

8/1984

Lutosiawa Richter

ANALIZA STATYSTYCZNA  
RYTMICZNEJ STRUKTURY WYPOWIEDZI  
W MOWIE POLSKIEJ

8/1984

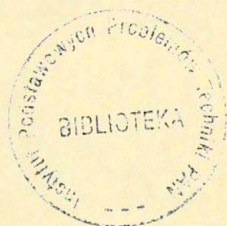
p. 269



WARSZAWA 1984

ISSN 0208-5658

Praca wpłynęła do Redakcji dnia 14 listopada 1983r.



56956



N a p r a w a c h   r ę k o p i s u

---

Instytut Podstawowych Problemów Techniki PAN  
Nakład 140 egz. Ark.wyd. 1,7. Ark. druk. 2,5  
Oddano do drukarni w lutym 1984 r.  
Nr zamówienia 149/84.

---

Warszawska Drukarnia Naukowa. Warszawa,  
ul. Sniadeckich 8



Lutosława Richter  
Pracownia Fonetyki Akustycznej  
IPPT PAN

ANALIZA STATYSTYCZNA RYTMICZNEJ STRUKTURY WYPOWIEDZI W MOWIE  
POLSKIEJ<sup>1/</sup>

Streszczenie

Możliwość kontrolowania struktury rytmicznej wypowiedzi w mowie syntetycznej wymaga przeprowadzenia szczegółowej analizy na materiale mowy naturalnej. Przejawem rytmu jest zjawisko izochronizmu polegające na skracaniu segmentów fonetycznych w dłuższych jednostkach rytmicznych. Jedno z możliwych matematycznych sformułowań tego zjawiska daje model regresyjny. Zastosowano model z dwiema lub trzema zmiennymi dla określenia siły izochronizmu zestrojowego w języku polskim. Poddano również analizie zjawisko rytmu na poziomie frazowym weryfikując hipotezę dotyczącą izochronizmu frazowego.

1. Wstęp.

Uzyskanie wysokiej jakości mowy syntetycznej wymaga uwzględnienia cech suprasegmentalnych. Zalicza się do nich iloczyn głoskowy, wyrażający się określoną rozciągłością czasową segmentów fonetycznych. Jednym z czynników, które decydują o czasie trwania głosek, jest rytm. W fonetyce przyjęł się podział języków ze względu na rytm na trzy grupy: z rytmem akcentowym, sylabicznym oraz morowym. W językach zaliczanych do pierwszej grupy (język angielski, niemiecki, rosyjski) obserwuje się tendencję, określaną mianem izochronizmu, do zachowywania w przybliżeniu stałej długości jednostek rytmicznych wyznaczanych przez akcent wyrazowy, którymi są stopy lub zestroje akcentowe.

<sup>1/</sup> Praca wykonana w ramach problemu węzłowego 06.9.

Zbliżone w sensie czasowym odległości pomiędzy akcentami wymagają skracania jednostek fonetycznych niższego rzędu (syłab, głosek, segmentów fonetycznych) w dłuższych jednostkach rytmicznych.

Wyniki zamieszczone w pracy autorki [11] upoważniają do stwierdzenia, że tego rodzaju tendencja występuje również w języku polskim. Zaproponowano interpretację matematyczną, w której model izochronizmu został wyrażony za pomocą funkcji potęgowej. W modelu tym zmienną niezależną stanowiła liczba sylab w zestroju akcentowym. Teoretyczny czas trwania samogłosek wyznaczono w oparciu o zależności :

$$V = \frac{D}{m^{\alpha} n^{\beta}} ,$$

gdzie D oznacza maksymalny czas trwania samogłoski, m - liczbę sylab pozostałych do wyprodukowania licząc od danej sylaby do końca zestroju, n - całkowitą liczbę sylab w zestroju.

Odmianą interpretację matematyczną zjawiska izochronizmu zaproponowali Jassem, Krzyśko, Stolarski [6] w pracy poświęconej rytmowi języka angielskiego. Posłużono się tu modelem, który wyrażał zależność regresyjną pomiędzy iloczasetem jednostki rytmicznej a liczbą głosek albo sumą średnich iloczasetów głosek składowych.

W literaturze fonetycznej [1], [8], [9] spotyka się pogląd, że zjawisko izochronizmu funkcjonuje nie tylko na poziomie wyrazowym, lecz również na poziomie frazowym, tzn. iloczas głosek uwarunkowany jest liczbą jednostek rytmicznych w obrębie frazy fonetycznej.

Praca niniejsza ma udzielić odpowiedzi na pytanie, czy zjawiska rytmiczne zachodzące w mowie polskiej można zinterpretować za pomocą modelu regresyjnego. Celem pracy jest również dokonanie analizy struktury czasowej wypowiedzi na poziomie frazy.

## 2. Struktura rytmiczna wypowiedzi na poziomie wyrazowym.

### 2.1. Model regresyjny.

Dla modelu regresyjnego zaproponowanego przez Jassemę, Krzyśkę, Stolarskiego [6] punkt wyjścia stanowiła interpretacja



izochronizmu za pomocą prostej funkcji  $d = np$ , gdzie  $d$  oznacza iloczyn jednostki rytmicznej,  $n$  - liczbę segmentów w jednostce,  $p$  - iloczyn segmentu. W rzeczywistości, na skutek oddziaływania różnych czynników, zależności te są bardziej skomplikowane i dają się wyrazić za pomocą modelu regresyjnego, w którym siłę izochronizmu określa współczynnik regresji  $b$  z równania regresji  $d = a + bn$ . Dla regresji liniowej dwuzmiennej przy odpowiednim wyskalowaniu osi rzędnych i odciętych linia regresji przechodzi przez środek układu, co pozwala łatwo zinterpretować nasilenie zjawiska. Jeśli linia regresji tworzy kąt  $45^\circ$  z osią odciętych, oznacza to brak izochronizmu, jeśli tworzy z osią odciętych kąt  $0^\circ$ , oznacza to ścisły izochronizm. Kąt nachylenia linii regresji określa wielkość  $\text{arc tg } b$ .

Wyznaczono wartość współczynnika  $b$  w oparciu o cztery modele regresyjne :

- 1/ regresja liniowa, model z dwiema zmiennymi :  $d, n$
- 2/ regresja kwadratowa, model z dwiema zmiennymi :  $d, n$
- 3/ regresja liniowa, model z dwiema zmiennymi :  $d$  i  $\tilde{d}$   
( $\tilde{d}$  oznacza skumulowany iloczyn średnich długości głosek wchodzących w skład jednostki rytmicznej)
- 4/ regresja liniowa, model z trzema zmiennymi :  $d, \tilde{d}, n$ .

### 2.3. Materiał doświadczalny.

Trzy osoby odczytały z naturalną prędkością mówienia fragment tekstu prasowego trwającego ok. 3 min. Fragment ten obejmował ponad 2100 głosek. Z materiału utrwalonego na taśmie magnetofonowej sporządzono oscylogramy przy użyciu oscylografu pętlicowego. W trakcie analizy stosowano szybkość przesuwu papieru 125 cm/s. Następnie przeprowadzono segmentację uzyskanego zapisu oscylograficznego w oparciu o zasady opisane przez autorkę [10], po czym dokonano pomiarów czasu trwania każdej głoski z dokładnością 5 ms. W niektórych przypadkach niemożliwe było przeprowadzenie segmentacji dla sąsiadujących ze sobą głosek, które charakteryzowały się takim samym rodzajem przebiegu akustycznego. Tego typu zbitki potraktowano jako jeden element przy wyznaczaniu długości jednostki rytmicznej, uwzględniając jednak ich dwusegmentową, w sporadycznych przypadkach trzusegmentową strukturę przy rozpatrywaniu liczebności



głosek w ramach jednostki rytmicznej. Ogólna liczba segmentów stanowiących podstawę do wyliczenia regresji wyniosła 2146 dla głosu PD, 2140 dla głosu ES oraz 2173 dla głosu JI.

Przed przystąpieniem do podziału tekstu na jednostki rytmiczne przeprowadzono testy odsłuchowe z udziałem trzech wyszkolonych fonetyków. Zadaniem ich było zaznaczenie sylab akcentowanych oraz granic zestrojów akcentowych, według definicji podanej przez Wierzchowską [12]. W nielicznych przypadkach, gdy postawienie granic nastroczało zbyt wiele wątpliwości, opuszczono odnośny fragment tekstu w analizie statystycznej.

Następnie przeprowadzono podział tekstu na jednostki rytmiczne będące stopami, a więc obejmujące odcinki od jednej sylaby akcentowanej do następnej sylaby akcentowanej lub końca frazy. Podwójny podział na jednostki rytmiczne wprowadzono w tym celu, by stwierdzić, czy dokładność opisu zjawisk rytmicznych w języku polskim, jaką daje model regresyjny, wykazuje powiązanie z rodzajem przyjętej w opisie jednostki rytmicznej w postaci bądź stopy, bądź zestroju akcentowego. W obrębie zestroju wydzielono jeszcze mniejsze jednostki - część przedakcentową, która nie zawsze jest obecna w zestroju, oraz część główną, liczoną od sylaby akcentowanej do końca zestroju, która jest jego integralną częścią. Wprowadzenie osobnych modeli dla części przedakcentowej i głównej zestroju było o tyle uzasadnione, że we wcześniejszej pracy autorki stwierdzono brak izochronizmu w sylabach przedakcentowych [11].

