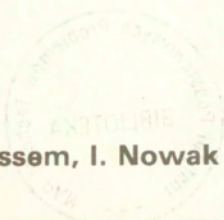


3.10. – akustyka (akustyka mowy)


W. Jassem, I. Nowak

WIELOWYMIAROWE SKALOWANIE
PRZESTRZENI PERCEPCYJNEJ
SAMOGŁOSEK PODSTAWOWYCH

21/1993

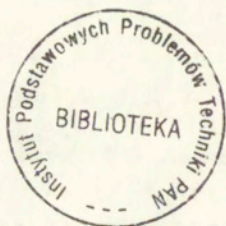
P. 269



W A R S Z A W A 1 9 9 3

<http://rcin.org.pl>

Praca wpłynęła do Redakcji dnia 15 grudnia 1992 r.



56670



Na prawach rękopisu

Instytut Podstawowych Problemów Techniki PAN
Nakład 100 egz. Ark.wyd.1,50 Ark.druk.2,00
Oddano do drukarni w maju 1993 r.

Wydawnictwo Spółdzielcze sp. z o.o.
Warszawa, ul.Jasna 1

Wiktor Jassem
Ignacy Nowak
Zakład Fonetyki Akustycznej
IPPT PAN

WIELOWYMIAROWE SKALOWANIE PRZESTRZENI PERCEPCYJNEJ SAMOGŁOSEK
PODSTAWOWYCH

Streszczenie

Dotychczasowe prace doświadczalne w zakresie artykulacji, cech akustycznych i percepcji samogłosek podstawowych są niewystarczające, a ich wyniki są kontrowersyjne. Szczególnie ostrą krytykę samogłosek podstawowych przedstawił Butcher (1982), który, stosując metodę skalowania wielowymiarowego do oceny doświadczeń percepcyjnych oraz oceniając inne, nieliczne zresztą, prace doświadczalne, doszedł do wniosku, że system samogłosek podstawowych powinien, jako naukowo niewłaściwy, ulec likwidacji. Przeprowadzone w ZFA, zreferowane tutaj eksperymenty oraz ich opracowanie metodą skalowania wielowymiarowego dały wyniki pozytywne. Krytyka Butchera, w świetle uzyskanych obecnie rezultatów, jest nieuzasadniona. Samogłoski podstawowe układają się w konsekwentną konfigurację w percepcyjnej przestrzeni dwuwymiarowej.

1. Samogłoski podstawowe

Samogłoski podstawowe powstały w okresie 1910-1920 z inicjatywy fonetyka angielskiego D. Jonesa oraz francuskiego P. Passy w oparciu o wcześniej już poczynione spostrzeżenie, iż samogłoski różnych języków różnią się między sobą pod trzema względami: (a) artykulacji bardziej przedniej lub tylnej, (b) artykulacji bardziej przykniętej lub otwartej oraz (c) zaokrąglenia warg lub jego braku. Wyznaczono najpierw cztery samogłoski podstawowe o krańcowych cechach artykulacyjnych: (1) najbardziej przykniętej i przedniej, (2) najbardziej otwartej i przedniej, (3) najbardziej otwartej i tylnej oraz (4) najbardziej przykniętej i tylnej. Z tych czterech ostatniej przypisano cechę labializacji, tj. zaokrąglenia warg, zgodnie z obserwacją poczynioną wówczas na wielu językach, która wskazywała, że istnieje powszechna tendencja zaokrąglenia warg

przy artykulacji tej ostatniej. Od samego początku zatem istniały wprawdzie trzy parametry opisowe, ale tylko dwa były niezależne. Potwierdzenie poprawności początkowo intuicyjnie założonych parametrów D. Jones znalazł już ok. roku 1920 wykonując zdjęcia rentgenowskie artykulacji czterech najbardziej pierwotnych samogłosek podstawowych. Miały one służyć stworzeniu systemu klasyfikacyjnego pozwalającego na niezależny („międzynarodowy”) opis samogłosek dowolnego języka. Te rentgenogramy (po jednym dla każdej z czterech skrajnych samogłosek) zostały zamieszczone w pracy podręcznikowej Jones 1956 a.

Po II Wojnie Światowej rentgenogramy czterech pierwotnych samogłosek podstawowych wykonał W. Jassem wraz z radiologiem J. Grądzkim i ogłosił je m. in. w pracach Jassem 1973 i 1983. Samogłoski wymawiała L. Frąckowiak (obecnie L. Richter, adiunkt w Zakładzie Fonetyki Akustycznej).

Celem uzyskania odniesień klasyfikacyjnych określonych wymienionymi wyżej trzema parametrami D. Jones we wczesnych latach dwudziestych wyznaczył system 8 samogłosek podstawowych, którym nadał oznaczenia w międzynarodowej transkrypcji fonetycznej i które ujął w schemat o następującej postaci:

i	u	przymknięte
e	o	półprzymknięte
ɛ	ɔ	półotwarte
a	ɑ	otwarte
przednie	tylne	

Powyższy schemat jest nieznacznie uproszczony: W pierwotnej formie górny bok czworokąta jest ukośny i opada lekko w prawo. Jednak Jones dopuszczał kilka uproszczeń (por. p. 1956 b:37), m. in. poziome położenie górnego boku, a także pionowe położenie prawego boku. Tym ostatnim kształtem posługuje się do dzisiaj wielu fonetyków, zwłaszcza dla celów dydaktycznych (por. np. Roach 1983 lub Rausch & Rausch 1991).

W podanej wyżej klasyfikacji nadal parametr zaokrąglenia nie był niezależny. Założono, że w obrębie samogłosek tylnych labializacja ulega wzmocnieniu od stopnia zerowego przy [ɑ]

poprzez pośrednie: słabą przy [ɔ], silniejszą przy [o] do najsilniejszej przy [u].

Ten system 8 samogłosek podstawowych stał się w latach dwudziestych i trzydziestych bardzo popularny, szczególnie w Europie, dzięki silnemu wpływowi na fonetykę europejską szkoły D. Jonesa skupionej w Department of Phonetics, University College London, gdzie mieściło się też prezydium Międzynarodowego Stowarzyszenia Fonetycznego (International Phonetic Association). Stowarzyszenie to, skupiające szerokie kręgi fachowców w zakresie fonetyki, głównie w Europie, przyjęło system 8 samogłosek podstawowych za obowiązujący w klasyfikacji „dźwięków mowy” (termin wówczas nieostro zdefiniowany). Zakładano i wykładano (a czyni się to do dzisiaj w podręcznikach fonetyki ogólnej i fonetyki poszczególnych języków), że pod względem artykulacyjnym położenie na obwodzie czworoboku samogłoskowego odpowiada położenie *najwyższego punktu obrysu języka na wzdłużnym przekroju toru artykulacyjnego* (por. np. Jones 1956, Jones b oraz Jones & Ward 1969, Armstrong & Jones 1951 i in).

Po II Wojnie Światowej czworobok samogłoskowy został uzupełniony. Parametr labializacji uznano za niezależny i stworzono system 16 samogłosek podstawowych plus dodatkowe 2, co do których nie podjęto decyzji w odniesieniu do ich statusu „kardynalności”, ale które dość powszechnie uznano za ważne punkty odniesienia. Łącznie 18 samogłosek, które dla uproszczenia ujmemy wspólną nazwą „podstawowych” (ang. *cardinal*), tworzy obecnie następujący system:

i y	i ɥ	ɯ u	przymknięte
e ø		ɤ o	półprzymknięte
ɛ œ		ʌ ɔ	półotwarte
	a ɶ	a ɒ	otwarte
	przednie	tylne	

W ciągu ostatnich ok. 40 lat podręczniki fonetyki oraz wiele prac badawczych stosuje dwa „czworoboki samogłoskowe”, w dwóch wersjach. W pierwszej jeden czworobok odnosi się do tzw.

samogłosek „pierwotnych” (ang. *primary cardinal vowels*), a drugi do „wtórnych” (ang. *secondary cardinal vowels*). Tworzą one dwa systemy (tutaj uproszczone do postaci prostokątów):

i	u	y	ɨ	ʉ	ɯ
e	o	ø			ɤ
ɛ	ɔ	œ			ʌ
a	ɑ	ɑ			ɒ
samogłoski pierwotne			samogłoski wtórne		

lub alternatywnie

i	ɨ	ɯ	y	ʉ	u
e		ɤ	ø		o
ɛ		ʌ	œ		ɔ
a		ɑ	ɑ		ɒ
samogłoski niezaokrąglone			samogłoski zaokrąglone		

Pierwszy układ miał uzasadnienie historyczne: Pierwotny system (por. wyż. str. 4) oparty był na obserwacji tendencji występującej w licznych językach w odniesieniu do typu samogłosek, które artykułuje się na ogół z labializacją.

