

Ryszard Ciarkowski

STEROWANA Z MNIKOMPUTERA MERA 303
SYNTEZA WYBRANYCH DIAD POLSKICH
I ICH PERCEPCJA

7/1984

P. 269



WARSZAWA 1984

Praca wpłynęła do Redakcji dnia 7 grudnia 1983 r.



56955



Na prawach rękopisu

Instytut Podstawowych Problemów Techniki PAN
Nakład 140 egz. Ark.wyd. 1,6. Ark.druk.2,75.
Oddano do drukarni w kwietniu 1984 r.
Nr zamówienia 264/84.

Warszawska Drukarnia Naukowa, Warszawa,
ul.Śniadeckich 8

Ryszard Ciarkowski
Pracownia Fonetyki Akustycznej
IPPT PAN

STEROWANA Z MINIKOMPUTERA MERA 303 SYNTEZA
WYBRANYCH DIAD POLSKICH I ICH PERCEPCJA^{1/}

Streszczenie

Przy użyciu zmodyfikowanego systemu programowego SPOS1 dokonano syntezy 84 diad polskich będących połączeniami 8 wybranych spółgłosek z samogłoskami. Uzyskany w ten sposób zbiór diad syntetycznych poddano ocenie percepcyjnej przez 25 słuchaczy. Praca stanowi kontynuację pracy [1] i zawiera przedstawienie modyfikacji wprowadzonych do systemu programowego SPOS, polegających na uzupełnieniu procesu syntezy równoczesną analizą mowy syntetycznej. Umożliwia to wizualną kontrolę syntetyzowanego sygnału przedstawionego w postaci spektrogramu. Dalsza część pracy stanowi opis uzyskanego zbioru diad syntetycznych i wyników oceny percepcyjnej.

1. Opis zmodyfikowanego systemu programowego SPOS1.

Zakres czynności operatorskich objętych Systemem Programowej Obsługi Syntetyzatora "SPOS", przeznaczonym dla syntetyzatora formantowego COMPUTALKER CT-1 i opisanym w pracy [1], został poszerzony o funkcję mającą w procesie syntezy zasadnicze znaczenie dla kontroli parametrów produkowanej mowy syntetycznej - analizę widmową mowy zarówno syntetycznej jak i naturalnej. Wykonywana równolegle z syntezą analiza widmowa produktu syntezy pozwala na szybkie uzyskanie spektrogramu mowy syntetycznej wyświetlanego na ekranie monitora graficznego MEMOSKOP. Wprowadzenie oddzielnej czynności operatorskiej ANALIZUJ umożliwia

^{1/} Praca wykonana w ramach planu C-1

dokonanie analizy widmowej i uzyskanie spektrogramu z wypowiedzi mówionej do mikrofonu lub zapisu na taśmie magnetofonowej.

1.1. Konfiguracja sprzętu wykorzystywanego w systemie SPOS 1

Dla potrzeb obsługi syntetyzatora wykorzystano w zmodyfikowanym systemie programowym SPOS1 zestaw minikomputerowy MERA 303 współpracujący z wyspecjalizowanymi urządzeniami peryferyjnymi. Konfigurację tego sprzętu przedstawia rys. 1.1.

W stosunku do konfiguracji sprzętu używanego dotychczas w systemie SPOS rozszerzeniem jest wprowadzenie wielokanałowego analizatora widmowego oraz tzw. kanału funkcji analogowych KF-1.

Wielokanałowy analogowy analizator widmowy wraz z komutatorem sygnałów wyjściowych opisany jest szczegółowo w pracy [3]. Analizator posiada 63 kanały. 43 kanały - każdy obejmujący pasmo częstotliwości 80 Hz - pokrywają łącznie zakres od 80 do 3560 Hz. W dalszych kanałach szerokość pasma przepustowego jest liniowo zależna od numeru kolejnego kanału zgodnie z zależnością

$$\Delta f_p = 80 + (k-43)\Delta f_d \quad (1)$$

gdzie

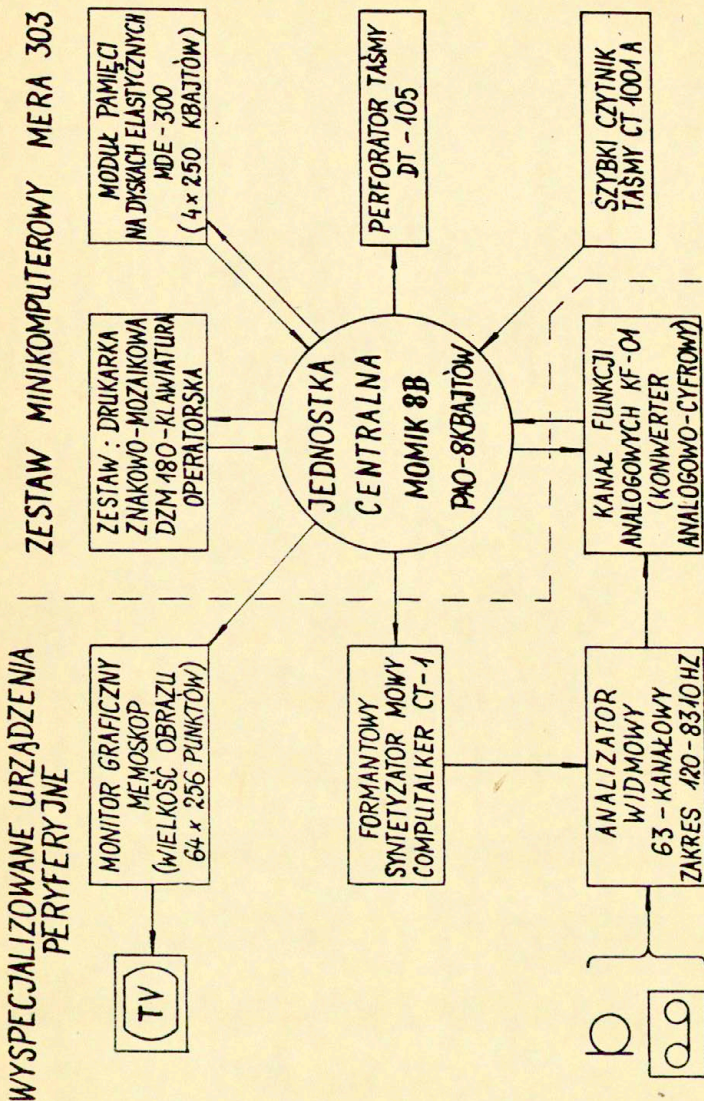
k - numer kolejny kanału,

$$\Delta f_d = 20 \text{ Hz.}$$

Szerokość pasma przepustowego pierwszego z dodatkowych kanałów wynosi więc 100 Hz, ostatniego - 480 Hz. Całkowity zakres analizatora obejmuje częstotliwości od 80 Hz do 8310 Hz.

Na każdy z kanałów analizatora składa się środkowo-przepustowy filtr aktywny o charakterystyce Butterworth'a trzeciego rzędu, wzmacniacz, układ prostujący, dolnoprzepustowy filtr aktywny RC o charakterystyce Butterworth'a trzeciego rzędu oraz sterowany komutacyjnie przełącznik elektroniczny łączący wyjście kanału z wyjściem analizatora na czas ok. 100 μ s raz na okres komutacji.