Określenie granic jednostek rytmicznych napotykało na przeszkody, jeśli na pograniczu dwóch jednostek znalazła się nierozsegmentowana zbitka głoskowa. W tych nielicznych przypadkach określony odcinek tekstu nie był brany pod uwagę przy obliczaniu regresji.

Dla każdego z przyjętych równań regresji zostały obliczone cztery modele odnoszące się do określonej jednostki rytmicznej: stopy, zestroju akcentowego, części przedakcentowej zestroju i części głównej zestroju.

#### 2.4. Klasy głosek jednorodnych statystycznie pod względem iloczynowym.

Czynnikiem systematycznym w języku, który nakłada ograniczenia na realizację izochronizmu, jest iloczyn właściwy głosek.



Poczyniono robocze założenie, że wyznaczenie dla materiału doświadczalnego klas głosek różniących się iloczasetem może zostać przeprowadzone niezależnie od uwarunkowań rytmicznych. W oparciu o kryteria fonetyczne ustalono maksymalną liczbę typów głoskowych, przy czym do jednego typu mogą należeć warianty dwóch różnych fonemów, ale charakteryzujące się zbliżoną średnią wartością czasu trwania, z drugiej strony warianty tego samego fonemu mogą należeć do dwóch różnych typów, o ile wstępne badania wykazały, że różnią się one średnim czasem trwania. Na tej podstawie wprowadzono 27 następujących typów głosek :

- 1 IA - samogłoska [i] akcentowana
- 2 IN - samogłoska [i] nieakcentowana
- 3 YA - samogłoska [ɨ] akcentowana
- 4 YN - samogłoska [ɨ] nieakcentowana
- 5 EA - samogłoska [e] akcentowana
- 6 EN - samogłoska [e] nieakcentowana
- 7 AA - samogłoska [a] akcentowana
- 8 AN - samogłoska [a] nieakcentowana
- 9 OA - samogłoska [o] akcentowana
- 10 ON - samogłoska [o] nieakcentowana
- 11 UA - samogłoska [u] akcentowana
- 12 UN - samogłoska [u] nieakcentowana
- 13 Y - samogłoska [ə]
- 14 P - zwarte bezdźwięczne [p, t, c, k] w pozycji śródgłosowej i wygłosowej
- 15 B - zwarte dźwięczne [b, d, ʃ, g]
- 16 S - trące bezdźwięczne [f, s, ʃ, ɸ, x]
- 17 Z - trące dźwięczne [v, z, ʒ, ʒ̣, ʒ̥]
- 18 T - zwarto-trące bezdźwięczne [t̪, t̪ʰ, t̪ʰ] w pozycji śródgłosowej i wygłosowej
- 19 D - zwarto-trące dźwięczne [d̪, d̪̃, d̪̃]
- 20 N - spółgłoski nosowe [m, n, ɲ, ŋ]
- 21 L - płynne [l, w], drżące [r], [j] nie poprzedzone spółgłoską wargową
- 22 J - [j] następujące po spółgłosce wargowej, spółgłoski zwarto-trące bezdźwięczne inicjalne (tzn. segment szumowy tych spółgłosek)



- 23 R - [r] uderzeniowe, spółgłoski zwarte inicjalne (tzn. segment pozycji tych spółgłosek)
- 24 NA - dyftongi [e<sup>~</sup>w, o<sup>~</sup>w] akcentowane w śródgłosie
- 25 NN - dyftongi [e<sup>~</sup>w, o<sup>~</sup>w] nieakcentowane w śródgłosie
- 26 V - samogłoski i spółgłoski w wygłosie przed pauzą
- 27 W - dyftongi [e<sup>~</sup>w, o<sup>~</sup>w] w wygłosie przed pauzą

W tablicy 1 zamieszczono średnie wartości czasu trwania dla wydzielonych typów głosek. Średnie zostały obliczone dla każdego głosu oddzielnie, następnie uszeregowane od wartości najniższych do najwyższych. Uzyskana w ten sposób kolejność nie jest jednako-  
kowa dla wszystkich głosek. Najmniej rozbieżności wystąpiło dla najniższych oraz najwyższych wartości iloczasu. Zróżnicowania wartości spowodowane są również niejednakowym dla poszczególnych osób tempem mówienia.

Tak samo we wszystkich głosach i zgodnie z regułami fonetycznymi układają się stosunki pomiędzy określonymi grupami głosek. Spółgłoski dźwięczne są krótsze od odpowiadających bezdźwięcznych, samogłoski nieakcentowane krótsze od odpowiadających akcentowa-  
nych.

Zwraca uwagę konsekwentnie utrzymywana w całym materiale pozycja samogłoski /ɨ/, która wykazuje najkrótszy czas trwania spośród wszystkich samogłosek. Przeczy to regule, według której iloczas samogłosek uzależniony jest od stopnia wzniesienia języka. W sylabach nieakcentowanych krótki czas trwania /ɨ/ może być związany z redukcją tej samogłoski w niektórych pozycjach. Najkrótszy czas trwania /ɨ/ również w sylabach akcentowanych przemawia za uznaniem za czynnik decydujący bardziej centralnego położenia języka w jamie ustnej w porównaniu z samogłoską /i/.

Różnice pomiędzy wartościami dla niektórych przynajmniej typów głosek okazały się istotne statystycznie na poziomie  $\alpha = 0.05$ , co wykazała analiza wariancji. Uzyskano następujące wartości F :

		Wartość F krytyczna
PD	94,936	
ES	73,014	1,501
JI	75,233	

Następnym krokiem było zbadanie, pomiędzy którymi średnimi



nie zachodzą różnice istotne, co pozwoli połączyć typy głosek w grupy jednorodnie statystycznie. Zastosowano test Gabriela [3], w którym decyzje dotyczące istotnego podziału na grupy oparte są na sumie kwadratów odchyleń pomiędzy średnimi przy zastosowaniu takich samych krytycznych wartości, jak w teście F. W wyniku uzyskano następujące skupienie jednorodnie statystycznie.

Głos PD

R, Y, J---L---YN, EN, IN, N, UN, ON, YA---IA, Z, B, EA, AN, UA, P, OA---D, S---  
---AA, T, ---NN, V---NA---W

Głos ES

R, Y---J, L---YN, IN, UN, ON, Z, N, YA, EN---IA, B, AN, EA, UA---OA, S, P, D---  
AA, T, V, NN---NA, W

Głos JI

R---J, Y, L---N, YN, B, UN, EN, IN, ON, Z, YA---AN, P, UA, D, EA-IA---OA, S---  
T, AA, NN---V, NA, W

Podział na grupy przebiega w niejednakowy sposób w badanych głosach. Wyraźna zgodność dotyczy jedynie wzajemnego stosunku samogłosek akcentowanych i nieakcentowanych. We wszystkich głosach samogłoski akcentowane z wyjątkiem /a/ znalazły się w innej grupie, aniżeli odpowiadające im samogłoski nieakcentowane. Przemawia to za uznaniem pozycji samogłoski w zestroju akcentowym za jeden z decydujących czynników wywierających wpływ na jej czas trwania.

Ponieważ średnie dla klas głosek stanowić będą dane wyjściowe w analizie regresji, znaczne zróżnicowania dotyczące podziału na grupy nie pozwalają traktować łącznie trzech głosów. Testowanie modeli regresyjnych będzie więc przeprowadzane oddzielnie dla każdego głosu.

#### 2.5. Dane wyjściowe do analizy regresji.

Tablice 2,3,4 zawierają podstawowe parametry statystyczne odnoszące się do klas głosek uzyskanych w wyniku grupowania średnich. Symbole literowe umieszczone w nawiasach mówią, jakie typy głoskowe wchodzi w skład danej klasy. Przeprowadzono test Studenta celem stwierdzenia, czy różnice pomiędzy średnimi ogólnymi dla różnych głosów są istotne. Wszystkie wartości  $t$  okazały się istotne na poziomie  $\alpha = 0,01$  lub  $\alpha = 0,001$ . Dla głosów PD - ES  $t = 7,709^{***}$ , dla głosów ES - JI  $t = 3,263^{**}$ , dla JI - PD  $t = 10,700^{***}$ . Jest to dodatkowy argument



przemawiający za tym, że analizę regresji należy przeprowadzać oddzielnie dla każdego głosu.

Następnie wartości średnie z tablic 2,3,4 rozbito na poszczególne typy jednostek rytmicznych. Rezultaty zamieszczono w tablicy 5. Będą one stanowiły dane wyjściowe przy estymowaniu długości jednostki rytmicznej z sumarycznego iloczasu głosek składowych.

