

A. BARTKOWIAKOWA (Wrocław)

*O ROZKŁADZIE I KOLEJNOŚCI ZDAŃ WSPÓŁRZĘDNYCH  
I PODRZĘDNYCH W UTWORACH POWIEŚCIOWYCH  
ŻEROMSKIEGO I SIENKIEWICZA*

**1. Opis zagadnienia i wyników.** W badaniach nad strukturą tekstu wiele uwagi poświęca się budowie zdania. Elementarne części, na które można rozbić zdanie, mogą być bardzo rozmaite. Mogą one, na przykład litery lub sylaby, tylko w bardzo niewielkim stopniu pozostawać w związku z treścią danej wypowiedzi. Inne znów, poprzez zasób informacji w nich zawartych, w sposób bezpośredni wiążą się z nią. Jednym ze sposobów rozbioru zdania jest badanie jego budowy pod względem logicznym — wyodrębnia się wtedy pewne części spójne pod względem treściowym. Takim podziałem „logicznym” jest na przykład porządkowanie wyrazów relacją: określający — określany, przy czym możemy wyodrębnić grupy wyrazów będących określeniami podmiotu lub orzeczenia. Statystyki tak rozumianych określeń dotyczy jedna z moich wcześniejszych prac.

Obecnie chciałabym się zająć budową zdania z punktu widzenia jego rozcłonkowania na zdania współrzędne i podrzędne, czyli *parataktyczne* i *hipotaktyczne*. O znaczeniu i interpretacji tych form składniowych pisze Z. Klemensiewicz [3]: „Parataksa jest łatwiejszą postacią kształtowania syntaktycznego, odznacza się większą prostotą i naturalnością, graniczącą nieraz z prymitywizmem. Parataksa zachowuje w wypowiedzeniu to, co najważniejsze, istotne i ekspresywne, odpowiada konkretnemu aspektowi rzeczywistości, pozwala przedstawić wrażenia, wizje, i dlatego jest chętnie wyzyskiwana przez poetów. Idzie w niej bowiem o szeregowanie, o nizanie stanów rzeczy, bez wnikania w ich wewnętrzne powiązania i zależności, a co najwyżej z uwydatnieniem ich najbardziej zewnętrznych przejawów. Ponadto w większości wypadków stosunek stanów rzeczy, które język ujmuje w parataksie, narzuca się autorowi jako gotowy, dany w obiektywnej rzeczywistości; jest to mianowicie zbieżność w czasie, podobieństwo lub identyczność...

...hipotaksa znamionuje wzrost intelektualizacji wypowiedzenia, przeważa też w przedstawieniu naukowym, ... czyni wypowiedzenie trudniejszym, nieraz wręcz skomplikowanym, ale też zapewnia większe

bogactwo komunikatywne, większą dokładność, precyzyjność, przejrzystość i jasność.

Rozwój i postęp stylistycznego kształtowania idzie od parataksy do hipotaksy. Rzecz wymaga szczegółowego zbadania, w jakiej mierze pewne typy hipotaksy są dla pewnych gatunków literackich, pewnych rodzajów kompozycyjnych, pewnych osobowości znamienne częstością lub rzadkością”.

Aż dziwną jest rzeczą, że zagadnienie kryjące takie bogactwo treści nie zostało dotychczas szczegółowo opracowane pod względem statystycznym. Znane są mi tylko dwie prace związane tematycznie z parataksą i hipotaksą, w których operuje się pewnymi wielkościami statystycznymi, zresztą najbardziej elementarnymi, bo jedynie średnimi. Są to: wyżej już cytowana praca Z. Klemensiewicza [3], w której autor podaje pewne dane liczbowe dotyczące zdań pojedynczych i złożonych, oraz ciekawa praca A. Wierzbickiej [6] nad tendencjami składni szesnastowiecznej, w której autorka bada rozwój języka od parataksy do hipotaksy, posługując się pewnymi danymi liczbowymi, pochodzącymi z prób odpowiednich tekstów.

Przedmiotem moich badań, których wyniki przedstawiam w niniejszej pracy, była ilość zdań współrzędnych i podrzędnych w utworach powieściowych Żeromskiego i Sienkiewicza ujęta z punktu widzenia statystyka. Stanowisko takie jest bez wątpienia inne, niż stanowisko syntaktyka-gramatyka lub syntaktyka-stylistyka, co nie znaczy jednak, że jest ono bez pożytku dla jednego lub drugiego.

W dalszym ciągu zajmować się będę ilością zdań współrzędnych — inaczej parataktycznych — w jednym zdaniu złożonym (cechę tę oznaczać będę symbolem *z*) oraz ilością zdań podrzędnych — inaczej hipotaktycznych — w jednym zdaniu głównym (cecha *t*). Okazuje się, że pewna statystyczna charakterystyka tych cech, mianowicie ich rozkład, wyraża się ściśle matematycznym prawem. Zgodność faktów empirycznych, uzyskanych z badań konkretnych utworów literackich z prawami otrzymanymi drogą dedukcji z założeń ściśle teoretycznych wskazuje na pewne mechanizmy rządzące powstawaniem zdań.

W dalszym ciągu zajmuję się sąsiedztwem i kolejnością zdań paraihipotaktycznych. Chodzi na przykład o to, czy zdania o ustalonej liczbie zdań współrzędnych są w tekście rozmieszczone przypadkowo, czy też można zauważyć pewne ich skupienia lub rozrzedzenia. Innym zagadnieniem jest wpływ rozbudowywania zdań w kierunku parataksy na ich rozbudowę w kierunku hipotaksy. Podaję również pewną klasyfikację różnych układów, konfiguracji zdań współrzędnych i podrzędnych z uwzględnieniem częstości ich występowania u dwu badanych autorów. Zajmuję się także sprawą jednorodności tekstu w różnych miejscach książki.

**2. Przedstawianie zdań w postaci dendrytów.** Badania moje ograniczyłam do dwu autorów: Sienkiewicza i Żeromskiego. Aby mieć materiał możliwie jednorodny, brałam pod uwagę tylko utwory powieściowe. Ponieważ budowa zdań dialogowych jest znacznie trudniejsza, bo więcej w niej skrótów i równoważników, uwzględniałam jedynie zdania opisowe, nie wchodzące w skład wypowiedzi bohaterów. Równoważniki zdań, w których nie występuje forma osobowa czasownika, traktowałam jako oddzielne zdania tylko wtedy, gdy występowały one samodzielnie, tzn. oddzielone były kropką od sąsiednich, normalnych zdań; w przeciwnym przypadku klasyfikowałam je jako określenia odpowiednich zdań mających formę osobową czasownika. Przypadków takich w badanym przeze mnie materiale nie było jednak wiele i dlatego wydaje mi się, że taka lub inna klasyfikacja równoważników nie wpływa tu zasadniczo na wyniki.

Uwzględniając relacje współrzędności i podrzędności można każde zdanie złożone przedstawić w postaci dendrytu, jak np. na rysunku 1. Zdania nadrzędne oznaczamy podwójnym kółkiem, zdania podrzędne w stosunku do nich wykreślamy w kierunku pionowym w dół i oznaczamy zwykłym kółkiem; jeśli to zdanie ma jeszcze jakieś zdanie podrzędne w stosunku do siebie, to prowadzimy jeszcze raz kreskę w dół i rysujemy na jej końcu zwykle kółko. Dwa zdania współrzędne oznaczamy dwoma kółkami połączonymi linią poziomą. Otrzymujemy w ten sposób wykres zdania, który nazywać będę *dendrytem*. Dendryt w ten sposób sporządzony nie uwzględnia „linearności” poszczególnych członów zdania i traci w ten sposób wiele walorów cennych dla stylistyka, ale za zadanie postawiłam sobie zbadanie najprostszych, najbardziej elementarnych związków.

Rysunek 1 przedstawia dendryty zdań dla dwóch urywków z pism Żeromskiego i Sienkiewicza:

a) „Namiestnik nie dostał się jednak tego wieczoru do zamku, bo pan Grodzicki zaprowadził taki porządek, że gdy przed zachodem słońca wybito hasło, nie puszczano nikogo z zamku i do zamku, i gdyby nawet sam król przyjechał, musiałby nocować w Słobutce stojącej pod wałami fortecy. Tak też uczynił namiestnik. Nocleg to był niezbyt wygodny, bo chaty w Słobutce, których znajdowało się kilkadziesiąt, ulepione z gliny tak były szczupłe, iż do niektórych okraciem trzeba było wlażyć. Innych też nie opłacało się budować, bo je forteca za każdym napadem tatarskim w perzynę obracała, a to dlatego, aby nie dawały napastnikom zasłony i bezpiecznego do wałów dostępu.” — *Ogniem i mieczem*, t. I.