Wprowadzanie do minikomputera informacji pochodzących z analizatora odbywa się poprzez tzw. kanał funkcji analogowych [3], [5]. Wykorzystuje się w tym celu logarytmujący konwerter analogowo-cyfrowy będący elementem składowym kanału funkcji analogowych. Praca tego konwertera opiera się na zależności logarytmicznej pomiędzy czasem t rozładowywania się kondensatora



Rys.1.1. Konfiguracja sprzętu wykorzystywanego w systemie SPOS1.

a napięciem U_c , do którego nastąpiło rozładowanie :

$$t = a \log U_c + b = -T \log U_c + T \log U_0 \quad (2)$$

gdzie

U_0 - napięcie na kondensatorze w chwili t_0

T - stała czasowa rozładowania kondensatora.

Od początku okresu konwersji ($t=0$) napięcie podlegające konwersji jest stale porównywane z napięciem na rozładowującym się kondensatorze. W czasie oczekiwania na zrównanie się obu tych napięć następuje zliczanie impulsów o częstotliwości ok. 1 MHz. Zrównanie napięć powoduje zatrzymanie zliczania. Stan licznika będący wynikiem konwersji danej analogowej do postaci cyfrowej z równoczesnym jej zlogarytmowaniem zostaje przepisany programowo do PAO minikomputera. Maksymalny czas trwania konwersji - 100 μ s.

1.2. Modyfikacje programowe w systemie SPCS1.

Włączenie w zakres czynności operatorskich objętych systemem SPOS1 analizy widmowej możliwe było po dokonaniu następujących zmian :

- maksymalna długość syntetyzowanego fragmentu mowy została skrócona z 315 do 91 fram, długość ta pozwala na syntetyzowanie niezbyt długich wyrazów,
- zrezygnowano z wyświetlania przebiegów parametrów syntezy na ekranie telewizyjnym przy użyciu urządzenia MEMOSKOP - komenda M została usunięta z języka komend systemu SPCS1,
- sporządzenie dokumentacji dla zgromadzonej w PAO minikomputera informacji o parametrach syntezy danego fragmentu mowy ograniczone zostało do tabelarycznego zestawienia wartości kompletu parametrów, zrezygnowano z kreślenia kształtu przebiegów parametrów syntezy (komenda K) oraz z tabelarycznego zestawienia wartości jednego wybranego parametru syntezy (ograniczenie komendy T),
- usunięto z systemu funkcję zerowania obszaru PAO przeznaczonego na zapis wartości parametrów syntezy (komenda ZP) oraz funkcję powtórzenia sekwencji wartości danego parametru syntezy zapisanej w PAO (komenda PS).

Wprowadzenie wyżej wymienionych zmian jest skutkiem :

- dążenia do zastąpienia mniej efektywnych lub rzadko wykonywanych czynności operatorskich innymi doskonalszymi,
- posiadanie zbyt małej pamięci operacyjnej w minikomputerze MERA 303, co pociąga za sobą konieczność ograniczenia realizowanych programowo funkcji do najniezbędniejszego minimum.

1.3. Realizacja modyfikacji programowych.

Rysunki 1.2. - 1.4. przedstawiają schematy blokowe programu i podprogramów wykonywanych podczas realizacji funkcji systemowej "Analiza sygnału mowy i wyświetlanie spektrogramu".

Rysunek 1.5. pokazuje schemat blokowy podprogramu OBSŁUGA SYNTETYZATORA, na którym uwzględniono zmiany związane z wprowadzeniem analizy widmowej mowy syntetycznej.

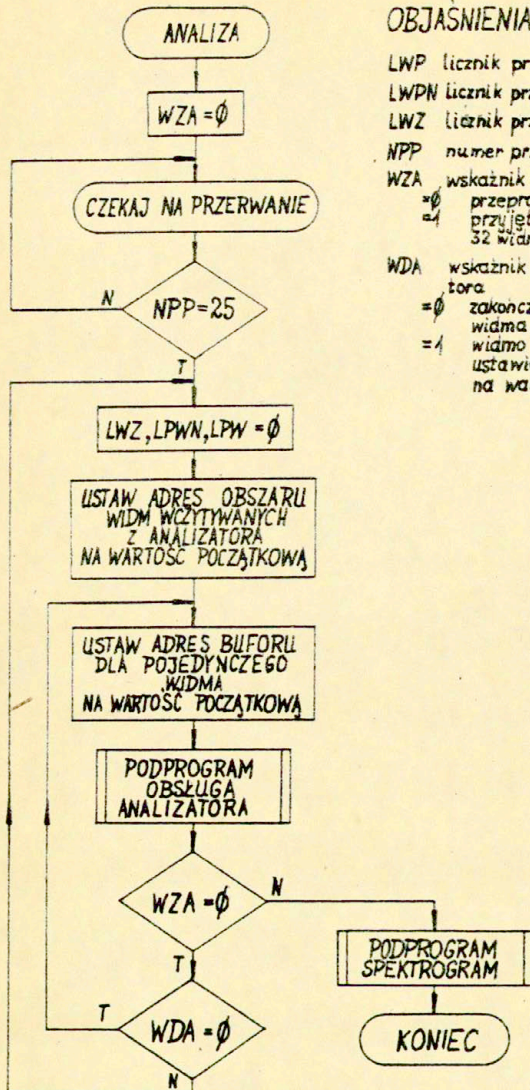
Odczyt z analizatora sekcji widmowych sygnału następuje co 20 ms. Przerwania o numerze 23 generowane w układzie konwertera analogowo-cyfrowego inicjują przesłanie każdej z sekcji widmowych sygnału do minikomputera. Generowane na zakończenie przesyłania kompletu 63 sekcji widmowych przerwanie o numerze 25 powoduje uruchomienie procedury wykrywania początku analizowanego sygnału. Zapamiętywanie uzyskiwanych co 20 ms widm sygnału rozpoczyna się w momencie przyjęcia pierwszego widma niezerowego. Widmem niezerowym nazywane jest widmo, którego co najmniej jedna rzędna ma wartość większą od zadanego progu. Jeżeli po widmie niezerowym wystąpi sześć kolejnych widm zerowych, to przyjęte widma zostają uznane za zakłócenie i anulowane. Po przyjęciu 32 widm analiza zostaje zakończona.

