

Agnieszka Kułacka  
Uniwersytet Wrocławski  
Wrocław

## WARUNKI ZACHODZENIA PRAWA MENZERATHA–ALTMANNA

### 1. Wstęp

Już Jan Baudouin de Courtenay (1927) pisał o potrzebie wykorzystania matematyki i statystyki do opisywania zjawisk świata językowego:

Jedyną metodą o podłożu matematycznym stosowaną dotychczas w językoznawstwie jest metoda statystyczna [podkr. J.B. de C.], tj. określenie empiryczne wzajemnego stosunku procentowego różnorodnych przejawów życia językowego (Baudouin de Courtenay 1927: 560).

W latach późniejszych nastąpiło znaczne spowolnienie badań nad statystycznymi prawami językowymi, chociaż badania ilościowe prowadzone były bardzo intensywnie, głównie w krajach anglosaskich (por. Köhler 1995). W Polsce również zajmowano się nimi, choć na mniejszą skalę (por. Sambor, Pawłowski 2005; Pawłowski 2006). Dopiero w latach 80. pojawiły się próby modelowania matematycznego dostrzeżonych wcześniej zależności w językach. Badania te były jednak rozproszone. Idąc za pracami Rolfa Hammerla i Jadwigi Sambor (Hammerl 1987, 1989; Hammerl, Sambor 1993), można wyróżnić cztery etapy rozwoju lingwistyki kwantytatywnej, z których dopiero ostatni charakteryzuje odkrywanie praw językowych za pomocą metody dedukcyjnej, polegającej na konstruowaniu określonych modeli funkcyjnych z wcześniejszym przyjęciem założeń dotyczących zależności między badanymi zmiennymi. Wychodząc od modeli statystycznych, relacje te wyjaśnia się w kategoriach lingwistycznych. Przyjęto, że odkryta przez George’a K. Zipfa zasada minimalnego wysiłku nadawcy (ujednolicenie jednostek językowych) oraz odbiorcy (zróznicowanie jednostek językowych) stanowi teoretyczne podstawy do wyjaśniania działania praw językowych.

## 2. Geneza prawa i zarys dotychczasowych badań

Prawo Menzeratha–Altmanna (prawo MA) zostało po raz pierwszy opisane przez Paula Menzeratha w 1928 r. na podstawie wcześniejszych obserwacji francuskiego językoznawcy A. Grégoire’a. Na początku Menzerath ograniczył swoje badania nad prawem do fonologii, by w 1954 r. sformułować relację MA dla całego systemu językowego. Prawo to mówi: „im dłuższa konstrukcja językowa, tym krótsze są jej składniki” (cyt. za: Hammerl, Sambor 1993: 23).

W 1980 r. Gabriel Altmann (Altmann 1980) podał dla prawa MA model funkcyjny, który został uzupełniony przez Agnieszkę Kułacką i Jana Maćutka w 2006 r. (Kułacka, Maćutek 2006). Obecnie istnieją dwie postaci modelu: ciągła i dyskretna.

Dotychczasowe badania obejmują wszystkie poziomy języka, począwszy od fonologii poprzez morfologię, składnię, semantykę, a skończywszy na poziomie dyskursu. Zajmowano się również znakami chińskimi. Podjęto też próbę wyjaśnienia konieczności działania prawa MA oraz postawiono hipotezy dotyczące współczynników we wzorach dla prawa. Badano je również w innych dziedzinach nauki: w muzykologii, genetyce, biologii ewolucyjnej itd.

Ze względu na stosowaną dotychczas metodę badań, to statystyczne prawo zdaje się zachodzić zawsze. Metoda ta została przeze mnie uzupełniona nowymi spostrzeżeniami, które skrótowo przybliżę poniżej. Dokładne jej omówienie podaję w artykule *Badania nad prawem Menzeratha–Altmanna* (Kułacka 2008). W niniejszej publikacji podsumuję moje badania nad czynnikami zakłócającymi działanie prawa MA, które zostały wskazane na zbiorze wybranych tekstów. Aspekt dwujęzyczny, uwzględniony tu dla wzbogacenia zbioru tekstów, stanowi materiał innego artykułu.