Jednakże od samego początku wiadomo było, że system ten nie był wystarczający, bo choćby w języku francuskim i niemieckim, które były brane pod uwagę już przez Jonesa i Passy'ego, istnieją samogłoski przednie przymknięte, półprzymknięte i półotwarte z labializacją. Od samego okresu powstawania transkrypcji międzynarodowej nadano im odpowiednio oznaczenia [y], [ø] oraz [œ], ale początkowo nie uznawano ich za „podstawowe”.

Drugi układ dzieli samogłoski ze względu na labializację.

Z punktu widzenia systemu mamy tu do czynienia z dwoma alternatywnymi układami, które w ciągu dziesięcioleci zdały praktyczny egzamin w postaci pomyślnych wyników w nauczaniu obcych języków oraz w niektórych pracach badawczych (np. z dialektologii). Jednakże naukowe i metodologiczne podstawy systemu samogłosek podstawowych są nadal nadzwyczaj kruche, w szczególności w obliczu od dawna już istniejących możliwości ich weryfikacji doświadczalnej.

Niezwykle ostrą krytykę systemu samogłosek podstawowych

przeprowadził Butcher (1982). Jego zdaniem „...continued use of the system has a positively dangerous effect in that it conditions phoneticians to an approach to their discipline which is theoretically inadequate and scientifically dishonest” (str. 50). Zdaniem Butchera jest najwyższy czas, by „wysłać samogłoski podstawowe na emeryturę” (w oryginale *retirement age*, loc. cit.). Krytykę swoją opiera Butcher przede wszystkim na braku jednoznacznych kryteriów metodologicznych określających podstawę klasyfikacyjną samogłosek podstawowych, początkowo rzekomo artykulacyjną, a następnie percepcyjną. W bardzo niewielkim stopniu uwzględnia Butcher zależności między czworobokiem (czworobokami) samogłosek podstawowych a ich *cechami widmowymi*.

Różni fonetycy w ostatnich 20 latach zakładali, że sam system jest poprawny, ale jego podstawy są *audytywne*, a nie artykulacyjne, jak twierdzili ich twórcy i pierwsi propagatorzy. Butcher (loc. cit.) wskazuje jednak, że nie tylko (sporadyczne zresztą) prace doświadczalno-artykulacyjne, lecz także badania percepcyjne przez niego przeprowadzone, wykorzystujące metody *skalowania wielowymiarowego*, dały wyniki negatywne, niezgodne z układem ani w pierwotnej postaci ze skośnymi bokami górnym, prawym i lewym, ani w żadnej uproszczonej.

Analizy *akustyczne* samogłosek podstawowych były dotychczas również bardzo ograniczone. Fant i Bladon (1978) zmierzili częstotliwości wszystkich ich czterech formantów przy okazji rozważań nad optymalizacją matematycznych sformułowań parametrów formantowych. Nieco wcześniej Jassem (1973) przedstawił również wyniki pomiarów czterech formantów samogłosek podstawowych. Jednakże w obu tych pracach materiał doświadczalny był bardzo ograniczony. Fant i Bladon (op. cit.) nie podają nawet, kto wymawiał analizowane przez nich samogłoski (z treści można się domyślić, że był to Bladon), ani czy wyniki pochodzą z pojedynczego wymówienia każdego typu, czy też podano wartości uśrednione (znowu domyślić się można, że analizowano pojedyncze wypowiedzi). Jassem (1973) oparł swoje pomiary na kilku powtórzeniach każdej samogłoski podstawowej, ale dana część pracy miała charakter pilotażowy.

Catford (1981) zaproponował bardzo oryginalny sposób naniesienia dwóch trapezoidów CV (będziemy w dalszym ciągu używać tego dogodnego skrótów z angielskiego *Cardinal Vowel(s)*), mianowicie osobno dla zbioru z labializacją i zbioru bez labializacji (por. wyż.) na nieortogonalne układy współrzędnych F_1 oraz F_2 . Zasugerował on stosowanie swoich dwóch „siatek” dla zobiektywizowanego przedstawienia samogłosek określonych języków i sam ich użył w swoich badaniach języków kaukaskich (Catford 1988). Nie wiadomo jednak, jak sobie Catford poradził z problemem efektu interakcji zmienności lingwistycznej i osobniczej. Już od kilku dziesięcioleci wiadomo, że bezwzględne wartości częstotliwości formantowych samogłosek zależą nie tylko od różnicowań lingwistycznych, lecz również osobniczych (por. np. Neary 1978, Papçun 1980), wobec czego analiza akustyczna CV musi te ostatnie uwzględnić.

Ryciny 1...8 ukazują porównawczo częstotliwości formantowe w jedynych dotychczas pracach podających wartości F_1 , F_2 , F_3 oraz F_4 dla wszystkich samogłosek podstawowych, mianowicie Fant i Bladon 1978, Jassem i Krzyško 1991 oraz Jassem 1973 (w tej kolejności). Ryciny te pozwalają wyprowadzić wnioski uogólniające w zakresie tych, jak na razie, szczupłych danych pomiarowych, tak w odniesieniu do zmienności lingwistycznej, jak i osobniczej.

Formant 1.

W obrębie samogłosek niezaokrąglonych zachodzi, dla wszystkich trzech źródeł, następująca regularność:

[i < e < ε < a > ɑ > ʌ > u].

Dla głosu WJ [w] > [i].

W terminach fonetycznych: częstotliwość F_1 wzrasta wraz ze stopniem otwarcia, przy czym [ɑ] w tym odniesieniu jest nieco bardziej przymknięte niż [a].

Powyższe relacje powtarzają się w samogłoskach zaokrąglonych:

[y < ø < œ < ɔ > ɔ > u ≅ u].

W głosie B [u < u].

Formant 2.

W samogłoskach zaokrąglonych zachodzi następująca regularność: Od pozycji [y] do pozycji [u], idąc po obrzeżu czworoboku

ruchem przeciwnym do ruchu wskazówki zegara, wartości F_2 konsekwentnie się zmniejszają:

[y > ø > œ > œ > a > o > u], ale [u < u].

Ta regularność załamuje się nieco w przypadku samogłosek niezaokrąglonych: [i > e > e > a > a]. W głosie WJ następuje dalszy, bardzo nieznaczny spadek do [u], natomiast w głosie B, F_2 samogłosek [ʌ, ʏ, u] jest niezmienna, ale wyższa niż w [a]. Ten szczegół wymagałby jeszcze wyjaśnienia. W obu głosach częstotliwość F_2 (i) > (u).

