

2.23 — Akustyka mowy — fonetyka akustyczna

33 / 1981

W. Jassen, M. Krzyśko, P. Stolarski

REGRESYJNY MODEL
IZOCHRONIZMU ZESTROJOWEGO
W SYGNALE MOWY

P. 269a



WARSZAWA 1981

<http://rcin.org.pl>

ISSN 0208-5658

Praca wpłynęła do Redakcji dnia 21 sierpnia 1981 r.

Zarejestrowana pod nr 33/1981



57080



Na prawach rękopisu

Instytut Podstawowych Problemów Techniki PAN
Nakład 180 egz. Ark. wyd. 1,6. Ark.druk.2,25
Oddano do drukarni w październiku 1981
Nr zamówienia S60/0

Warszawska Drukarnia Naukowa, Warszawa,
ul. Śniadeckich 8

<http://rcin.org.pl>

Wiktor Jassem

Pracownia Fonetyki Akustycznej IPPT PAN

Mirosław Krzyśko

Przemysław Stolarski

Ośrodek Informatyki UAM

REGRESYJNY MODEL ISOCHRONIZMU ZESTROJOWEGO W SYGNALE MOWY¹

Streszczenie

Izochronizm zestrojowy związany jest z akcentuacją i rytmem mowy. Polega on na wyrównywaniu czasu trwania zestroju akcentowego bez względu na liczbę sylab lub głosek. Matematyczne sformułowanie tego zjawiska jest konieczne dla celów syntezy mowy sterowanej z komputera i opartej na regułach związanych z jednostkami o rozciągłości fonemu. Do języków wykazujących izochronizm zestrojowy zalicza się angielski, dla którego istnieją dwa teoretyczne modele rytmiczne. Praca ma na celu doświadczalne zbadanie słuszności impresyjnego opisu izochronizmu poprzez pomiar naturalnego sygnału mowy, a w przypadku jej potwierdzenia - sformułowanie matematyczne tego zjawiska, które umożliwiłoby jego uwzględnienie w algorytmizacji reguł syntezy mowy. Posługując się modelem regresyjnym sformułowano proste ilościowe ujęcie izochronizmu i jego siły posługując się dwiema lub trzema zmiennymi. Stwierdzono rzeczywiste jego występowanie w sygnale mowy na przykładzie języka angielskiego oraz wskazano, które z dwóch jego sformułowań fonetycznych należy uznać za prawidłowe.

1. Wstęp.

Większość systemów softwarowych elektronicznej syntezy sygnału mowy ze sterowaniem z maszyny cyfrowej przyjmuje jako podsta-

¹ Praca wykonana w ramach problemu międzyresortowego MR I-24

wową jednostką segment lub ciąg segmentów o rozciągłości fonemu. Pozwala to ekonomicznie wykorzystywać pamięć operacyjną tak, iż nawet małe komputery o pojemności 8 lub 16 K bajtów i długości słowa 16 lub 24 bitów pozwalają z klawiatury syntetyzować mowę o dużej naturalności brzmienia. Synteza segmentalna jest w szczególności niezbędna dla celów odczytywania tekstu drukowanego w formie syntetycznej mowy w przypadku tych języków, które stosują pismo alfabetyczne, tzn. przede wszystkim tzw. języków konferencyjnych (angielski, francuski, rosyjski, niemiecki) oraz języków z nimi spokrewnionych. W tych przypadkach poszczególne fonemy lub allofony opisane są w pamięci wartościami ich parametrów fonetyczno-akustycznych, najczęściej w terminach cech widmowych. Poza tym jednak określona być musi rozciągłość czasowa każdego segmentu, czyli w terminologii fonetycznej - jego iloczas.

Iloczas głoski zależy od kilku współdziałających czynników, wśród których dotychczas zauważono: 1. parametry fonetyczne głoski (np. stopień przewężenia w torze ustnym, lokalizacja tego przewężenia, nazalizacja), 2. kontekst (np. sąsiedztwo głosek dźwięcznych w opozycji do kontekstu bezdźwięcznego), 3. akcentuacja i rytm, 4. średnie tempo, 5. pozycja w wypowiedzi, 6. elementy emocjonalne (por. Frackowiak-Richter 1973 i 1976). Liczne prace z zakresu iloczasu głoskowego w różnych językach zajmują się najczęściej oddzielnie któryś z wymienionych czynników, a tylko w nielicznych z nich materiałem lingwistycznym jest swobodna (prawie) spontaniczna mowa.

Niniejsza praca dotyczy zależności iloczasu głoskowego od rytmu mowy w powiązaniu ze zróżnicowaniami segmentalno-fonetycznymi.

Jako jednostkę rytmiczną mowy przyjmuje się bądź sylabę bądź zestrój akcentowy, przy czym oba te pojęcia podlegają znacznym zróżnicowaniom definicyjnym. W każdym razie odróżnia się w fonetyce rytm sylabiczny od rytmu akcentowego, a poszczególne języki określa się jako podlegające prawom bądź jednego bądź drugiego typu rytmicznego (por. np. Ladefoged 1975, str. 222). Rytm jednego i drugiego typu objawia się izochronizmem, tj. wyrównywaniem iloczasu suprasegmentalnego w obrębie

bądź sylaby bądź zestroju akcentuacyjnego. Tak więc w przypadku ścisłego izochronizmu zestrojowego iloczyn elementów linearnych niższego rzędu podlega prawu odwrotnej proporcjonalności względem liczby tych elementów w zestroju.

Przyjmuje się na ogół, iż rytm języka angielskiego w jego formie nieregionalnej (tak brytyjskiej, jak i amerykańskiej) jest typu zestrojowego ("stress timing"). W przeglądzie prac dotyczących rytmu angielskiego Lehište (1977) wskazuje jednak na różnice zdań co do realności izochronizmu w tym języku. Z jednej strony opisy podręcznikowe jednoznacznie izochronizm stwierdzają (np. Pike 1954, str. 34, Ladefoged 1975, str. 102-103), z drugiej jednak, na podstawie wyników pomiarów oraz doświadczeń wysuwano słabsze lub silniejsze zastrzeżenia co do trafności tych opisów, mających charakter impresyjny (Udall 1971, O'Connor 1965, Bolinger 1965, a szczególnie Lea 1974 oraz Shen i Peterson 1962). Własne doświadczenia i pomiary Lehište (1973, 1975, 1977) doprowadziły ją do wniosku, że "Niektóre aspekty danych przemawiają za obecnością izochronizmu, a inne przeciwko niemu (Lehište 1977, str. 256). Autorka zwraca uwagę w swych pracach również na ważne współzależności między rytmem a składnią języka angielskiego. Adams (1973) przypomina, iż rytmiczność zestrojowa w angielszczyźnie zauważył już John Hart w latach sześćdziesiątych XVI wieku. Ostatnio wysunięto hipotezę, iż izochronizm nie jest zjawiskiem kategoriowym i że jego nasilenie jest prawdopodobnie zależne od stylu (Jassem, Gibbon 1980, str. 9).

2. Dwie teorie rytmu w mowie angielskiej.

Zaproponowano dotychczas dwie szczegółowe teorie rytmu mówionej angielszczyzny brytyjskiej. Pierwszą, którą oznaczać będziemy przez A sformułował Abercrombie (1964, 1973), a druga, B, zaproponował Jassem (1952, w zmodyfikowanym ujęciu Jassem 1979 a i b). Wspólna cecha obu teorii ma charakter definicyjny i polega na stwierdzeniu, iż zachodzi tendencja wyrównywania przedziału czasowego pomiędzy akcentami (sylabami akcentowanymi) w wypowiedzi. Poza tym obie teorie mają również ważną wspólną cechę metodologiczną polegającą na uznaniu implicite niezależności jednostek suprasegmentalnych od jednostek składniowych przy uwzględnieniu związków zachodzących pomiędzy fonetycznymi

jednostkami suprasegmentalnymi a jednostkami syntaktycznymi. Pod tym względem obie teorie różnią się w zasadniczy sposób od założeń fonologii generatywnej. Teoria Abercrombiego została rozwinięta przez Hallidaya (1970) i Wittena (1977). Obie teorie można zreasumować następująco :

Teoria A :

1. Jednostka rytmiczna zwana STOPĄ zawsze rozpoczyna się od sylaby akcentowanej, wobec czego każda sylaba nieakcentowana występuje po akcentowanej w obrębie stopy. Stąd wszystkie nieakcentowane sylaby można określić jako postakcentuacyjne (postiktyczne).