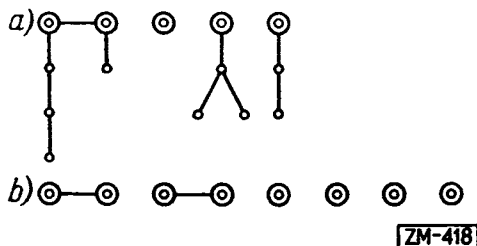
b) „Powóz wjechał w obszerny dziedziniec i zatrzymał się przed lustrzanymi drzwiami parterowego domu. Wybiegł służący w liberii i pomógł panom wysiąść. Za chwilę znalazł się Rafał w mieszkaniu dość

obszernym, a urządzonym w sposób nieco dziwny. Obok mebli wykwintnych stały tam graty najpospolitsze. Nie były to salony mieszkalnego domu, lecz męskie, kawalerskie pokoje. Niektóre z nich były brudne jak numery zajezdni.” — *Popioły*, t. II.

Jako bardzo proste i narzucające się niejako charakterystyki tak otrzymanych dendrytów przyjął:

*w* — równe liczbie zdań podrzędnych w stosunku do jednego zdania głównego, inaczej składowa hipotaksy. Patrząc na rysunek 1 można powiedzieć, że jest to liczba punktów leżących poniżej jednego zdania głównego oznaczonego podwójnym kółkiem i tworzących z nim jedną gałąź dendrytu.

*z* — równe liczbie zdań złożonych współrzędnie w jednym zdaniu złożonym. Jest to składowa parataksy. Biorę tu pod uwagę tylko zdania współrzędne pierwszego stopnia, to jest takie, które tworzą zdania główne



Rys. 1. Przykłady dendrytów zdań: a) dla fragmentu z *Ogniem i mieczem*, b) dla fragmentu z *Popiołów*

zamieszczone w pierwszym wierszu dendrytu (oznaczone na rysunku 1 podwójnym kółkiem).

Dla fragmentów powyżej zacytowanych wartości odpowiednich składowych wynoszą:

dla urywka z *Ogniem i mieczem*  $z = 2, 1, 1, 1$ ,  $w = 3, 1, 0, 3, 2$ ,

dla urywka z *Popiołów*  $z = 2, 2, 1, 1, 1, 1$ ,  $w = 0, 0, 0, 0, 0, 0, 0$ .

Oczywiście nie badałam wszystkich zdań opisowych wyjętych z wybranych utworów, lecz operowałam jedynie reprezentacyjną próbą. Próba została pobrana w sposób gronowy, tzn. losowałam stronicę książki i z niej i następnych brałam po kolei pewną ustaloną liczbę zdań opisowych (około 60). W ten sposób uzyskałam kilkaset zdań wybranych „gronami” z różnych miejsc książki.

**3. O rozkładzie liczby zdań współrzędnych i podrzędnych.** Licząc w wybranej próbce zdania złożone składające się z jednego tylko zdania nadrzędnego ( $n_0$ ), z dwu zdań złożonych współrzędnie ( $n_1$ ), z trzech zdań

złożonych współrzędnie ( $n_2$ ), itd., otrzymujemy empiryczny rozkład liczby zdań parataktycznych

$$(1) \quad p_k = \frac{n_k}{\sum n_k}, \quad k = 0, 1, 2 \dots$$

Wartość  $p_k$  mówi nam, jaka frakcja ogólnej liczby zdań złożonych składa się z  $k$  zdań złożonych współrzędnie.

W podobny sposób, oznaczając przez  $n_0$  liczbę zdań głównych nie mających żadnego zdania podrzędnego,  $n_1$  — liczbę zdań głównych mających jedno zdanie podrzędne, itd., otrzymujemy empiryczny rozkład liczby zdań hipotaktycznych.

Okazuje się, że uzyskane w ten sposób rozkłady, czyli zależności  $p_k$  od  $k$ , dają się ująć matematyczną formułą:

$$(2) \quad p_k = \left( \frac{t}{1+bt} \right)^k \cdot \frac{1 \cdot (1+b) \dots [1+(k-1)b]}{k!} \cdot p_0, \quad k \geq 1,$$

$$p_0 = (1+bt)^{-1/b},$$

gdzie  $t$  i  $b$  są parametrami rozkładu:  $t$  jest wartością średnią,  $b$  nosi nazwę współczynnika przyciągania.

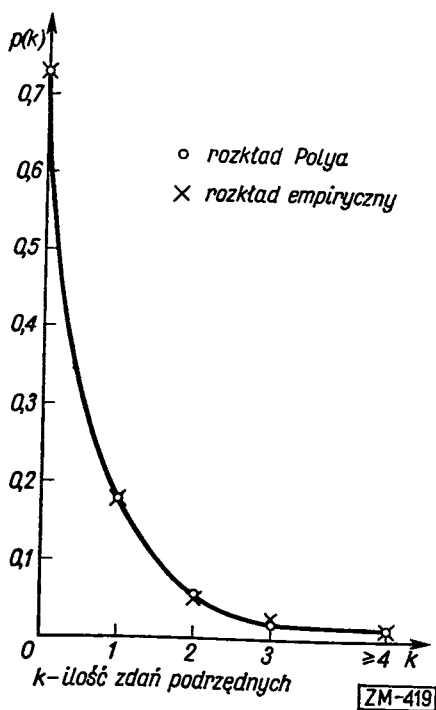
Wartość oczekiwana i wariancja tego rozkładu wynoszą odpowiednio:

$$(3) \quad \bar{k} = \sum_{k=0}^{\infty} k p_k = t,$$

$$s_k^2 = \sum_{k=0}^{\infty} (k - \bar{k})^2 p_k = t(1+bt).$$

Rozkład opisywany powyższym wzorem znany jest jako graniczny rozkład Pólyi lub ujemny dwumianowy.

Na rys. 2 porównuję przykładowo rozkład empiryczny z teoretycznym, uzyskanym z wzoru Pólyi (2) dla  $t$  i  $b$  oszacowanych metodą momentów według wzorów (3). Jest to rozkład liczby zdań podrzędnych dla *Ogniem i mieczem*. Wartość  $\chi^2$  obliczona dla tego przykładu wynosi 5,24 przy dwóch stopniach swobody. Prawdopodobieństwo, że



Rys. 2. Rozkład liczby zdań podrzędnych dla *Ogniem i mieczem*

wskutek przypadkowych, losowych odchyień otrzymamy jeszcze większe rozbieżności, jest większe niż 0,05 ( $\chi^2_{0,05} = 5,99$ ), zatem odchylenia rozkładu otrzymanego empirycznie od dopasowanego do niego rozkładu Pólyi uznajemy za nieistotne.

W tablicy 1 zamieszczam otrzymane rozkłady empiryczne dla sześciu badanych utworów; obok nich podaję rozkłady teoretyczne obliczone według wzoru Pólyi dla parametrów  $b$  i  $t$  oszacowanych metodą momentów.