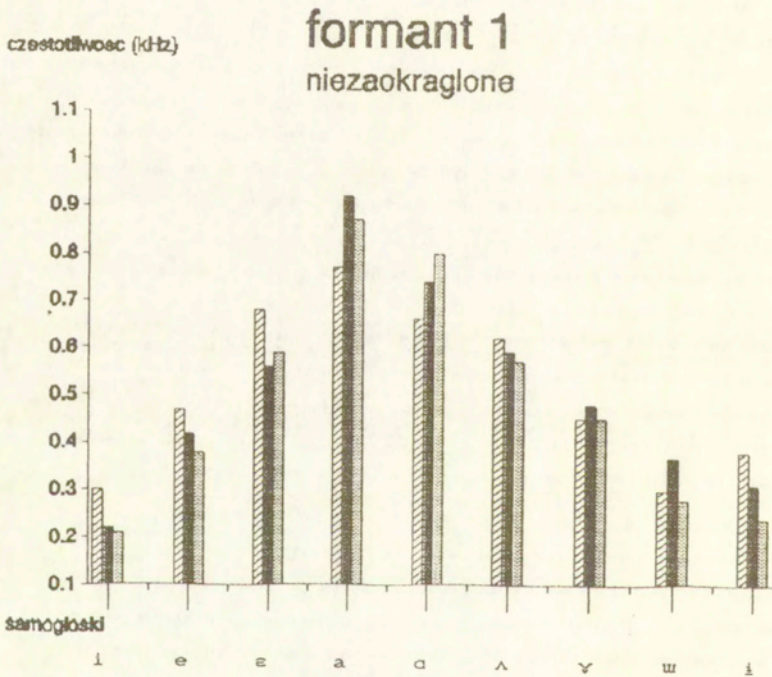
Formant 3.

Ani w obrębie samogłosek zaokrąglonych ani w zakresie niezaokrąglonych nie ma prostej zależności między częstotliwością formantu trzeciego a położeniem samogłoski na czworoboku. W zakresie niezaokrąglonych oba głosy są zgodne w odniesieniu do trzech samogłosek: [i > e > e]. W pozostałych niezaokrąglonych nie można dopatrzeć się regularności, a różnice są tak niewielkie, że mogą okazać się nieistotne. Podczas gdy w obrębie zaokrąglonych, F_3 samogłosek [o, o, u, u] wydaje się być bez znaczenia, ciekawe zależności widać w samogłoskach pozostałych, ale tutaj głosy wyraźnie się różnią: W głosie B [y < ø < œ < œ < o]. Natomiast u WJ zależności są następujące: [y > ø > œ < o]. Wobec szczupłości materiału doświadczalnego nie można obecnie nawet postawić hipotezy co do tego, czy mamy tu do czynienia ze zróżnicowaniem wyłącznie międzyosobniczym, czy też B i WJ wykazują tu nieznaczne różnice lingwistyczne.

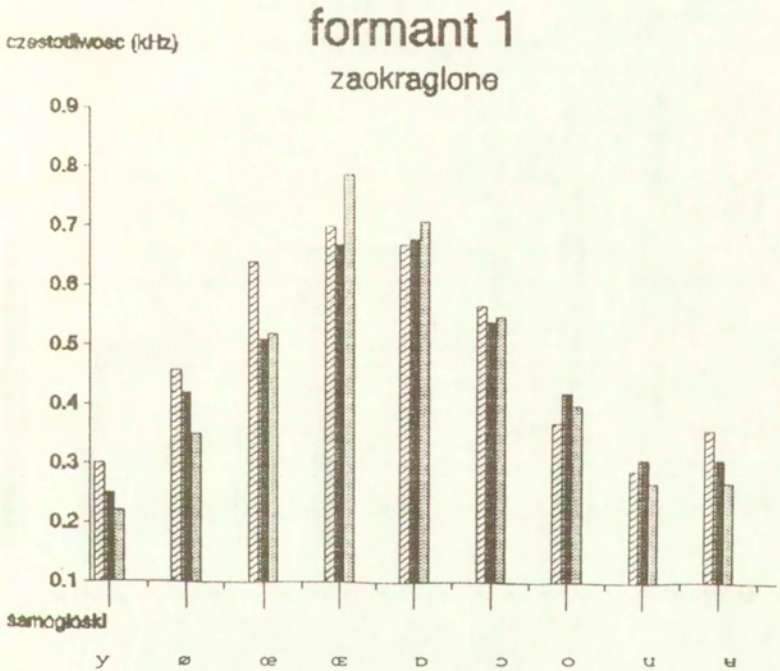
Formant 4.

Wyniki pomiarów nie wykazują żadnych wyraźnych zależności między częstotliwością tego formantu a położeniem samogłosek na czworoboku. Ponadto oba głosy znacznie się różnią w odniesieniu do względnych wartości F_4 . Uderzające są różnice pomiędzy B a WJ pod względem częstotliwości F_4 w samogłoskach przednich zaokrąglonych. Wartości w głosie B są tu znacznie niższe.

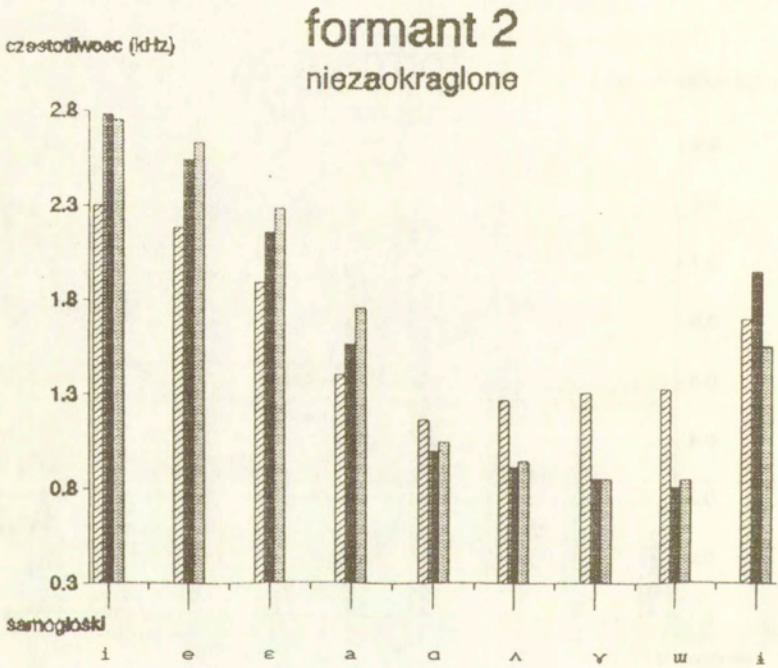
Wpływ labializacji na częstotliwości F_1 i F_2 omówimy krótko w dalszym ciągu, tutaj zwrócimy uwagę na fakt, że wyjąwszy parę [a, o], labializacja obniża wartość tak F_3 , jak i F_4 dla danej pozycji na czworoboku.



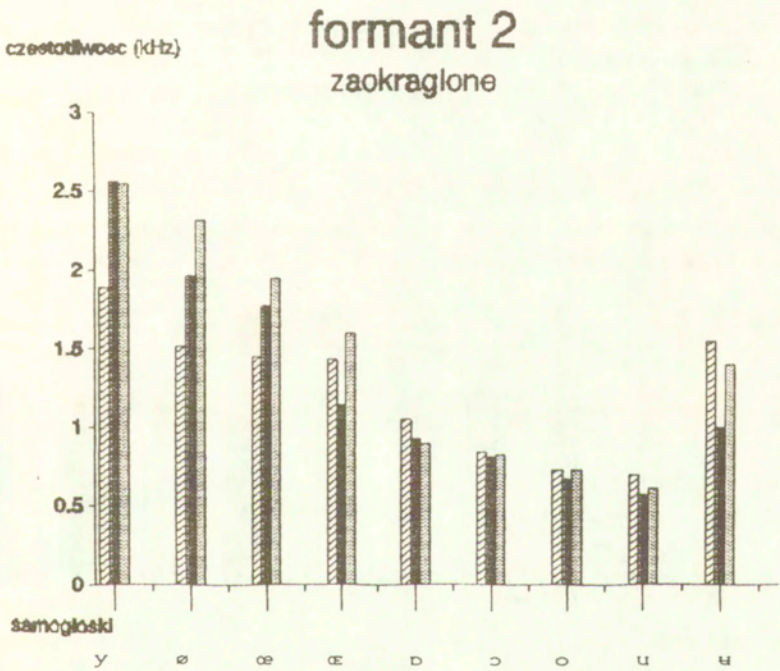
Ryc.1. Formant 1, samogłoski niezaokrąglone, wg Fant i Bladon 1978, Jassem i Krzyśko 1991, Jassem 1973.