Poszczególne widma uzyskane z analizatora charakteryzują się nieregularną obwiednią. Nieregularność ta powodowana jest periodycznym charakterem dźwięków mowy jak też przypadkowością wzajemnego położenia układu pasm analizatora i składowych harmonicznych sygnału. W celu uzyskania regularnej obwiedni widma zastosowano przekształcenie wygładzające [4] wyrażone formułą :

$$P_{wi} = P_i + \sum_{j=1}^2 (P_{i+j} + P_{i-j}) \cdot W_j \quad (3)$$

gdzie :

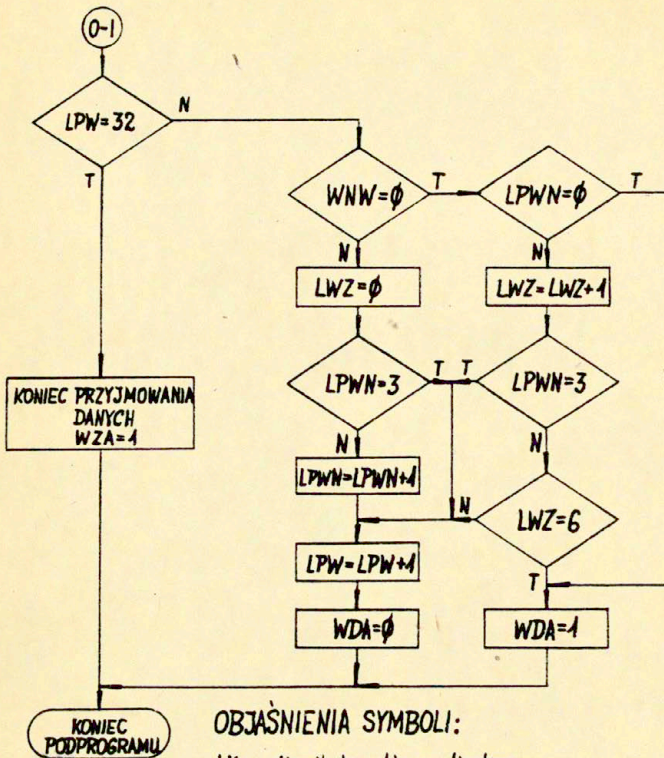
P_{wi} - i-ta rzędna widma wygładzonego



OBJASNIENIA SYMBOLI:

- LWP licznik przyjętych widm
- LPWN licznik przyjętych widm niezerowych
- LWZ licznik przyjętych widm zerowych
- NPP numer przyjętego przerwania
- WZA wskaźnik zakończenia analizy
 - = 0 przeprowadzanie analizy
 - = 1 przyjęte zostały wartości 32 widm (zakończenie analizy)
- WDA wskaźnik wpisu danych z analizatora
 - = 0 zakończono wpisanie pojedynczego widma, które zostało przyjęte
 - = 1 widmo nie zostało zaakceptowane, ustawienie adresu PAO i buforu na wartość początkową

Rys. 1.2. Schemat blokowy programu ANALIZA.



OBJAŚNIENIA SYMBOLI:

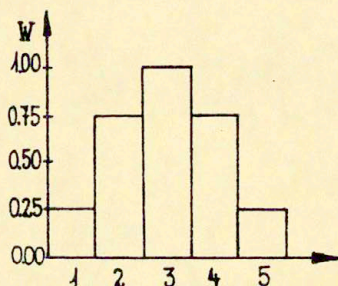
- LK licznik kanałów analizatora
- LPW licznik przyjętych widm
- LPWN licznik przyjętych kolejno widm niezerowych
- LWZ licznik przyjętych widm zerowych
- WDA wskaźnik wpisu danych z analizatora
=φ zakończono wpisanie pojedynczego widma, które zostało przyjęte
=1 widmo nie zostało zaakceptowane, ustawienie adresu PAO i bufora na wartość początkową.
- WNW wskaźnik niezerowości widma
- WZA wskaźnik zakończenia analizy (przyjęto 32 widma = 1)

Rys.4.3b. Schemat blokowy podprogramu OBSŁUGA ANALIZATORA.

$P_i, P_{i \pm j}$ - rzędna widma pierwotnego

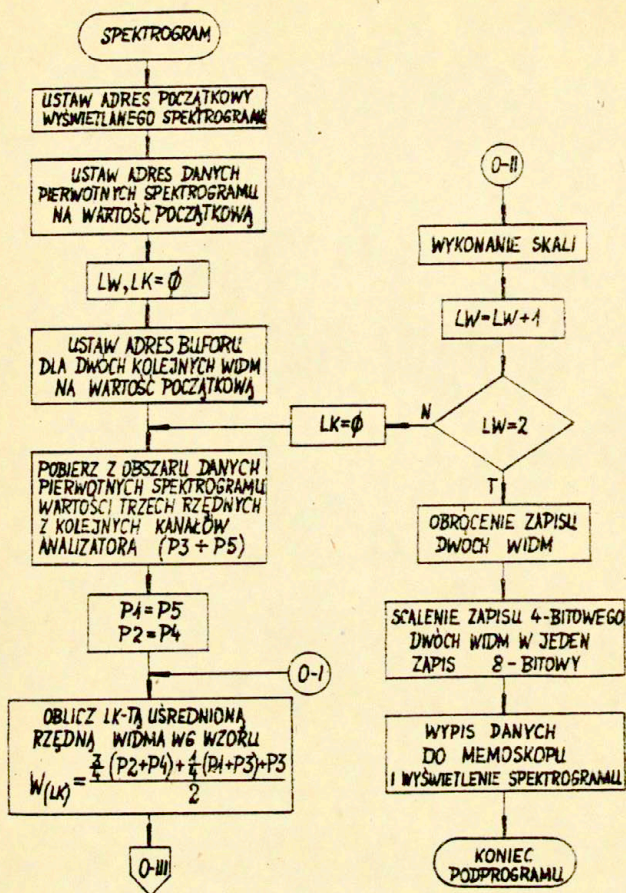
w_j - współczynnik wagowy

Ilustracją współczynnika wagowego jest okienko, którego wymiary zależą od założonego zakresu wygładzania widma. Okienko zapewniające optymalne wygładzenie obwiedni widma sygnału harmonicznego, którego rzędne są odległe od siebie o nie więcej niż 3 pasma analizy, przedstawia rys. 1.4.

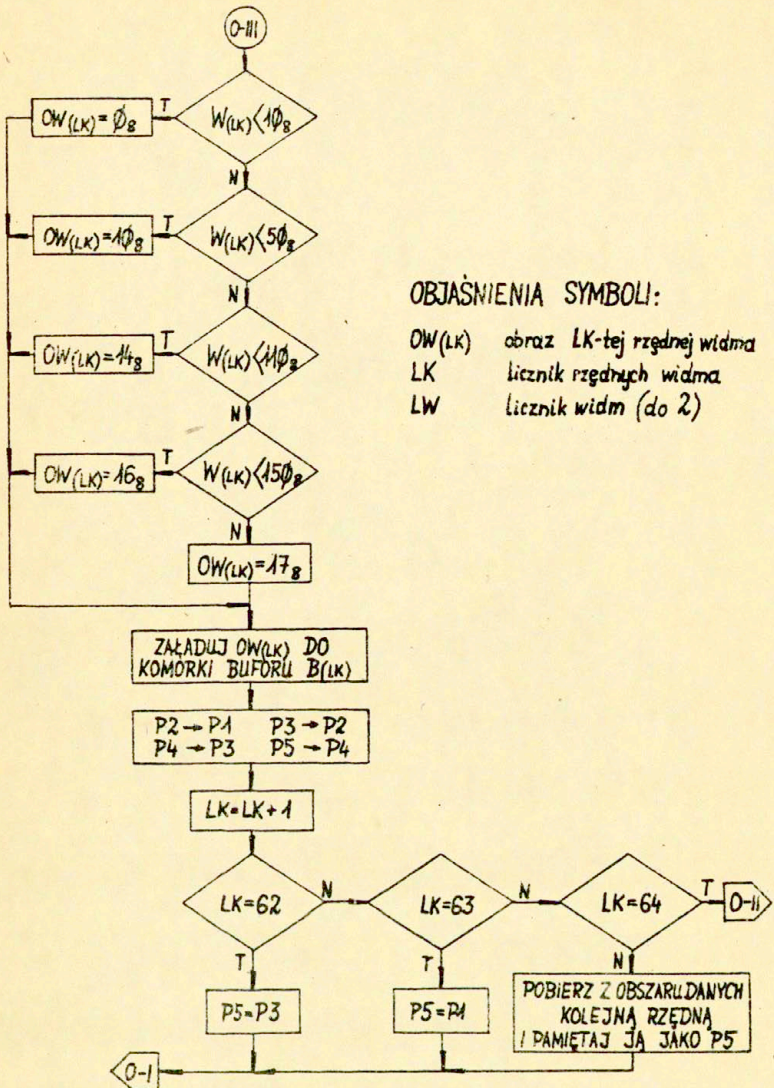


Rys. 1.4. Wymiary okienka wagowego.