TABLICA Ia

Rozkład liczby zdań złożonych współrzędnie

 $n_k$  emp – liczebności otrzymane empirycznie,  $n_k$  Pólyi – liczebności obliczone według wzoru Pólyi

## Utwory Żeromskiego

Popioły			Uroda życia			Syzyfowe prace		
$k$	$n_k$ emp	$n_k$ Pólyi	$k$	$n_k$ emp	$n_k$ Pólyi	$k$	$n_k$ emp	$n_k$ Pólyi
0	770	771	0	450	449	0	320	327
1	280	262	1	129	131	1	136	126
2	42	56	2	28	30	2	46	45
3	15	10	3	11	8	3	14	15
4	3	9				4	1	} 8
						5	2	
						6	1	
						8	1	
$\bar{k} = t = 0,375, b = 0,273,$ $n = 1108$			$\bar{k} = t = 0,353, b = 0,586,$ $n = 618$			$\bar{k} = t = 0,572, b = 0,842,$ $n = 521$		

## Utwory Sienkiewicza

Rodzina Polanieckich			Ogniem i mieczem			W pustyni i w puszczy		
$k$	$n_k$ emp	$n_k$ Pólyi	$k$	$n_k$ emp	$n_k$ Pólyi	$k$	$n_k$ emp	$n_k$ Pólyi
0	286	295	0	349	357	0	361	365
1	171	151	0	200	191	1	216	208
2	38	50	2	89	82	2	69	70
3	15	14	29	29	33	3	13	18
4	1	} 4	4	5	12	4	5	} 5
5	2		5	4	} 7	5	2	
6	1		6	4				
			7	2				
$\bar{k} = t = 0,607, b = 0,310,$ $n = 514$			$\bar{k} = t = 0,796, b = 0,610,$ $n = 682$			$\bar{k} = t = 0,637, b = 0,174,$ $n = 666$		

TABLICA Ib

Rozkład liczby zdań podrzędnych w jednym zdaniu głównym

Żeromski

Popioły			Uroda życia			Syzyfowe prace		
$k$	$n_k \text{emp}$	$n_k \text{Pólyi}$	$k$	$n_k \text{emp}$	$n_k \text{Pólyi}$	$k$	$n_k \text{emp}$	$n_k \text{Pólyi}$
0	608	606	0	648	652	0	492	489
1	86	91	1	143	134	1	112	120
2	28	22	2	32	36	2	32	27
3	3	6	3	7	10	3	6	}7
4	2	}3	4	3	}4	4	1	
5	1		5	3				
$\bar{k} = t = 0,225, n = 728,$ $b = 2,1927$			$\bar{k} = t = 0,305, n = 836,$ $b = 1,5978$			$\bar{k} = t = 0,308, n = 643,$ $b = 0,8375$		

Sienkiewicz

Rodzina Połanieckich			Ogniem i mieczem			W pustyni i w puszczy		
$k$	$n_k \text{emp}$	$n_k \text{Pólyi}$	$k$	$n_k \text{emp}$	$n_k \text{Pólyi}$	$k$	$n_k \text{emp}$	$n_k \text{Pólyi}$
0	492	486	0	895	893	0	342	334
1	187	202	1	220	218	1	110	126
2	95	82	2	59	72	2	47	44
3	29	33	3	34	26	3	23	15
4	13	14	4	13	10	4	3	}8
5	7	}9	5	1	}6	5	2	
6	1		6	3				
7	2							
$\bar{k} = t = 0,691, n = 826,$ $b = 0,9602$			$\bar{k} = t = 0,420, n = 1225,$ $b = 1,7165$			$\bar{k} = t = 0,560, n = 527,$ $b = 0,8656$		

Oba rozkłady dla poszczególnych utworów wykazują dużą zgodność. Można sprawdzić, że we wszystkich zamieszczonych przypadkach różnice między obu rozkładami są nieistotne. Jest rzeczą charakterystyczną, że zarówno liczba zdań współrzędnych, jak i podrzędnych wyraża się tym samym prawem.

Czy to wszystko jednak może mieć jakieś praktyczne znaczenie? Rzeczywistość dała się w tym przypadku ująć w ramy matematycznych formuł, a więc statystyk jest zadowolony, że potrafi tak ściśle opisać zjawiska — wydawałoby się — zupełnie spontaniczne. Ale na tym nie kończy się rola matematyka-statystyka. Potrafi on powiedzieć znacznie więcej. Bo oto odpowiednie rozkłady otrzymuje się w ściśle określonych warunkach. Z faktu, że rozkład liczby zdań współrzędnych i podrzędnych opisuje się właśnie wzorem Pólyi, możemy wyciągnąć pewne wnioski co do sposobu, mechanizmu powstawania zdań złożonych.

Nie podając opisu szczegółów matematycznego modelu (który można znaleźć np. w monografii O. Lundberga [4] lub częściowo w podręczniku M. Fisza [2] <sup>(1)</sup>), można powiedzieć ogólnie, że wzór Pólyi przedstawia rozkład „z przyciąganiem”. Jego cechą charakterystyczną jest to, że wykazuje on w stosunku do rozkładu Poissona nadwyżkę zdań o bardzo małej (zerowej) oraz o dużej liczbie zdań podrzędnych i współrzędnych.

Jak można sobie wyobrazić proces powstawania zdań? — Każde zdanie mówi o czymś, przedstawia pewien fakt, obraz cząstkowy utrwalony przez autora. Łączenie zdań odbywa się na zasadzie kojarzenia wrażeń, jakiegoś porządkowania przedstawień. Każde określenie rodzi następne. Im więcej jest przedstawień, tym łatwiej pojawiają się dalsze określenia. Pisarz albo poprzestaje na wypowiedziach jednozdaniowych, albo też rozbudowuje je w sposób lawinowy.

Jest rzeczą ciekawą, że rozkład Pólyi (2) ma dualną interpretację (por. M. Fisz, [2], str. 152). Oprócz schematu zaraźliwego, w którym pojawienie się zdarzenia zwiększa prawdopodobieństwo jego wystąpienia w następnym doświadczeniu, rozkład (2) otrzymujemy jeszcze w innych okolicznościach, w których zdarzenia są niezależne, lecz materiał jest niejednorodny. Wyobraźmy sobie, że liczba zdań współrzędnych w danym zdaniu jest ściśle poissonowska z pewnym parametrem  $\lambda$ , lecz parametr ten zmienia się od zdania do zdania w sposób losowy, czyli jest również zmienną losową. Jeśli się przyjmie, że gęstość parametru  $\lambda$  wyraża się rozkładem gamma, wówczas prawdopodobieństwo, że liczba zdań współrzędnych w danym zdaniu złożonym wynosi dokładnie  $k$ , określone będzie przez wyrażenie postaci

$$(4) \quad P\{X = k\} = (-1)^k \binom{-v}{k} p^k q^v,$$

czyli przez tzw. rozkład ujemny dwumianowy. Lundberg [4] zauważył, że jest on równoważny granicznemu rozkładowi Pólyi (2).

Łatwo tłumaczyć sobie niejednorodność tekstu pod względem ilości zdań podrzędnych lub współrzędnych. Znaczy to, że autor mniej lub więcej świadomie posługuje się w różnych partiach tekstu zdaniami niejednakowo rozwiniętymi. Może on w ten sposób przystosowywać swój sposób pisania do treści wypowiedzi, aby wywołać zamierzoną reakcję czytelnika. Tak więc dla uzyskania jednego efektu używa zdań krótkich, mało rozbudowanych, dla innego — zdań obszernych, dobrze rozwiniętych. Celowe może być również umieszczenie zdania krótkiego na tle długich i rozmaite inne kombinacje.

Do uporządkowania zdań złożonych powrócimy niebawem, teraz zaś chciałabym dodać kilka słów o rozkładzie Pólyi. Rozkład ten jest często używany w statystyce; opisuje on takie zjawiska, jak zapadanie

<sup>(1)</sup> Dalszą bibliografię można znaleźć np. w pracy [7].



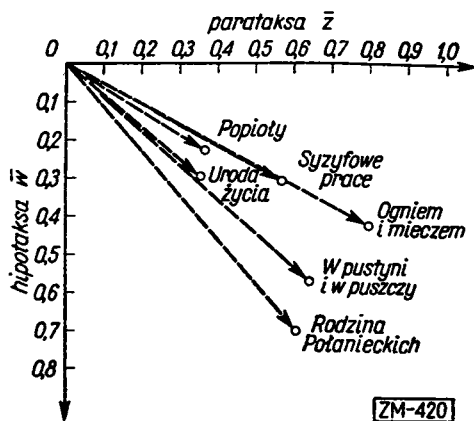
na zaraźliwą chorobę, uleganie nieszczęśliwemu wypadkowi, rozpad cząstek kosmicznych, powstawanie braków w seryjnej produkcji narzędzi, itd. Tak więc jeden matematyczny wzór może opisać wiele różnych postaci rzeczywistości, wskazując na istotę mechanizmu zjawiska. Dla statystyka jest rzeczą obojętną, czy liczy on ludzi, cząstki kosmiczne, czy też słowa lub zdania; dla niego ważne jest jedynie to, by była to zbiorowość cząstek elementarnych podlegających tym samym pracom masowym.

