low publishing level. The numbers are given for both books (squares) and journal pieces (ovals). In terms of publishing we see a steady rise of small publishing centres, which reach a similar level as the large ones in 2002. In terms of journals, we see a steady growth in both types of cities. The graph below uses the same features but arranges them according to the actual number of cities falling under given category. Here again journals behave differently as we see a small rise and subsequent decline of the number of smaller locations which host literary journals. In terms of books, both levels seem to rise and there are more locations with both smaller and larger output.
Fig. 7  Number of publications in Polish cities
In order to shed more light on those numbers, let us once again turn to economy. Herfindahl–Hirschman Index (HHI) is usually used to describe market concentration, hence to answer the very question we are dealing with here. I computed HHI using DescTools package for R.\textsuperscript{54} As in the case of Gini coefficient, described above, I generated a table containing the number of books and journal pieces published in a given city each year. Cities without publications in a given year were omitted. The results are presented in Fig. 8. HHI usually takes values from 0 to 1000, with the highest value representing monopoly. As it is a sum of squares of all entities’ shares in the market we should not be surprised to find that distribution of literary production among few hundred cities is highly deconcentrated, because many cities have a very low publishing output (cf. Fig. 7). Hence, we should not concentrate on the absolute values, but rather on trends and change overtime.

HHI plot in Fig. 8 shows a clear trend towards deconcentration in terms of both number of books and journal pieces published in cities. It seems that almost each

year new actors emerge and take their small share of the market. Although we hardly see any trend towards re-centralisation, it should be noted that in both cases we can distinguish the radical deconcentration during the first 5 years, from the rather steady downward movement in the last five years. In between, i.e. around 1995, we even witness a small movement towards concentration.

4.5 Network analysis

So far I have been using various measurements of single variables to account for the processes in question. In the concluding study I borrow tools of network analysis to look at literary life as a network of relationships and interdependencies. On the basis of such relationships, like being published by the same publisher, covered in a journal, receiving the same prize, one can build a network of literary life and explore its shape, which could be centralised or dispersed.55

The network of literary life that emerges from PBL data is complex because it features different types of actors and relationships between them. I decided to work with the author-centered network and linked them with other entities: publishers (published by), journals (published in, or reviewed in), reviewers (reviewed by). One can of course oppose to using reviews twice—to, first, mark the author’s relationship with a reviewer and then with a journal—but it is providing an additional layer of connections (a reviewer usually publishes in other journals, and any journal usually publishes many reviewers). I did not link reviewers to journals (to keep the network author-centred) and I did not use cities (in order not to introduce another entity class). The network contains the total of 339,287 unweighted connections (including duplicate entries). For each year I prepared a list of relationships between authors on one side and other entities on another. Those tables where then transformed into lists of nodes and edges in Gephi.56 Each node, i.e. relationship between two entities was assigned a weight, which indicates the number of times they were connected in a given year.

This network contains many different kinds of entities but I will simplify its interpretation and will treat it as a two-mode undirected network of writers and literary institutions. In order to deepen the understanding of this network I also use a one-mode ‘projection’ of this network, prepared with the use of bibliographic


coupling in Sci2. A ‘projection’ is a translation of a two-mode network into relationships among the entities of one kind. For instance, authors A and B, published novels with publisher X, and authors B and C printed their poems with publisher Y. In a two-mode network there are no connections between authors, but only between them and publishers (A–X, B–X, B–Y, C–Y) in a one-mode network relationships with publishers are translated into connections between authors on the basis of co-occurrence. So, B is connected to both A and C, but A and C are not connected directly because they did not publish together. Transforming network into a one-mode projection is often essential to proceed with network statistics as most measures are designed for this kind of networks.

The network concentration was measured with the average degree statistics. A node’s degree is the sum of its connections, while average degree is a mean of all degrees in the network. The concept of degrees differs in one- and two-mode networks. If a writer published only with two publishers, both of whom published books of 10 other authors, her two-mode degree would be 2 (2 publishers), while one-mode degree would amount to 20 (2 × 10 authors). Hence the latter is larger than the former. Figure 9 shows the average degree of one-mode and two-mode networks—the former was computed with Gephi, the latter, with tnet package for R.

For reasons described above, both networks operate on different scales but in the graph they were adjusted by the starting point, to allow for visual comparison. The one-mode network of relationships between writers (orange) seems to be perfectly fitting to the dispersion-recentralisation narrative, as after 1994 writers seem to increase their connections on average. Quite interestingly, the two-mode network (red), focused more on institutional relationships, tells a similar, yet not exactly same narrative. It seems as if it was more prone to the extremes (1993, 1996, 2001), as the general lower number of connections makes ‘stars’ more influential for the overall score. The sudden drop in the end corresponds with the general growth of book publishing, visible on earlier graphs. So, although big players may actually publish more books they remain unnoticed on the two-node network because of the wave of new publishers. The one-mode network rises on the contrary, probably because large publishers produce more books in general what allows for establishing new connections through co-occurrence.