Rzędne wygładzonego i uśrednionego widma poddawane są pięcipoziomowemu kwantowaniu warunkowanemu przez sposób przedstawienia widma na monitorze graficznym. Obszar ekranu monitora został podzielony na 64×64 punkty o kształcie zbliżonym do kwadratu. Wysokość punktu jest sztywno ustalona przez rozwiązanie konstrukcyjne monitora i odpowiada szerokości trzech kolejnych linii obrazu telewizyjnego, szerokość natomiast równą 4 punktom elementarnym ustalono programowo. Wartości poszczególnych rzędnych widma zostają wyrażone na ekranie monitora szerokością czerni w odpowiednich miejscach czynnego pola ekranu. Najniższemu poziomowi kwantowania odpowiada punkt całkowicie biały, najwyższemu - całkowicie zaczerniony. Rysunek 1.6. pokazuje przykładowy spektrogram uzyskany z monitora graficznego.



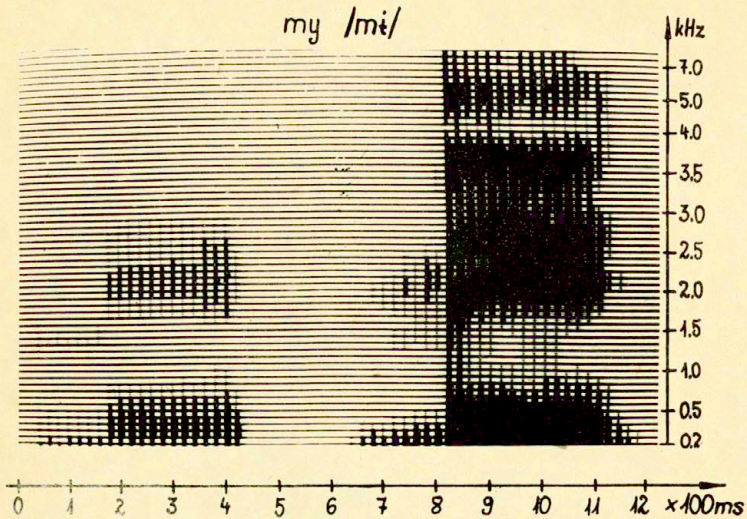
Rys.4.5a. Schemat blokowy podprogramu SPEKTROGRAM.



OBJAŚNIENIA SYMBOLI:

- OW(LK) obraz LK-tej rzędnej widma
- LK licznik rzędnych widma
- LW licznik widm (do 2)

Rys. 4.5b. Schemat blokowy podprogramu SPECTROGRAM.

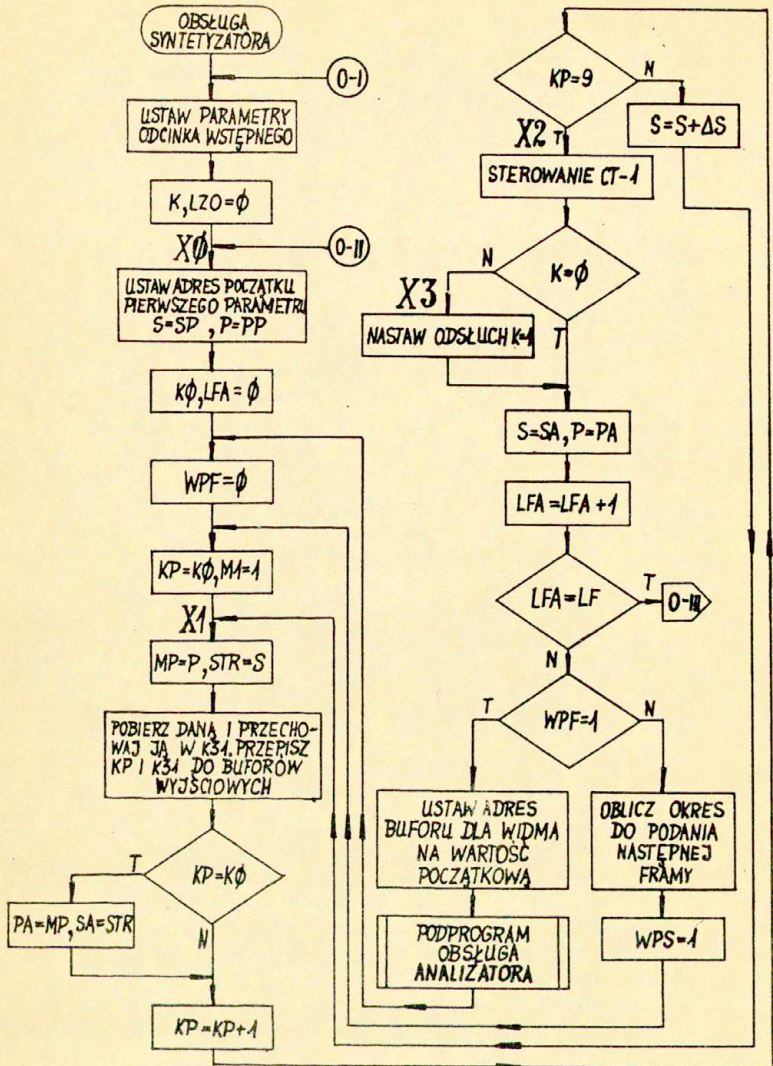


Rys. 1.6. Spektrogram mowy naturalnej /po stronie prawej/ i produktu jej syntetycznej symulacji /po stronie lewej/.

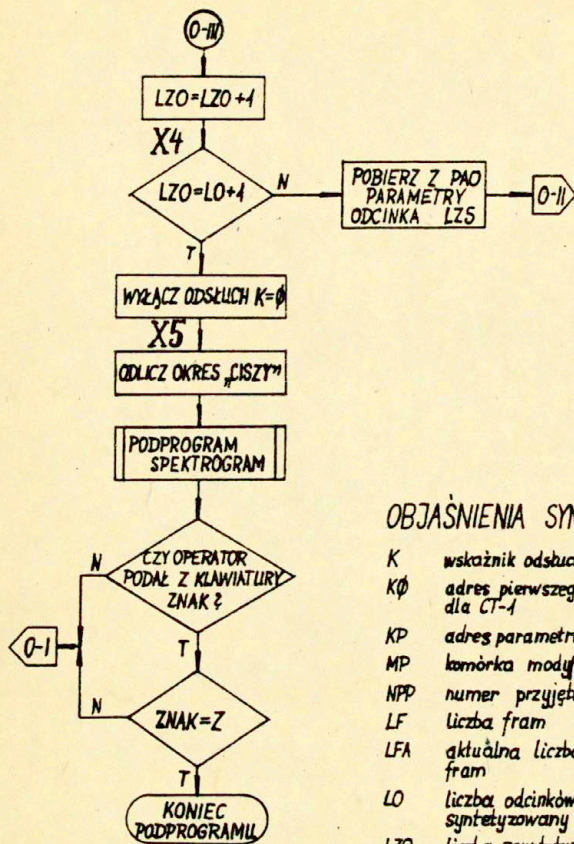
Ekran monitora podzielony został na dwie części. Lewą przyporządkowano spektrogramowi mowy syntetycznej, prawą spektrogramowi mowy naturalnej. Na długość każdego ze spektrogramów składają się 32 widma - odpowiada to sygnałowi o czasie trwania ok. 640 ms.