4. **Korelacja między składowymi parataksy i hipotaksy.** Bardzo łatwo można obliczyć średnią liczbę zdań podrzędnych i współrzędnych przypadających na jedno zdanie złożone. W tablicy II podaję wartości liczbowe, uzyskane dla badanych dzieł Żeromskiego i Sienkiewicza:

TABLICA II

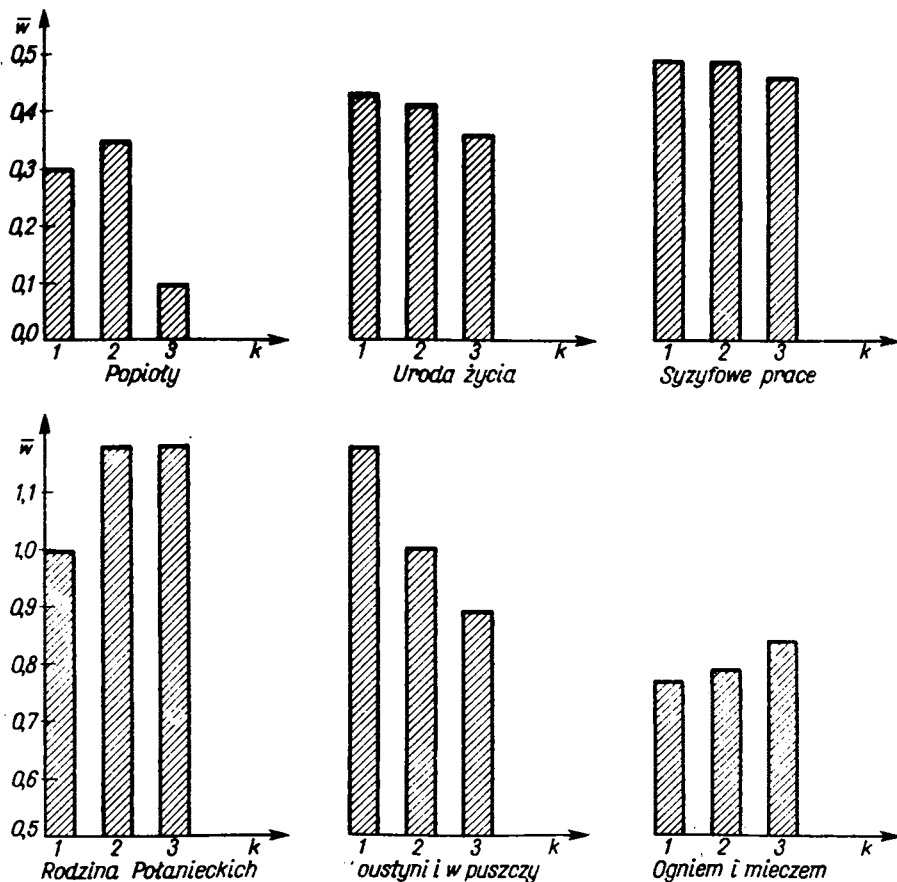
	średnia liczba zdań podrzędnych dla jednego zdania nadrzędnego	średnia liczba zdań złożonych współrzędnie z jednym zdaniem głównym
<i>Popioły</i>	0,225	0,375
<i>Uroda życia</i>	0,305	0,353
<i>Szybyłowe prace</i>	0,308	0,572
<i>Rodzina Polanieckich</i>	0,691	0,607
<i>Ogniem i mieczem</i>	0,420	0,796
<i>W pustyni i w puszczy</i>	0,560	0,635

Bardzo charakterystyczne jest to, że wszystkie trzy utwory Żeromskiego mają wartości średnie mniejsze niż utwory Sienkiewicza. Jeszcze wyraźniej można to przedstawić na rysunku. Charakteryzujemy mianowicie każde dzieło parą liczb: średnia liczba zdań współrzędnych, średnia liczba zdań podrzędnych, i tak uzyskane dane dla punktów indywidualnych zaznaczamy na płaszczyźnie. Przedstawia to rysunek 3. Widzimy na nim zupełnie wyraźnie, że utwory obu pisarzy tworzą dwie odrębne grupy. Zdania Sienkiewicza są bardziej rozbudowane niż zdania Żeromskiego. Na tym przykładzie widzimy, że długość wektora rozbudowy zdania  $\sqrt{\bar{z}^2 + \bar{w}^2}$  jest przejawem indywidualnym i mogłaby być cechą dyskryminacyjną dla obu pisarzy.



Rys. 3. Wektor rozbudowy zdania

Można tu jeszcze dodać, że językoznawcy przypisują składowym hipo- i parataksy znaczenie w określaniu stopnia rozwoju języka: „Przyjęło się sądzić, że rozwój języka postępuje od parataksy do hipotaksy, i że te dwie zasadnicze tendencje — do luźnego współrzednego lub ścisłego podrzędnego porządkowania elementów wypowiedzi — wyznaczają kierunek doskonalenia się języka, jego usprawniania jako narzędzia myśli. Zamiłowanie do parataksy byłoby więc wyrazem pewnego prymity-



[ZM-421]

Rys. 4. Średnia liczba zdań podrzędnych  $\bar{w}$  w zdaniach o różnej liczbie zdań współrzednych  $k$

wizmu językowego, skłonność zaś do hipotaktycznego wyrażania myśli cechowałaby prozę intelektualną, wysoce rozwiniętą, precyzyjnie ukształtowaną. Rzecz jasna, chodzi tu tylko o wytyczenie najbardziej ogólnego kierunku rozwojowego...” ([6], str. 135).

Widzimy, że język Żeromskiego i Sienkiewicza, w sensie cytatu przytoczonego powyżej, jest na tym samym stopniu rozwoju — u obu bowiem

pisarzy hipotaksa i parataksa są rozwinięte w tym samym stosunku.

Nie znaczy to jednak, że autor, rozbudowując jedno konkretne, pojedyncze zdanie w kierunku parataksy, rozbudowuje je zarazem pod względem hipotaktycznym. Przeprowadzając badania, obliczyłam średnią liczbę zdań podrzędnych w zdaniu złożonym z jednego tylko zdania nadrzędnego, z dwu zdań nadrzędnych złożonych współrzędnie, oraz z trzech zdań złożonych współrzędnie. Zdań złożonych z więcej niż trzech zdań parataktycznych było bardzo mało i dlatego ich nie uwzględniłam. Wyniki przedstawiłam na rysunku 4. Okazuje się, że liczba zdań podrzędnych w zdaniu o dwu i trzech zdaniach współrzędnych nie jest bynajmniej dwu i trzykrotnie wyższa niż w pojedynczym zdaniu nadrzędnym, nie mającym żadnych zdań złożonych z nim współrzędnie, a wprost przeciwnie: dość często nawet spada ze wzrostem liczby zdań połączonych parataktycznie. Znaczy to, że pisarz nie rozbudowuje zdania jednocześnie w obu kierunkach: zdania złożone współrzędnie nie zawierają tylu zdań podrzędnych co zdania stojące pojedynczo.

**5. Klasyfikacja dendrytów zdań.** Posiadany przeze mnie materiał dendrytów dla 2802 zdań Żeromskiego i 2355 zdań Sienkiewicza sklasyfikowałam według liczby punktów (czyli form osobowych czasownika) wchodzących w skład dendrytu. Tablica III pokazuje różne konfiguracje dendrytów z podaniem ich częstości występowania (w %). Widzimy, że najliczniejszą klasę tworzą dendryty jednopunktowe — stanowią one 45,6 % ogółu zdań opisowych Żeromskiego i 27,3 % ogółu zdań opisowych Sienkiewicza. Dla porównania podaję, że wynikiem częściowych badań Klemensiewicza [3] była liczba 30 jako średni procent zdań pojedynczych we współczesnej polszczyźnie. Wierzbicka dla fragmentów prozy z XIX i XX w. otrzymała 39,5 %. Jak widać z przytoczonych przykładów, wielkość ta może być różna dla poszczególnych autorów.

W następnej klasie dendrytów dwupunktowych konfiguracja: „dwa pojedyncze zdania złożone współrzędnie” ma pewną przewagę nad konfiguracją „jedno zdanie główne z jednym zdaniem podrzędnym”. Zjawisko to występuje u obu autorów, a test  $\chi^2$  wykazuje wysoką istotność. W następnej klasie dendrytów trzypunktowych możliwe są cztery uporządkowania trójki punktów. Okazuje się, że nie są one jednakowo prawdopodobne, przy czym znowu spotykamy dla obu autorów tę samą kolejność poszczególnych uporządkowań ze względu na ich częstość występowania w tekście. Odkrywamy tu ponownie pewną przewagę parataksy nad hipotaksą; trzy zdania pojedyncze połączone relacją współrzędności są bardziej prawdopodobne, niż trzy zdania z podwójnym stopniem złożoności.