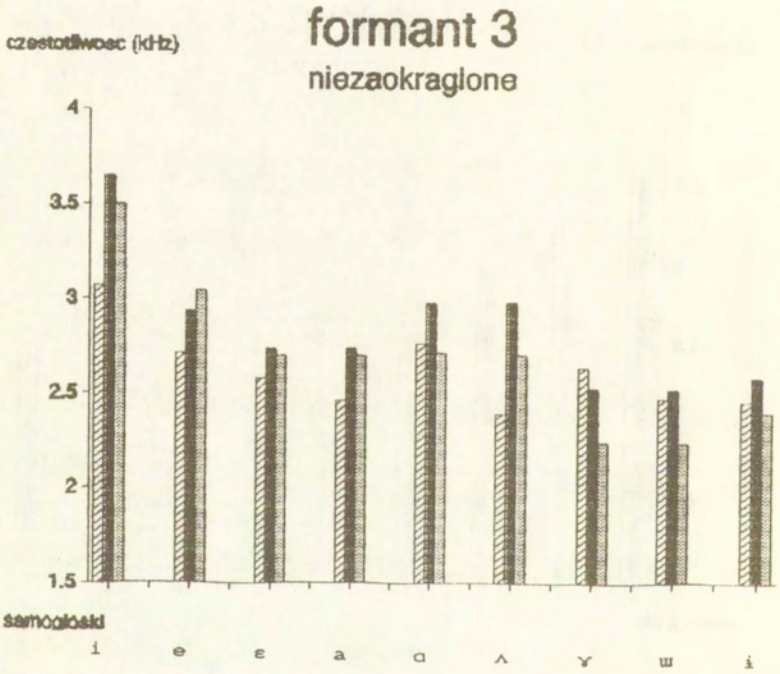
Ryc.2. Formant 1, samogłoski zaokrąglone, wg Fant i Bladon 1978, Jassem i Krzyżko 1991, Jassem 1973.



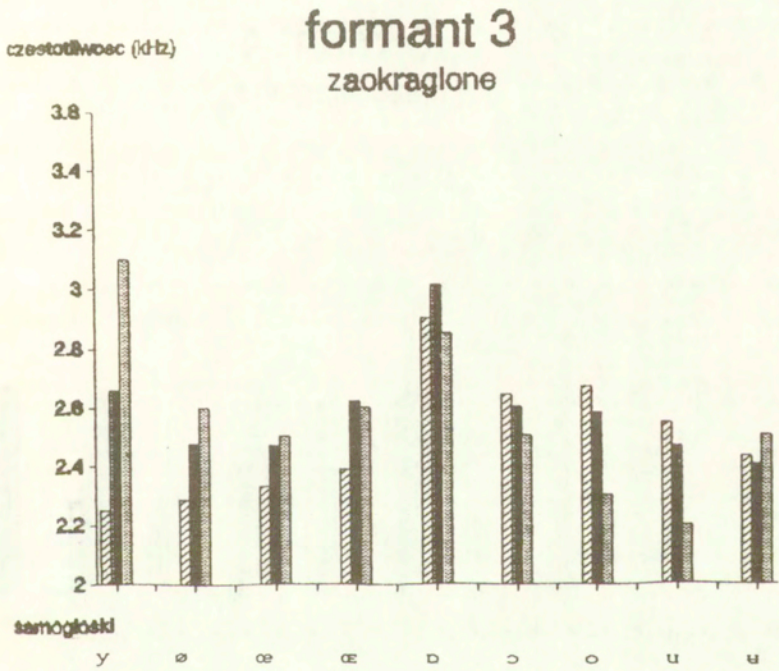
Ryc. 3. Formant 2 samogłoski niezaokrąglone, wg Fant i Bladon 1978, Jassem i Krzyśko 1991, Jassem 1973.



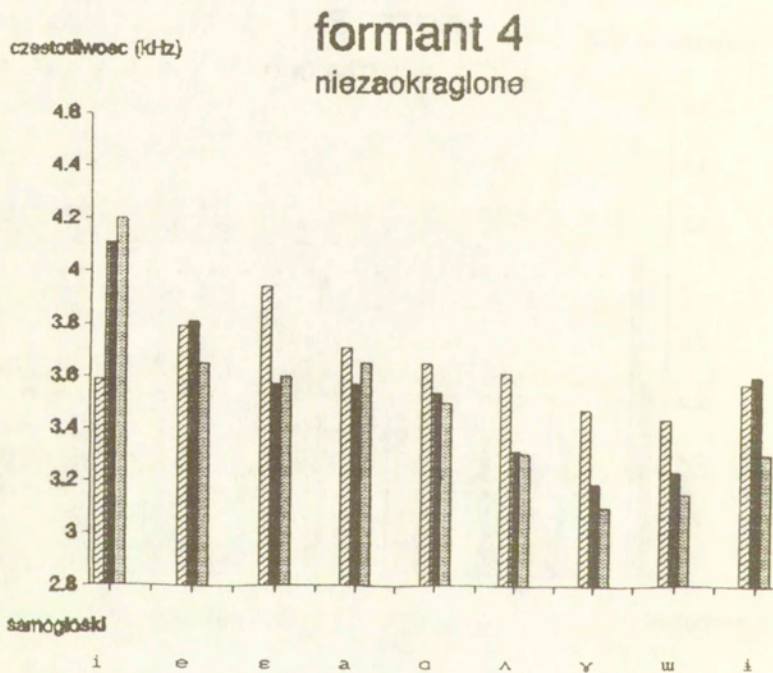
Ryc. 4. Formant 2, samogłoski zaokrąglone, wg Fant i Bladon 1978, Jassem i Krzyżko 1991, Jassem 1973.



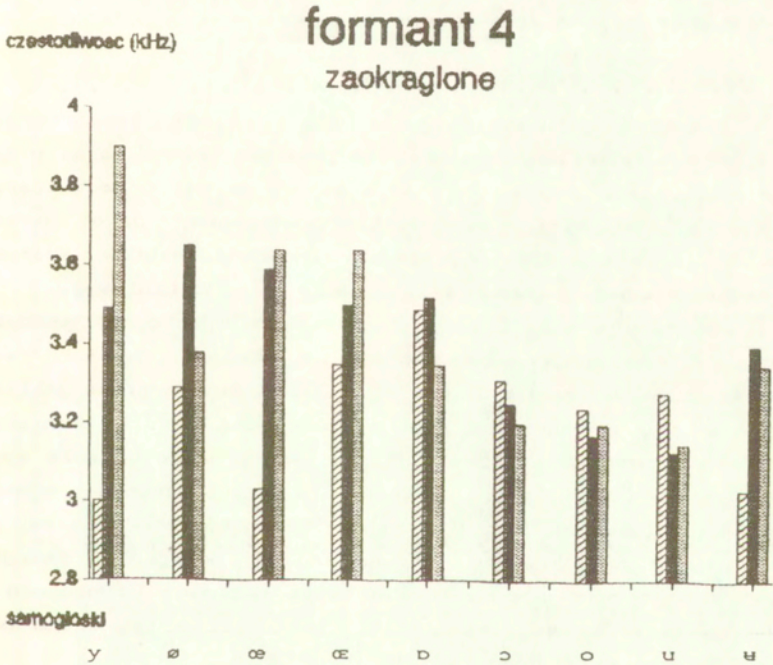
Ryc.5. Formant 3, samogłoski niezaokrąglone, wg Fant i Bladon 1978, Jassem i Krzyško 1991, Jassem 1973.



Ryc. 6. Formant 3, samogłoski zaokrąglone, wg Fant i Bladon 1978, Jassem i Krzyśko 1991, Jassem 1973.



Ryc.7. Formant 4, samogłoski niezaokrąglone, wg Fant i Bladon 1978, Jassem i Krzyśko 1991, Jassem 1973.