## 3. Zarys metody badań nad prawem

Moje badanie przeprowadziłam na wybranych tekstach polskich i angielskich, napisanych w dwóch różnych stylach: literackim i naukowym. Metoda próbkowania, którą zastosowałam, to nieprobabilistyczna<sup>1</sup> technika wyboru próby. W moim badaniu na pierwszym etapie dokonałam wyboru zbioru tekstów metodą przypadkową (ang. *convenience sampling*), tj. wybrałam teksty, które były dostępne. Następnie doboru zdań dokonałam metodą kwotową (ang. *quota sampling*), tj. założyłam górną granicę liczby zdań z danego tekstu. Badania na danym tekście kończyłam, gdy jego długość osiągała liczbę: KONIEC = min(M, ≈2000), gdzie M jest minimalną długością odcinka danego tekstu, dla którego prawo MA zachodzi, natomiast 2000 jest liczbą badanych zdań (analizie poddano całe rozdziały książek; więcej na ten temat zob. Kułacka 2008).

Trzeci etap wyboru próby realizowałam, stosując metodę celową (ang. *judgemental sampling*). Do analizy włączyłam 98% populacji zdań danego tekstu, stanowiących zdania o najmniejszym stopniu złożoności. Rozbioru składniowego zdań dokonałam na podstawie schematów zdaniowych zaproponowanych przez Renatę

<sup>1</sup> W literaturze nie ma zgody co do sposobu wyboru próby do badania prawa MA.

Grzegorzyczkową (2004), porównując je z klasyfikacją Roberta M. W. Dixona (2005), by ujednoczyć w miarę możliwości analizę składniową zdań.

Prawo w odniesieniu do poziomu składni mówi o odwrotnie proporcjonalnej relacji między stopniem złożoności zdania a długością jego składników. Weryfikowałam zachodzenie prawa MA za pomocą współczynnika korelacji rangowej Spearmana. Prawo MA opisuje nieliniową korelację między stopniem złożoności zdania złożonego a średnią długością zdania składowego i dlatego wskazane jest stosowanie tego współczynnika. Przyjrzyjmy się tej metodzie na poniższym przykładzie.

Przyjmuję następujące oznaczenia:  $x$  – liczba zdań składowych,  $y_{ang}$  – średnia długość zdania składowego,  $n_{ang}$  – liczba przebadanych zdań.

Tabela 1. Dane na podstawie angielskiej wersji *Kodu Leonarda da Vinci* D. Browna

x	$y_{ang}$	$n_{ang}$	ranga	$d_i$
1	6,7584	745	1	0
2	6,1677	465	3	1
3	6,1976	275	2	1
4	5,7442	86	4	0
5	5,5928	28	5	0
6	5,3627	17	–	–
7	5,8571	5	–	–
8	6,4583	3	–	–
9	6,4444	2	–	–
suma		1626	–	–

98% z liczby 1626 to, w zaokrągleniu do liczb całkowitych, 1593. Suma zdań złożonych, gdzie  $n=1,2,3,4,5$ , wynosi 1599. Do analizy włączyłam więc zdania o liczbie zdań składowych 5 lub mniej.

Rozważmy hipotezę zerową  $H_0$  i jej alternatywę  $H_1$ :

$H_0$  :  $\rho = 0$  – nie istnieje zależność między liczbą zdań składowych  $x$  a ich średnią długością  $y_{ang}$ .

$H_1$  :  $\rho > 0$  – istnieje dodatnia zależność, co oznacza, że prawo MA zachodzi.

W powyższym zapisie grecka litera  $\rho$  (*rho*) oznacza poziom korelacji. Jeśli wynosi on 0, mamy do czynienia z brakiem korelacji między badanymi rangami zmiennych. Jeśli jest dodatni, korelacja jest dodatnia, czyli wraz ze wzrostem wartości jednej rangi rośnie wartość drugiej. W wypadku prawa MA wraz ze wzrostem złożoności zdań maleje ich długość (por. rangi w tabeli 1).