## 2.6. Analiza regresji.

### 2.6.1. Model z dwiema zmiennymi.

#### 2.6.1.1. Model I, zmienne : d, n

Analizę regresji przeprowadzono w oparciu o równanie liniowe

$$\frac{d - \bar{d}}{\bar{p}} = a + b(n - \bar{n})$$

Zmienna niezależna  $n$  (liczba głosek wchodzących w skład jednostki rytmicznej) występuje tu w postaci znormalizowanej :  $n - \bar{n}$ , zaś zmienna zależna  $d$  (iloczas jednostki rytmicznej) w postaci znormalizowanej  $\frac{d - \bar{d}}{\bar{p}}$  ( $\bar{p}$  oznacza średni iloczas głoskowy dla określonego typu jednostki rytmicznej). Dzięki takiemu wyskalowaniu osi rzędnych i odciętych stała wartość  $a$  jest równa 0, a linia regresji przechodzi przez środek układu i tworzy kąt  $45^\circ$  z osią odciętych dla braku izochronizmu, a  $0^\circ$  dla ścisłego izochronizmu.

Silę izochronizmu określa współczynnik regresji  $b$  i związany z nim  $\text{arc tg } b$ .

Wyniki analizy zamieszczone w tablicach 7,8,9 prowadzą do następujących wniosków :

1/ Najwyższe wartości  $\text{arc tg } b$  zbliżone do  $45^\circ$  występują we wszystkich głosach dla części przedakcentowej zestrojów i wskazują na słabą tendencję do izochronizmu.

2/ Najniższe wartości  $b$  występują dla części głównej zestroju.  $\text{Arc tg } b$  wskazuje na dość silną tendencję do izochronizmu. Wartości  $b$  dla części głównej zestroju są w głosie PD : 0,68, w głosie ES : 0,73, w głosie JI : 0,74 raza mniejsze od  $b$  dla części przedakcentowej.

3/ Wyniki analizy regresji potwierdzają wcześniejsze spostrzeżenia [11], że zjawisko izochronizmu w języku polskim występuje przede wszystkim w części zestroju akcentowego



obejmującej sylabę akcentowaną i następujące sylaby poakcentowe.

4/ Współczynnik regresji dla całego zestroju akcentowego przyjmuje wartości pośrednie pomiędzy  $b$  dla części przedakcentowej i  $b$  dla części głównej zestroju. Arc  $\text{tg } b$  dla trzech głosów wynosi średnio  $37^\circ$ . Tak więc zestrój akcentowy traktowany całościowo wykazuje również dość wyraźną tendencję do izochronizmu.

5/ Nieco niższe wartości  $b$  wystąpiły dla stopy. Wynika to prawdopodobnie z faktu, że jeśli po pauzie znalazły się sylaby nieakcentowane, nie były zaliczane do żadnej stopy jako tzw. "niemy akcent" natomiast wchodziły w skład zestroju jako jego część przedakcentowa, co wywarło wpływ na wielkość współczynnika  $b$ .

6/ Wartości współczynnika determinacji wykazują zadowalającą siłę predykcji, przy czym w największym stopniu wyjaśniają zmienność  $d$  względem  $n$  w części przedakcentowej zestroju. Niższe wartości  $r$  dla głównej części zestroju w porównaniu z częścią przedakcentową (w każdym z głosów stosunek  $r$  w części głównej do  $r$  w części przedakcentowej wynosi  $0,7$ ) pozwalają przypuszczać, że w części głównej zestroju oddziałują na iloczas głoski jakieś dodatkowe czynniki.

7/ Zgodność poszczególnych wartości  $r$  oraz  $b$  jest bardzo wysoka dla wszystkich głosów.

8/ Ze względu na strukturę rytmiczną odpowiednikiem części przedakcentowej zestroju w języku polskim jest anakruza w języku angielskim, zaś części głównej zestroju - ścisła jednostka rytmiczna. Porównanie wyników analizy regresji liniowej  $d$  względem  $n$  przeprowadzonej dla języka angielskiego i języka polskiego nasuwa wniosek następujący: O ile można mówić o jednakowym nasileniu tendencji do izochronizmu w części przedakcentowej w obu językach (dla języka polskiego średnia z trzech głosów wartość  $\text{arc tg } b$  wynosi  $42,7^\circ$ , dla języka angielskiego  $\text{arc tg } b$  w anakruzie wynosi  $42,5^\circ$ ), o tyle w pozostałej części zestroju język angielski wykazuje silniejszą tendencję do izochronizmu ( $\text{arc tg } b$  dla ścisłej jednostki rytmicznej w języku angielskim wynosi  $29,3^\circ$ ,  $\text{arc tg } b$  dla głównej części zestroju w języku polskim wynosi średnio  $33,5^\circ$ ). Różnica ta ulega zatarciu przy porównaniu stóp ( $\text{arc tg } b$  dla języka angielskiego =  $32,8^\circ$ , dla języka polskiego =  $34,3^\circ$ ).



Nasilenie tendencji do izochronizmu przedstawiono w formie graficznej na ryc. 1, gdzie siłę izochronizmu określa kąt nachylenia prostej regresji względem osi odciętych.

2.6.1.2. Model II, zmienne :  $d, \tilde{d}$ .

Przeprowadzono analizę regresji  $d$  względem  $\tilde{d}$ , tj. skumulowanego iloczasu średnich długości głosek wchodzących w skład jednostki rytmicznej. Wartość  $\tilde{d}$  uzyskiwano sumując średnie dla określonych klas głosek z tablicy 6. W przypadku braku izochronizmu  $d$  powinno być równe  $\tilde{d}$ . Zależność pomiędzy rzeczywistą a estymowaną długością jednostki rytmicznej wyrażono za pomocą liniowego równania regresji, w którym zmienna niezależna  $d$  oraz zmienna zależna  $\tilde{d}$  występują w postaci znormalizowanej.

$$\frac{d-\bar{d}}{\bar{p}} = a + b \frac{\tilde{d}-\bar{\tilde{d}}}{\bar{p}}$$

Dla braku izochronizmu arc  $\operatorname{tg} b$  wynosi  $45^\circ$ .

Wyniki analizy regresji zamieszczono w tablicach 10, 11, 12. Prowadzą one do następujących wniosków :

1/ Współczynnik  $r$  jest najwyższy dla części przedakcentowej i równy wartości  $r$  w modelu I. Również dla części przedakcentowej najwyższy jest arc  $\operatorname{tg} b$  przybierając wartości podobne, jak w modelu I. Wyniki analizy regresji odnoszące się do części przedakcentowej zestroju są więc takie same bez względu na zastosowany model.

2/ Znacznemu podwyższeniu uległy wartości  $r$  dla części głównej zestroju, zaś w mniejszym dla zestroju i stopy. Współczynniki determinacji mówią, że znacznie mniejsza część wariancji pozostaje niewyjaśniona w przypadku szacowania długości jednostki rytmicznej z sumarycznego iloczasu głosek składowych niż z liczby głosek, a zatem większa jest siła predykcji modelu II.

3/ Współczynnik  $b$  i odpowiadający mu arc  $\operatorname{tg} b$  jest wyższy dla stopy, zestroju oraz jego części głównej, aniżeli w modelu I (podobnie jak miało to miejsce dla języka angielskiego). Wiadomo, że w przypadku regresji  $d$  względem  $\tilde{d}$  oddziałuje jakiś czynnik, który nie pojawił się w regresji  $d$  względem  $n$ . Powoduje on, że tendencja do izochronizmu wystąpiła w sposób mniej wy-



razisty w modelu II.

Graficzną ilustrację wyników zamieszczonych w tablicach 10, 11, 12 stanowi ryc. 2.

### 2.6.2. Model III z trzema zmiennymi : $d, \tilde{d}, n$ .

Przeprowadzono analizę regresji z uwzględnieniem dwóch zmiennych niezależnych  $\tilde{d}$  i  $n$  celem stwierdzenia, czy estymacja iloczasu jednostki rytmicznej z tych wielkości pozwoli uzyskać większą siłę predykcji modelu. Posłużono się następującym równaniem liniowym regresji wielokrotnej :

$$\frac{d-\bar{d}}{\bar{p}} = a + b \frac{\tilde{d}-\bar{\tilde{d}}}{\bar{p}} + c (n-\bar{n})$$

Na podstawie wyników zamieszczonych w tablicach 13, 14, 15 nasuwają się następujące wnioski :

1/ Najwyższe wartości  $r$  uzyskano dla korelacji  $(n-n, \frac{\tilde{d}-\bar{\tilde{d}}}{\bar{p}})$ , co wynika ze ścisłej zależności pomiędzy liczbą głosek składowych, a ich skumulowanym iloczusem.