Ryc.8. Formant 4, samogłoski zaokrąglone, wg Fant i Bladon 1978, Jassem i Krzyśko 1991, Jassem 1973.

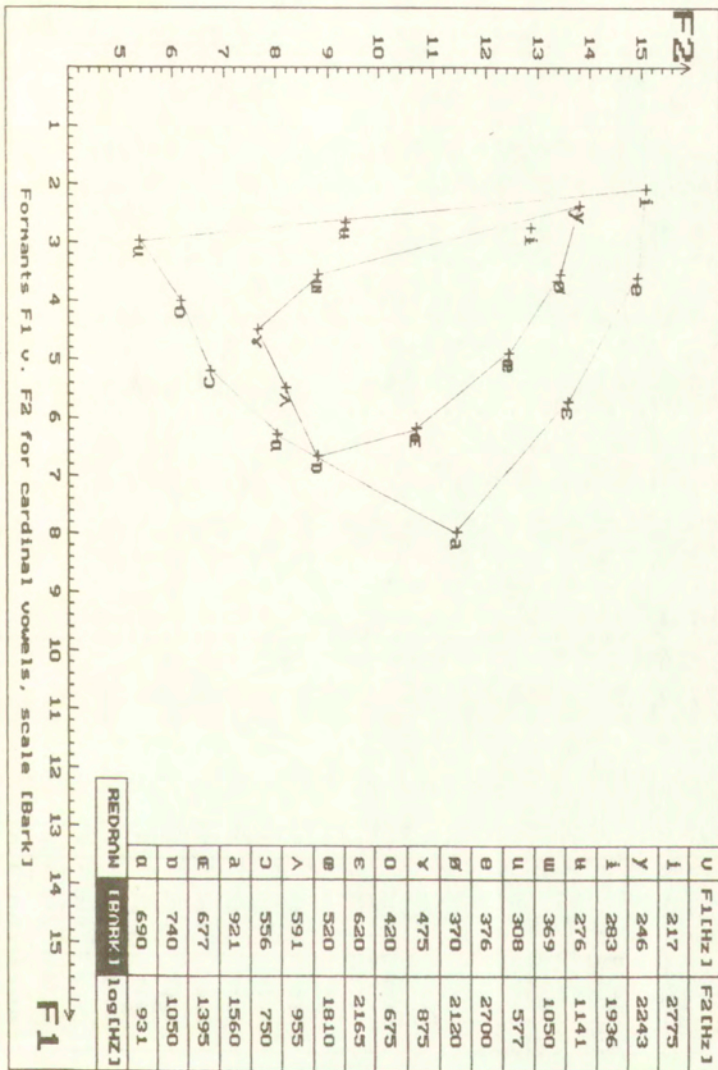
Niektóre z powyższych regularności były już wcześniej zauważane przez różnych badaczy (zwłaszcza w odniesieniu do F_1 i F_2) bądź to przy sposobności analiz samogłosek poszczególnych języków, bądź też - jak np. u Fanta 1960 - w związku z modelami akustycznymi artykulacji samogłoskowej. Formułujemy je tutaj po to, by ukazać, że czworoboki samogłosek podstawowych mają regularne odzwierciedlenie co najmniej w częstotliwościach F_1 i F_2 . Tym samym okazuje się, że krytyka Butchera w każdym razie nie stosuje się do relacji między podstawowymi cechami widmowymi a układem samogłosek podstawowych na czworoboku (czworobokach). Istnieje tutaj duża zbieżność, która w pełni uzasadnia stosowalność samogłosek podstawowych w klasyfikacji samogłosek konkretnych języków.

2. Materiał doświadczalny.

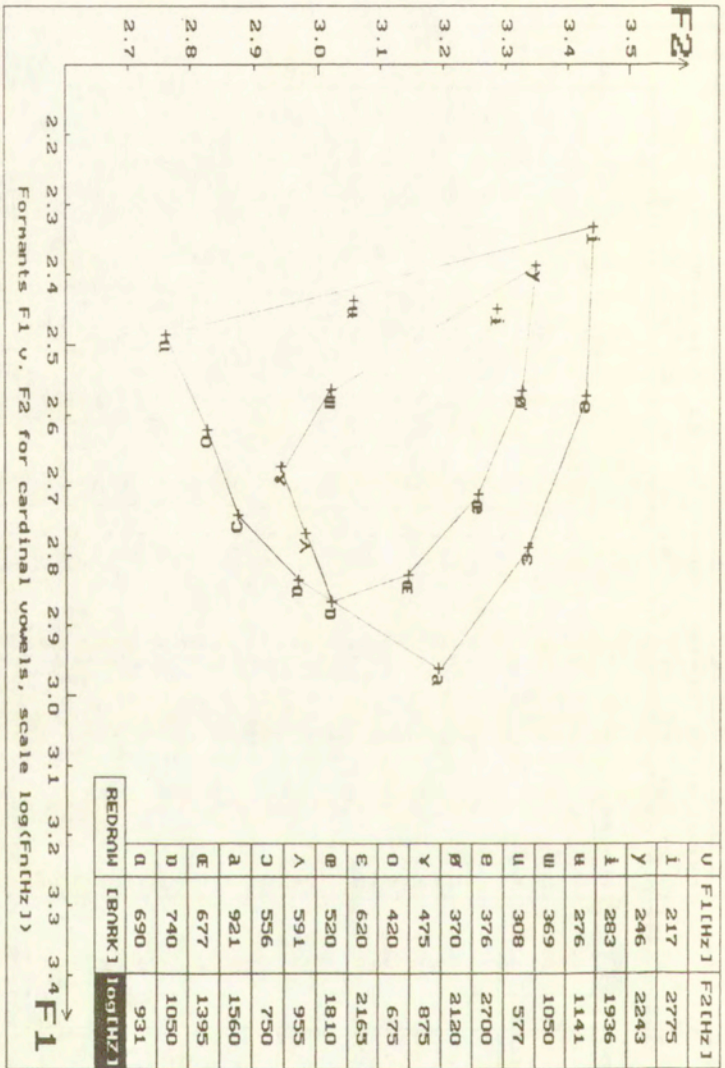
Wykorzystując syntetyzator cyfrowy SMOK konstrukcji M. Owsianego wytworzono sygnały samogłoskowe, które mają, w sposób nie dający się w obecnym stanie wiedzy określić jednoznacznie, symulować wszystkie samogłoski podstawowe. Jako parametry syntezy przyjęto częstotliwości czterech formantów, szerokości wstęgi formantów oraz określone cechy tonu krtaniowego.

Wartości częstotliwości formantowych oparto w zasadzie na danych uzyskanych dla głosu jednego z współautorów (por. Jassem, Krzyżko 1992). Ponieważ jednak brak dla danego głosu jakichkolwiek danych dotyczących źródła pobudzania, zaszła konieczność dokonania pewnych dopasowań, które dałyby w rezultacie sygnały samogłoskowe spełniające optymalnie w danych warunkach wymaganie maksymalnej naturalności brzmienia. Uwzględniono jeszcze, przynajmniej w przybliżeniu, warunek równej odległości samogłosek tego samego porządku na bokach lewym i prawym czworoboku (tj. odległości np. między elementami zbioru [i, e, ɛ, a] na płaszczyźnie o współrzędnych $(\log F_1, \log F_2)$).

Ryc. 9 i 10 ukazują samogłoski zsyntetyzowane dla celów poniższego doświadczenia, na płaszczyźnie stworzonej przez częstotliwości formantów pierwszego i drugiego, w skali [Bark] i [log(Hz)]. Tabl.1 podaje wszystkie wartości częstotliwości oraz szerokości wstęgi formantów dla wszystkich samogłosek w doświadczeniu.



Ryc.9. F1 i F2 samogłosek podstawowych - skala [Bark].