Zmiany dokonane w podprogramie OBSŁUGA SYNTETYZATORA polegają na zsynchronizowaniu pracy syntetyzatora i analizatora. W tym celu wyeliminowane zostało sterowanie przesyłaniem danych do syntetyzatora za pomocą generatora przerwania układu pośredniczącego zestawu : syntetyzator - minikomputer /patrz rozdz. 2.2. poz. [1]/. Funkcję tę powierzono generatorowi przerwania o numerze 25 wykorzystywanemu w procesie analizy. Sterowanie syntetyzatorem ma więc przebieg następujący :

Po przyjęciu widma ostatnio zsyntetyzowanej framy i jego wstępnej analizie zostają przesłane do układu syntetyzatora wartości parametrów sterujących kolejnej framy. Ponieważ jednak framy powinny być przesyłane do syntetyzatora co 10 ms /dla



Rys. 1.7a. Schemat blokowy podprogramu OBŚLUGA SYNTETYZATORA.



OBJAŚNIENIA SYMBOLI:

- K wskaźnik odsłuchu
- $K\emptyset$ adres pierwszego parametru dla CT-1
- KP adres parametru dla CT-1
- MP komórka modyfikująca
- NPP numer przyjętego przerwania
- LF liczba fram
- LFA aktualna liczba wyprowadzonych fram
- LO liczba odcinków tworzących syntetyzowany fragment
- LZO liczba zsintetyzowanych odcinków
- S, P adres komórki PAO (strona i numer komórki na tej stronie)
indeks A - aktualny
indeks P - początkowy
- STR rejestr strony
- W wskaźnik gotowości do przyjęcia przerwania
- WPF wskaźnik parzystości numeru framy
=0 numer nieparzysty, =1-parzysty

Rys. 1.76. Schemat blokowy podprogramu OBSŁUGA SYNTETYZATORA.

normalnego tempa mowy/, a próbkowania widma w ramach analizy dokonuje się co 20 ms, następuje programowy podział czasu pomiędzy dwoma próbkowaniami na odcinki 10-milisekundowe i powtórne przesłanie ramy do układu syntetyzatora. W ten sposób po każdym przesłaniu pary kolejnych fram następuje jedno próbkowanie widma dla potrzeb analizy.

Syntetyzowana wypowiedź jest powtarzana cyklicznie. Spektrogram zostaje wykonany jednorazowo po pierwszym cyklu syntezy i jest stale wyświetlany w trakcie następnych cykli aż do momentu ponownego zainicjowania syntezy /ponownego użycia komendy S/.

1.4. Język komend systemu SPOS1.

Język komend systemu SPOS1 obejmuje sześć komend o następujących symbolach :

A - analizuj /komenda nowo wprowadzona/

C - czytaj

DO - dysk - odczyt

DZ - dysk - zapis

S - syntetyzuj

T - tablicuj /komenda zmodyfikowana/

Zasady ogólne dotyczące wszystkich komend systemu SPOS /patrz rozdz. 4 poz. [1]/ pozostają niezmienione. Również opisy komend C, DO, DZ i S /patrz rozdz. 5.1. poz. [1]/ są w pełni aktualne. Poniżej przedstawione są opisy komend A i T. Przez podkreślenie oznaczono w nich parametry komend podawane przez operatora.

Komenda A

Funkcja komendy : Analizuj materiał mówiony do mikrofonu lub zapisany na taśmie magnetofonowej.

Postać ogólna komendy :

A
A > p

p - dopuszczalne dwie możliwości

A - akceptacja wykonania komendy

B - odwołanie komendy

Wykonanie :

Po akceptacji wykonania komendy następuje zatrzymanie programu przy stanie licznika rozkazów minikomputera 140.05. Po nastaniu gotowości do przesłania sygnału mowy z mikrofonu

lub magnetofonu, przez wyciśnięcie i wcisnięcie klawisza SS inicjuje się wykonanie programu analizy. Wykonanie programu zostaje zakończone wyświetleniem na ekranie monitora graficznego MEMOSKOP spektrogramu cyfrowego analizowanego materiału.

Komenda T

Funkcja komendy : Tablicuj wartości kompletu parametrów syntezy.
Postać ogólna komendy :

T

T ? np ? nk.

> p

Znaczenie symboli :

nk - numer ramy końcowej

np - numer ramy początkowej

p - dopuszczalne dwie możliwości :

A - akceptacja parametrów komendy

B - błąd w podaniu parametrów komendy

Wykonanie :

Wyzkawszy akceptację system drukuje w postaci tabelarycznej zestawienie wartości wszystkich parametrów syntezy. Przed przystąpieniem do wykonywania komendy należy pamiętać o założeniu na drukarkę szerokiego papieru obrzeźnie perforowanego - w systemie SPOS1 brak jest programowego zatrzymania dla wykonania tej czynności.

2. Opis zbioru diał syntetycznych i ich ocena percencyjna^{1/}.

W wyniku eksperymentów z mową syntetyczną uzyskano zbiór 84 polskich diał syntetycznych typu CV i VC powstałych z połączeń spółgłosek /j/, /l/, /w/, /m/, /n/, /p/, /r/ oraz /v/ z sześcioma samogłoskami /i/, /ɛ/, /e/, /a/, /o/, /u/. Na rysunku 2.1. pokazano parametry akustyczne diał, natomiast na rysunku 2.2. spektrogramy kilku reprezentatywnych diał w wersji naturalnej i syntetycznej. Wszystkie diały zrealizowano przy wykorzystaniu wyłącznie toru formantowego syntetyzatora, a więc liczbę parametrów sterujących ograniczono do pięciu - amplitudy i częstotliwości sygnału podstawowego /AV i FO/ oraz trzech

^{1/} Tabele 1 - 5 oraz rysunki 2.1. i 2.2., do których odwołano się w tym rozdziale zamieszczone są na końcu pracy.

częstotliwości formantowych F_1, F_2, F_3 .

Dla oceny jakości uzyskanych diad syntetycznych przeprowadzony został eksperyment percepcyjny z 25 słuchaczami. Każdemu ze słuchaczy odtworzono jednokrotnie trzy zrandomizowane listy zestawione z diad syntetycznych. Lista 1 zawierała wszystkie 84 diady typu zarówno CV jak i VC, lista 2 - 43 diady typu CV, a lista 3 - 41 diad typu VC. W każdej z list po trwającym ok. 0.5 s bodźcu następowała 3.5-sekundowa przerwa przeznaczona na zapis przez słuchacza wyniku rozpoznania. Przed rozpoczęciem eksperymentu słuchaczy informowano o zbiorze prezentowanych diad. Każdą listę poprzedzał komentarz o typie diad, które dana lista zawierała.