Największej średniej długości zdań przypisałam rangę 1. Jest to test jednostronny, ponieważ sprawdza on nie tylko to, czy istnieje korelacja między rangami (dodatnia lub ujemna), jak w wypadku testu dwustronnego, ale sprawdza również, jaki jest jej kierunek (w tym wypadku odpowiadamy na pytanie, czy korelacja jest dodatnia). Mamy tu korelację nieliniową, a więc celowe jest stosowanie współczynnika

korelacji rangowej Spearmana, który daje nam odpowiedź, którą z hipotez (zerową czy alternatywną) przyjąć. Jeśli jego wartość jest wystarczająco duża, tj. większa lub równa wartości krytycznej, uwzględniając poziom istotności i liczbę danych, to przyjmuje się hipotezę alternatywną, mówiącą o dodatniej zależności między rangami, czyli że prawo MA zachodzi<sup>2</sup>.

Wartość współczynnika korelacji rangowej Spearmana oblicza się, stosując następujący wzór (1):

$$(1) \quad r = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)},$$

gdzie  $d_i$  to różnica między rangami, a  $n$  to liczba danych.

Obliczę dla przykładu wartość tego współczynnika  $r$  dla danych z tabeli 1:

$$\sum d_i^2 = 2, \text{ a więc } r = 0,9.$$

Wartość krytyczna dla tekstu jednostronnego na poziomie 5% dla 5 danych wynosi 0,9. Obliczona wartość współczynnika korelacji rangowej Spearmana dla tego tekstu wynosi 0,9. Przyjmujemy więc hipotezę alternatywną, czyli prawo MA zachodzi.

#### 4. Badane teksty

Badania przeprowadziłam na materiale następujących tekstów w wersjach polskiej i angielskiej: *Kod Leonarda da Vinci* Dana Browna, *Zagadnienia teorii składni* Noama Chomsky'ego, *Hobbit* J. R. R. Tolkiena, *Semantyka 2* Johna Lyonsa, *Zarys logiki matematycznej* Andrzeja Grzegorzcyka, *Zabić drozda* Harper Lee, *Demony dobrego Dextera* Jeffa Lindsaya. W tablicach 2–8 umieściłam wyniki badań empirycznych na powyższych tekstach wraz z przypisanymi zdaniom rangami. Oznaczenia zastosowano jak w tablicy 1.

Tabela 2. Dane na podstawie *Kodu Leonarda da Vinci* D. Browna

$x$	$y_{ang}$	$n_{ang}$	ranga	$y_{pol}$	$n_{pol}$	ranga
1	6,7584	745	1	5,8664	711	1
2	6,1677	465	3	5,4780	499	3
3	6,1976	275	2	5,5000	260	2
4	5,7442	86	4	4,9932	110	4
5	5,5928	28	5	4,8176	34	5
6	5,3627	17	–	4,8529	17	–
7	5,8571	5	–	5,7857	4	–

<sup>2</sup> Nie wystąpiły czynniki zakłócające zachodzenie prawa MA.

8	6,4583	3	–	5,8125	2	–
9	6,4444	2	–	6,2778	2	–
Suma		1626	–	–	1639	–

Wartość krytyczna dla testu jednostronnego na poziomie 5% dla 5 danych wynosi 0,9. Obliczona wartość współczynnika korelacji rangowej Spearmana dla obu wersji jest równa i wynosi 0,9. Widać zatem, że prawo MA zachodzi na obu tekstach.

W tekście naukowym *Zagadnienia teorii składni* N. Chomsky'ego założone 98% zdań złożonych włączonych do badań obejmuje także zdania sześcioskładnikowe. Bardziej złożona struktura tekstu jest charakterystyczna dla tego typu dyskursu, chociaż nie jest to konieczne, co zobaczymy na przykładach poniżej.

Tabela 3. Dane na podstawie *Zagadnień teorii składni* N. Chomsky'ego

<i>x</i>	<i>y<sub>ang</sub></i>	<i>n<sub>ang</sub></i>	ranga	<i>y<sub>pol</sub></i>	<i>n<sub>pol</sub></i>	ranga
1	14,6505	103	1	13,8121	165	1
2	12,5000	167	2	11,0925	173	2
3	10,3273	111	4	9,1667	100	3
4	10,1992	64	5	9,0991	53	4
5	10,4000	24	3	8,9667	18	5
6	8,2083	8	6	7,6250	8	6
7	8,6190	3	–	7,9429	5	–
8	9,1250	1	–	8,3750	2	–
9	9,8889	3	–	9,1111	1	–
10	8,7500	2	–	–	0	–
Suma		486	–	–	525	–

Wartość krytyczna dla testu jednostronnego na poziomie 5% dla 6 danych wynosi 0,7714. Obliczona wartość współczynnika korelacji rangowej Spearmana dla tekstu angielskiego wynosi 0,8286, a dla tekstu polskiego 1. Przyjmujemy hipotezę alternatywną, czyli prawo MA zachodzi na obu tekstach.