2/ Wartości współczynnika korelacji wielokrotnej i odpowiadających im współczynników determinacji są wyższe niż w modelach dwuzmiennych, z tym że nieznacznie tylko przekraczają wartości dla modelu ze zmienną niezależną  $\tilde{d}$ . Tak więc model z trzema zmiennymi posiada największą siłę predykcji, podobnie jak miało to miejsce dla języka angielskiego.

### 3. Struktura rytmiczna wypowiedzi na poziomie frazy.

#### 3.1. Izochronizm frazowy.

Skracanie czasu trwania segmentów fonetycznych uwarunkowane rytmem akcentowym przejawia się zdaniem niektórych autorów nie tylko na poziomie wyrazowym, lecz również na poziomie frazowym. Tezę taką wysunęli Carlson, Granström, Lindblom, Rapp [1] w oparciu o analizę struktury czasowej wypowiedzi języka szwedzkiego. Stwierdzili, że skracanie iloczasu głosek wraz ze zwiększaniem liczby wyrazów we frazie można ująć w postaci modelu zaproponowanego dla poziomu wyrazowego, w którym wielkości  $m$  i  $n$  będą oznaczały liczbę wyrazów we frazie. Model dla poziomu wyrazowego miał postać następującą :

$$V = \frac{D}{\frac{m}{n} \cdot c}$$



gdzie  $D$  oznacza maksymalny czas trwania segmentu,  $n$  - liczbę sylab w wyrazie,  $m$  - liczbę sylab pozostałych do wyprodukowania wyrazie.

W późniejszych pracach zależność funkcyjną pomiędzy czasem trwania segmentów fonetycznych a liczbą sylab w wyrazie przedstawił Lindblom w zmodyfikowanej postaci [8]:

$$S = \frac{D}{(a+1)^a (b+1)^b} ,$$

Gdzie  $a$  oznacza liczbę sylab następujących po sylabie rozpatrywanej, zaś  $b$  - liczbę sylab wyprodukowanych przed tą sylabą. Autor podkreśla, że analogiczne zależności stwierdzono na poziomie frazy. Lyberg [9] na podstawie wyników przeprowadzonego przez siebie doświadczenia również na materiale języka szwedzkiego uznał, że model Lindbloma z dużą dokładnością oddający zależności iloczynowe w obrębie wyrazu, podlega znacznym ograniczeniom w odniesieniu do frazy. Powyższa reguła stosuje się dla wyrazów w pozycji końcowej we frazie (przed pauzą) i wówczas można przyjąć model

$$D_p = \frac{D_w}{(b_p+1)^{b_p}} ,$$

natomiast nie znajduje zastosowania dla wyrazów w pozycji niekońcowej we frazie, w których iloczyn samogłosek wykazuje żadne albo słabe uzależnienie od liczby wyrazów następujących. W tej pozycji, przynajmniej dla celów syntezy, można przyjąć stały czas trwania samogłoski.

### 3.2. Materiał.

Materiał badawczy obejmował trzy grupy wypowiedzi. Każda grupa składała się z siedmiu fraz o zmieniającej się liczbie zestrojów akcentowych. Najkrótsza wypowiedź obejmowała jeden zestroj, po czym wydłużano stopniowo wypowiedzi poprzez dodawanie kolejnych zestrojów. Najdłuższa fraza obejmowała cztery zestroje akcentowe.

Grupa pierwsza : Stach.

Stach wiedział.



Stach wiedział wszystko.

Stach wiedział wszystko zawsze.

W celu zbadania ewentualnego wpływu akcentu zdaniowego wypowiedzi 1 - 4 zostały wymówione z akcentem zdaniowym padającym na wyraz Stach oraz wypowiedzi 2 - 4 z akcentem padającym na wyraz wiedział. Łącznie w tej grupie znalazło się siedem fraz.

Na grupę drugą składały się następujące wypowiedzi :

Staszek.

Staszek wie.

Staszek wie wszystko.

Staszek wie wszystko zawsze.

Akcent zdaniowy padał na wyrazy Staszek bądź wie, co dało łącznie siedem fraz.

Grupa trzecia obejmowała wypowiedzi :

Marysia.

Marysia napisze.

Marysia napisze do ciebie.

Marysia napisze do ciebie we wtorek.

Akcent zdaniowy padał na wyraz Marysia bądź napisze. W grupie znalazło się siedem fraz.

Łącznie materiał doświadczalny składał się z 21 wypowiedzi. Bezpośrednią podstawę do zbadania wpływu długości frazy na czas trwania głosek stanowią wyrazy : Stach, wiedział, Staszek, wie, Marysia, napisze.

Wszystkie wypowiedzi z trzech grup zostały ułożone w porządku losowym, a następnie wymówione z odpowiednim akcentem zdaniowym przez 10 osób. Czas trwania segmentów mierzono na oscylogramach.

### 3.3. Wyniki pomiarów.

W celu wyeliminowania różnic spowodowanych przez zmienność osobniczą, a wynikających między innymi z różnic tempa mówienia, obliczono względne wartości iloczasu samogłosek lub dyftongów przyjmując w każdym głosie wartość 1 dla czasu trwania w najkrótszej frazie, w jakiej dana samogłoska wystąpiła. Dla samogłosek w wyrazach : Stach, Staszek, Marysia, na które pada akcent zdaniowy, jest to fraza jednozestrojowa, dla samogłosek w wyrazach : Stach, Staszek, Marysia, na które pada akcent zdaniowy oraz w wyrazach : wie, wiedział, napisze w obu tych



pozycjach - fraza dwuzestrojowa.

W tabelicy 16 zamieszczono uśrednione dla wszystkich głosów względne wartości iloczasu z zaznaczeniem pozycji akcentu wyrazowego oraz zdaniowego. Niektóre z samogłosek wykazują tendencję do skracania się wraz z wydłużaniem frazy. Odnosi się to przede wszystkim do samogłosek wymówionych w wyrazach, które w jednej wypowiedzi zajmują pozycję końcową. W tabeli pozycję we frazie określa liczba zestrojów następujących: 0 - pozycja końcowa, 1,2,3 - pozycja niekończąca. W odniesieniu do wyrazów zajmujących wyłącznie pozycję niekończącą we frazie tendencja do skracania samogłosek przejawia się w słabym stopniu lub wogóle nie dochodzi do głosu.

W badanym materiale można wykluczyć wpływ akcentu zdaniowego na sposób realizacji tej tendencji. Stwierdzono jedynie różnice długości samogłosek/dyftongów wymówionych w identycznym kontekście fonetycznym i zajmującym taką samą pozycję w zestroju akcentowym (w sylabie akcentowanej lub nieakcentowanej) w zależności od tego, czy padał na nie akcent zdaniowy. Iloczas samogłosek wymówionych w wyrazach bez akcentu zdaniowego wynosi średnio 0,9 iloczasu takich samych samogłosek znajdujących się w wyrazach z akcentem zdaniowym.

#### 3.4. Realizacja tendencji do skracania samogłosek w poszczególnych głosach.

Wartości iloczasu dla poszczególnych samogłosek wymówionych we frazach o różnej długości sklasyfikowano na skali porządkowej według rang (dla każdej z osób oddzielnie). Następnie zbadano, jak wysoka jest zgodność tych uporządkowań, tzn. jakie zachodzą korelacje pomiędzy rangami dla poszczególnych osób. Jako miara zgodności uporządkowań wielokrotnych posłużył współczynnik konkordancji  $r_w$  [4]. W tabelicy 17 zamieszczono wartości współczynnika  $r_w$ . Sformułowanie zawarte w tabeli: pozycja końcowa oznacza, że w jednej z wypowiedzi dany wyraz zajmuje pozycję końcową we frazie, zaś pozycja niekończąca, że w żadnej wypowiedzi wyraz nie zajmuje pozycji końcowej.

Wysoka wartość współczynnika  $r_w$  (powyżej 0,50) wystąpiła dla samogłosek/dyftongów wymówionych w wyrazach, które w jednej z wypowiedzi zajęły pozycję końcową we frazie. Dotyczy to wszystkich samogłosek w sylabie akcentowanej oraz w sylabie poakcentowej,



która jest jednocześnie sylabą wygłosową. Jedynie w sylabach przedakcentowych tych wyrazów uzyskano niskie wartości  $r_w$  (0,02 - 0,31). Dla samogłosek/dyftongów wymówionych w wyrazach, które w żadnej wypowiedzi nie zajęły pozycji końcowej, wartość  $r_w$  jest niska i wynosi od 0 do 0,33. Dotyczy to samogłosek w sylabach akcentowanych oraz nieakcentowanych.

Wielkość współczynnika  $r_w$  nie wykazuje związku z pozycją akcentu zdaniowego.