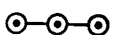



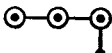

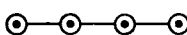
Konfiguracje opisane w powyższych trzech wyszczególnionych klasach są większością ogółu wszystkich spotykanych; dla Żeromskiego stanowią one 92,3 %, dla Sienkiewicza 75,9 % zdań złożonych.

W tablicy III są również podane w pewnym rozbiu frekwencje dendrytów czteropunktowych: i tu widzimy, że nie wszystkie konfiguracje są jednakowo prawdopodobne; jest ich zresztą bardzo niewiele: 5% u Żeromskiego i 12 % u Sienkiewicza.

TABLICA III

## Różne konfiguracje dendrytów zdań

(liczby podane przy klasach wskazują na frakcję zdań poszczególniej klasy w stosunku do całkowitej liczby zdań badanych)

Rodzaj klasy	Żeromski	Sienkiewicz
A)  jednopunktowe	0,456	0,273
B) dwupunktowe		
 dwa zdania współrzędne	0,176	0,155
 jedno zdanie nadrzędne i jedno zdanie podrzędne	0,137	0,118
C) trzypunktowe		
 dwa zdania współrzędne jedno podrzędne	0,0585	0,0912
 trzy zdania współrzędne	0,0485	0,0526
 jedno zdanie główne z dwoma podrzędnymi	0,0332	0,0403
 jedno zdanie główne z podwójnym stopniem złożoności	0,0139	0,0285
D) czteropunktowe		
 dwa zdania współrzędne z dwoma zdaniami podrzędnymi	0,0150	0,0447
 trzy zdania współrzędne z jednym zdaniem podrzędnym	0,0107	0,0285
 jedno zdanie główne z trzema podrzędnymi	0,0103	0,0289
 cztery zdania współrzędne	0,0121	0,0148
E) inne	0,0290	0,125

Ogólnie możemy z tego zestawienia wysnuć następujące wnioski:

1. Zdania pojedyncze stanowią największą frakcję ogółu zdań.
2. Nie wszystkie konfiguracje są jednakowo prawdopodobne. Pisarze mają preferencje do pewnych form, schematów.

## 6. Badanie sąsiedztwa (kolejności) zdań prostych i złożonych.

W sekcji tej interesować nas będzie zagadnienie, czy zdania proste i złożone rozmieszczone są w tekście w sposób zupełnie przypadkowy, czy też można zaobserwować w ich kolejności pewne odstępstwa od losowości. Można przypuszczać, że kolejność zdań prostych i złożonych jest cechą stylistyczną tekstu; wykrycie jakichś odstępstw od losowości będzie świadczyć o pewnej ingerencji autora, który dla celów ekspresywnych gromadzi w jakimś miejscu zdania długie, unikając krótkich, lub na odwrót. W dalszym ciągu przez zdanie proste rozumiem zdanie posiadające tylko jedno orzeczenie; oznaczam je symbolem  $0$ ; zdania posiadające więcej niż jedno orzeczenie w formie osobowej czasownika zaliczam do zdań złożonych i oznaczam symbolem  $1$ . W danym tekście wyodrębniamy pary zdań oraz badamy częstość par:  $\langle \text{zdanie proste} - \text{zdanie proste} \rangle = \langle 0, 0 \rangle$ ,  $\langle \text{zdanie proste} - \text{zdanie złożone} \rangle = \langle 0, 1 \rangle$ ,  $\langle \text{zdanie złożone} - \text{zdanie proste} \rangle = \langle 1, 0 \rangle$ ,  $\langle \text{zdanie złożone} - \text{zdanie złożone} \rangle = \langle 1, 1 \rangle$ . Tak np. w ciągu  $0010$  wyodrębniamy pary  $\langle 0, 0 \rangle$ ,  $\langle 0, 1 \rangle$ ,  $\langle 1, 0 \rangle$ . Niech  $p_0$  oznacza prawdopodobieństwo wystąpienia zdania prostego, a  $p_1$  — zdania złożonego; oczywiście  $p_0 + p_1 = 1$ . Jeśli elementy wchodzące w skład par są niezależne, prawdopodobieństwa odpowiednich par wynoszą:

$$(5) \quad \begin{aligned} p_{\langle 0, 0 \rangle} &= p_0^2, \\ p_{\langle 0, 1 \rangle} &= p_0 p_1, \\ p_{\langle 1, 0 \rangle} &= p_1 p_0, \\ p_{\langle 1, 1 \rangle} &= p_1^2. \end{aligned}$$

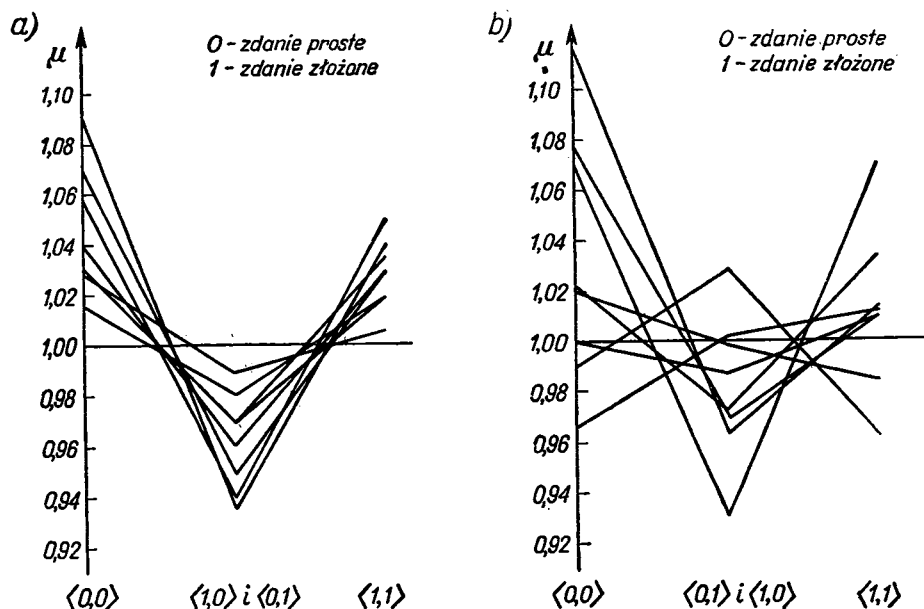
Uzyskany wynik dla pojedynczego utworu nie jest ciekawy — otrzymane częstości par nie różnią się istotnie od odpowiednich częstości teoretycznych, obliczonych na podstawie wzoru (5). Jeśli jednak obliczenia przeprowadzimy wielokrotnie, dla różnych utworów, to otrzymamy bardzo charakterystyczne rezultaty. Obliczenia takie powtórzyłam ośmiokrotnie, a wyniki przedstawiam w tablicy IV.

TABLICA IV  
Empiryczne częstości par zdań opisowych

$0$  oznacza zdanie proste,  $1$  oznacza zdanie złożone, liczby w nawiasach oznaczają częstość oczekiwaną przy założeniu niezależności elementów w parach

	$\langle 0, 0 \rangle$	$\langle 0, 1 \rangle, \langle 1, 0 \rangle$	$\langle 1, 1 \rangle$	Razem
<i>Nawracanie Judasza</i>	256 (239)	539 (566)	344 (334)	1139
<i>Zwierzenia</i>	395 (384)	701 (724)	354 (342)	1450
<i>Popioły</i>	152 (143)	243 (258)	122 (116)	617
<i>Uroda życia</i>	166 (159)	294 (306)	152 (147)	612
<i>Syzyfowe prace</i>	173 (159)	490 (523)	449 (430)	1112
<i>Ogniem i mieczem</i>	99 (96)	456 (470)	588 (577)	1143
<i>W pustyni i w puszczy</i>	40 (39)	240 (243)	379 (377)	659
<i>Rodzina Polanieckich</i>	40 (40)	201 (206)	275 (270)	516

Zamieszczone są w niej otrzymane empirycznie licznosci par  $\langle 0, 0 \rangle$ ,  $\langle 0, 1 \rangle$  lub  $\langle 1, 0 \rangle$  i  $\langle 1, 1 \rangle$ , a obok nich w nawiasach częstości oczekiwane przy założeniu niezależności zdań w parach — według wzoru (5). Porównując dane empiryczne z wartościami oczekiwanymi widzimy, że we wszystkich 8 przypadkach pary:  $\langle$ zdanie proste — zdanie proste $\rangle$  występują częściej, niżby to wynikało z rozkładu przypadkowego. Podobnie we wszystkich 8 przypadkach pary  $\langle$ zdanie złożone — zdanie złożone $\rangle$  występują częściej, niż to wynika z założenia niezależności elementów w parach. Możemy oszacować istotność otrzymanych wyników. Przy założeniu niezależności zdań w parach powinniśmy otrzymać 8 nadwyżek i 8 niedoborów na 16 wyników, my tymczasem otrzymaliśmy 16 nadwyżek i 0 niedoborów. Wartość  $\chi^2$  takiego wyniku wynosi 16, przy jednym stopniu swobody, przy czym wiadomo, że  $\Pr\{\chi^2 \geq 6,635\} = 0,01$ , a więc otrzymane różnice są wysoce istotne.