2. Pojedyncza sylaba stopy jednosylabowej jest akcentowana z definicji.

3. Jeśli wypowiedź rozpoczyna się od sylaby nieakcentowanej, postuluje się "niemy akcent", abstrakcją uzewnętrzną (ujawnianą, albo realizowaną) jako "zero dźwięku", lecz realną psychologicznie (subiektywnie).

4. Dwusylabowa stopa ma rytm trójkowy i może być reprezentowana przez jedną z następujących struktur : $u - -$ krótka-długa , $o o$ średnia-średnia , lub $- u$ długa-krótka . Pierwotna wersja teorii nie uwzględniała stóp więcej niż dwusylabowych, lecz została odpowiednio rozszerzona przez Wittena (1977). Tę część teorii, opartą na jednostce mowy występującej w poetyce języków klasycznych, pominął Halliday (1970).

5. Wewnętrzna struktura rytmiczna stopy jest ściśle związana z jej strukturą segmentalną. Na przykład $(C)V^1CV(C)$ stanowią $u -$, zaś $(C)VC(C)V(C)$ oraz $(C)V^2(C)V(C)$ tworzą $o o$ itd¹.

6. Zachodzą pewne, nie zbadane jeszcze, relacje między rytmem a składnią. Na przykład typ $- u$ związany jest z występowaniem granicy międzywyrazowej (ibid. str. 219). W ten sposób eksplikuje Abercrombie różnicę fonetyczno-rytmiczną między np. (take) Grey to London w opozycji do Greater London (segmentalnie w obu przypadkach /'grejtə'landən/). Por. także the way to cut it w opozycji do the waiter cut it (segmentalnie w obu

¹ V^1 oznacza krótką samogłoskę, V^2 - długą samogłoskę lub dyftong, C - spółgłoskę. Abercrombie zastrzega się, że "w niektórych przypadkach struktura fonematyczna sylaby może być nieistotna" (Abercrombie 1964, str. 217).

przypadkach /ðə'wejtəkətit/). Tak więc iloczyn sylaby akcentowanej może zależeć od tego, czy jest ona enklityczna, czy proklityczna.

7. Zakłada się, (Abercrombie 1967, str. 35, Ladefoged 1975, str. 222) że akcent polega na zwiększeniu wysiłku artykulatoryjnego (lub energii). Pojęcie akcentu jest pierwotne względem rytmu.

Teoria B.

1. W wypowiedzi wystąpić mogą dwa typy jednostek supra-segmentalnych: (a) ścisła jednostka rytmiczna, stanowiąca konstytutywny element rytmu oraz (b) anakruza będąca elementem nierelwantnym rytmicznie. Dla danego tempa iloczyn ścisłej jednostki rytmicznej jest zależny od liczby tworzących ją sylab. Ten iloczyn, stały dla danego tempa, w jednostce jednosylabowej można zapisać przez Y. Z wzrostem liczby sylab w ścisłej jednostce rytmicznej jej iloczyn wzrasta również, lecz nie w sposób proporcjonalny. Dwusylabowa jednostka rytmiczna jest dłuższa niż jednosylabowa, lecz znacznie krótsza niż 2Y. Trzysylabowa ścisła jednostka rytmiczna jest dłuższa niż dwusylabowa, lecz jej długość jest istotnie mniejsza od 3Y. Iloczyn dłuższych jednostek rytmicznych jest określony analogicznie.

2. Poszczególne sylaby wewnątrz wiasylabowych ścisłych jednostek rytmicznych mają w przybliżeniu równą długość¹.

3. Akcent jest efektem organizacji czasowej wypowiedzi: Jedyna lub pierwsza sylaba ścisłej jednostki rytmicznej jest akcentowana.

4. Iloczyn poszczególnych głosek podlega z konieczności zmianom wynikającym ze zmian iloczynu sylabowego w ścisłych jednostkach rytmicznych. Tak na przykład /ij/ w read jest dłuższe niż /ij/ w reader, a to z kolei krótsze niż /ij/ w reading it (por. Ladefoged 1975, str. 103). Nie wiadomo jeszcze, w jakim stopniu rytm wpływa na iloczyn segmentalny

¹ Z p. 1 i 2 wynika, iż nieformalny graficzny zapis iloczynu ścisłych jednostek rytmicznych i sylab składowych można zapisać:

1 sylaba —————
2 sylaby ————
3 sylaby ————
4 sylaby ————
itd.

poszczególnych fonemów lub typów głoskowych.

5. oprócz ścisłych jednostek rytmicznych wypowiedź może zawierać sylabę lub ciąg sylab charakteryzujący się szczególnie krótkim iloczasem. Taka sylaba lub ciąg sylab tworzy ANAKRUZĘ. Iloczas segmentalny zbliża się w obrębie anakruzy do minimum wynikającego z możliwości artykulacyjnych. Stąd średni czas trwania anakruzy jest w przybliżeniu proporcjonalny do liczby składowych głosek. Ewentualna anakruza zawsze poprzedza ścisłą jednostkę rytmiczną, tworząc z nią całkowitą jednostkę rytmiczną. Wpowiedź składająca się z dwóch lub większej liczby całkowitych jednostek rytmicznych może zawierać odpowiednią liczbę anakruz, najwyżej po jednej przed ścisłą jednostką rytmiczną.

6. Pomiedzy opisanymi wyżej jednostkami rytmicznymi a jednostkami składniowymi wypowiedzi zachodzą określone relacje. Opisano je szczegółowo w pracy : Jassem 1952.

7. Opis rytmu oparty na teorii B pozwala oznaczyć go w bieżącej transkrypcji tekstu, która wówczas w sposób scalony opisuje strukturę segmentalną i suprasegmentalną wypowiedzi, wraz z akcentuacją i intonacją. Szczegóły i przykłady takiej transkrypcji znajdują się w pracach Jassema 1980 i 1981.

Opartą na teorii B transkrypcję ukazującą linearnie elementy segmentalne w połączeniu z rytmem zastosował m.in. O'Connor (1967, str. 125 - 127, 160 - 162).¹

Różnicę pomiędzy teorią A i B można zilustrować rozpatrując następujące wypowiedzi : (1) John's pleased, (2) John was pleased, (3) John would be pleased, (4) John would have been pleased (5) John would have been extremely pleased. Według teorii A sylaba /dʒɒn/ musiałaby ulegać stopniowemu skracaniu od (1) do (5), tak aby w stopie mogły pomieścić się sylaby nie akcentowane. Według teorii B sylaby nie akcentowane następujące po /dʒɒn/ w wypowiedziach (2) do (5) stanowią anakruzę i należą do całkowitej jednostki rytmicznej z następującym pleased, wobec czego nie wpływają istotnie na iloczas sylaby /dʒɒn/.

¹ Allen (1968) wysunął trzecią teorię rytmu, która pominiemy w niniejszych rozważaniach, gdyż ma ona charakter bardzo tentatywny, nie została przez autora eksperymentalnie poparta i dotychczas nie spotkała się z szerszym odzewem.

3. Sylaba czy głoska?

Za najmniejszy element rytmu uważa się sylabę, co jest niewątpliwie słuszne w przypadku poezji i co znalazło swój wyraz w przedstawionych powyżej teoriach A i B. Również w opisach podręcznikowych rytmu języka angielskiego przyjmuje się taką zasadę (por. np. Jones 1956, 238-244, O'Connor 1967, 123-125). O'Connor najpierw (1965) poddał w wątpliwość obiektywny fakt istnienia izochronizmu. Trzy lata później (O'Connor 1968) wykonał on eksperyment mający na celu sprawdzenie izochronizmu w zakresie jednosylabowego zestroju akcentowego. Materiał pomiarowy stanowiły sylaby o różnej liczebności składowych jednostek segmentalnych (fonematycznych): od 3 do 9. Stwierdzono pewną tendencję do kompensacji w postaci skracania iloczasu segmentalnego, lecz całkowity iloczyn jednosylabowego zestroju wzrastał wraz z liczebnością składowych elementów segmentalnych. 6-segmentalna sylaba jest 1.67 raza dłuższa, a 9-segmentalna 2.14 raza dłuższa od 3-segmentalnej. Zatem już na jednosylabowej jednostce rytmicznej stanowiącej zestrój akcentowy stwierdzono zależność od liczebności głosek. Uzasadnia to konieczność analogicznego traktowania dłuższych jednostek rytmicznych.