Na dużą zgodność uporządkowań w wyrazach, które zajęły w którejś z wypowiedzi pozycję końcową we frazie wpłynął fakt, że we wszystkich głosach iloczasa samogłosek w tej właśnie pozycji (gdzie liczba zestrojów następujących wynosiła 0) przyjmowały najwyższe wartości, co za tym idzie uzyskiwały zawsze najwyższą rangę. Ponieważ liczba obiektów na skali porządkowej była niewielka -  $N=4$ , konsekwencja, z jaką najwyższą rangę na każdej ze skal przyjmowały takie same obiekty (iloczas samogłoski w zestroju końcowym we frazie) musiała rzutować na wynik korelacji wielokrotnych.

Mała zgodność uporządkowań w przypadku wyrazów, które we wszystkich wypowiedziach zajęły pozycję niekończącą, poddaje w wątpliwość istnienie w języku polskim reguły, w myśl której iloczasa samogłosek zwiększałby się wraz z wydłużaniem frazy.

Zasadnicze znaczenie dla iloczasu samogłosek posiada w świetle uzyskanych wyników pozycja wyrazu we frazie.

### 3.5. Weryfikacja modelu matematycznego izochronizmu na poziomie frazowym.

W pracach na rytmem języka szwedzkiego [1], [8], [9] podkreśla się rekurencyjny charakter reguł iloczasowych, które mają taką samą postać na poziomie wyrazowym oraz na poziomie frazowym. Zgodnie z tym model wyrażający się funkcją potęgową, który zastosowano dla zestrojów akcentowych w języku polskim [1] powinien odnosić się również do poziomu frazowego. Posłużono się wzorem :

$$V = \frac{D}{(mn)^{\alpha}}$$

na podstawie którego wyliczone wartości teoretyczne iloczasu



samogłoskowego były w takim samym stopniu zgodne z wartościami empirycznymi, jak wartości teoretyczne wyliczone na podstawie wzoru

$$V = \frac{D}{m^{\alpha} n^{\beta}}$$

We wzorze zastosowanym na poziomie frazowym wielkość  $m$  oznacza liczbę zestrojów akcentowych liczonych od danego do końca frazy, zaś  $n$  oznacza całkowitą liczbę zestrojów we frazie.  $D$  jest to czas trwania samogłoski/dyftongu w najkrótszej frazie, w jakiej dana samogłoska wystąpiła.

Wyliczono wartości  $\alpha$  dla materiału doświadczalnego. Najniższe i bardzo zbliżone do siebie wartości uzyskano dla samogłosek, które wystąpiły w wyrazach zajmujących wyłącznie pozycję niekończącą we frazie, jak też dla samogłosek w sylabach przedakcentowych pozostałych wyrazów - średnio = 0,03. Natomiast  $\alpha$  obliczone dla samogłosek wymówionych w wyrazach zajmujących pozycję końcową w jednej z fraz przybierało bardzo różne wartości. W przypadku samogłosek w sylabach akcentowanych rozbieżności te mieściły się w granicach : 0,06 - 0,18 dla różnych wyrazów. Największe zróżnicowanie obliczonych wartości  $\alpha$  wystąpiło w sylabach wygłosowych tych wyrazów. Co więcej, znaczne różnice pojawiły się nie tylko dla samogłosek w różnych wyrazach, lecz również dla samogłosek w tych samych wyrazach znajdujących się w różnych frazach. Np.  $\alpha$  dla samogłoski /a/ w wyrazie Stach ma wartość 0,27 we frazie dwuzestrojowej, 0,21 we frazie trzyzestrojowej, 0,16 we frazie czterozestrojowej.  $\alpha$  dla /a/ wygłosowego w wyrazie Marysia wynosi odpowiednio : 0,52, 0,32, 0,26. Przy tak zmieniających się wartościach  $\alpha$  ryzykowne jest przyjęcie określonej, stałej dla niej wartości i podstawienie jej do wzoru na  $V$ , ponieważ z góry można spodziewać się, że obliczone wartości teoretyczne będą dość znacznie odbiegać od wartości empirycznych.

Wydaje się, że zaobserwowane rozbieżności zostały wywołane w znacznym stopniu przez interakcje zachodzące pomiędzy długością frazy, a pozycją badanej samogłoski we frazie.

Na podstawie wyników doświadczenia należy uznać, że model izochronizmu dla języka polskiego wyrażający się funkcją



potęgowa, przyjęty na poziomie wyrazowym, nie znajduje zastosowania na poziomie frazowym.

### 3.6. Wpływ pozycji we frazie.

Zjawisko wydłużania iloczasu głoskowego na końcu frazy, na które zwracano uwagę w wielu pracach fonetycznych (m.in. [2], [5], [7], [9]) występuje konsekwentnie w badanym materiale doświadczalnym. Jest to tendencja na tyle silna, że wybija się jako zasadniczy czynnik związany z rytmem. Słabo zaznaczony wpływ długości frazy na iloczas samogłosek odgrywał w materiale doświadczalnym drugorzędą rolę. Maksymalne wartości średniego czasu trwania bez wyjątku występują w wyrazach zajmujących pozycję końcową w zdaniu, a różnice pomiędzy iloczasetem samogłosek w tej pozycji oraz w pozostałych pozycjach są znacznie większe, niż różnice pomiędzy iloczasetem samogłosek w wyrazach zajmujących pozycję niekończącą we frazach o różnej długości. W tej sytuacji wydaje się wystarczające określić iloczas samogłoski na poziomie frazowym za pomocą prostej zależności

$$V = f \cdot D$$

D oznacza czas trwania samogłoski/dyftongu w wyrazie zajmującym pozycję końcową we frazie,  $f$  przybiera różne wartości w zależności od tego, jaką pozycję przyjmuje dana samogłoska/dyftong w wyrazie oraz we frazie.

$f = 1$  dla wyrazów końcowych

$f < 1$  dla wyrazów niekończących

Dla materiału doświadczalnego  $f$  przybrało wartości następujące :

1/  $f = 0,55$  gdy samogłoska/dyftong znajduje się w wygłosie wyrazu "w<sub>1</sub>ie, v<sub>1</sub>ie, "mar<sub>1</sub>ica, "nap<sub>1</sub>ise, nap<sub>1</sub>ise, "v<sub>1</sub>iedzaw, v<sub>1</sub>iedzaw

2/  $f = 0,69$  gdy samogłoska leży w sylabie wygłosowej wyrazu przed spółgłoską "st<sub>1</sub>ax, "st<sub>1</sub>afek

3/  $f = 0,81$  gdy samogłoska/dyftong znajduje się w sylabie akcentowanej niewygłosowej wyrazu "st<sub>1</sub>afek, "mar<sub>1</sub>ica, "v<sub>1</sub>iedzaw, v<sub>1</sub>iedzaw, "nap<sub>1</sub>ise, nap<sub>1</sub>ise

4/  $f = 0,93$  gdy samogłoska leży w sylabie przedakcentowej "nap<sub>1</sub>ise, nap<sub>1</sub>ise, "mar<sub>1</sub>ica

W oparciu o bezwzględne wartości czasu trwania samogłosek przeprowadzono test Studenta w celu stwierdzenia, czy różnice spowodowane pozycją we frazie (kończąca/niekończąca) są istotne pod



względem statystycznym. Dla wszystkich samogłosek/dyftongów w sylabach akcentowanych i poakcentowych (grupa 1), 2), 3)) uzyskano wartości  $t$  przekraczające wartość krytyczną na poziomie istotności 0,001 (13 przypadków) oraz na poziomie 0,01 (2 przypadki). Dla wszystkich samogłosek w sylabie przedakcentowej (grupa 4) obliczone wartości  $t$  są mniejsze od wartości krytycznych.

Można więc przyjąć, że w wyrażeniu  $V = f \cdot D$   $f$  przyjmuje wartości mniejsze od 1 dla samogłosek w sylabach akcentowanych oraz poakcentowych wyrazów zajmujących pozycję niekończącą we frazie.

#### 4. Podsumowanie.

W oparciu o analizę regresji dokonano oszacowania długości jednostek rytmicznych z liczby głosek składowych oraz z ich skumulowanego średniego iloczasu. Współczynnik  $b$  z równania regresji pozwolił sformułować w sposób matematyczny siłę izochronizmu. Prosta interpretacja wyników w modelu dwuzmiennym opiera się na wielkości  $\arctg b$ , której odpowiada kąt nachylenia linii regresji do osi odciętych. Dla trzech głosów wytestowano trzy liniowe równania regresji: równanie ze zmienną niezależną  $n$  odpowiadającą liczbie głosek w jednostce rytmicznej, równanie ze zmienną niezależną  $\tilde{d}$  odpowiadającą sumarycznemu średniemu iloczynowi głosek składowych oraz równanie z dwiema zmiennymi  $n$  i  $\tilde{d}$ . Wyniki pozwalają wnioskować o tendencji do izochronizmu w języku polskim, chociaż wykazuje ona słabsze nasilenie, niż stwierdzono to dla języka angielskiego. Najmniejszą podatność na zmiany długości uwarunkowane rozciągłością jednostki rytmicznej wykazują głoski w części przedakcentowej zestroju. Najlepsze oszacowanie iloczasu jednostki rytmicznej uzyskano w modelu III, to jest regresji  $d$  względem  $\tilde{d}$  oraz  $n$ . Niewiele mniejszą siłą predykcji wykazuje model II wyrażający się regresją  $d$  względem  $\tilde{d}$ .