Jak już wspomniałam wyżej, również teksty beletrystyczne mogą zawierać zdania o większej liczbie zdań składowych niż teksty naukowe. W *Hobbicie* J. R. R. Tolkiena w zakres 98% zdań złożonych weszły zdania siedmioskładnikowe, a w samym tekście w wersji angielskiej występują zdania nawet trzynastokładnikowe.

Tabela 4. Dane na podstawie *Hobbita* J. R. R. Tolkiena

<i>x</i>	<i>y<sub>ang</sub></i>	<i>n<sub>ang</sub></i>	ranga	<i>y<sub>pol</sub></i>	<i>n<sub>pol</sub></i>	ranga
1	6,9929	282	1	6,1987	302	1
2	6,2154	311	2	5,7607	351	2
3	6,1474	251	3	5,2227	244	4

4	5,9562	137	4	5,0968	142	5
5	5,6205	88	7	5,0444	81	6
6	5,9133	50	6	5,3372	43	3
7	5,9388	35	5	4,6897	29	7
8	5,6083	15	–	4,8958	12	–
9	5,5556	9	–	5,2889	5	–
10	6,0000	2	–	5,4000	1	–
11		0	–	6,4091	2	–
12		0	–	5,2500	1	–
13	5,2308	2	–	–	0	–
Suma		1182	–	–	1213	–

Wartość krytyczna dla testu jednostronnego na poziomie 5% dla 7 danych wynosi 0,6786. Obliczona wartość współczynnika korelacji rangowej Spearmana dla tekstu angielskiego wynosi 0,8571, a dla tekstu polskiego 0,7857. Widzimy więc, że prawo MA zachodzi na obu tekstach.

W trzech powyżej omówionych tekstach obserwujemy podobną strukturę składniową w wersji oryginalnej i tłumaczeniu. Rozkłady złożoności zdań *Semantyki 2* J. Lyonsa w angielskim oryginale i w wersji polskiej wskazują na mniej wierne tłumaczenie pod względem zachowania struktury składniowej tekstu. W wersji angielskiej zdania są wielokrotnie złożone, natomiast w tłumaczeniu na język polski mamy zdania złożone o co najwyżej pięciu składnikach.

Tabela 5. Dane na podstawie *Semantyki 2* J. Lyonsa

x	$y_{ang}$	$n_{ang}$	ranga	$y_{pol}$	$n_{pol}$	ranga
1	16,3731	67	1	14,6395	172	1
2	11,6364	66	2	9,5400	100	2
3	10,6269	67	3	8,0714	42	3
4	9,9565	46	4	7,6500	15	4
5	8,8741	27	5	8,5429	7	–
6	7,7292	16	7	–	0	–
7	7,7449	14	6	–	0	–
8	7,5357	7	8	–	0	–
9	6,9556	5	–	–	0	–
10	10,0000	1	–	–	0	–
20	8,0500	1	–	–	0	–
Suma		317	–	–	336	–

Wartość krytyczna dla testu jednostronnego na poziomie 5% dla ośmiu danych wynosi 0,6429, a dla czterech 1. Obliczona wartość współczynnika korelacji ran-

gowej Spearmana dla tekstu angielskiego wynosi 0,9762, a dla tekstu polskiego 1. Przyjmujemy hipotezę alternatywną, a więc prawo MA zachodzi na obu tekstach.

W badanych tekstach naukowych wykluczyłam z analizy wszystkie wzory, liczby i stosowane formuły. Jeśli stanowiły one część zdania, były liczone jako jeden wyraz. Ze względu zapewne na trudność merytoryczną badanego fragmentu *Zarysu logiki matematycznej* A. Grzegorzcyka zdania złożone są co najwyżej ośmioskładnikowe, a do analizy weszły tylko te, które mają 5 lub mniej składników.