Długość sylaby w wielu językach jest niezawsze wielkością określoną, gdyż w pewnych przypadkach nie jest możliwe wyznaczenie granicy pomiędzy sylabami. W języku polskim, na przykład, wyraz "iskra" można interpretować dowolnie jako składający się z sylab i-skra, is-kra lub isk-ra. Taki podział jak po-strach opiera się na kryteriach morfologicznych, nie fonetycznych, które zresztą i tutaj uzasadniałyby równie po-strach, jak pos-trach lub post-rach. W języku angielskim, według Hocketta (1955, str. 52) zachodzą przypadki, w których nie ma momentu podziału na sylaby, czyli brak jest granicy międzysylabowej, ponieważ dwa szczyty sylabiczne przedzielone są jedynie tzw. interludiami. Zachodzi to na przykład w "smaller" /'smələ/, które fonologicznie można podzielić jedynie na sylaby /smo/ oraz /ə/ z interludium w postaci segmentu (głoski) /l/¹.

¹ "China" i "finer" stanowią dokładne rymy w nieregionalnej angielszczyźnie brytyjskiej. Podział chi-na oraz fin-er przyjęty w pisowni jest morfologiczny. W płaszczyźnie fonologicznej w obu przypadkach /n/ stanowi interludium.

Wyniki wstępnej pracy O'Connora (1968) oraz powyższe przesłanki fonologiczno-strukturalne skłaniają do oparcia teorii izochronizmu na ciągach segmentalnych, a nie na ciągach sylabowych.

4. Interpretacja funkcyjna i regresyjna izochronizmu w rytmie zestrojowym.

a. interpretacja funkcyjna

Przyjmujemy jako zmienne :

d - iloczyn jednostki rytmicznej

p - iloczyn segmentu

n - liczba segmentów w jednostce rytmicznej.

W braku izochronizmu długość segmentu $p = \text{const}$ oraz $d = np$.

Ścisły izochronizm w ujęciu funkcyjnym wyraża $d = \text{const}$ przy $p = \frac{d}{n}$. Jeśli p przyjmiemy jako jednostkę długości d, wówczas prosta reprezentująca ścisły izochronizm jest równoległa do osi n i przecina oś d w punkcie $p(1)$, tzn. w punkcie odpowiadającym długości segmentu dla monosegmentalnej jednostki rytmicznej. Całkowity brak izochronizmu reprezentuje wtedy prosta nachylona do osi n pod kątem 45° przechodząca przez punkt $(d = p(1), n = 1)$ oraz początek układu. Interpretacja fizyczna i fonetyczna dotyczy tylko półprostej ograniczonej punktem $(d = p(1), n = 1)$, jak na ryc. 1.

b. interpretacja regresyjna

Funkcyjny związek między długością jednostki rytmicznej d, a liczbą głosek n ulega zaburzeniu na skutek występowania czynnika losowego (dokładność realizacji artykulacyjnej) oraz czynników systematycznych, z których (zakładając stałe tempo ogólne) zasadniczą rolę odgrywa zróżnicowanie iloczasu głoskowego (segmentalnego) wynikające z przynależności do określonego typu głosek. Nie było dotychczas danych określających średni iloczyn poszczególnych typów głoskowych, lecz wiadomo, iż takie systematyczne zróżnicowania zachodzą. Pomijając ewentualne inne źródła zmienności iloczasu segmentalnego, wymieniona tutaj zmienność losowa i międzyklasowa powodują, iż w materiale eksperymentalnym otrzyma się dla każdej wartości zmiennej n określony rozrzut wartości d, dla którego w pierwszym przybliżeniu przyjmujemy rozkład normalny

ze zmniejszającym się odchyleniem standardowym wraz z oddaleniem się od średniej wartości n .

Siła izochronizmu nie jest interpretowalna w terminach korelacji. W przypadku całkowicie losowego rozkładu wielkości d względem n otrzymuje się rozrzut taki, jak na ryc. 3.

Współczynnik korelacji

$$r = \frac{\text{cov}(n,d)}{\sqrt{\text{var}(n) \cdot \text{var}(d)}} \quad (1)$$

jest w przypadku zilustrowanym na ryc. 3 równy zeru dokładnie tak samo jak przy $d = \text{const}$ (ryc. 1), gdyż kowariancja

$$\text{cov}(n,d) = \frac{\sum_{i=1}^n (n_i - \bar{n})(d_i - \bar{d})}{n - 1} \quad (2)$$

jest w obu przypadkach zerowa.

Siłę izochronizmu można wyrazić poprzez współczynnik regresji liniowej, jeśli korelacja osiąga żądany poziom istotności. Poprzez odpowiedni dobór skal dla d i n można przedstawić izochronizm w formie linii prostej przechodzącej przez początek układu. Wówczas dla braku izochronizmu $\arctg a = 45^\circ$, a dla ścisłego izochronizmu $\arctg a = 0^\circ$, gdzie

$$a = \frac{b-d}{n} \quad \text{z równania regresji}$$

$$d = an + b.$$

4. Materiał doświadczalny.

Jako materiał doświadczalny wykorzystano komercyjnie dostępne zapisy na taśmie magnetofonowej dwóch tekstów z Hallidaya (1970). Praca ta ma charakter podręcznikowy, ale zawiera teksty określone przez autora jako spontaniczne lub czytane w stylu swobodnym, naturalnym i przy zachowaniu stałego, nieemocjonalnego, średnio szybkiego tempa. Wybrano lekcję 30 oraz 39, pierwszą wypowiedzianą przez jeden głos, a drugą stanowiącą dialog. Wstępne analizy (Hill, Jassen, Witten 1978, Hill, Witten, Jassen 1978) wskazały, iż statystycznie oba teksty można uznać za pochodzące z tej samej populacji, wobec czego potraktowano je obecnie łącznie. Iloczas segmentalny mierzono na podstawie analizy

spektrograficznej wykonanej przy użyciu aparatury Kay Electric Sona-Graph. Iloczas jednostek rytmicznych określono jako sumę iloczasów głosek składowych. W obliczeniach pominięto kilka przypadków oczywistych zawahań i paralingwistycznych wzdłużeń. Jednostki rytmiczne obu tekstów wyznaczońo osobno według teorii A oraz teorii B. W pierwszym przypadku obliczano wyłącznie iloczas tzw. stóp wypełnionych, tzn. nie zawierających "niemego akcentu", z definicji bowiem iloczas stóp zawierających taki akcent jest niemierzalny. W obliczeniach dotyczących poszczególnych typów głosek pominięto 79 przypadków, w których pomiar iloczasu segmentalnego był niepewny z powodu trudności segmentacyjnych. Łącznie wyznaczono iloczas 2499 głosek. Ścisłe jednostki rytmiczne będą w dalszym ciągu gwołi zwiężłości nazywane "taktami".

5. Iloczas segmentalny.

Wstępnie wykonano szereg prób wyznaczenia klas głosek, które stanowiłyby jednorodnie statystycznie grupy przy zachowaniu kryteriów fonetycznych. Próby te wskazały na zasadność pogrupowania głosek w następujący sposób :

1. Głoski uderzeniowe oraz inicjalne zwarte słabe monosegmentalne złożone z samej plosji. Typ F.
2. Trące słabe niesybilantne /ʃ, v/. Typ D.
3. Płynne /j, w, r, l/. Typ G.
4. Samogłoski /ɪ, e, ʌ, ə, ə/. Typ E.
5. Spółgłoski zwarte słabe /b, g, d/ nieinicjalne. Typ B.
6. Nosowe /m, n, ŋ/. Typ N.
7. Aspirata /h/. Typ H.
8. Zwarte słabe /p, t, k/ nieaspirowane. Typ K.
9. Słabe trące sybilantne /z, ʒ/. Typ Z.
10. Spółgłoski sylabiczne /l, ŋ, r, m/. Typ SC.
11. Zwarte mocne /p, t, k/ aspirowane nieakcentowane. Typ KH.
12. Mocne trące /f, θ, s, ʃ/. Typ S.
13. Słabe afrykaty /dʒ, dr/. Typ S.
14. Samogłoski /i, u, ɛ, ɔ/. Typ O.
15. Zwarte mocne /p, t, k/ aspirowane akcentowane. Typ KHA.
16. Mocne afrykaty /tr, tʃ/. Typ AF.
17. Monoftongi /a, ə, ɜ/ oraz wszystkie dyftongi. Typ A.
18. Finalne mocne zwarte. Typ FTH.

Oznaczenia powyższych klas mają charakter mnemotechniczny.

Wyniki obliczeń dotyczących iloczasu powyższych typów głoskowych (klas) z pominięciem zróżnicowania na stopy, takty i anakruzy podaje poniższa Tablica.