Ryc.10. F1 i F2 samogłosek podstawowych - skala [log(Hz)]

Samo- głoska	Formanty i szerokości wstęgi							
	F ₁	B ₁	F ₂	B ₂	F ₃	B ₃	F ₄	B ₄
[i]	217	45	2775	70	3645	170	4107	250
[e]	376	45	2700	110	2944	170	3805	300
[ɛ]	620	45	2165	70	2743	170	3689	300
[a]	921	80	1560	100	2741	200	3571	300
[ɑ]	690	45	931	50	3007	100	3516	150
[ɔ]	556	45	750	50	2660	100	3253	150
[o]	420	45	675	50	2582	100	3167	150
[u]	308	45	577	50	2467	100	3133	150
[y]	246	45	2243	50	2663	100	3494	200
[ø]	370	45	2120	50	2479	100	3649	250
[œ]	520	45	1810	50	2466	100	3586	200
[ɛ̃]	677	45	1395	50	2615	100	3501	250
[ɔ̃]	740	45	1050	50	2982	100	3539	200
[ʌ]	591	45	955	50	2770	100	3313	200
[ɤ]	475	45	875	50	2520	100	3172	170
[ɯ]	369	45	1050	50	2525	100	3240	150
[ɨ]	283	45	1936	50	2594	100	3693	200
[ɘ]	276	45	1141	50	2399	100	3400	200

Tablica 1. Wartości formantów i szerokości wstęg wzorców syntetycznych samogłosek podstawowych.

Iloczas wszystkich wzorców syntetycznych samogłosek był stały i wynosił 350 ms, a $F_0 \cong 105$ Hz.

Traktując zbiór powyższych wzorców jako bazę, skonstruowano następnie program komputerowy, przeprowadzający test oceny odległości percepcyjnej, w układzie wszystkich par samogłosek syntetycznych wymienionych w tablicy 1 (zatem 324 pary), z oceną stopnia "niepodobieństwa" (dissimilarity) dla każdej pary w skali 0 - 9, oraz losową kolejnością par bodźców. Program przeprowadzał najpierw instruktaż osoby biorącej udział w teście, a następnie prezentował kolejno wylosowane pary bodźców - każdą parę tyle razy, aż uczestnik testu zdecydował się podać ocenę odległości percepcyjnej dla tej pary.

Test z każdym uczestnikiem przeprowadzany był trzykrotnie: w każdej serii testu powstawała macierz kwadratowa odległości percepcyjnych (18 x 18), a dodatkowo po trzeciej serii - macierz zbiorcza, będąca sumą macierzy z serii 1 - 3 .

Każdy przebieg pojedynczego testu był czasowo niezależny od pozostałych. Odstępy czasowe między kolejnymi seriami testu dla danego słuchacza wynosiły 4 do 6 dni.

Po przeprowadzeniu z 6 osobami testów opisaną metodą, utworzona została macierz "globalna", powstała przez dodanie macierzy zbiorczych wszystkich 6 testów (p. tablica 2).

Macierze wyjściowe: z każdej serii testu, zbiorcze dla kolejnych uczestników testu oraz globalna, były plikami tekstowymi, utworzonymi w formacie umożliwiającym bezpośrednio wprowadzenie ich jako danych wejściowych do programu skalowania wielowymiarowego KYST.

0	34	74	121	131	145	43	74	142	149	108	112	141	152	156	110	150	152
41	0	53	107	115	142	65	50	137	138	99	97	140	148	157	102	151	145
80	54	0	59	67	111	69	39	115	123	111	80	123	131	149	88	138	122
135	112	65	0	19	54	129	102	78	76	138	115	106	94	133	106	128	125
147	119	68	17	0	43	115	91	57	54	134	124	77	96	139	103	106	95
149	138	108	41	45	0	126	125	77	39	143	133	93	90	142	127	135	117
69	63	79	124	118	136	0	42	140	128	56	68	124	140	119	74	133	136
86	65	51	97	90	139	45	0	119	117	78	60	123	130	143	69	131	123
146	145	122	88	42	88	137	118	0	21	125	114	29	25	103	81	89	69
149	143	118	73	62	68	137	123	21	0	129	126	44	31	108	91	105	87
116	108	101	131	129	149	46	74	126	124	0	36	123	127	103	45	114	111
119	114	80	117	112	135	62	51	121	109	38	0	90	120	114	33	106	104
148	142	127	78	76	103	124	115	27	40	114	105	0	29	95	65	57	20
150	143	125	99	73	97	139	128	28	20	137	108	24	0	105	97	89	53
159	158	142	135	127	136	119	134	106	102	107	120	100	102	0	72	46	60
123	112	87	124	99	133	77	65	94	98	56	45	82	106	96	0	85	85
156	150	131	121	103	128	131	136	93	98	112	102	62	93	55	80	0	13
152	151	128	115	96	118	129	120	61	82	123	103	35	76	72	84	12	0

Tablica 2. Test odległości percepcyjnych syntetycznych samogłosek podstawowych - macierz globalna wyników.

Uwaga: Wzorce przyporządkowane są wierszom i kolumnom macierzy w następującej kolejności:

[i],[y],[i],[u],[w],[u],[e],[ø],[v],[o],[e],[æ],[ʌ],[ɔ],
[a],[æ],[o],[ɔ].

3. Skalowanie wielowymiarowe. Wnioski

Metodą statystyczną szczególnie dobrze nadającą się do oceny doświadczeń, w których dominuje czynnik psychologiczny, jest skalowanie wielowymiarowe, któremu poświęcono szereg prac monograficznych, np. Davison (1984), Kruskal & Wish (1984), Young & Hamer (1987) lub Biela (1992). Jeden z licznych wariantów tej metody zastosowano tutaj do oceny doświadczenia nad percepcją samogłosek podstawowych przez polskich słuchaczy. Poniżej podano najistotniejsze parametry określające zastosowany wariant. Jak wynika ze skali ocen (0 = „identyczność”; 9 = „maksymalna subiektywna różnica”), zastosowano zasadę rosnącego „niepodobieństwa” (ang. *dissimilarity*).

W doświadczeniu wzięło udział sześć osób. Dwie z nich odbyły ok. 25 lat wcześniej kurs fonetyki doświadczalnej, w obrębie którego opanowały samogłoski podstawowe w sposób bierny i czynny. Potem jednak te dwie osoby nie miały z tymi samogłoskami praktycznie nic do czynienia. Pozostałe osoby znały samogłoski podstawowe w zasadzie tylko teoretycznie, tj. z opisów w podręcznikach fonetyki ogólnej lub fonetyki poszczególnych języków. Materiał jest zbyt szczupły, by móc pokusić się o próbę zbadania, czy ta różnica miała wpływ na oceny dokonane przez poszczególnych słuchaczy zwłaszcza z uwagi na rozmiar macierzy ocen. Każda osoba wzięła udział w doświadczeniu trzykrotnie. Program sterujący doświadczeniem umożliwiał wyprowadzenie wyników pojedynczych cykli odsłuchowych, sumarycznych wyników dla każdego słuchacza oraz wyników łącznych dla wszystkich słuchaczy i wszystkich cykli odsłuchowych. Optyczna ocena tych różnych macierzy ocen nie wskazywała na istotne różnicowania między słuchaczami.

Analizę wyników doświadczenia przeprowadzono na wyjściowym materiale w postaci macierzy ocen łącznej dla wszystkich słuchaczy i wszystkich cykli odsłuchowych, którą przedstawiono w Tab. 2.

W analizie statystycznej przyjęto następujące parametry:

- (1) Maksymalna liczba wymiarów: 8, minimalna liczba wymiarów: 1.
- (2) Maksymalna liczba iteracji: 50.

- (3) Minimalny stres: 0.01.
- (4) Formuła dla stresu: I.
- (5) Odległości euklidesowe
- (6) Modele regresji: (a) monotoniczny, (b) wielomian 2 stopnia bez czynnika stałego, (c) wielomian 3 stopnia bez czynnika stałego.
- (7) Konfiguracja końcowa z rotacją na składowe główne.