Tabela 6. Dane na podstawie *Zarysu logiki matematycznej* A. Grzegorzcyka

<i>x</i>	<i>y<sub>ang</sub></i>	<i>n<sub>ang</sub></i>	ranga	<i>y<sub>pol</sub></i>	<i>n<sub>pol</sub></i>	ranga
1	12,0686	160	1	10,0722	194	1
2	8,5351	114	2	7,5743	101	2
3	7,8490	64	3	7,4065	41	4
4	7,3125	32	4	7,4107	28	3
5	7,2000	20	5	7,0833	12	5
6	8,5417	4	–	7,1111	6	–
7	6,6667	3	–	5,0000	1	–
8	6,7500	2	–	–	0	–
Suma		399	–	–	383	–

Wartość krytyczna dla testu jednostronnego na poziomie 5% dla 5 danych wynosi 0,9. Obliczona wartość współczynnika korelacji rangowej Spearmana dla tekstu angielskiego wynosi 1, a dla tekstu polskiego 0,9. Widać zatem, że prawo MA zachodzi na obu tekstach.

Niżej omówione dwa teksty obu wersji zdają się kwestionować uniwersalność prawa MA. Postawiłam sobie za zadanie odpowiedzieć na pytanie, jakie czynniki mogą wpływać na zakłócenie jego działania na materiale tych tekstów.

Tabela 7. Dane na podstawie *Zabić drozda* H. Lee

<i>x</i>	<i>y<sub>ang</sub></i>	<i>n<sub>ang</sub></i>	ranga	<i>y<sub>pol</sub></i>	<i>n<sub>pol</sub></i>	ranga
1	5,7181	525	1	5,5487	698	1
2	4,5789	583	5	4,3236	683	6
3	4,2289	401	6	4,6414	382	5
4	4,8837	215	3	5,1871	175	3
5	4,8075	134	4	4,7630	92	4
6	4,9259	54	2	5,2197	44	2
7	5,1173	28	–	5,4714	20	–
8	5,4737	19	–	5,5694	9	–
9	5,2778	10	–	6,0370	3	–
10	4,9500	4	–	5,5750	4	–

11	–	0	–	4,7273	3	–
12	4,2500	1	–	–	0	–
13	3,4231	2	–	4,2307	1	–
16	4,0625	1	–	–	0	–
Suma		1977	–	–	2114	–

Wartość krytyczna dla testu jednostronnego na poziomie 5% dla 6 danych wynosi 0,7714. Obliczona wartość współczynnika korelacji rangowej Spearmana dla tekstu angielskiego wynosi -0,0286, a dla tekstu polskiego -0,0857. Przyjmujemy hipotezę zerową, czyli prawo MA nie zachodzi na obu tekstach.

Podobny wniosek możemy wyciągnąć, analizując dane zebrane na badanym odcinku *Demonów dobrego Dextera J. Lindsaya*.

Tabela 8. Dane na podstawie *Demonów dobrego Dextera J. Lindsaya*

<b>x</b>	<b><math>y_{ang}</math></b>	<b><math>n_{ang}</math></b>	<b>ranga</b>	<b><math>y_{pol}</math></b>	<b><math>n_{pol}</math></b>	<b>ranga</b>
1	5,6301	1030	1	4,8042	1231	1
2	4,8431	596	5	4,3962	549	4
3	5,1373	221	2	4,6563	193	2
4	5,0625	88	3	4,5184	68	3
5	4,9590	39	4	4,7292	28	–
6	5,1500	10	–	5,5278	6	–
7	4,9524	6	–	4,4857	5	–
8	5,1563	4	–	4,1250	2	–
9	–	0	–	4,5556	1	–
Suma		1994	–	–	2083	–

Wartość krytyczna dla testu jednostronnego na poziomie 5% dla pięciu danych wynosi 0,9, a dla czterech 1. Obliczona wartość współczynnika korelacji rangowej Spearmana dla tekstów angielskiego i polskiego wynosi 0,4. Widzimy więc, że prawo MA nie zachodzi na obu tekstach.

Porównam teraz wartości współczynników korelacji rangowej Spearmana dla pięciu<sup>3</sup> z badanych tekstów w obu wersjach. W tym celu wyliczyłam wartości współczynnika dla odcinka tekstu powiększonego o kolejny rozdział badanej książki. W tabeli 9,  $n$  jest liczną zdań składowych analizowanych 98% zdań złożonych danego tekstu.