Nr	Typ	Wartość średnia [ms]	Wariancja [ms ²]	Odchylenie standardowe [ms]	Procentowy współczynnik zmienności %	Liczebność
1	F	17	33	6	34.3	20
2	D	48	303	17	36.6	148
3	G	56	501	22	40.2	213
4	E	59	643	25	42.9	584
5	B	60	487	22	36.7	142
6	N	62	809	28	46.1	283
7	H	65	1166	34	52.6	38
8	K	65	524	23	35.1	200
9	Z	70	1145	34	48.1	83
10	SC	76	685	26	34.3	45
11	KH	85	539	23	27.2	53
12	S	88	1158	34	38.7	181
13	AVF	94	833	29	30.7	19
14	O	96	1551	39	40.9	155
15	KHA	117	1075	33	28.0	43
16	AF	127	836	29	22.8	31
17	A	132	2972	55	41.3	38
18	FTH	137	709	27	19.4	6

Porównując kolejne średnie z wartościami odpowiednich odchyżeń standardowych lub bezpośrednio oceniając procentowy współczynnik zmienności, należy oczekiwać, iż nie wszystkie różnice między średnimi są statystycznie istotne, wobec czego można typy głoskowe pogrupować w skupienia jednorodne statystycznie tak, aby otrzymać zbiory statystycznie różne w sposób istotny.

Celem przeprowadzenia takiej procedury grupującej zastosowano metodę opisaną w pracy : Gabriel 1964, zalgorytmizowaną

i zaprogramowaną w języku Algol 68 JAS.

Jest to jedna z metod skupień według średnich grupowych i opiera się na ogólnej zasadzie maksymalizacji wariancji międzygrupowej przy równoczesnej minimalizacji wariancji wewnątrzgrupowej. Procedura ta dała dla naszego materiału następujące skupienia :

Nr typu	Kod typu	Średnia grupowa [ms]
1	F	16.64
2	D	47.59
3	G	59.27
4	E	
5	B	
6	N	
7	H	67.74
8	K	
9	Z	
10	SC	
11	KH	91.07
12	S	
13	AFV	
14	O	
15	KHA	129.79
16	AF	
17	A	
18	FTH	

Średnia ogólna wynosi $\bar{p} = 74.12$ ms.

6. Głoski i jednostki rytmiczne : dane wyjściowe.

Na podstawie wstępnych wyników testów statystycznych głoski podzielono na 18 klas, jak następuje :

F - uderzeniowe oraz nagłosowe zwarte słabe

D - słabe trące niesybilantne /ð, v/

G - niesylabiczne samogłoski /w, j, r/

E - samogłoski /e, i, ə, ə, ʌ/

- B - nienagłosowe słabe zwarte
- N - nosowe niesylabiczne
- H - aspirata i nagłosowe zwarte mocne aspirowane
- K - mocne zwarte nieaspirowane
- Z - sybilantne słabe trące /z,ʒ/
- SC - spółgłoski sylabiczne
- KH - aspirowane mocne zwarte nieakcentowane
- S - mocne trące
- AFV - słabe afrykаты /dʒ,dr/
- O - samogłoski /i:,u:,ɜ (:),v/
- KHA - akcentowane mocne zwarte
- AF - mocne afrykаты /tʃ, tr/
- A - samogłoski /ɜ:,a:,ɔ:/ oraz dyftongi
- FTH - aspirowane wygłosowe zwarte mocne.

Należy zwrócić uwagę, że poszczególne klasy nie odnoszą się do fonemów, lecz do głosek. Wstępna analiza statystyczna pokazała bowiem, że z jednej strony w obrębie fonemu zachodzą systematyczne różnicowania iloczasu, podczas gdy z drugiej strony określone typy głóskowe mają ten sam iloczas mimo przynależności do różnych fonemów.

Poczyniono najpierw upraszczające założenie, że średnie czasy trwania klas głóskowych można obliczyć nie biorąc pod uwagę różnicowań rytmicznych określonych w teorii B. Tablica 1 zawiera średnie iloczasy, wraz z wielkościami wariancji, odchylenia standardowego oraz współczynników zmienności¹.

Różnice między wartościami średniej dla poszczególnych typów głóskowych mogą niewszystkie być znaczące statystycznie. Analiza wariancji wykazała, że co najmniej niektóre różnice są istotne na poziomie $\alpha = 0.01$, jak to ukazuje Tablica 2.

Hipoteza zerowa o równości wszystkich średnich zostaje zatem odrzucona, wobec czego przejść można do zbadania, które średnie różnią się między sobą w sposób statystycznie istotny. Celem zbadania tych różnic zastosowano test opracowany przez Gabriela (1964). Jest on oparty na zasadzie minimalizacji

¹ Współczynnik zmienności określony jest ilorazem odchylenia standardowego i średniej arytmetycznej.

wariancji wewnątrzgrupowej przy równoczesnej maksymalizacji wariancji międzygrupowej. Test ten doprowadził w pięciu krokach do następującego pogrupowania średnich :

F----D----G E B N----H K Z SC----KH S AFV O----KHA AF A FTH.

Różnice między średnimi w obrębie wskazanych grup nie okazały się w badanym materiale istotne na poziomie $\alpha = 0.05$, natomiast każda różnica między średnią z dowolnej grupy a średnią z dowolnej innej grupy jest na tym poziomie istotna.

Dane dotyczące iloczasu głoskowego rozbito następnie według typu jednostki rytmicznej, jak to ukazano w Tabelicy 3.

Nieznaczne różnice pomiędzy wartościami średnimi w kolumnie STOPA a odpowiednimi średnimi w Tabelicy 1 wynikają z nieco odmiennego zbioru danych (np. głoski w "stopach" z "niemym akcentem" zostały pominięte w obliczeniach średnich w kolumnie STOPA).

Liczby w Tabelicy 3 prowadzą do następujących wniosków :

1. Z nielicznymi wyjątkami kolejność średnich jest taka sama w każdej kolumnie i jest w zasadzie zgodna z kolejnością w Tabelicy 2.
2. Przesunięcia w kolejności występują przeważnie w kolumnie ANAKRUZA, a tutaj wartości średnie są nieco mniej wiarygodne z uwagi na mniejszą liczebność prób.
3. Jeśli średnie występują we wszystkich kolumnach dla danej klasy głosek, to z kilku wyjątkami średnia dla ŚCISŁEJ JEDNOSTKI RYTMICZNEJ jest największa, a dla ANAKRUZY - najmniejsza, przy czym średnia dla STOPY jest bliższa tej pierwszej.

Tabelica 4 podaje względne wartości średnich odniesione do wartości dla ścisłej jednostki rytmicznej wziętej jako jedność. Tabelica 5 zawiera średnie wartości szczególnie ważne dla podstawowego problemu tutaj rozpatrywanego : Wprawdzie dane nie są wystarczające dla postawienia hipotezy co do tego, czy różnica między ścisłą jednostką rytmiczną a anakruzą wpływa w podobny sposób na wartości średnie wszystkich klas głosek, lecz średni iloczyn ścisłej jednostki rytmicznej jest przeszło dwukrotnie większy niż iloczyn anakruzy. Liczba głosek w anakruzie jest średnio mniejsza od liczby głosek w ścisłej jednostce rytmicznej ze współczynnikiem 0.69, natomiast średni iloczyn głoskowy - ze współczynnikiem 0.48. Ostatecznie więc średni

Iloczas głoskowy w anakрузie wynosi 0,7 średniego iloczasu w ścisłej jednostce rytmicznej. Różnica ta jest istotna na poziomie $\alpha = 0.001$ i oznacza ona, iż prędkość przepływu fonemów jest w anakрузie 1,43 razy większa niż w ścisłej jednostce rytmicznej : 17,7 głoski w anakрузie wobec 12,4 głoski w ścisłej jednostce rytmicznej.

7. Matematyczne modele izochronizmu.

7.1. Model z dwiema zmiennymi.

Oznaczmy iloczas jakiegokolwiek jednostki rytmicznej przez d , jej wielkość, wyrażoną liczbą wchodzących w jej skład głosek, przez n , zaś iloczas średni każdej głoski przez p . Pomińmy dla uproszczenia różnice wynikające z przynależności głosek do różnych klas iloczasowych. Z definicji $d = np$. Przyjmując funkcyjny związek między d oraz n , można rozważyć dwa skrajne przypadki :

A. Całkowity brak izochronizmu : $d = \text{const.} = np$.

B. Ścisły izochronizm : $d = \text{const.} = np$. W obu przypadkach przyjmujemy, że d jest funkcją n . Te dwa skrajne przypadki ilustruje ryc. 1.