Dane wyjściowe do analizy miały, jak widać z Tab.2, postać macierzy kwadratowej 18×18 z zerową przekątną główną. Widoczne jest, że macierz nie jest symetryczna, ale wobec jej dość znacznego wymiaru nie jest obecnie możliwe zbadanie ewentualnych czynników systematycznych wywołujących asymetrię. W analizie uwzględniono zatem obie części trójkątne macierzy. Wartości w poszczególnych klatkach mieszczą się w przedziale 0...162 (6 słuchaczy \times 3 cykle osłuchowe \times 9 (maksymalna ocena).

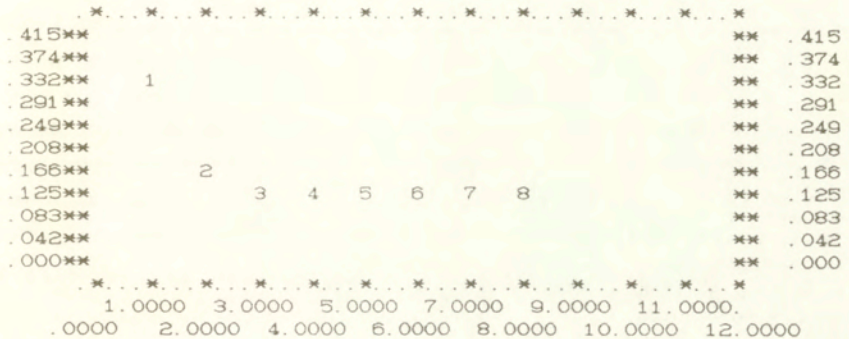
Zastosowany program obliczeniowy KYST pozwala prześledzić cały proces tworzenia konfiguracji. Dla każdej założonej liczby wymiarów program pokazuje tzw. wykres Sheparda, tj. rzut wartości δ_{ij} (niepodobieństwa empiryczne) względem wartości \hat{a}_{ij} (odległości estymowanej z danego typu regresji), wartości współrzędnych dla konfiguracji elementów badanego zbioru w przestrzeni n -wymiarowej oraz, dla danej liczby n wymiarów, wszystkie $n \times (n - 1) / 2$ wykresów na płaszczyźnie dwóch zmiennych (w skrajnym przypadku na jednej osi przy jednej zmiennej). Wyprowadzane są również wartości stresu dla każdej liczby wymiarów w każdym zadanym modelu regresyjnym.

Wielkość stresu dla poszczególnych modeli regresyjnych oraz poszczególnych liczb wymiarów przedstawia się następująco:

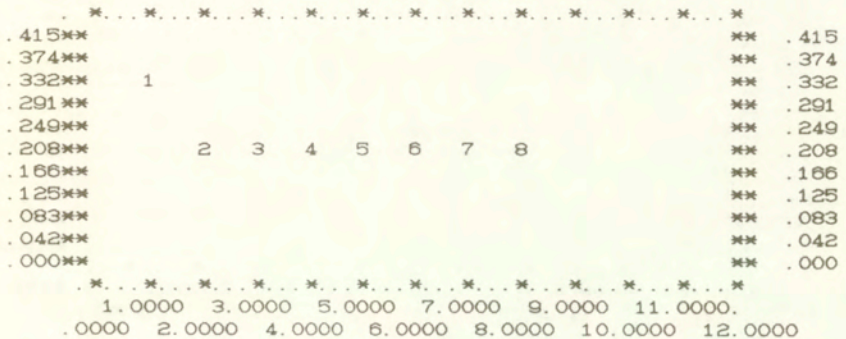
liczba wym. \rightarrow regresja \downarrow	8	7	6	5	4	3	2	1
mono	.122	.126	.125	.131	.131	.140	.152	.313
2 st	.192	.190	.188	.192	.192	.196	.202	.357
3 st	.188	.188	.188	.190	.190	.194	.199	.349

Ryc.11 i 12 ukazują wielkość stresu w funkcji liczby wymiarów dla modelu monotonicznego i wielomianu 3 st. Z tabeli oraz ryci-

ny widać bardzo wyraźnie, że przestrzeń percepcji samogłosek podstawowych jest dwuwymiarowa.



Ryc.11. Wielkość stresu jako funkcja liczby wymiarów, KYST, regresja monotoniczna.



Ryc.12. Wielkość stresu jako funkcja liczby wymiarów, KYST, regresja wielomianowa 3 stopnia.

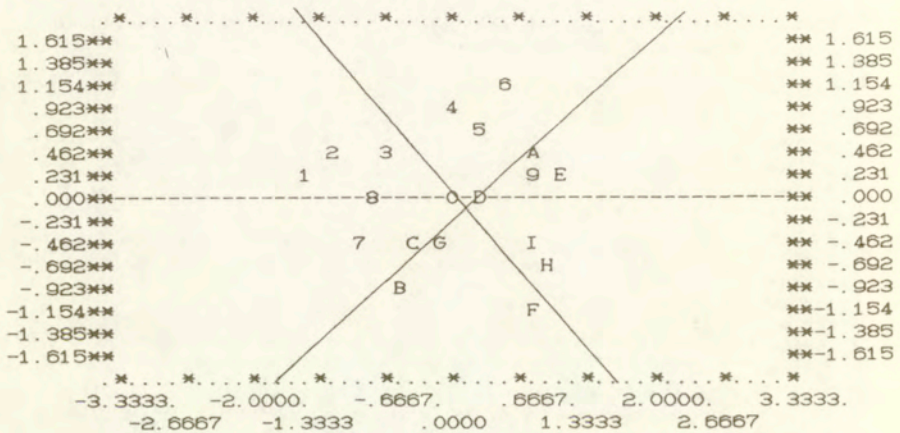
Wartości stresu dla 8 do 2 wymiarów są dla wszystkich trzech modeli praktycznie takie same i utrzymują się w strefie „dobrych”.

Program KYST przyjmuje własną symbolikę dla poszczególnych elementów zbioru bodźców na wykresach konfiguracji przed transformacją na składowe główne oraz dowolną symbolikę dwu- lub jednoliterową dla końcowej konfiguracji po rotacji na składowe

główne, ale w obrębie podstawowych znaków klawiatury. Poniższa Tabela przedstawia symbolikę zastosowaną w wykresach przed oraz po transformacji na składowe główne:

Przed transf.	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	G	H	I
Po transformacji	A1	B1	C1	C2	B8	A8	A2	B2	B7	A7	A3	B3	B6	A6	A4	B4	B5	A5
Znak transkr.	i	y	i	u	u	u	e	e	y	o	e	e	Λ	o	a	α	o	o

Ryc.13 ukazuje konfigurację samogłosek na płaszczyźnie w modelu monotonicznym.



Ryc.13. KYST, regresja monotoniczna, konfiguracja w 2 wymiarach.

Po rotacji osi i niewielkim przesunięciu początku układu otrzymuje się następujące wyniki:

1. Samogłoski przymknięte i półprzymknięte znajdują się w górnej połowie z wyjątkiem [y], które znalazło się tuż pod osią, wraz z samogłoskami półotwartymi i otwartymi.

2. Po lewej stronie znajdują się wszystkie samogłoski przednie, natomiast po stronie prawej - wszystkie samogłoski tylne.

Ponieważ wymieniony wyjątek znajduje się bardzo blisko

odpowiednich osi, można przyjąć, że albo należy nieznacznie zmodyfikować zadany układ samogłosek przez niewielkie przesunięcie tego wyjątku, albo jeszcze szerzej zakrojone doświadczenie odsłuchowe wskaże, że tutaj otrzymane wyniki obarczone są niewielkim błędem.

Ryc.14 i 15. ukazują wyniki końcowe, po rotacji na składowe główne, odpowiednio dla regresji monotonicznej i wielomianowej 3 stopnia.

Widoczne są (ryc.14, dla regresji wielomianowej 3 stopnia otrzymano odbicie zwierciadlane względem osi Y), następujące podziały:

1. Samogłoski przymknięte i półprzymknięte znajdują się powyżej osi poziomej, z wyjątkiem [e], położonego tuż poniżej osi wraz z półotwartymi i otwartymi, spośród których [o] jest przesunięte tuż ponad oś.