Przeprowadzono analizę struktury rytmicznej wypowiedzi na poziomie frazowym. Rezultaty pomiarów przeprowadzonych na 10 głosach nie potwierdziły hipotezy o występowaniu w języku polskim izochronizmu frazowego, to jest zależności pomiędzy iloczynem głosek, a liczbą zestrojów akcentowych we frazie. Stwierdzono natomiast wpływ pozycji wyrazu we frazie, który ujęto w postaci prostej zależności matematycznej.



Wyniki uzyskane w pracy świadczą o konieczności uwzględnienia czynnika rytmu przy opracowywaniu reguł dotyczących cech suprasegmentalnych dla celów syntezy mowy.

Tablica 1. Wartości iloczasu dla typów głosek (w ms).

FD			ES			JI		
Typ	Liczeb- ność	Średnia	Typ	Liczeb- ność	Średnia	Typ	Liczeb- ność	Średnia
R	18	19,17	R	18	18,33	R	23	19,13
Y	17	35,59	Y	10	31,00	J	32	37,66
J	37	39,32	J	33	37,88	Y	2	45,00
L	194	49,72	L	173	45,46	L	171	46,20
YN	58	64,91	YN	62	57,74	N	201	53,86
EN	104	67,07	IN	47	59,15	YN	66	55,45
IN	44	67,39	UN	34	61,18	B	129	57,48
N	217	68,04	ON	116	61,81	UN	29	59,48
UN	35	69,42	Z	155	62,58	EN	109	59,82
ON	113	70,13	N	202	62,87	IN	46	60,76
YA	31	71,61	YA	27	63,89	ON	113	61,90
IA	20	75,00	EN	107	64,77	Z	148	63,71
Z	159	75,35	IA	14	67,86	YA	30	63,83
B	131	76,91	B	123	71,34	AN	105	70,24
EA	56	80,36	AN	105	73,24	P	209	71,87
AN	100	80,45	EA	50	76,10	UA	32	73,44
UA	32	81,88	UA	28	77,50	D	33	76,06
P	211	83,69	OA	57	80,44	EA	56	77,05
OA	61	85,98	S	179	82,23	IA	17	78,24
D	33	90,61	P	201	83,53	OA	55	82,91
S	179	93,58	D	32	84,53	S	179	84,33
AA	104	101,97	AA	102	97,21	T	97	96,39
T	103	104,27	T	93	104,52	AA	99	96,91
NN	30	124,50	V	69	106,16	NN	23	104,13
V	78	133,91	NN	33	121,52	V	77	126,30
NA	15	164,00	NA	10	162,00	NA	14	128,93
W	8	211,25	W	3	193,33	W	4	167,50



Tablica 2. Podstawowe parametry statystyczne dotyczące klas głosek. Głos PD.

Range	Typ	Srednia [ms]	Odczylenie standardowe [ms]	Współczynnik zmienności [%]	Liczebność
1	R (R, Y, J)	32,8	13,6	41	67
2	L (L)	49,8	15,8	32	190
3	N (N, EN, IN, N, UN, ON, YA)	68,3	18,3	27	577
4	A (IA, Z, B, EA, AN, UA, P, OA)	80,6	25,1	31	734
5	S (D, S)	92,9	26,4	28	208
6	T (AA, T)	103,8	27,0	26	202
7	V (IN, V)	132,1	31,8	24	116
8	O (NA)	166,7	30,1	18	12
9	W (W)	211,3	62,3	30	8
		$\bar{x} = 80,2$			



Tablica 3. Podstawowe parametry statystyczne dotyczące klas glosek. Głos ES.

Ranga	Typ	Średnia [ms]	Dochylenie standardowe [ms]	Współczynnik zmienności [%]	Liczebność
1	R (R, Y)	22,9	10,2	45	28
2	L (J, L)	44,2	14,8	34	206
3	N (YN, IN, UN, ON, Z, N, YA, EN)	62,2	16,9	27	750
4	A (IA, B, AN, EA, UA)	73,1	17,6	24	320
5	S (OA, S, P, D)	82,7	24,6	30	469
6	T (AA, T, V, NI)	104,3	28,2	27	297
7	W (NA, W)	169,2	37,8	22	13
		$\bar{x} = 72,9$			

Tablica 4. Podstawowe parametry statystyczne dotyczące klas głosek. Głos JI.

Range	Typ	Średnia [ms]	Odczylenie standardowe [ms]	Współczynnik zmienności [%]	Liczebność
1	R (R)	19,1	5,2	27	23
2	L (J, Y, L)	44,9	14,0	31	205
3	N (N, YN, B, UN, EN, IN, ON, Z, YA)	58,9	16,3	28	871
4	A (AN, P, UA, D, EA, IA)	72,8	20,9	29	452
5	S (OA, S)	84,0	24,5	29	234
6	T (T, AA, NN)	97,4	27,6	28	219
7	W (V, NA, W)	128,4	40,8	32	95
		$\bar{x} = 70,0$			



Tablica 5. Średni iloczyn głošek w jednostkach rytmicznych (w ms).

Klasa	FD				ES				JI					
	Stopa	Zestroj	Przed-akc.	Główna	Klasa	Stopa	Zestroj	Przed-akc.	Główna	Klasa	Stopa	Zestroj	Przed-akc.	Główna
R	34,0	32,8	22,1	36,0	R	24,5	22,8	17,5	28,5	R	18,9	19,3	21,0	18,2
L	49,9	50,1	47,3	50,9	L	44,5	44,5	42,0	45,6	L	45,0	44,5	44,6	45,2
N	68,4	68,2	66,6	69,2	N	62,6	62,6	59,1	64,6	N	59,0	59,0	57,1	60,0
A	80,5	79,9	75,8	81,9	A	73,6	73,2	67,7	75,8	A	73,5	73,1	65,8	76,6
S	91,6	92,2	86,1	95,0	S	83,3	83,2	74,5	86,9	S	83,5	84,5	77,6	96,6
T	103,1	103,2	106,2	102,6	T	104,3	104,6	89,7	105,0	T	97,1	98,0	89,7	98,1
V	132,7	131,7	126,7	132,2	V	169,2	169,2	-	169,2	V	127,0	127,3	-	127,0
O	166,7	166,7	-	166,7	O					O				
W	211,3	211,3	-	211,3	W					W				

Tablice 6. Wartości średnie dla jednostek rytmicznych.

Głos	Jednostka rytmiczna	Średni czas trwania głoski [ms]	Średnia liczba głosek	Średni czas trwania jednostki rytmicznej [ms]	Estymowany średni czas trwania jednostki rytmicznej [ms]	Średnia prędkość przepływu głosek (głoski/sek)
PD	Stopa	80,7	6,6	529,0	530,3	12,4
	Zestrój	79,9	7,1	567,2	570,6	12,5
	Przedakc.	70,3	3,7	260,3	263,6	14,2
	Główna	84,1	5,0	415,3	416,8	11,9
ES	Stopa	74,1	6,7	488,2	491,6	13,5
	Zestrój	73,2	7,2	521,4	524,6	13,7
	Przedakc.	62,2	3,8	235,8	237,1	16,1
	Główna	77,9	5,0	381,6	384,5	12,8
JI	Stopa	70,8	6,5	451,6	455,5	14,1
	Zestrój	70,0	7,3	505,6	510,3	14,3
	Przedakc.	60,4	3,8	228,1	229,9	16,6
	Główna	74,2	5,1	370,4	373,6	13,5



Tablica 7. Model I :  $\frac{d-\bar{d}}{p} = a+b(n-\bar{n})$ . Gros FD.

Jednostka rytmiczna	Srednia $n-\bar{n}$	Srednia $\frac{d-\bar{d}}{p}$	$\bar{d} (n-\bar{n})$	$\bar{d} \frac{d-\bar{d}}{p}$	r	$r^2, 100\%$	a	b	arc tg b
Stopa	0,0000014	0,0000015	2,584	2,095	0,85	71,22	0,0000	0,684	34,4°
Zestrój	-0,00000005	0,0000017	2,662	2,338	0,86	74,13	0,0000	0,756	37,1°
Przedakc.	0,0000010	0,0000004	1,999	1,957	0,95	90,03	-0,0000	0,929	42,9°
Główna	0,0000003	0,0000023	1,292	1,298	0,63	39,07	0,0000	0,628	32,2°

Tablica 8. Model I :  $\frac{d-\bar{d}}{p} = a+b(n-\bar{n})$ . Gros ES.