Funkcje na ryc. 1 mają sens fizyczny tylko dla n należącego do zbioru liczb rzeczywistych, całkowitych i dodatnich. Górny kres dziedziny funkcji nie jest znany.

W rzeczywistości nie zachodzi związek funkcyjny między d oraz n , z dwóch co najmniej przyczyn : 1. Istnieją systematyczne różnicowania w iloczasię głosek, tak że jednostki rytmiczne o tej samej rozciągłości w sensie liczby głosek mogą mieć znacząco różne długości. 2. Z natury neurofizjologicznej wytwarzania mowy wynika pewien zakres przypadkowości w jednostkach rytmicznych lingwistycznie ekwiwalentnych. Zresztą zachodzą zapewne także różnice systematyczne, które nie zostały jeszcze bliżej zbadane (należy do nich zapewne większy iloczas jednostek rytmicznych wygłosowych w stosunku do pozawygłosowych). Tak więc, realistyczny model izochronizmu można przedstawić w postaci regresyjnej $d = a + bn$, gdzie iloczas jednostki rytmicznej jest estymowany z jej głoskowej rozciągłości, jak to pokazuje ryc. 2. Jeżeli współczynnik regresji $b = 0$, wówczas izochronizm jest regresyjnie nieokreślony, jak to widać z ryc. 3 i 4.

7.2. Model regresyjny z trzema zmiennymi.

Należy jeszcze zbadać zależność wielkości d (iloczas zestroju) od \bar{d} (skumulowanej wartości średnich iloczynów głošek składowych) równocześnie ze zmianą n (liczby głošek składowych). Można oczekiwać wysokiej korelacji między \bar{d} a n , lecz mimo to estymacja d z wielkości \bar{d} oraz n w regresji wielokrotnej może być lepsza niż w modelach dwuzmiennych. W ogólnym ujęciu otrzymujemy równanie regresji

$$d = a + bd + an$$

8. Analiza regresji.

8.1. Dwie zmienne : d , n .

8.1.1. Regresja liniowa.

W dwuzmiennym modelu regresyjnym ze zmiennymi d oraz n linia regresji będzie przechodzić przez środek układu i tworzyć kąt 45° z osią odciętych, dla braku izochronizmu, jeśli oś odciętych wyskalujemy w jednostkach względnych ($n - \bar{n}$) tj. biorąc miarę odchylenia od średniej wartości n (rozciągłości jednostki rytmicznej), zaś oś rzędnych w jednostkach równych \bar{p} , tj. średniej wartości ogólnej dla głoški. Równanie regresji przyjmie wówczas postać :

$$\frac{d - \bar{d}}{\bar{p}} = a + b(n - \bar{n}). \quad (3)$$

Idealnie, przy takiej transformacji a powinno być równe zero, ale w rzeczywistości $a \approx 0$, ponieważ \bar{p} jest obliczane z nieco innych danych wyjściowych. Jak wspomniano wcześniej, niektóre jednostki rytmiczne trzeba było pominąć, gdyż były niemierzalne. Wyniki obliczeń dla zmiennych d oraz n ukazuje Tablica 6. Wielkością, która jest bezpośrednio związana z izochronizmem jest bądź współczynnik regresji b , bądź $\arctg b$. Następujące wnioski nasuwają się na podstawie wyników zamieszczonych w Tablicy 6 :

(1) Współczynnik regresji dla ANAKRUZY jest zbliżony do jedności, a więc odpowiedni kąt nachylenia prostej regresji jest zbliżony do 45° . Oznacza to, że anakruza wykazuje tylko bardzo nieznaczną tendencję do izochronizmu.

(2) Ścisła jednostka rytmiczna ma współczynnik regresji, który jest około połowę mniejszy (dokładnie 0.613) od współczyn-

nika dla anakruzy, zaś odpowiedni kąt nachylenia ścisłej jednostki rytmicznej wynosi 0.69 kąta dla anakruzy. Tak więc istnieje zupełnie wyraźna tendencja do izochronizmu w ścisłej jednostce rytmicznej /por. ryc.(5)/.

(3) Współczynnik regresji dla "stopy" jest pośredni, ale bliższy współczynnikowi dla ścisłej jednostki rytmicznej.

(4) Współczynnik determinacji $r^2 \cdot 100\%$ wskazuje, że chociaż wyraźnie więcej wariancji pozostaje nie wyjaśnionej przez regresję dla ścisłej jednostki rytmicznej niż dla anakruzy, oba współczynniki można uznać za posiadające zadowalającą siłę predykcji.

Fakt, iż współczynnik korelacji i związany z nim współczynnik determinacji są wysokie dla anakruzy, a mniejsze dla ścisłej jednostki rytmicznej nasuwa przypuszczenie, że istnieje jeszcze prawdopodobnie jakiś czynnik od którego uzależniony jest iloczyn ścisłej jednostki rytmicznej, a który nie działa w anakruzie. Przypuszczalnie takim czynnikiem jest wspomniane wyżej zróżnicowanie pozycji wygłosowej wobec pozawygłosowej, jako że anakruza z definicji ma wyłącznie pozycję pozawygłosową. Współczynnik korelacji dla "stopy" jest niemal dokładnie pośredni między odpowiednimi współczynnikami na anakruzy i stopy. Pozwala to przypuszczać, iż model izochronizmu oparty na stopie nie jest wprawdzie niedopuszczalny, lecz zamazuje on istotną statystycznie różnicę między anakruzą a ścisłą jednostką rytmiczną, które to rozróżnienie pozwala, jak wykazano w pracy: Jassem 1952, w sposób silnie wyjaśniający skojarzyć płaszczyznę fonologiczną z płaszczyzną syntaktyczną.

6.1.2. Regresja kwadratowa.

Tablica 7 zawiera wyniki analizy regresji d względem n oraz n^2 . Po normalizacji zmiennych równanie regresji kwadratowej przyjmuje postać :

$$\frac{d - \bar{d}}{\bar{p}} = a + b(n - \bar{n}) + c(n - \bar{n})^2 \quad (4)$$

Dane w Tablicy 7 ukazują, że współczynniki determinacji dla wszystkich trzech rodzajów jednostek rytmicznych są tylko nieznacznie lepsze niż w przypadku modelu liniowego, który jest prostszy w interpretacji. Dlatego dodatkowa korzyść

z zastosowania równania kwadratowego jest wątpliwa.

8.2. Zmienne \bar{d} i \bar{d} .

Możliwe jest również estymowanie długości jednostki rytmicznej ze skumulowanego iloczynu średnich długości głosek wchodzących w skład jednostki. Innymi słowy bierzemy teraz średnie iloczyny głosek tworzących daną jednostkę rytmiczną według danych w Tabelicy 3, dodajemy te średnie i otrzymujemy ustabilizowany estymator \bar{d} szacujący wartość d . Jeśli brak jest izochronizmu, wówczas, średnio biorąc, $d = \bar{d}$. Jeśli jednak izochronizm zachodzi, wówczas d jest stałe, wobec czego musi istnieć współczynnik kompresji, którego iloczyn z wartością \bar{d} dawałby właściwy estymator wartości d . Zależność d od \bar{d} wyrażono następującym równaniem liniowym :

$$\frac{d}{\bar{p}} = a + b \frac{\bar{d}}{\bar{p}} \quad (5)$$

gdyż wtedy kąt nachylenia prostej regresji do osi odciętych jest łatwo interpretowalny.

Wyniki analizy regresji zawarte są w Tabelicy 8. Ponieważ średnie dla d są znormalizowane przez odpowiednie wartości \bar{p}_{Stopa} , \bar{p}_{SJR} , $\bar{p}_{\text{Anakruza}}$, reprezentują one oszacowanie średniej długości jednostek rytmicznych w terminach liczby składowych głosek. Współczynnik korelacji jest obecnie, podobnie jak w przypadku regresji d względem n , najwyższy dla anakruzy, a najmniejszy dla SJR, ale wszystkie współczynniki korelacji, a co za tym idzie, wszystkie współczynniki determinacji są wyraźnie wyższe niż odpowiednie wartości w Tabelicy 6. Tak więc siła predykcji modelu regresji d względem \bar{d} jest większa niż w poprzednich modelach dwuzmiennych. Wszystkie współczynniki determinacji są bardzo wysokie. Ryc. 6 ukazuje krzywe regresji w tym modelu w układzie współrzędnych $\frac{d}{\bar{p}}$ oraz $\frac{\bar{d}}{\bar{p}}$, zaś na ryc. 7 przeniesiono odpowiednie proste umieszczając je z zachowaniem kąta we wspólnym początku układu.