<sup>3</sup> Dwa pierwsze teksty, *Kod Leonarda da Vinci* oraz *Zagadnienia teorii składni*, były częścią badań wstępnych i dane szczegółowe nie zachowały się.



Tabela 9. Wartości współczynnika korelacji rangowej Spearmana

Książka	rozdziały	wersja polska	n danych	wersja angielska	n danych
<i>Hobbit</i>	1	0,6786	7	0,6786	7
	1–2	0,3214	7	0,6071	7
	1–3	0,4643	7	0,7500	7
	1–4	0,7857	7	0,8571	7
	1–5	0,7857	7	0,7143	7
<i>Semantyka 2</i>	10.1	0,8	4	0,8571	7
	10.1–10.2	0,8	4	0,9833	9
	10.1–10.3	1	4	0,9762	8
	10.1–10.4	1	4	1	7
<i>Zarys logiki matematycznej</i>	1.0–1.1	0	5	0,3	5
	1.0–1.2	0,1	5	-0,4	6
	1.0–1.3	0,7	5	1	5
	1.0–1.4	0,9	5	1	5
	1.0–1.5	0,9	5	0,9	5
<i>Zabić drozda</i>	1	0,2857	6	0,5429	6
	1–2	0,1429	6	0,5714	6
	1–3	0,0286	6	0,3714	6
	1–4	-0,0857	6	0,1429	6
	1–5	0,0286	6	0,0857	6
	1–6	0,0286	6	0,0857	6
	1–7	0,0286	6	0,0857	6
	1–8	-0,0857	6	-0,0857	6
<i>Demony dobrego Dextera</i>	1	0,1	5	0,6	5
	1–2	0,3	5	0,6	5
	1–3	0,4	4	0,6	5
	1–4	0,4	4	0,6	5
	1–5	0,4	4	0,7	5
	1–6	0,4	4	0,7	5
	1–7	0,4	4	0,4	5

W tekstach, dla których prawo MA zachodzi, wartości współczynnika korelacji rangowej Spearmana sukcesywnie rosły, aż osiągnęły wartość większą od wartości krytycznej. W *Hobbicie* 98% badanych zdań obu wersji to zdania siedmiokrotnie złożone. Wartość krytyczna współczynnika dla siedmiu danych wynosi 0,6786. Dla polskiej wersji książki obserwujemy, że prawo MA zachodzi dla pierwszego rozdziału oraz zbioru rozdziałów 1–4 i 1–5. Dla angielskiego oryginału mamy również

wartość współczynnika powyżej wartości krytycznej dla rozdziału 1 i ponownie prawo zaczyna zachodzić po dołączeniu rozdziału 3 i następnych.

Przyjrzymy się teraz wartościom współczynnika obliczonych dla tekstu *Semantyka 2*. Wartości krytyczne dla współczynnika są mniejsze od 0,6786 dla liczby 7, 8 i 9 danych w wersji angielskiej i dla wszystkich rozdziałów wartości tego współczynnika są większe od tej największej wartości krytycznej. Jest to jedyna wersja, dla której obserwuje się, że prawo zachodzi już od rozdziału 1 i nie zmienia się to po dodaniu następnych fragmentów książki. Dla wersji polskiej 98% zdań stanowią zdania co najwyżej czterokrotnie złożone i prawo zachodzi od zbioru pierwszych trzech podrozdziałów.

Jak już wiemy, wartość krytyczna współczynnika dla 4 danych równa się 1. Natomiast dla 5 i 6 danych wartości krytyczne współczynnika to odpowiednio 0,9 i 0,7714. Taką maksymalną liczbę zdań składowych mają zanalizowane zdania złożone z *Zagadnień logiki matematycznej*. Dla wersji polskiej długość tekstu musiała przekroczyć trzy rozdziały, by można było obserwować zachodzenie prawa MA. Natomiast dla tłumaczenia angielskiego prawo zachodziło już dla pierwszych rozdziałów.