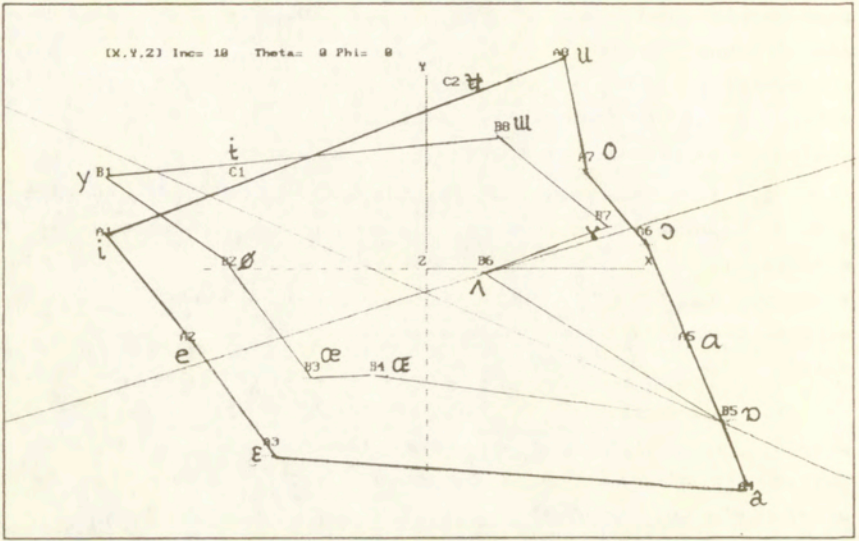
2. Samogłoski dzielą się na przednie - po lewej stronie osi oraz tylne - po prawej, z wyjątkiem [a], które znalazło się po niewłaściwej stronie.

Tak przed transformacją, jak i po transformacji na składowe główne widoczna jest bardzo silna tendencja występowania samogłosek peryferyjnych [i e e a o o u] w dalszej odległości od środka układu niż samogłosek wtórnych [e e a y u i u]. Bardzo nieznaczne odchylenie od tej zasady wykazują [y] oraz [o].

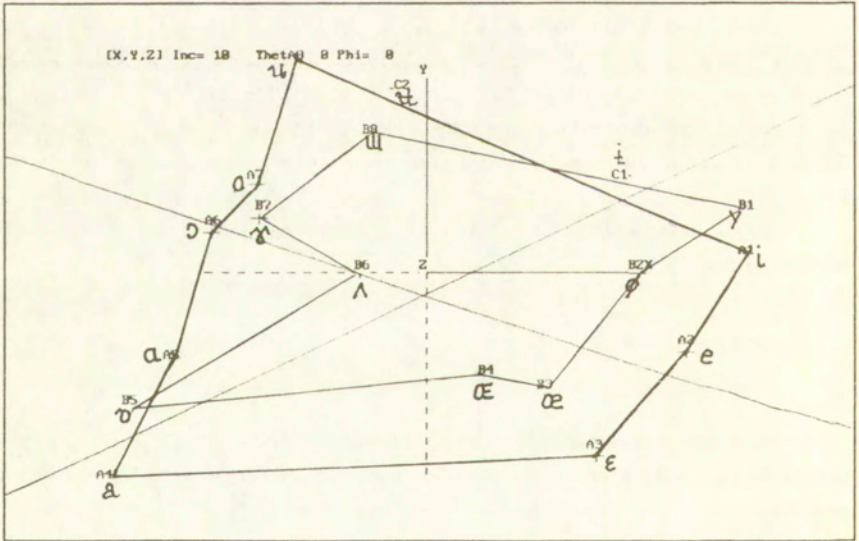
Układ systematyczny polskich fonemów samogłoskowych jest następujący:

	przednie	centralne	tylne
przymknięte	/i/	/i/	/u/
otwarte	/e/	/a/	/o/

Ani przedstawione wyżej konfiguracje, ani nieznaczne odchylenia od układu na czworoboku (czworobokach) samogłosek podstawowych nie ukazują jakiegokolwiek wpływu polskiego systemu wokalicznego na uzyskane wyniki.



Ryc.14. Mapa samogłosek podstawowych - KYST, regresja monotoniczna.



Ryc.15. Mapa samogłosek podstawowych - KYST, regresja wielomianowa 3 stopnia.

Wnioski końcowe są następujące:

Przeprowadzone doświadczenie ukazało, że przestrzeń percepcyjna samogłosek podstawowych jest dwuwymiarowa z wymiarami: artykulacja przymknięta ↔ otwarta oraz przednia ↔ tylna. W tej przestrzeni pozwalają się odróżnić samogłoski peryferyjne od nieperyferyjnych, wobec czego nie ma potrzeby postulowania dwóch oddzielnych czworoboków samogłoskowych, a tym mniej figury trójwymiarowej. Doświadczenie potwierdziło poprawność topologiczną pojedynczego czworoboku samogłosek podstawowych przedstawionego na str. 3.

Wbrew twierdzeniom Butchera (1984) istnieje pełne uzasadnienie konfiguracji samogłosek podstawowych, a - również wbrew Butcherowi - przeprowadzone doświadczenie nie wykazało żadnego wpływu ojczystego języka na percepcję tych samogłosek.

BIBLIOGRAFIA

- ARMSTRONG, L.E. & JONES, D. (1951): *The Phonetics of French*, G. Bell & Sons, London.
- BIELA, A. (1992): *Skalowanie wielowymiarowe jako metoda badań naukowych*, Tow. Naukowe Katol. Uniwersytetu Lubelskiego, 1992.
- BLADON, R.A.W. & FANT, G. (1978): A two-formant model and the cardinal vowels, *Speech Transmission Laboratory Quarterly Progress and Status Report 1/1978*, 1-8.
- BUTCHER, A (1982): Cardinal Vowels and other Problems, in: *Linguistic Controversies* (D. Crystal, ed.), E.1 Arnold, London, 50-71.
- CATFORD, I.C. (1981): Observations on the Recent History of Vowel Classification, in: *Towards a History of Phonetics* (R.E. Asher & E.J. Henderson, eds.), Edinburgh University Press, Edinburgh, 19-31.
- CATFORD, I.C.: *A Practical Introduction to Phonetics*, Oxford University Press, Oxford.
- DAVISON, M.L. (1983): *Multidimensional Scaling*, J. Wiley & Sons, New York.
- FANT, G. (1960): *Acoustic Theory of Speech Production*, Mouton & Co., 's-Gravehage.
- JASSEM, W. (1973): *Podstawy fonetyki akustycznej*, PWN, Warszawa.
- JASSEM, W. (1983): *The Phonology of Modern English*, PWN, Warszawa.
- JASSEM, W. & KRZYŹSKO, M. (1992): In Quest for Cardinal Vowels, *Archives of Acoustics* 17, 343-356.
- JONES, D (1956a): *The Pronunciation of English*, 4th ed. Cambridge Univ. Press, Cambridge.
- JONES, D. (1956b): *An Outline of English Phonetics*, 8th ed., Heffer & Sons, Cambridge.
- JONES, D. & WARD, D. (1969): *The Phonetics of Russian*, Cambridge University Press, Cambridge.
- KRUSKAL, J.B. & M. WISH (1984): *Multidimensional Scaling*, Sage Publications, Beverly Hills.

- NEARLEY, M. (1978): *Phonetic Feature Systems for vowels*, Indiana Univ. Linguistics Club, Bloomington Indiana.
- PAPÇUN, G (1980): *How do different People say the same Vowels?*, *UCLA Working Papers in Phonetics* No. 49, Los Angeles.
- RAUSCH, R. & RAUSCH, I. (1991): *Deutsche Phonetik für Ausländer*, Langenscheidt Verlag Enzyklopädie, Berlin.
- ROACH, P. (1983): *English Phonetics and Phonology*, Cambridge University Press, Cambridge.
- YOUNG, F.W., & R.M. Hamer, (1987): *Multidimensional Scaling: History, Theory and Applications*. L. Erlbaum Associates, Hillside, N.J.