8.3. Regresja z trzema zmiennymi.

Chociaż modele dla dwóch zmiennych są zupełnie zadowalające, szczególnie w przypadku predykcji iloczynu zestroju akcentowego z sumy średnich czasów głosek składowych, należy zbadać, czy nie uzyska się dalszego polepszenia modelu przyjmując

regresję d względem \tilde{d} , jak i n. Wytestowano następujące równanie regresji wielokrotnej :

$$\frac{d}{p} = a + b \frac{\tilde{d}}{p} + cn, \quad (6)$$

a Tablica 9 podaje wyniki analizy regresyjnej. W idealnym przypadku wartości \bar{n} , \bar{d}/\bar{p} oraz \tilde{d}/\bar{p} powinny być równe. Nieznaczące różnice wynikają z pewnych różnic w danych wyjściowych. Współczynnik korelacji dla $(n, \tilde{d}/\bar{p})$ jest z natury rzeczy bardzo wysoki, jako że sumaryczna długość jednostki rośnie w przybliżeniu proporcjonalnie do liczby składowych głosek. Dla $(n, d/\bar{p})$ oraz $(d/\bar{p}, \tilde{d}/\bar{p})$ współczynniki korelacji są najwyższe dla anakruzy, a najniższe dla ścisłej jednostki rytmicznej, przy czym wszystkie wartości są wysokie. Jak widać z wartości współczynników determinacji, model z trzema zmiennymi wyjaśnia znacznie więcej zmienności niż którykolwiek z modeli dwuzmiennych, ale jest porównywalny z modelem dla d i \tilde{d} .

9. Aspekty lingwistyczne.

Jak to wyjaśniono w pracach Jassema 1949 i 1952, jeśli przyjmie się model rytmu tam zaproponowany, wówczas rytm można bardzo prosto zaznaczyć w sposób jednoznaczny w transkrypcji fonematycznej angielskiego tekstu przez zastosowanie kilku reguł (p. wyż. str. 7-8). Początek tekstu 30 w Halliday 1970 przetranskrybowany zgodnie z podanymi regułami przedstawia się następująco :

'ɪzðæt 'tʃɪz tə'ɪt ə'ɪzɪt
 tə'pɑdɪn ə'maʊs ,træp
 ɪfɑ:d'nɪʒən 'ɔ:lɪər ə'baʊt ʒə'wedɪŋ aɪdɔv,sent
 ə'keɪbl̩.

Początek tekstu 39 :

aɪ'læɪk ʒæt'pɜ:m baɪdæt'levnɪzr 'bɔ:ld p'streɪlɪən
 bɔ:l. 'dɪdʒə 'sɪlt.

Zestroje is that, cheese oraz earlier są przykładami jednostek rytmicznych ścisłych, które pokrywają się z pełnymi jednostkami rytmicznymi, jako że nie zawierają anakruzy. To act, or is it, to put in, if I'd known są przykładami pełnych jednostek rytmicznych rozpoczynających się od anakruzy. W systemach proponowanych przez Abercrombiego (1964) oraz Wittena (1977) nie zaproponowano oznaczenia rytmu w bieżącej

transkrypcji.

Jak wspomniano wyżej, obie propozycje, A oraz B, są niezależne od składni, lecz obie wskazują na związki między rytmem a składnią. Model Abercrombiego w interpretacji Hallidaya (1970) prowadzi do postulowania bardzo osobliwych zestrojów wyższego rzędu (zestrojów tonalnych), takich jak //if I'd known earlier about the wedding I'd have // sent a cable//. Pierwszy zestrój zawiera całe zdanie poboczne wraz z podmiotem i częścią orzeczenia zdania głównego, a drugi zestrój - pozostałość zdania głównego. Wiele przykładów takich dziwnych podziałów można znaleźć w pracy Hallidaya (1970). Oto kilka dalszych przykładów : // he was gay and he was woolly and his // pride was inordinate. he //danced on a sand-bank in the // middle of Australia and he // went to the Big God Ngong // (str. 121). // on the Isle of Man you can // still ride in a horse drawn trap// (str. 117). Takie dziwne podziały powstają na skutek przyjęcia zasady, że wszelkie sylaby niekcentowane należą do jednego zestroju akcentowego i intonacyjnego wraz z poprzedzającą sylabą akcentowaną, a także przyjęcia pauzy jako wyznacznika granicy między zestrojami tonalnymi¹. Model zaproponowany w ramach teorii B nie prowadzi do tak rażącej dyskrepancji między strukturą fonologiczną a syntaktyczną.

10. Podsumowanie i wnioski.

Celem określenia izochronizmu w terminach kwantytatywnych zastosowano metodę statystyczną analizującą iloczas zestrojów akcentowych (jednostek rytmicznych) za pomocą kilku modeli regresyjnych, w których zmienną niezależną była liczba głosek w zestrój albo suma średnich iloczasów głosek tworzących zestrój. Materiał stanowiły zapisy na taśmie dwóch tekstów mówionych w stylu standardowej angielszczyzny brytyjskiej o łącznej rozciągłości ok. 2500 głosek, których iloczas określono na podstawie pomiarów spektrograficznych. Głoski podzielono na statystycznie jednorodne klasy, a średnie wartości iloczasu w każdej klasie były wykorzystywane w modelach regresyjnych.

Poddano sprawdzeniu empirycznemu dwie teorie rytmu mówionej

¹ O dystrybucyjnym wyznaczaniu granicy między zestrojami tonalnymi, zob. Jassem 1978.

angielszczyzny. Pierwsza, lansowana przez Abercrombiego, postuluje jeden rodzaj jednostki rytmicznej zwanej "stopą". Druga, postulowana przez Jassemę, przewiduje dwa rodzaje jednostek : anakruzę, która nie wykazuje izochronizmu oraz ścisłą jednostkę rytmiczną, która jest nacechowana izochronizmem. W czterech modelach regresyjnych przyjęto zależność jednostki rytmicznej (a) liniową od liczby głosek, (b) liniowo-kwadratową od liczby głosek, (c) liniową od iloczasu sumarycznego głosek tworzących jednostkę oraz (d) liniową wielokrotną od iloczasu sumarycznego i liczby głosek. W podanej kolejności współczynniki determinacji rosną dla wszystkich trzech typów jednostek : stopy, ścisłej jednostki oraz anakruzy, przy czym regresja (c) daje zarazem w wyniku wysokie współczynniki determinacji i prostą interpretację umożliwiającą ilościowe określenie siły izochronizmu w postaci kąta odpowiadającego nachyleniu prostej regresji względem osi odciętych. Biorąc ten kąt jako miarę siły izochronizmu otrzymuje się następujące, zasadnicze wyniki analizy : Nasilenie izochronizmu w anakрузі jest nieznaczne, zaś w ścisłej jednostce rytmicznej bardzo wyraźne, przy czym różnica jest istotna statystycznie na poziomie $\alpha = 0.0001$. Stopa wykazuje wprawdzie też izochronizm, ale wyróżnienie w sygnale mowy tylko stopy zacierą tę ważną różnicę.

Teoria rytmu mowy postulująca anakruzę i ścisłą jednostkę rytmiczną pozwala w pełni jednoznacznie określić rytm w transkrypcji fonematycznej tekstu. Ponadto teoria ta umożliwia proste określenie relacji między płaszczyzną fonologiczną a płaszczyzną syntaktyczną, która to relacja ulega silnemu zachwianiu przy zastosowaniu teorii Abercrombiego i Hallidaya.

Teoria rytmu i izochronizmu zaproponowana przez Jassemę umożliwia stworzenie prostych reguł s y n t e z y opartej na wytwarzaniu tekstu z konkatenacji głosek.

Tablica 1.

Analiza statystyczna iloczasu głoskowego.

Ranga	Typ	Srednia [ms]	Wariancja [ms ²]	OchYLENIE standardowe [ms]	Współczynnik zmienności [%]
1	F	17	33	6	34
2	D	48	303	17	37
3	G	56	501	22	40
4	E	59	643	25	43
5	B	60	487	22	37
6	N	62	809	28	46
7	H	65	1165	34	53
8	K	65	524	23	35
9	Z	70	1145	34	48
10	SC	76	685	26	34
11	KH	85	539	23	27
12	S	88	1158	34	39
13	AFV	94	833	29	31
14	O	96	1551	39	41
15	KHA	117	1075	33	28
16	AF	126	836	29	23
17	A	132	2972	55	41
18	FTH	137	709	27	19

Tablica 2

Jednokierunkowa analiza wariancji iloczasu głoskowego

Wariancja	SS	df	MS	F _{obl.}	F _{.05}	F _{.01}
między	1 636 860	17	96 286.1	98.985	1.62	1.96
błąd	2 413 350	2 481	972.7			
całkowita	4 050 210	2 498				

Tablica 3.