Tabela 1 prezentuje listę diad syntetycznych będących przedmiotem eksperymentu. Dla każdej z diad podano liczbę poprawnych rozpoznań w liście 1 oraz 2 lub 3 wyrażoną w procentach. Średnia dla 25 słuchaczy globalna liczba poprawnych rozpoznań dla listy 1 /diady typu CV i VC/ wynosi 75.4 %, dla listy 2 /diady typu CV/- 74.2 % oraz dla listy 3 /diady typu VC/ - 81.5 %.

Z względu na zbyt duże wymiary tabeli szczegółowych wyników rozpoznawania dla listy 1 oraz znaczne podobieństwo występujących prawidłowości w wynikach dla listy 1 oraz 2 i 3, podanie szczegółowych wyników rozpoznawania diad syntetycznych ograniczono do listy 2 i 3 /patrz tabele 2 i 3/. Ogólna liczba poprawnych rozpoznań samogłosek występujących w diadach wynosząca 98.8 % /tabela 4/ pozwoliła na zbiorcze przedstawienie wyników odbioru spółgłosek syntetycznych dla listy 1, 2 i 3 w tabeli 5. Błędy, które wystąpiły w odbiorze spółgłosek syntetycznych są efektem :

- a/ subiektywnego podobieństwa pomiędzy założoną treścią fonetyczną nadanego bodźca a treścią zidentyfikowaną przez słuchacza - błędy tego rodzaju miały miejsce w przypadku diad zawierających głoski nosowe /m/, /n/ i /ɲ/, które są trudniej wzajemnie rozróżnialne,
- b/ ograniczenia liczby parametrów mowy syntetycznej do pięciu /amplituda i częstotliwość sygnału podstawowego, częstotliwości trzech pierwszych formantów/ i związanego z tym nietypowego kształtowania przebiegu tych parametrów

w torze samogłoskowym

błędy polegające na zamianie spółgłosek nosowych /n/ i /ŋ/ ze spółgłoską /l/ oraz słabo szumowej głoski w ze spółgłoską /w/

lub

c/ działania obu tych czynników równocześnie :

- głoski /p/ i /j/ są głoskami miękkimi /palatalnymi/ i w przypadku syntezy /p/ bez użycia toru nosowego brak jest parametru zdecydowanie je różnicującego,
- głoska drżąca /r/ w przypadku zbyt płytkiej, periodycznej modulacji amplitudowej oraz słabego trzeciego formantu /co jest cechą sygnału mowy produkowanego przy użyciu układu CT-1/ wykazuje duże podobieństwo do głoski /l/, co jest zgodne z subiektywną interpretacją tych głosek.

Inną prawidłowością dającą się zauważyć w wynikach odbioru spółgłosek jest większa ilość błędów występująca dla połączeń spółgłosek z samogłoskami /i/, /ɨ/ oraz /u/. Samogłoski te mają niską częstotliwość pierwszego formantu i są gorzej percypowane. Natomiast samogłoski /e/, /a/ oraz /o/ będące bardziej otwartymi samogłoskami, charakteryzuje wysoka częstotliwość pierwszego formantu, czym zdecydowanie różnią się od występujących w analizowanych połączeniach spółgłosek. Łatwość odbioru tych samogłosek pozwala na koncentrację uwagi na spółgłosce i lepszą jej percepcję.

3. Wnioski końcowe.

Przedstawione w tej pracy wyniki eksperymentów z syntezą wybranych diał polskich i ich percepcją należy traktować jako etap wstępny badań mających na celu :

- a/ maksymalne ujednoczenie przebiegów parametrów akustycznych w stadiach ustalonych dla samogłosek i spółgłosek oraz w stadiach przejściowych pomiędzy nimi przy zapewnieniu możliwie najlepszej jakości mowy syntetycznej,
- b/ syntezę wyrazów i prostych zdań bazującą na uzyskanym zasobie diał,
- c/ opracowanie na tej podstawie zasad syntezy według reguł bazującej na uzyskanym eksperymentalnie zasobie elemen-

tarnych fragmentów mowy syntetycznej.

Uzyskane wyniki można uznać za dobre i wskazujące na możliwość dokonania poprawnej syntezy mowy w zakresie badanego materiału fonetycznego.

Za pomoc i cenne uwagi udzielone podczas realizacji niniejszej pracy dziękuję prof. W. Jassemowi, dr inż. H. Kubzdeli i dr P. Łobacz.

BIBLIOGRAFIA

- [1] CIARKOWSKI, R., Algorytmizacja i oprogramowanie działań operatorskich dla parametrycznej syntezy mowy za pomocą układu CT-1 sterowanego z minikomputera MERA 303, Prace IPPT 14/83, Warszawa, 1983.
- [2] CIARKOWSKI, R., Parametryczna synteza mowy z zastosowaniem cyfrowo-analogowego syntetyzatora COMPUTALKER CT-1 sterowanego z minikomputera MERA 303, Prace XXX Otwartego Seminarium z Akustyki, t.1, str. 7-10, Gdańsk, 1983.
- [3] KUBZDELA, H., Weryfikacja i optymalizacja metody rozpoznawania wyrazów w skończonych zbiorach hasłowych w oparciu o spektrogramy binarne, Prace IPPT 10/82, Warszawa, 1982.
- [4] KUBZDELA, H., Metoda wyznaczania położenia formantów w widmie sygnału mowy, Prace XXX Otwartego Seminarium z Akustyki, t.1, str. 27-30, Gdańsk, 1983.
- [5] MYTKOWSKI, K., Kanał funkcji analogowych typ KF-01 do wprowadzania i wyprowadzania informacji w systemie "on-line" do/z pamięci minikomputera MOMIK 8B/100, Prace IPPT 39/76, Warszawa, 1976.

Diady typu CV	Poprawne rozpoznania [%]		Diady typu VC	Poprawne rozpoznania [%]	
	lista I	lista II		lista I	lista III
de	100	92	ej	92	92
da	100	96	ej	92	100
do	92	100	aj	100	100
du	84	88	oj	100	100
li	76	68	uj	60	56
le	96	96	il	96	84
la	100	100	il	68	68
lo	96	84	el	100	92
lu	68	72	al	96	92
wi	88	100	ol	96	92
wi	96	92	ul	88	72
we	100	100	iw	96	100
we	100	100	iw	84	100
wo	100	100	ew	100	100
wu	64	64	aw	100	96
mi	88	72	ow	100	100
mi	88	88	ul	40	48
me	100	100	im	84	84
me	100	100	im	100	100
mo	100	92	em	100	100
mu	96	92	am	100	100
ni	40	60	om	92	92
ne	64	48	um	92	92
na	60	68	in	40	44
no	56	52	in	56	52
nu	24	24	en	72	84
pi	68	50	an	84	88
pe	44	64	on	64	60
pa	68	56	un	12	16
po	64	76	ip	64	44
pu	76	68	ip	76	88
ri	48	64	ep	80	96
ri	80	58	ap	100	84
re	64	84	op	32	20
ra	72	76	up	64	68
ro	76	92	ir	44	64
ru	52	60	ir	88	96
vi	32	24	er	76	96
vi	40	16	er	72	96
ve	16	24	or	100	96
va	20	76	or	96	96
vo	64	72	ur	96	96
vu	92	84			