Dla powieści *Zabić drozda* wartości te mają tendencję do zmniejszania się. Być może, wskazuje ona na to, że istnieją warunki zakłócające zachodzenie prawa MA na danym tekście. Pomimo wydłużania badanego tekstu wartość współczynnika korelacji rangowej Spearmana nigdy nie przekroczyła wartości krytycznej, która dla 6 danych równa się 0,7714. Obserwujemy nawet, że jego wartości są coraz mniejsze. Podobną sytuację można zauważyć dla danych zebranych na tekście *Demony dobrego Dextera*. Dla 4 i 5 danych wartości krytyczne współczynnika wynoszą odpowiednio 1 i 0,9. Dla wersji polskiej wartość współczynnika rosła, aż osiągnęła wartość 0,4, natomiast dla wersji angielskiej obliczona wartość dla pierwszych czterech i mniejszej liczby rozdziałów równa się 0,6. Później wartość ta wzrosła do 0,7 po dołączeniu rozdziału 5, a następnie 6, by spaść do 0,4 po wydłużeniu badanego tekstu o kolejny rozdział.

## 5. Wnioski

Czym różnią się od pozostałych te dwa teksty, dla których nastąpiło zakłócenie działania prawa MA?

Przed wszystkim obserwujemy w tabeli 9, że prawo MA działa, gdy dany tekst osiągnie minimalną<sup>4</sup> długość, która może być różna dla różnych tekstów, ale podobna dla zbliżonych strukturalnie wersji językowych. Na przykładzie tekstu *Semantyka 2* widzimy, że wersje różniące się budową składniową mają różne minimalne długości tekstu potrzebne do tego, by prawo MA miało odpowiednie warunki, aby zachodzić. Dla pozostałych dwóch tekstów, *Hobbita* i *Zarysu logiki matematycznej*, te minimalne długości są zbliżone. W wypadku tekstów o zakłóconym działaniu prawa MA, liczba 2000 zdań okazała się niewystarczająca. Być może nawet tak, że prawo

<sup>4</sup> Mówiąc o minimalnych długościach, mam na myśli liczbę rozdziałów.

to nigdy nie zajdzie na całości tych tekstów, dostrzegłam bowiem tendencję zmniejszania się wartości współczynnika korelacji rangowej Spearmana lub jego ustalenia się na poziomie mniejszym od wartości krytycznej. Jeśli prawo MA nie zachodzi na danym tekście, to jakie mogą być tego przyczyny?

W tabeli 10 umieściłam średnie długości zdań pojedynczych dla ostatnich badanych rozdziałów. Zdania pojedyncze dla wszystkich tekstów miały rangę 1, czyli ich średnie długości były największe.

Tabela 10. Średnie długości zdań pojedynczych

książka	wersja polska	wersja angielska
<i>Kod Leonarda da Vinci</i>	5,8664	6,7584
<i>Zagadnienia teorii składni</i>	13,8121	14,6505
<i>Hobbit</i>	5,8004	6,6142
<i>Semantyka 2</i>	14,8514	15,8654
<i>Zarys logiki matematycznej</i>	10,0218	11,8708
<i>Zabić drozda</i>	5,5487	5,7181
<i>Demony dobrego Dextera</i>	4,8042	5,6301

Możemy zauważyć, że średnia minimalna długość zdania pojedynczego dla tekstów polskich, dla których prawo MA zachodzi, to 5,8004. Liczba ta jest większa od średniej maksymalnej długości zdania pojedynczego dla tekstów, dla których prawo MA zdaje się nie zachodzić, czyli liczby 5,5487. Dla tekstów angielskich średnia minimalna długość zdań, dla których prawo zachodzi, to 6,6142, a średnia maksymalna długość zdań pojedynczych, dla których prawo nie zachodzi, wynosi 5,7181. Być może, istnieje graniczna wartość średniej długości zdania pojedynczego, dla których zachodzenie prawa MA jest niezakłócone. Nie przeczy to konieczności zachodzenia prawa MA, ponieważ, jak to wykazał R. Köhler (1984), istotne jest obciążenie pamięci krótkotrwałej. Jeśli zdania są wystarczająco krótkie, założona hierarchia średniej długości zdań złożonych może zostać zakłócona (wniosek własny).