Sredni iloczas głosek w jednostkach rytmicznych.

Typ	STOPA		SJR		ANA	
F	17	16.8	9	18.5	10	14.5
D	135	48.4	73	49.8	74	45.2
G	193	56.2	166	57.9	37	43.2
E	513	59.8	303	66.0	244	50.1
B	131	60.4	108	60.5	26	59.6
N	259	62.2	208	62.9	65	60.5
H	37	65.8	33	69.3	5	35.6
K	179	66.1	144	68.5	43	55.4
Z	69	73.2	57	77.2	20	55.4
SC	44	75.9	38	77.6	6	65.0
KH	49	86.3	37	88.5	12	77.4
S	165	90.4	144	93.4	27	65.5
AFV	18	93.8	16	95.3	1	80.6
O	141	97.3	123	100.0	25	78.8
KHA	41	118.3	42	118.0	-----	
AF	30	127.7	30	127.7	-----	
A	231	134.2	199	139.9	47	100.2
FTH	6	137.5	5	136.5	-----	

Tablica 4.

Względny iloczas głosek.

Typ	STOPA	ANA
F	0.91	0.78
D	0.97	0.91
G	0.97	0.75
E	0.91	0.76
B	1.00	0.99
N	0.99	0.96
H	0.95	0.51
K	0.96	0.81
Z	0.95	0.72
SC	0.97	0.84
KH	0.98	0.87
S	0.97	0.70
AFV	0.98	0.85
O	0.97	0.79
KHA	1.00	----
AF	1.00	----
A	0.96	0.72
PTH	1.01	----

Tablica 5

Analiza statystyczna jednostek rytmicznych

Jednostka rytmiczna	Ogólna średnia wartość iloczasu głoskowego [ms]	Średni iloczyn jednostki rytmicznej [ms]	Średnia liczba głosek (średnia długość)	Średnia prędkość przepływu głosek (głoski/sek)
Stopa	75.32	427	5.7	13.3
SJR	80.90	339	4.2	12.4
ANA	56.57	162	2.9	17.7

Tablica 6

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

	$n - \bar{n}$	$\frac{d - \bar{d}}{p}$	$\delta(n - \bar{n})$	$\frac{d - \bar{d}}{p}$	r	$r^2 \cdot 100\%$	a	b	arctg b
Stopa	0.00386	-0.00007	2.21	1.98	0.72	51.7	-0.0026	0.644	32.8°
SJR	0.00349	0.00019	1.63	1.47	0.62	39.0	-0.0018	0.561	29.3°
ANA	0.00383	0.00013	1.53	1.68	0.83	69.6	-0.0036	0.916	42.5°

Tablica 7

$$\frac{d - \bar{d}}{p} = a + b(n - \bar{n}) + c(n - \bar{n})^2$$

	$(n - \bar{n})^2$	$\delta(n - \bar{n})^2$	$r[(n - \bar{n})^2, d]$	$r[(n - \bar{n}), d]$	$r[(n - \bar{n}), (n - \bar{n})^2]$	a	b	c	$r^2 \cdot 100\%$
Stopa	4.90	7.14	0.29	0.72	0.27	-0.141	0.619	0.0282	52.7
SJR	2.67	4.52	0.37	0.62	0.47	-0.078	0.523	0.0287	39.6
ANA	2.34	3.81	0.55	0.83	0.56	0.117	0.844	0.0515	70.3

Tablica 8

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

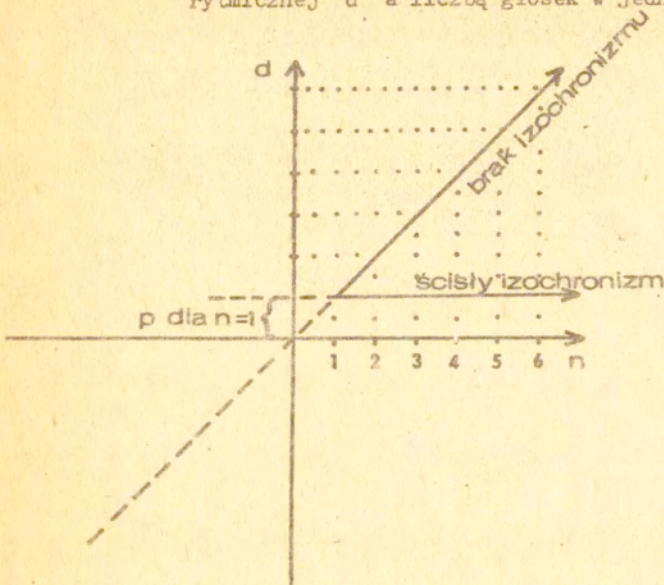
	$\frac{\tilde{d}}{\bar{p}}$	$G\left(\frac{\tilde{d}}{\bar{p}}\right)$	r	r ² , 100%	a	b	arctg b
Stopa	5.74	2.09	0.80	63.8	-4.35	0.758	37.1°
SJR	4.25	1.48	0.73	52.8	-3.07	0.724	35.9°
ANA	2.92	1.58	0.90	80.4	-2.78	0.951	43.6°

Tablica 9

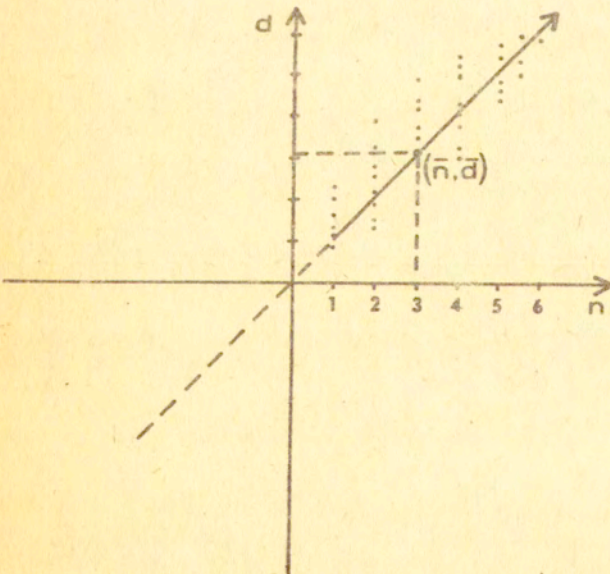
$$\frac{\tilde{d}}{\bar{p}} = a + b \frac{\tilde{d} + cn}{\bar{p}}$$

	$\frac{\tilde{d}}{\bar{p}}$	\bar{n}	$G(n)$	$G\left(\frac{\tilde{d}}{\bar{p}}\right)$	$r\left(n, \frac{\tilde{d}}{\bar{p}}\right)$	$r\left(n, \frac{\tilde{d}}{\bar{p}}\right)$	$r\left(\frac{\tilde{d}}{\bar{p}}, \frac{\tilde{d}}{\bar{p}}\right)$	a	b	c	r ² , 100%
Stopa	5.76	5.76	2.21	1.98	0.93	0.72	0.80	1.33	-0.187	0.942	64.4
SJR	4.25	4.25	1.63	1.47	0.91	0.62	0.73	1.11	-0.221	0.950	53.8
ANA	2.94	2.92	1.53	1.68	0.96	0.83	0.90	0.17	-0.378	1.302	81.3

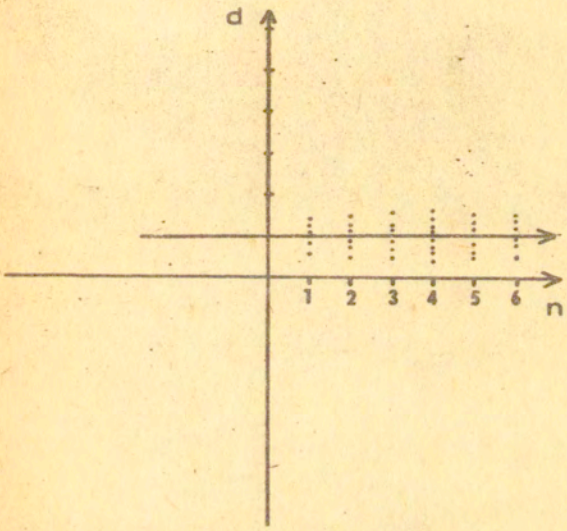
Ryc. 1. Związek funkcyjny między iloczasetem jednostki rytmicznej d a liczbą głosek w jednostce n .