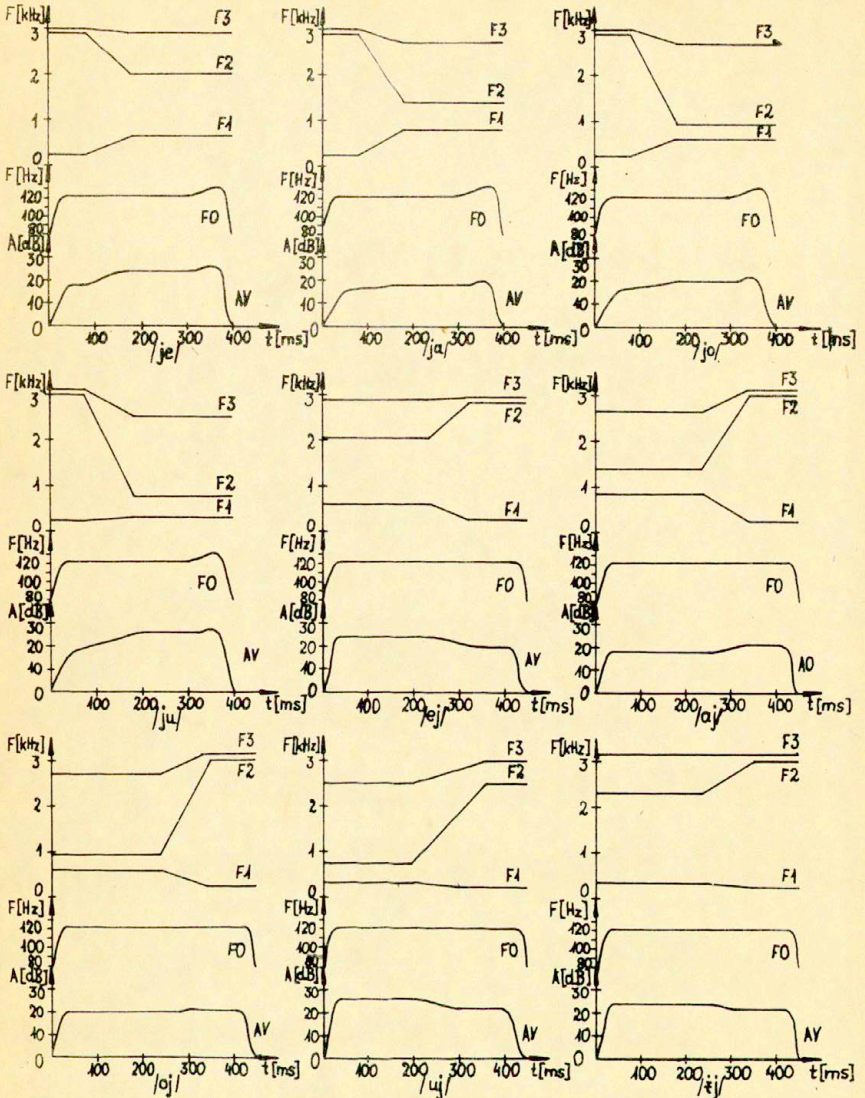
Tabela 1. Poprawność rozpoznawania syntetycznych diad polskich.

Samogłoska	Poprawność rozpoznawania samogłoski (w %) w połączeniach			
	typu CV		typu VC	
	lista 1	lista 2	lista 1	lista 3
/i/	97.3	100.0	98.0	97.3
/ɛ/	96.0	100.0	92.6	99.4
/e/	100.0	100.0	100.0	99.4
/a/	100.0	100.0	100.0	100.0
/o/	99.5	98.0	98.3	98.9
/u/	100.0	98.0	99.4	98.9

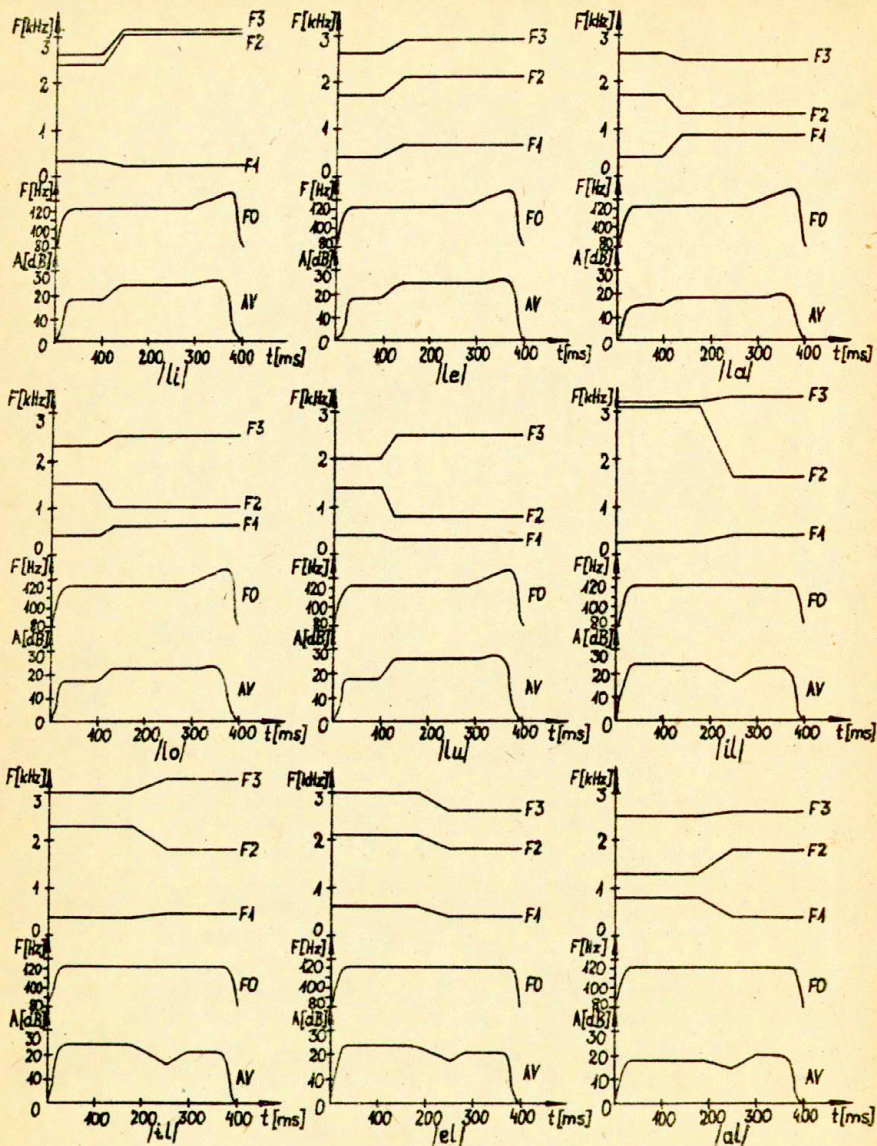
Tabela 4. Poprawność rozpoznawania samogłosek syntetycznych.