Jednostka rytmiczna	Srednia $n-\bar{n}$	Srednia $\frac{d-\bar{d}}{p}$	$\bar{d} (n-\bar{n})$	$\bar{d} \frac{d-\bar{d}}{p}$	r	$r^2, 100\%$	a	b	arc tg b
Stopa	0,0000012	0,0000029	2,630	2,080	0,86	73,13	0,0000	0,676	34,1°
Zestrój	0,0000007	0,0000026	2,633	2,278	0,85	72,07	0,0000	0,734	36,3°
Przedakc.	-0,00000006	0,00000006	2,121	1,990	0,96	91,13	0,0000	0,895	41,8°
Główna	0,0000010	0,0000008	1,252	1,297	0,63	39,49	0,0000	0,651	33,1°

Tablica 9. Model I :  $\frac{d-\bar{d}}{\bar{p}} = a+b(n-\bar{n})$ . Głos JI.

Jednostka rytmiczna	Srednia $n-\bar{n}$	Srednia $\frac{d-\bar{d}}{\bar{p}}$	$\sigma(n-\bar{n})$	$\sigma \frac{d-\bar{d}}{\bar{p}}$	r	$r^2$ 100%	a	b	arc tg b
Stopa	0,0000001	0,0000012	2,323	1,979	0,80	64,68	0,0000	0,685	34,4°
Zestrój	0,0000006	0,0000034	2,698	2,456	0,85	72,20	0,0000	0,774	37,7°
Przedakc.	0,0000000	-0,0000006	2,070	2,056	0,96	91,32	-0,0000	0,949	43,5°
Główna	0,0000004	0,0000024	1,362	1,495	0,64	41,41	0,0000	0,706	35,2°

Tablica 10. Model II :  $\frac{d-\bar{d}}{\bar{p}} = a+b \frac{\bar{d}-\bar{d}}{\bar{p}}$ . Głos PD.

Jednostka rytmiczna	Srednia $\frac{d-\bar{d}}{\bar{p}}$	Srednia $\frac{\bar{d}-\bar{d}}{\bar{p}}$	$\sigma \frac{d-\bar{d}}{\bar{p}}$	$\sigma \frac{\bar{d}-\bar{d}}{\bar{p}}$	r	$r^2$ 100%	a	b	arc tg b
Stopa	0,0002554	0,0000015	2,308	2,095	0,91	82,76	-0,0002	0,826	39,6°
Zestrój	0,0002433	0,0000017	2,548	2,338	0,92	83,78	-0,0002	0,840	40,0°
Przedakc.	0,0000278	0,0000004	2,121	1,957	0,95	90,35	-0,0000	0,876	41,3°
Główna	0,0001768	0,0000023	1,223	1,298	0,81	65,22	-0,0001	0,857	40,6°



Tablica 11. Model II :  $\frac{d-\tilde{d}}{p} = a + b \frac{\tilde{d}-\tilde{d}}{p}$  . Głos ES.

Jednostka rytmiczna	Srednia $\frac{\tilde{d}-\tilde{d}}{p}$	Srednia $\frac{d-\tilde{d}}{p}$	$\sigma \frac{\tilde{d}-\tilde{d}}{p}$	$\sigma \frac{d-\tilde{d}}{p}$	r	$r^2$ 100%	a	b	arc tg b
Stopo	0,0002497	0,0000029	2,401	2,080	0,90	81,51	-0,0002	0,782	38,0°
Zestroj	0,0002739	0,0000026	2,448	2,278	0,90	80,72	-0,0002	0,836	39,9°
Przedakc.	0,0000317	0,0000006	2,149	1,990	0,96	92,53	-0,0000	0,891	41,7°
Główna	0,0001789	0,0000008	1,173	1,297	0,78	60,36	-0,0002	0,859	40,7°

Tablica 12. Model II :  $\frac{d-\tilde{d}}{p} = a + b \frac{\tilde{d}-\tilde{d}}{p}$  . Głos JI.

Jednostka rytmiczna	Srednia $\frac{\tilde{d}-\tilde{d}}{p}$	Srednia $\frac{d-\tilde{d}}{p}$	$\sigma \frac{\tilde{d}-\tilde{d}}{p}$	$\sigma \frac{d-\tilde{d}}{p}$	r	$r^2$ 100%	a	b	arc tg b
Stopa	0,0002327	0,0000012	2,114	1,979	0,88	77,84	-0,0002	0,826	39,6°
Zestroj	0,0002594	0,0000034	2,502	2,456	0,90	81,52	-0,0002	0,886	41,5°
Przedakc.	0,0000297	-0,0000006	2,131	2,056	0,96	92,80	-0,0000	0,929	42,9°
Główna	0,0001818	0,0000024	1,259	1,495	0,81	65,16	-0,0002	0,959	43,8°

Tablica 13. Model III :  $\frac{d-\bar{d}}{p} = a + b \frac{\tilde{d}-\bar{\tilde{d}}}{\tilde{p}} + c(n-\bar{n})$ . Głos PD.

	Stopa	Zestrój	Przedakcen.	Główna
średnia $(n-\bar{n})$	0,0000014	-0,00000005	0,0000010	0,0000003
średnia $\frac{\tilde{d}-\bar{\tilde{d}}}{\tilde{p}}$	0,0002554	0,0002433	0,0000278	0,0001768
średnia $\frac{d-\bar{d}}{p}$	0,0000015	0,0000017	0,0000004	0,0000023
$\sigma (n-\bar{n})$	2,584	2,662	1,999	1,292
$\sigma \frac{\tilde{d}-\bar{\tilde{d}}}{\tilde{p}}$	2,308	2,548	2,121	1,223
$\sigma \frac{d-\bar{d}}{p}$	2,095	2,338	1,957	1,298
$r(n-\bar{n}, \frac{\tilde{d}-\bar{\tilde{d}}}{\tilde{p}})$	0,96	0,96	0,98	0,88
$r(n-\bar{n}, \frac{d-\bar{d}}{p})$	0,85	0,86	0,95	0,65
$r(\frac{\tilde{d}-\bar{\tilde{d}}}{\tilde{p}}, \frac{d-\bar{d}}{p})$	0,91	0,92	0,95	0,81
a	-0,0003	-0,0003	-0,0000	-0,0002
b	-0,2773	-0,2092	0,4384	-0,3475
c	1,1233	1,0500	0,4742	1,1783
r - współcz. korelacji wielokrotnej	0,92	0,92	0,96	0,85
$r^2 \cdot 100\%$	83,71	84,22	91,35	68,03



Tablica 14. Model III :  $\frac{\hat{d}-\bar{d}}{p} = a + b \frac{\hat{d}-\bar{d}}{p} + c(n-\bar{n})$ . Głos ES.

	Stopa	Zestroj	Przedakcen.	Główna
średnia $(n-\bar{n})$	0,0000012	0,0000009	-0,00000006	0,0000010
średnia $\frac{\hat{d}-\bar{d}}{p}$	0,0002497	0,0002739	0,0000317	0,0001789
średnia $\frac{d-\bar{d}}{p}$	0,0000029	0,0000026	0,00000006	0,0000002
$\sigma (n-\bar{n})$	2,630	2,633	2,121	1,252
$\sigma \frac{\hat{d}-\bar{d}}{p}$	2,401	2,448	2,149	1,173
$\sigma \frac{d-\bar{d}}{p}$	2,080	2,278	1,990	1,297
$r(n-\bar{n}, \frac{\hat{d}-\bar{d}}{p})$	0,97	0,97	0,99	0,89
$r(n-\bar{n}, \frac{d-\bar{d}}{p})$	0,86	0,85	0,96	0,63
$r(\frac{\hat{d}-\bar{d}}{p}, \frac{d-\bar{d}}{p})$	0,90	0,90	0,96	0,78
a	-0,0002	-0,0003	-0,0000	-0,0002
b	-0,2171	-0,2323	0,1826	-0,3170
c	1,0122	1,0774	0,7127	1,1604
r - współcz. korelacji wielokrotnej	0,91	0,90	0,96	0,79
$r^2 \cdot 100\%$	82,00	81,21	92,62	62,30

Tablica 15. Model III :  $\frac{d-\bar{d}}{\bar{p}} = a + b \frac{\hat{d}-\bar{d}}{\bar{p}} + c(n-\bar{n})$ . GrosJI.