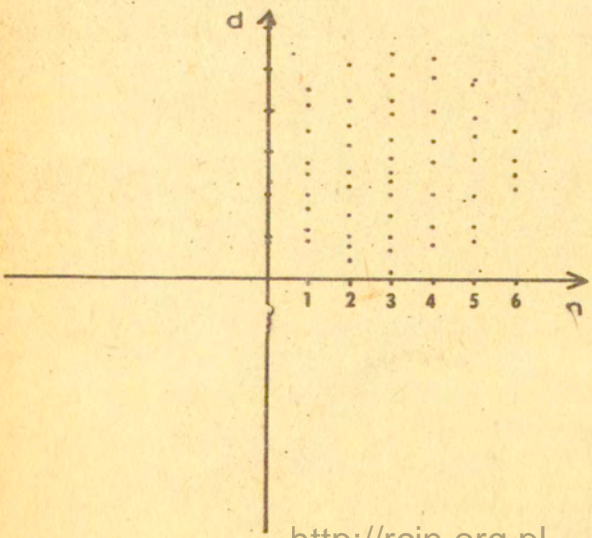
Ryc. 2. Model regresyjny związku między iloczasetem jednostki rytmicznej d a liczbą głosek w jednostce n .



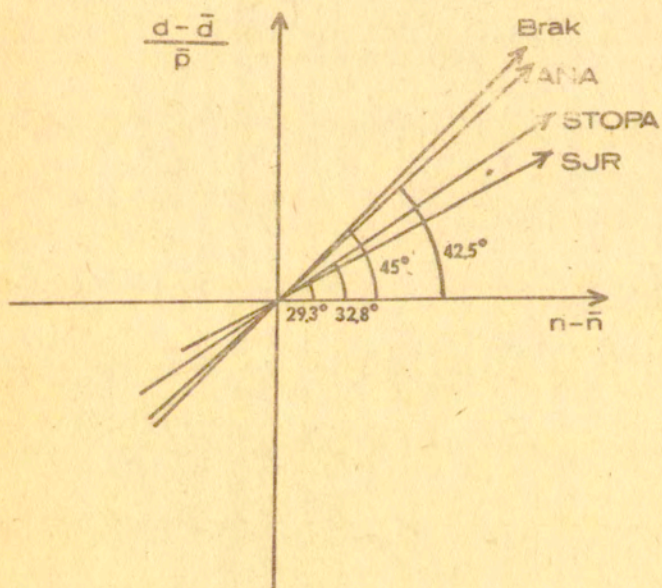
Ryc. 3. Współczynnik regresji $b = 0$. Wartości d w pobliżu linii ścisłego izochronizmu.



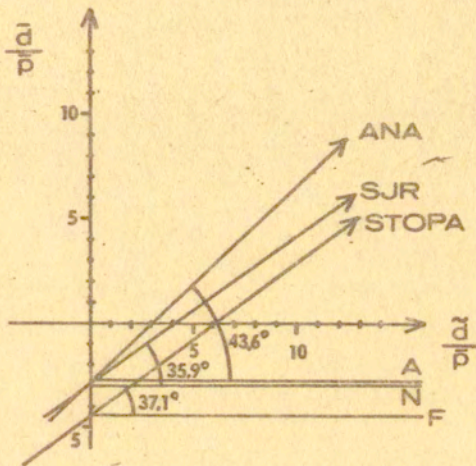
Ryc. 4. Współczynnik regresji $b=0$. Wartości d całkowicie przypadkowe.



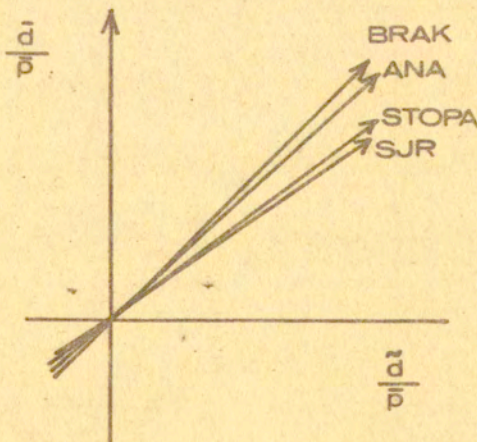
Ryc. 5. Liniowa regresja d względem n w stopie, ścisłej jednostce rytmicznej i anakruzie.



Ryc. 6. Liniowa regresja \bar{d} względem \bar{d} w stopie, ścisłej jednostce rytmicznej i anakruzie wg Tablicy 8.



Ryc. 7. Liniowa regresja \bar{d} względem \bar{d} . Linie regresji przesunięte do wspólnego punktu.



BIBLIOGRAFIA

- 1 ABERCROMBIE, D. : Syllable quantity and enclitics in English, in : In Honour of Daniel Jones, D.Abercrombie et al. eds., Longmans, London, 1964, 216-222.
- 2 ABERCROMBIE, D. : Elements of General Phonetics, Edinburgh Univ. Press, 1967.
- 3 ABERCROMBIE, D. : A phonetician's view of verse structure, in : Phonetics in Linguistics, W.E. Jones a. J. Laver eds., Longman, London, 1973, 6-13.
- 4 ADAMS, C. : English Speech Rhythm and the Foreign Learner, Mouton Publ., The Hague, 1979.
- 5 ALLAN, G.D. : On testing for certain stress-timing effects, Working Papers in Phonetics, UCLA No. 10, 47-59. 1968.
- 6 BOLINGER, D. : Pitch accent and sentence rhythm in : D.L. Bolinger, Forms of English, Hokuou Publ. Co., Tokyo, 1965, 139-180.
- 7 GABRIEL, K.R. : A procedure for testing the homogeneity of all sets of means in analysis of variance, Biometrics v. 20 3 459-477, 1964.
- 8 HALLIDAY, M.A.K. : A Course in Spoken English : Intonation, Oxford Univ. Press, Oxford, 1970.
- 9 HOCKETT, Ch.F. : A Manual of Phonology, Indiana Univ. Publications in Anthropology and Linguistics, Memoir 11, 1955.
- 10 JASSEM, W. : Indication of rhythm in the transcription of Educated Southern English, Le Maitre phonetique III/92, 22-24, 1949.
- 11 JASSEM, W. : Stress in Modern English, Bulletin de la Société Linguistique Polonaise XII, 189-194, 1952.
- 12 JASSEM, W. : Fonetyka języka angielskiego, PWN, Warszawa, 1980.
- 13 JASSEM, W. : Podręcznik wymowy angielskiej, PWN, Warszawa, 1981.
- 14 JONES, D. : Outline of English Phonetics, Heffer and Sons Ltd., Cambridge 9th ed. repr., 1976.
- 15 LADEFOGED, P. : A course in Phonetics, Harcourt, Brace, Jovanovich Incl. New York, 1975.
- 16 LEA, W.A. : Prosodic aids to speech recognition : IV. A general strategy for phonologically-guided speech understanding, Univac Rep. No. PK 10791, 1974.
- 17 LEHISTE, I. : Rhythmic units and syntactic units in production and perception, Journ. Acoust.Soc.Am. vol. 54, 1228-1234, 1973.
- 18 LEHISTE, I. : The role of temporal factors in the establishment of linguistic units and boundaries, Phonologica 1972, W.U. Dressler a.F.V.Maros, eds.115-122.

- 19 LEHISTE, I. : Isochrony reconsidered, *Journal of Phonetics*, vol. 5, 253-263, 1977.
- 20 O'CONNOR, D. : The perception of time intervals, *Progress Report*, Sept. 1965. *Phonetica Lab.*, UCL 11-13, 1965.
- 21 O'CONNOR, D. : *Better English Pronunciation*, Cambridge Univ. Press, Cambridge, 1967.
- 22 O'CONNOR, D. : The duration of the feet in relation to the number of component sound-segments, *Progress Report* June 1968, *Phonetics Lab.* UCL 1-13, 1968.
- 23 JASSEM, W. : On the distributional analysis of pitch phenomena, *Language and Speech*, vol. 21, 362-372, 1978.
- 24 PIKE, K.L. : *The Intonation of American English*, Ann Arbor, Univ. of Michigan Press, 1945.
- 25 SHEN, Y., PETERSON, G.G. : Isochronism in English, *Univ. of Buffalo Studies in Linguistics*, Occasional Papers No. 9, 1-36, 1972.
- 26 ULDALL, E., *Isochronous stresses in R.P.*, in : *Form and Substance*, L.L. Hammerich et al. eds., Akademisk Forlag, Odense, 205-210.
- 27 WITTEN, I.R. : Flexible scheme for assigning timing and pitch to synthetic speech, *Language and Speech*, vol. 20, 240-260, 1977.