5 Conclusions

Probably the most obvious conclusion of this study is that there is no one clear way to operationalise the claims made by the critics. Once we start to disentangle the complex nature of literary life it appears that certain measures correspond with critical predictions and some do not. Interestingly the patterns which seem to corroborate the critical claims almost always contained the reception variable: patterns of critical attention given to big and small publishers, reception of debutants and old masters, *Gini coefficient* and the network analysis. The only time the decentralisation pattern was visible without the reception component was in the study of concentration of literary production with *HHI*.

Claims of the critics are based on various materials they have at hand: literary texts, critical discourse, scholarly articles. Some of those claims are based on empirical data. Their interpretation of macro processes is then an individual interpretation of the events. However, the sharper the diagnosis, the more question

---

61 E.g. Przemysław Czapliński/Maciej Leciński/Eliza Szybowicz et al. (eds.), *Kalendarium Życia Literackiego 1976–2000: Wydarzenia, Dyskusje, Bilanse*, Kraków 2003, use number of journals having been shut down in the early 1990s, new literary prizes, number of debuts, etc.
it raises: what are the phenomena taken into account? Which ones were left out? Were all texts treated equally? Yet, similar questions may be directed towards the bibliographical material which, as the methodological section of this paper pointed out, is far from being an “objective representation” and is rather quite a comprehensive sample of literary material, gathered according to a certain methodology and by no means free from errors.

Furthermore, critics do not work in vacuum, they engage in literary debates and take stands on issues raised by other critics and they are not mere observes but active participants of the events in question. Hence, it is hard to draw a clear border between descriptive and prescriptive activity. Actually, critics quoted in this volume authored 1.6% of the reviews analysed in this study.

Quantitative approaches allow us to go beyond the critical writings and take a look at different dimensions of literary life, especially at those which usually escape critical scrutiny, like the ‘long tail’ of literary production. The obvious question that arises is which construct of literary life is the ‘right’ one: quantitative analysis or critical synthesis? Are we actually talking about the same literary culture? If literary processes are visible most clearly in secondary literature, which is itself a certain interpretation of literary facts, attempted reconstruction of critical diagnoses may be in fact a reconstruction of a certain aspect of literary life, namely actual critical debates.

The main goal of this study was to corroborate certain claims concerning the processes of literary life and the results show that certain type of generalisation is possible. This is why the central, yet problematic variable in this study is time. As we all know too well, the fact that a piece was published in a given year and not the previous or the following one is often a matter of chance. One can blur the time boundaries by using a moving average but this in turn would soften the impact of individual literary events, as the case of Szymborska’s Nobel prize shows. This study marks a point of departure for the further work with PBL data, which—freed from the constraints imposed by the goal of the present study—would look at further avenues of macroanalytical research into literary with the use of bibliographic data.

Acknowledgements Research presented in this study was made possible by the National Programme for the Development of Humanities (Research grant: The Polish Literary Bibliography—a knowledge lab on contemporary Polish culture, NR 0061/NPRH3/H11/82/2014), Polish-American Fulbright Commission (Fulbright Senior Award: Literary transition as viewed from above. Quantitative history of Polish literature (1989–2000). Research visit at Stanford University), and COST Association (COST Action CA16213: New Exploratory Phase in Research on East European Cultures of Dissent). The author is greatly indebted to colleagues from the Stanford Literary Lab (Mark Algee-Hewitt, Ryan Heuser, Erik Fredner, J.D. Porter, and Hannah Walser) for their insightful feedback on this project, as well as to the participants of the DFG conference at Villa Vigoni. Special thanks to Cezary Rosiński for his help with the data dump. All mistakes remain my own.
References


Bode, Katherine, Reading by numbers. Recalibrating the literary field, London 2014.


Czapliński, Przemysław, Powrót centrali: Literatura w nowej rzeczywistości, Kraków 2007.

Czapliński, Przemysław, Ruchome Marginesy: Szkice O Literaturze Lat 90, Kraków 2002.

Czapliński, Przemysław, Ślady Przelomu, Kraków 1997.


Moretti, Franco, Distant Reading, London 2013a.


---

62 All digital references were last viewed on April 20, 2018.

Ramsay, Stephen, Reading Machines. Towards an Algorithmic Criticism, Urbana 2011.


Online Resources


Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (http://creativecommons.org/licenses/by/4.0/deed.de) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungsliegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.