	Stopa	Zestroj	Przedakcen.	Główna
średnia $(n-\bar{n})$	0,0000001	0,0000006	0,0000000	0,0000004
średnia $\frac{\hat{d}-\bar{d}}{\bar{p}}$	0,0002327	0,0002594	0,0000297	0,0001818
średnia $\frac{d-\bar{d}}{\bar{p}}$	0,0000012	0,0000034	0,00000006	0,0000024
$\sigma(n-\bar{n})$	2,323	2,698	2,070	1,362
$\sigma \frac{\hat{d}-\bar{d}}{\bar{p}}$	2,114	2,502	2,131	1,259
$\sigma \frac{d-\bar{d}}{\bar{p}}$	1,979	2,456	2,056	1,495
$r(n-\bar{n}, \frac{\hat{d}-\bar{d}}{\bar{p}})$	0,95	0,96	0,99	0,89
$r(n-\bar{n}, \frac{d-\bar{d}}{\bar{p}})$	0,80	0,85	0,96	0,65
$r(\frac{\hat{d}-\bar{d}}{\bar{p}}, \frac{d-\bar{d}}{\bar{p}})$	0,88	0,90	0,96	0,81
a	-0,0003	-0,0003	-0,0000	-0,0002
b	-0,3208	-0,2723	0,1751	-0,3891
c	1,1616	1,1696	0,7613	1,3331
r - współcz. korelacji wielokrotnej	0,89	0,91	0,97	0,82
$r^2 \cdot 100\%$	79,17	82,15	92,88	67,79



Tablica 16. Średnie względne wartości czasu trwania samogłosek dyftongów .  
 'akcent wyrazowy  
 "akcent zdaniowy  
 -głoska, której dotyczy dane w kolumnie

Stach

liczba zestr. następ.	'stax	stax
0	1	-
1	0,69	1
2	0,63	1,07
3	0,65	0,96

Staszek

liczba zestr. następ.	'stąjek	stąjek	'stąjek	stąjek
0	1	-	1	-
1	0,91	1	0,83	1
2	0,83	0,94	0,65	0,93
3	0,85	0,94	0,67	0,99

Marysia

liczba zestr. następ.	'mar <sub>1</sub> ęa	mar <sub>1</sub> ęa	'mar <sub>2</sub> ęa	mar <sub>2</sub> ęa	'mar <sub>3</sub> ęa	mar <sub>3</sub> ęa
0	1	-	1	1	-	-
1	0,78	1	0,93	0,49	1	1
2	0,78	1,03	0,89	0,50	0,98	1
3	0,73	0,94	0,90	0,49	0,93	1

wie

liczba zestr. następ.	'vje	vje
0	1	1
1	0,59	0,57
2	0,56	0,54
3	-	-

wiedział

liczba zestr. następ.	vjedzaw	vjedzaw	vjedzaw	vjedzaw
0	1	1	1	1
1	0,83	0,78	0,70	0,65
2	0,81	0,78	0,65	0,66
3	-	-	-	-

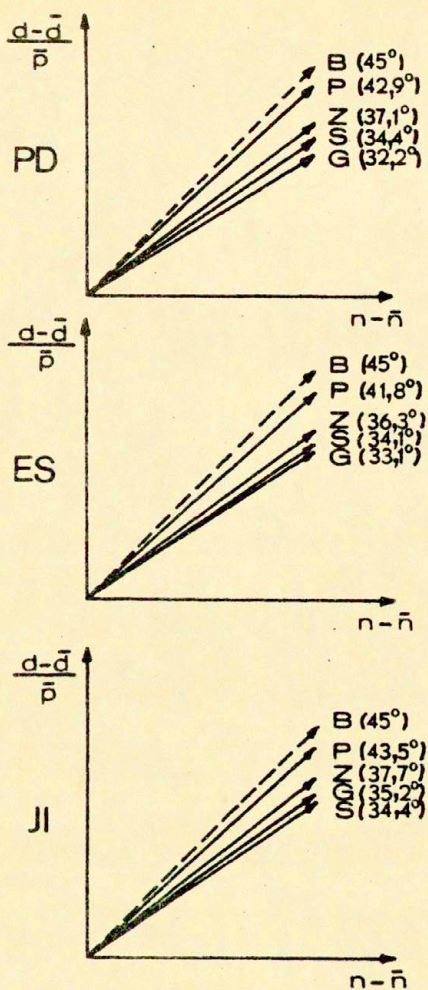
napisze

liczba zestr. następ.	"napise	napise	"napise	"napise	napise	napise
0	1	1	1	1	1	1
1	0,80	0,82	0,98	0,46	0,97	0,44
2	0,82	0,78	0,96	0,45	0,86	0,45
3	-	-	-	-	-	-



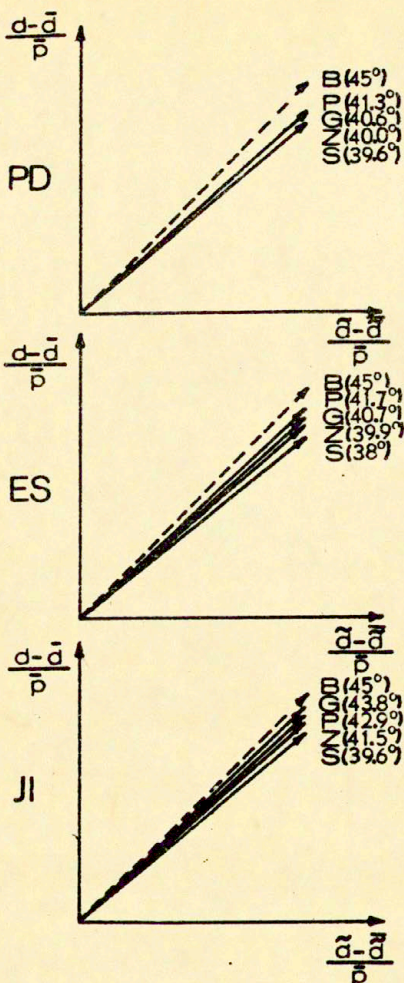
Tablica 17. Wartości współczynnika r<sub>w</sub>.

Zestroj akcentowy	Wyraz z akcentem zdaniowym. Pozycja końcowa we frazie.		Wyraz bez akcentu zdaniowego					
	Sylaba akcentowana		Sylaba nieakcentowana		Sylaba akcentowana		Sylaba nieakcentowana	
	Przedakcent.	Poakcentowa	Poz. końcowa	Poz. niekończąca	Poz. końcowa	Poz. niekończąca		
Stach	-	-	-	/a/ 0,33	-	-	-	
wiedzia?	-	/aw/ 0,91	/je/ 0,79	-	/aw/ 0,76	-	-	
Steszek	-	/e/ 0,55	-	/a/ 0,23	-	-	/e/ 0,11	
wie	-	-	/je/ 0,79	-	-	-	-	
Marysia	/a/ 0,31	/a/ 0,61	-	/z/ 0,13	-	-	przedakc. /a/ 0,10 poakcent. /a/ 0	
napisze	/a/ 0,02	/e/ 0,77	/i/ 0,65	-	przedakc. /a/ 0,02 poakcent. /e/ 0,75	-	-	



Ryc. 1. Regresja liniowa  $d$  względem  $n$  w różnych jednostkach rytmicznych.  
B - brak izochronizmu, S - stopa, Z - zestroj akcentowy, P - część przedakcentowa zestroju, G - główna część zestroju.





Ryc. 2. Regresja liniowa  $d$  względem  $\bar{d}$  w różnych jednostkach rytmicznych.  
B - brak izochronizmu, S - stopa, Z - zestrój akcentowy, P - część przedakcentowa zestroju, G - część główna zestroju.

BIBLIOGRAFIA

- [1] CARLSON, R., GRANSTRÖM, B., LINDBLÖM, B., RAPP, K. : Some timing and fundamental frequency characteristics of Swedish sentences, Speech Transm. Lab. Quarterly Progress and Status Report, 4, 1972, 11-20, Stockholm.
- [2] FISCHER-JØRGENSEN, E. : Segment duration in Danish words in dependency on higher level phonological units, Ann. Rep. Inst. Phon. Univ. Copenhagen, 16, 1982, 137-189.
- [3] GABRIEL, K.R. : A procedure for testing the homogeneity of all sets of means in analysis of variance, Biometrics, v. 20, 3, 1964, 459-477.
- [4] GORALSKI, A. : Metody opisu i wnioskowania statystycznego w psychologii, Warszawa, 1974.
- [5] HOEQUIST, C.J. : Durational correlates of linguistic rhythm categories, Phonetica, v. 40, 1, 1983, 19-32.
- [6] JASSEM, W., KRZYSKO, M., STOLARSKI, P. : Regresyjny model izochronizmu zestrojowego w sygnale mowy, Prace IPPT 33/1981, Warszawa, 1981.
- [7] KLATT, D.H., COOPER, W.E. : Perception of segment duration in sentence contexts, [in]: Structure and Process in Speech Perception, eds. A.Cohen, S.G.Nooteboom, Berlin 1975, 69-89.
- [8] LINDBLÖM, B. : Some temporal regularities of spoken Swedish, [in]: Auditory Analysis and Perception of Speech, (ed. by G.Fant and M.A.A.Tatham.), London, 1975, 387-396.
- [9] LYBERG, B. : Some observations on the timing of Swedish utterances, Journal of Phonetic, vol. 5, 1, 1977, 49-69.
- [10] RICHTER, L. : The duration of Polish vowels, [in]: Speech Analysis and Synthesis, vol. 3, 1973, 87-117.
- [11] RICHTER, L. : Wstępna charakterystyka izochronizmu zestrojowego w języku polskim, Prace IPPT 4/1983, Warszawa, 1983.
- [12] WIERZCHOWSKA, B. : Wymowa polska, Warszawa, 1971.