Najkrótsze średnie długości zdań pojedynczych spotyka się w powieściach, jeśli autor wykorzystuje dużą liczbę dialogów, które nie występują w tekstach naukowych. To właśnie dialogi charakteryzują się krótkimi zdaniami pojedynczymi i złożonymi. Warunek minimalnej średniej długości zdania pojedynczego może być przyczyną, dla której prawo MA może nie działać w tekstach mówionych. Oczywiście wtedy występuje też inny problem, o którym piszą Udo Strauss, Gabriel Altmann, K.-H. Best (2007): mianowicie fragmentacja tekstu i określenie granic zdań są wykonalne tylko w tekstach pisanych.

Podsumowując: istnieją więc przynajmniej dwa warunki działania prawa MA. Jeden z nich podaje w artykule *Badania nad prawem Menzeratha-Altmana* (Kulacka 2008): prawo działa tylko w zamkniętych całościach, np. rozdziałach książki. Drugim jest opisany w tym artykule warunek dotyczący minimalnej długości zdania złożonego (zdanie pojedyncze traktuję jako zdanie złożone z jednego zdania składowego), by prawo MA zachodziło na danym tekście.

**Literatura**

- ALTMANN G., 1980, *Prolegomena to Menzerath's Law*, „Glottometrika” 2, Bochum: Studienverlag Dr. N. Brockmeyer, s. 1–10.
- BAUDOIN DE COURTENAY J., 1927, *Ilościowość w myśleniu językowym*, [w:] J. N. Baudouin de Courtenay, *Dzieła wybrane*, t. IV, Warszawa, s. 546–563.
- DIXON R. M. W., 2005, *A Semantic Approach to English Grammar*, Oksford.
- GRZEGORCZYKOWA R., 2004, *Wykłady z polskiej składni*, Warszawa.
- HAMMERL R., 1987, *Prawa językowe we współczesnej kwantytatywnej lingwistyce modelowej (Na przykładzie tzw. Prawa Martina)*, „Poradnik Językowy” nr 6, s. 414–428.
- HAMMERL R., 1989, *Cztery etapy rozwoju lingwistyki kwantytatywnej*, [w:] W. Lubaś (red.), *Wokół współczesnego języka polskiego*, t. II: *Studia leksykograficzne*, cz. 3, Wrocław, 115–126.
- HAMMERL R., SAMBOR J., 1993, *O statystycznych prawach językowych*, Warszawa.
- KÖHLER R., 1984, *Zur Interpretation Des Menzerathschen Gesetzes*, [w:] J. Boy, R. Köhler, *Glottometrika* 6, s. 177–183.
- KÖHLER R., 1995, *Bibliography of Quantitative Linguistics*, Amsterdam.
- KUŁACKA A., 2008, *Badania nad prawem Menzeratha–Altmanna*, „LingVaria” nr 2(6), s. 167–174.
- KUŁACKA A., MAĆUTEK J., 2006, *A Discrete Formula for the Menzerath–Altmann Law*, „Journal of Quantitative Linguistics” vol. 13, nr 2–3, s. 217–226.
- PAWŁOWSKI A., 2006, *Bibliografia polskiej lingwistyki kwantytatywnej*. <http://www.lingwistyka.uni.wroc.pl/bql/> [dostęp 27 VII 2007].
- SAMBOR J., PAWŁOWSKI A., 2005, *Quantitative Linguistics in Poland*, [w:] R. Köhler, G. Altmann, R. G. Piotrowski, *Quantitative Linguistics. An International Handbook*, Berlin, s. 115–129.
- STRAUSS U., ALTMANN G., BEST K.-H., 2007, *Sentence and Clause Length*, [http://www.uni-trier.de/uni/fb2/ldv/lql\\_wiki/index.php/Sentence\\_and\\_clause\\_length](http://www.uni-trier.de/uni/fb2/ldv/lql_wiki/index.php/Sentence_and_clause_length) [dostęp 26 VII 2007].

### Necessary conditions of the Menzerath–Altmann Law Summary

In the paper the author discusses conditions necessary for the Menzerath–Altmann law to hold. The research was conducted on Polish and English syntax and a suitable comparison between texts under the research has been made. In the opening of the article, the statistical language law is presented and the technique of its research has been outlined. They are followed by all necessary assumptions deciding whether the law would hold and the conditions for it to hold.