ZM-422

Rys. 5. Następstwa zdań prostych i złożonych: a) Wykres wartości  $\mu$  dla par zdań sąsiednich, b) Wykres wartości  $\mu$  dla par zdań nie sąsiadujących ze sobą

Aby uzyskać łatwe wyobrażenie o wielkości zaobserwowanych odstępstw, obliczamy stosunek licznosci zaobserwowanych do teoretycznych (oczekiwanych)

$$\mu = \frac{n_{\text{emp}}}{n_{\text{teor}}}.$$

Odpowiednie wielkości  $\mu$  dla par  $\langle 0, 0 \rangle$ ,  $\langle 0, 1 \rangle$  lub  $\langle 1, 0 \rangle$  i  $\langle 1, 1 \rangle$  przedstawiono wykreślnie na rysunku 5a. Dla większej pogładowości wielkości odpowiadające parom pochodzącym z tego samego utworu połączono linią łamaną. Otrzymany obraz jest bardzo znamieny. Wszystkie łamane mają ten sam kształt.

Dla porównania obliczono analogiczne wartości dla par nie sąsiednich, lecz przedzielonych jednym elementem (zdaniem); z ciągu 0010 otrzymamy dwie takie pary:  $\langle 0, 1 \rangle$  i  $\langle 0, 0 \rangle$ . Jak wykazuje rysunek 5b, harmonia łamanych odpowiadających odpowiednim utworom jest już zakłócona. Znaczy to, że przyciąganie się zdań prostych i złożonych nie sięga zbyt daleko i daje się zauważyć jedynie w sąsiednich elementach.

Dokładniejsze wyobrażenie o sposobie i wielkości przyciągania się czy też odpychania zdań prostych lub złożonych daje nam badanie rozkładu serii. Jeśli mamy na przykład ciąg składający się z dwóch rodzajów elementów zerowych i jedynkowych, których kolejność jest następująca: 0, 1, 1, 1, 0, 0, to ciąg ten zawiera dwie serie elementów zerowych, mianowicie jedną o długości 1 i jedną o długości 2, oraz jedną serię elementów jedynkowych o długości 3.

Gdyby elementy zerowe wykazywały tendencję do przyciągania się, to w otrzymanym ciągu nastąpiłyby pewne zgrupowania elementów zerowych, większe niż to przewiduje ich rozkład obliczony przy założeniu ściśle przypadkowej kolejności. W takim przypadku liczba serii złożonych z elementów zerowych byłaby mniejsza niż ich oczekiwana wartość. Można sobie również wyobrazić taką sytuację: zdania proste w niektórych partiach tekstu wykazują tendencję do skupiania się, w innych zaś partiach do odpychania się. W sumie wszystkich serii mogłoby być tyle ile trzeba, wspomnianą zaś prawidłowość można wykryć tylko badając cały rozkład serii.

Badania rozkładu serii zostały przeprowadzone na dwóch specjalnych próbach. Chodziło o to, aby uzyskać względnie długi ciąg zdań opisowych (w próbach badanych uprzednio zdania opisowe mogły być przeplatane zdaniami dialogowymi). Za jedną próbę wzięto cały rozdział z *Ludzi bezdomnych* — mianowicie *Zwierzenia (Pamiętnik Joasi)*; za drugą próbę wzięto możliwie długie partie opisowe z *Nawracania Judasza Żeromskiego*.

Posługiwano się następującymi wzorami: Załóżmy, że było  $n_0$  elementów zerowych i  $n_1$  elementów jedynkowych, tak że  $n_0 + n_1 = n$ .

Wprowadźmy oznaczenie potęgi faktorowej zmiennej  $x$ :

$$x^{(a)} = x(x-1)(x-2)\dots(x-a+1) \quad (x^{(0)} = 1, \quad a \text{ całkowite}).$$

Niech  $r_{0i}$  oznacza liczbę serii składających się z elementów zerowych o długości  $i$ .

Wówczas wartość oczekiwana i wariancja wynoszą odpowiednio:

$$(6) \quad E(r_{0i}) = \frac{(n_1+1)^{(2)}n_0^{(i)}}{n^{(i+1)}},$$

$$(7) \quad \sigma^2(r_{0i}) = \frac{n_1^{(2)}(n_1+1)^{(2)}n_0^{(2i)}}{n^{(2i+2)}} + \frac{(n_1+1)^{(2)}n_0^{(i)}}{n^{(i+1)}} \left(1 - \frac{(n_1+1)^{(2)}n_0^{(i)}}{n^{(i+1)}}\right).$$

Oznaczmy przez  $s_{0k}$  ilość serii elementów zerowych o długości większej niż  $k$ :

$$s_{0k} = \sum_{j=k+1}^{n_0} r_{0j}.$$

Wartość oczekiwana i wariancja zmiennej losowej  $s_{0k}$  wynoszą odpowiednio

$$(8) \quad E(s_{0k}) = \frac{(n_1+1)n_0^{(k)}}{n^{(k)}},$$

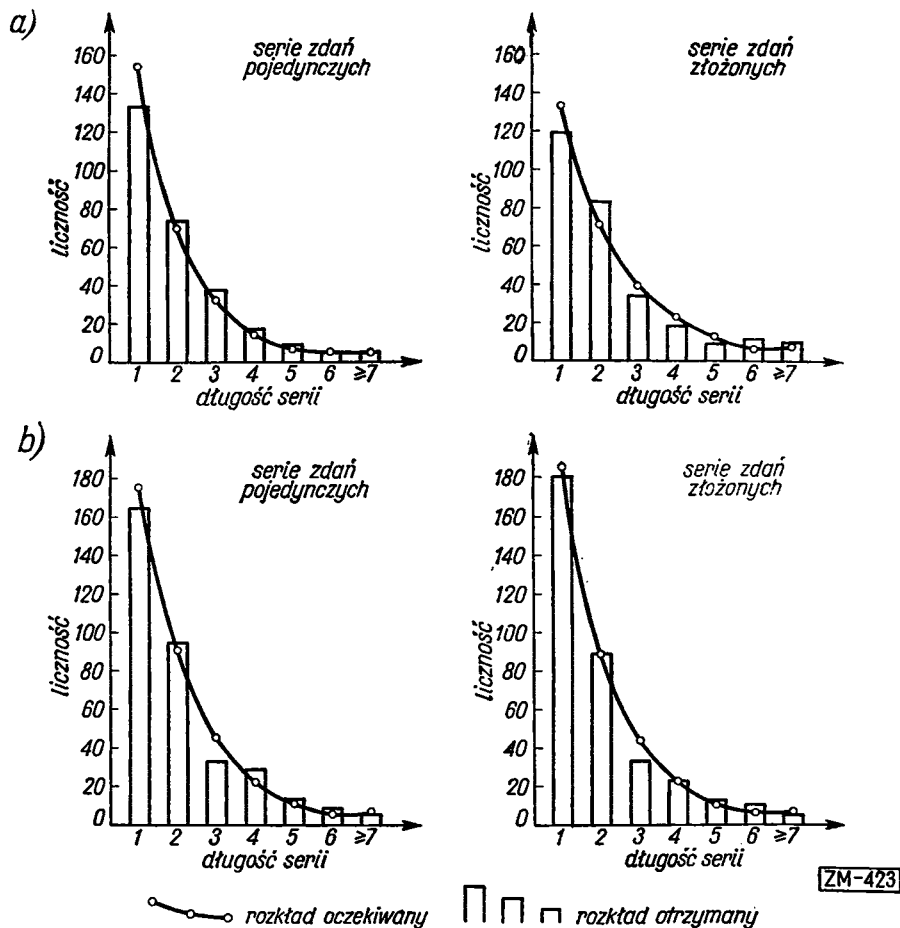
$$(9) \quad \sigma^2(s_{0k}) = \frac{(n_1+1)^{(2)}n_0^{(2k)}}{n^{(2k)}} + \frac{(n_1+1)n_0^{(k)}}{n^{(k)}} \left(1 - \frac{(n_1+1)n_0^{(k)}}{n^{(k)}}\right).$$

Wyprowadzenie tych wzorów znajdzie czytelnik w pracy Mooda [5]. Wzory na  $r_{1i}$  i  $s_{1i}$  uzyskuje się przez odpowiednią permutację zer i jedynek. Otrzymane rozkłady serii dla dwóch badanych utworów są przedstawione graficznie na rysunku 6. Policzone na podstawie prób empiryczne częstości serii o odpowiedniej długości przedstawione są na wykresie w postaci słupków; oczekiwane (teoretyczne) licznosci, obliczone według wzorów (6) i (8), zostały zaznaczone punktami i dla większej poglądowości połączono je linią ciągłą. Widzimy, że rozkład empiryczny nie odbiega na ogół od rozkładu teoretycznego. Zastanawia tylko fakt, że we wszystkich czterech przypadkach serii o długości 1 jest mniej niż wynosi ich wartość oczekiwana. Różnice te nie są istotne i mieszczą się w odchyleniu  $2\sigma$ , niemniej jednak są one zgodne z wynikami otrzymanymi uprzednio przy badaniu par sąsiednich elementów. Otrzymaliśmy tam mianowicie wprowadzić niewielką, ale istotną nadwyżkę par typu  $\langle 0, 0 \rangle$  i  $\langle 1, 1 \rangle$ , skąd powinno wynikać, że elementów zerowych i jedynkowych stojących samotnie powinno być trochę mniej, niżby to wynikało z losowego rozkładu. Tak więc można sądzić, że gdybyśmy badanie serii przeprowadzili na odpowiednio większym materiale, otrzymalibyśmy pewną istotną chociaż niewielką różnicę na odpowiednich miejscach.

Jak można zaobserwowany powyżej fakt przetłumaczyć na język stylistów? Chciałabym wysunąć tu pewną hipotezę: Zaobserwowaliśmy mniej sytuacji takich, że po zdaniu złożonym następuje zdanie proste, a potem znowu złożone, lub też sytuacji odwrotnych: po zdaniu zło-



zonym następuje zdanie proste i potem znowu złożone. Znaczy to, że nie zachodzą raptowne zmiany w wielkości rozbudowy zdania, że stopień rozbudowy zdania posiada jak gdyby pewną bezwładność. Daje się ona zauważyć właśnie w małej ilości raptownych przejść od jednego „nastroju” do drugiego (o ile odpowiednio rozwiniętemu zdaniu przypiszemy jakiś „nastój” pisarza).



Rys. 6: a) Rozkład serii dla *Nawracania Judasza*, b) Rozkład serii dla *Zwierzeń*

Wszystko to przypomina pewne zjawisko, zaobserwowane przy badaniu określeń [1]. Okazało się wówczas, że ilość słów w określeniu ulega — jak to wówczas nazwałam — jak gdyby „przyciąganiu na przekątną”. Po określeniu jednoelementowym występowała pewna nadwyżka określeń jednoelementowych, po określeniach dwuelementowych — pewna nadwyżka określeń dwuelementowych, itd. (Jedynie w klasach liczących 6-10 i więcej elementów nie zaobserwowano wyraźnie żadnych prawidłowości, były to jednak klasy bardzo nieliczne i dlatego słabo repre-

zentowane.) Świadczy to znów o jakiejś „bezwładności” wrażeń czy przedstawień konkretyzowanych w postaci słów w określeniu. W przypadku zdań prostych i złożonych bezwładność ta odnosi się do większych jednostek tekstowych, ale opisuje to samo zjawisko.

**7. Badanie sąsiedztwa (kolejności) zdań parataktycznych i hipotaktycznych.** Badania przeprowadzono na tym samym materiale, co badanie serii zdań prostych i złożonych. Otrzymane wyniki nie wnoszą nic specjalnie nowego, potwierdzają jedynie pewne nasze przypuszczenia.

Badanie liczby serii dla zdań składających się z jednego, dwu i więcej orzeczeń połączonych współrzędnie (podrzędnie) wykazuje, że zawsze otrzymujemy łączną liczbę serii mniejszą niż oczekiwaną (patrz tablica V).

TABLICA V

Liczba serii w zdaniach

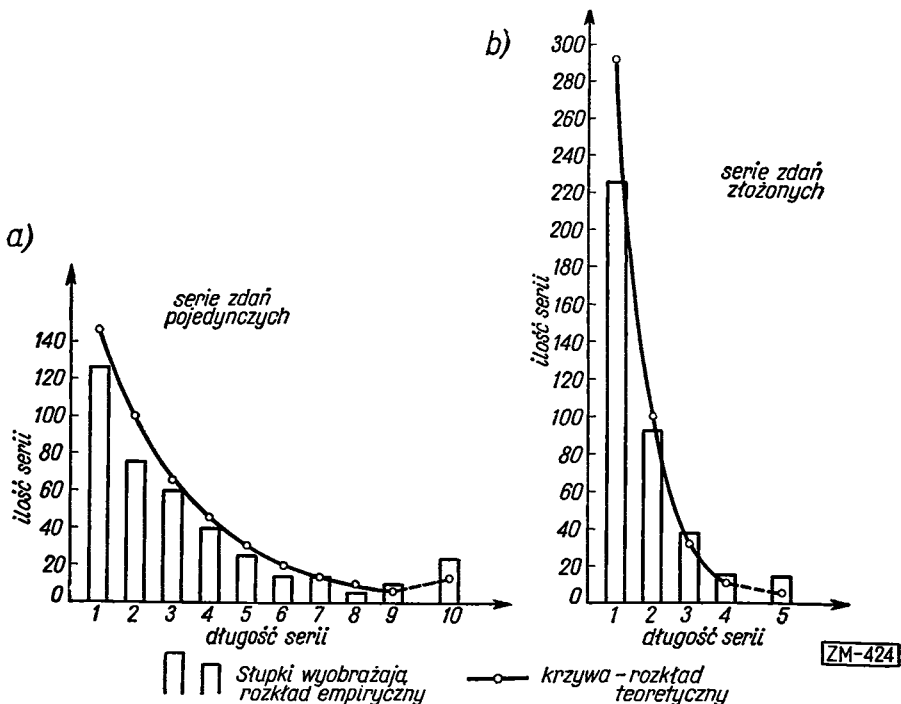
(Liczby w nawiasach oznaczają oczekiwaną liczbę serii)

	<i>Zwierzenia</i>		<i>Nawracanie Judasza</i>	
	zдания współrzędne	zдания podrzedne	zдания współrzędne	zдания podrzedne
zдания o jednym orzeczeniu	241 (247)	349 (386)	208 (214)	397 (445)
zдания o dwóch orzeczeniach	201 (205)	277 (298)	182 (183)	323 (332)
zдания o więcej niż dwóch orzeczeniach	66 (66)	153 (156)	45 (45)	195 (216)
łączna liczba serii	508 (518)	779 (840)	435 (442)	916 (993)

Rzeczą uderzającą w powyższym zestawieniu jest to, że w obydwu badanych utworach dla zdań współrzędnych ogólna liczba serii jest wprawdzie większa niż jej wartość oczekiwana, ale różnica ta jest niewielka i statystycznie nieistotna. Rozkład serii dla zdań współrzędnych jest bardzo podobny do badanego już rozkładu serii dla zdań prostych i złożonych i dlatego nie przytaczam go.

Rozkład serii zdań podrzędnych (rys. 7) zarówno dla zdań prostych, jak i dla złożonych wykazuje pewne charakterystyczne odchylenia na początku i końcu rozkładu. Na początku mianowicie rozkład empiryczny przyjmuje wartości mniejsze od wartości oczekiwanych, na końcu rozkładu wartości większe. Znaczy to, że zdań prostych, stojących między dwoma zdaniem o rozbudowanej hipotaksie jest mniej, niżby wynikało z przypadkowego ich rozmieszczenia; istnieje więc jakiś powód przeszkadzający takiej konfiguracji. Może to być owa „bezwładność” wyobrażeń, gdy pisarz „nastawiony” na budowanie zdań złożonych podrzędnie niełatwo przechodzi do zdania prostego i znowu do zdań złożonych, chyba że czyni to z pewnych względów impresyjno-stylistycznych.

Na końcu rozkładów uwidocznionych na rysunku 7 spotykamy się ze zjawiskiem odwrotnym: dane empiryczne są większe, niżby to wynikało z krzywej teoretycznej. Świadczy to o pewnej nadwyżce długich serii, czyli grupowaniu zdań pojedynczych i o rozbudowanej hipotaksie. Jest to drugie ogniwo wypowiedzianej hipotezy o „bezwładności” wrażeń. O bezwładności świadczył fakt, że raptownych przejść: zdanie proste-złożone-proste było mniej niż można by oczekiwać, częściej natomiast spotykaliśmy układy: zdanie złożone-złożone-złożone lub: zdanie proste-proste-proste; świadczy to o pewnej preferencji utrzymywania przez



Rys. 7. Rozkład serii dla zdań hipotaktycznych z *Nawracania Judasza*

autora nastroju do zdań złożonych lub prostych. Jak widać jednak z przytoczonych danych, zjawisko to występuje w bardzo nikłym stopniu i dlatego w niewielkich próbach może być niedostrzegalne.

**8. Badanie jednorodności gron.** Interesujące jest, czy pisarz w różnych miejscach utworu zmienia parametry rozwinięcia zdania, czy na przykład w niektórych partiach książki przeważają zdania dobrze rozwinięte, w innych zaś mniej. Materiał, jakim dysponowałam, składał się z kilku prób pobranych z różnych miejsc książki, po kilkadziesiąt zdań każda. Aby uprościć obliczenia, sklasyfikowałam zdania na dwa typy: zdania proste i zdania złożone (posiadające więcej niż jedno orzeczenie).

Wyniki dla każdego utworu zestawiałam w tabelkach, jak na przykład:

nr grona typ zdania	I	II	III	IV	V	VI	VII
0	20	24	21	35	16	17	12
1	52	48	51	51	61	55	60

Stosując test  $\chi^2$  jako kryterium jednorodności testuję hipotezę, że dane próbkowe przedstawione w kolumnach powyższej tabelki są wszystkie pobrane z tej samej populacji.

Wartość  $\chi^2$  dla powyższego zestawienia, które przedstawia skład kolejnych gron dla próby *Rodziny Połanieckich*, wynosi w przybliżeniu 15,4 przy 6 stopniach swobody. Z tablic odczytujemy, że  $\Pr\{\chi^2 \geq 12,59\} = 0,05$ ,  $\Pr\{\chi^2 \geq 16,81\} = 0,01$ . Wynik uzyskany przez nas leży więc na granicy istotności.

Podobne wyniki (wartości  $\chi^2$  zbliżają się do granic istotności) otrzymałam dla trzech innych badanych utworów, pozostałe dwie badane próby wykazują dużą jednorodność w gronach.

Tak więc, w świetle uzyskanych wyników nie udało nam się wykazać, że zdania w różnych miejscach utworu mają różny stopień rozbudowy.

#### Prace cytowane

- [1] A. Bartkowiakowa, *O rozkładzie określeń w zdaniach opisowych Żeromskiego i Sienkiewicza*, Zastosow. Mat. 6 (1962), str. 285-301.
- [2] M. Fisz, *Rachunek prawdopodobieństwa i statystyka matematyczna*, Warszawa 1958.
- [3] Z. Klemensiewicz, *Problematyka składniowej interpretacji stylu*, Pamiętnik Literacki 42 (1951), str. 102-157.
- [4] O. Lundberg, *On random processes and their application to sickness and accident statistics*, Uppsala 1940.
- [5] A. M. Mood, *The distribution theory of runs*, Annals of Math. Stat. 11 (1940), str. 367.
- [6] A. Wierzbicka, *Okres retoryczny a ogólne tendencje składni szesnastowiecznej*, Pamiętnik Literacki 52 (1961), str. 125-138.
- [7] C. J. Bliss and R. A. Fisher, *Fitting the negative binomial distribution to biological data and note on the efficient fitting of the negative binomial*, Biometrics 9 (1953), str. 176.

INSTYTUT MATEMATYCZNY POLSKIEJ AKADEMII NAUK

Praca wpłynęła 9. 6. 1962

А. БАРТКОВЯКОВА (Вроцлав)

О РАСПРЕДЕЛЕНИИ И ОЧЕРЕДНОСТИ СОЧИНЁННЫХ  
И ПОДЧИНЁННЫХ ПРЕДЛОЖЕНИЙ В ПРОИЗВЕДЕНИЯХ  
ЖЕРОМСКОГО И СЕНКЕВИЧА

## РЕЗЮМЕ

В работе показывается, что число подчинённых предложений по отношению к одному главному предложению и число сложно-сочинённых с одним главным предложением выражается предельным распределением Поля (Pólya) (отрицательно-биномиальным). Распределения, полученные эмпирическим образом и соответствующие им распределения Поля (рассчитанные по формуле (2)) приведены в таблице I. Тот факт, что рассматриваемые распределения выражаются именно распределением (2), помогает в понимании механизма, управляющего применением простых и сложных предложений.

В дальнейшем исследуется количество подчинённых предложений в предложениях с различным числом сочинённых предложений (рис. 4). Оказывается, что если данное предложение построено сильнее в паратактическом направлении, то отдельные главные предложения связаны слабее с гипотактической точки зрения, а другими словами, соотношение между этими двумя величинами — отрицательное.

В таблице III представлены некоторые конфигурации сложных предложений и число их выступления в исследуемом тексте.

Исследовалась также последовательность применения простых и сложных предложений, особенно внимание обращая на соседние элементы (пара предложений) и распределение серии. Выяснено, что простые предложения в некоторой степени отталкивают сложные, и аналогично сложные предложения — притягивают простые. Это же явление относится как для паратакса, так и для гипотакса. Из распределения серии следует, что предложения с большой степенью следственности (по крайней мере содержащее одно подчинённое предложение) избегают соседства с простыми предложениями (не составленными гипотактически), и наоборот обладают тенденцией группирования в длинной серии.

При использовании теста  $\chi^2$ , как критерия однородности, не было обнаружено существенных различий для простых и сложных предложений, взятых из разных мест книги.

A. BARTKOWIAKOWA (Wrocław)

ON THE DISTRIBUTION AND ORDER OF HYPOTACTIC AND PARATACTIC  
SENTENCES IN THE NOVELS OF ŻEROMSKI AND SIENKIEWICZ

## SUMMARY

It is proved in the paper that the number of subordinate (paratactic) clauses per one compound sentence and the number of compound sentences with one main clause is expressed by a limit (negative binomial) Pólya distribution. The distributions obtained empirically and the corresponding Pólya distributions (found according to formula (2)) are put together in Table I. The very fact that the distributions in

question are expressed by distribution (2) throws some light of the mechanism controlling the formation of simple and compound sentences.

The author then investigates the number of subordinate clauses in sentences with different numbers of coordinate clauses (fig. 4). It is shown that the more a given sentence is developed paratactically the weaker is the hypotactical development of the individual main clauses, i.e. that the correlation between these two quantities is negative.

Table III lists certain configurations of compound sentences and gives the frequencies of their occurrence in the text under consideration.

The author also investigates the order of appearance of simple and compound sentences, taking into consideration neighbouring elements (pairs of sentences) and the distribution of series. It is shown that simple sentences attract, in a slight degree, simple sentences and repel compound sentences, and, analogously, compound sentences attract compound sentences while repelling simple ones. A similar phenomenon occurs for parataxis and hypotaxis. It follows from the distribution of a series (fig. 7) that sentences with a greater degree of subordination (having at least one subordinate clause) avoid contact with simple sentences (undeveloped hypotactically) and have a tendency to concentrate into longer series.

With the use of test  $\chi^2$  as a criterion of homogeneity, no significant differences have been obtained for the numbers of simple and compound sentences in clusters taken from different parts of a given book.

---