Spółgłoska	Poprawność rozpoznawania spółgłoski (w %) w połączeniach			
	typu CV		typu VC	
	lista 1	lista 2	lista 1	lista 3
/j/	94.0	94.0	88.8	89.6
/ɹ/	87.2	84.0	93.3	84.0
/w/	92.0	92.7	88.0	90.7
/m/	95.3	91.3	96.7	97.3
/n/	48.8	50.4	56.0	57.3
/ɲ/	64.0	64.8	71.3	70.0
/r/	65.3	74.0	80.0	90.7
/v/	44.0	49.3	-	-

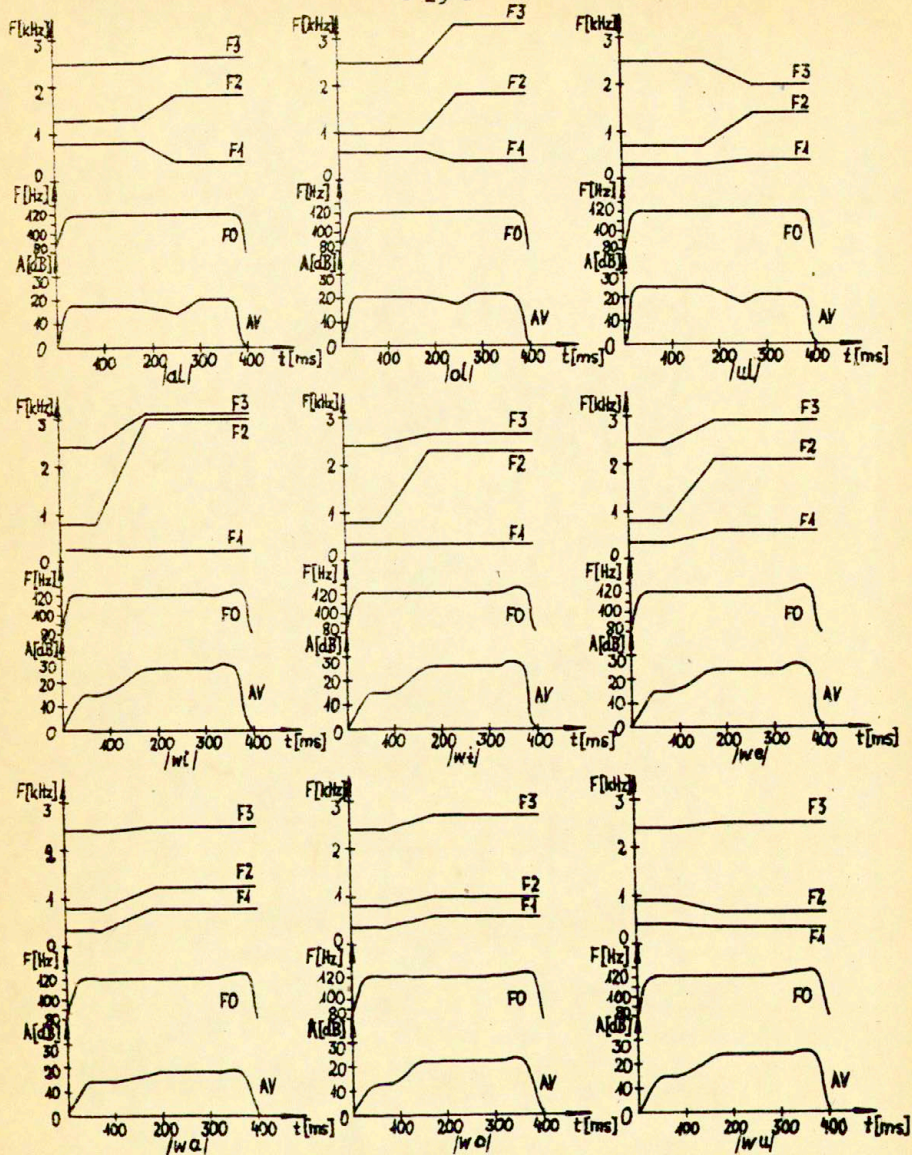
Tabela 5. Poprawność rozpoznawania spółgłosek syntetycznych.



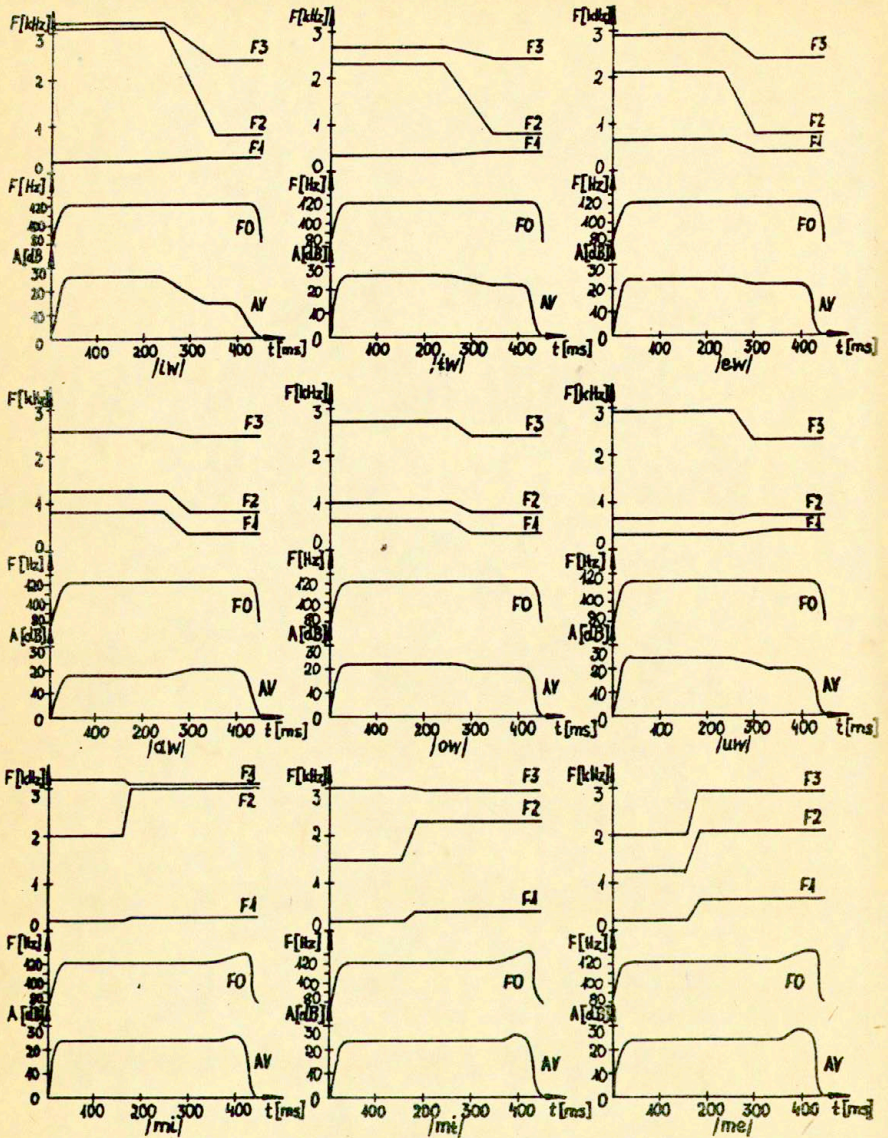
Rys. 2.1.a. Wykresy przebiegów parametrów akustycznych dźwięków syntetycznych.



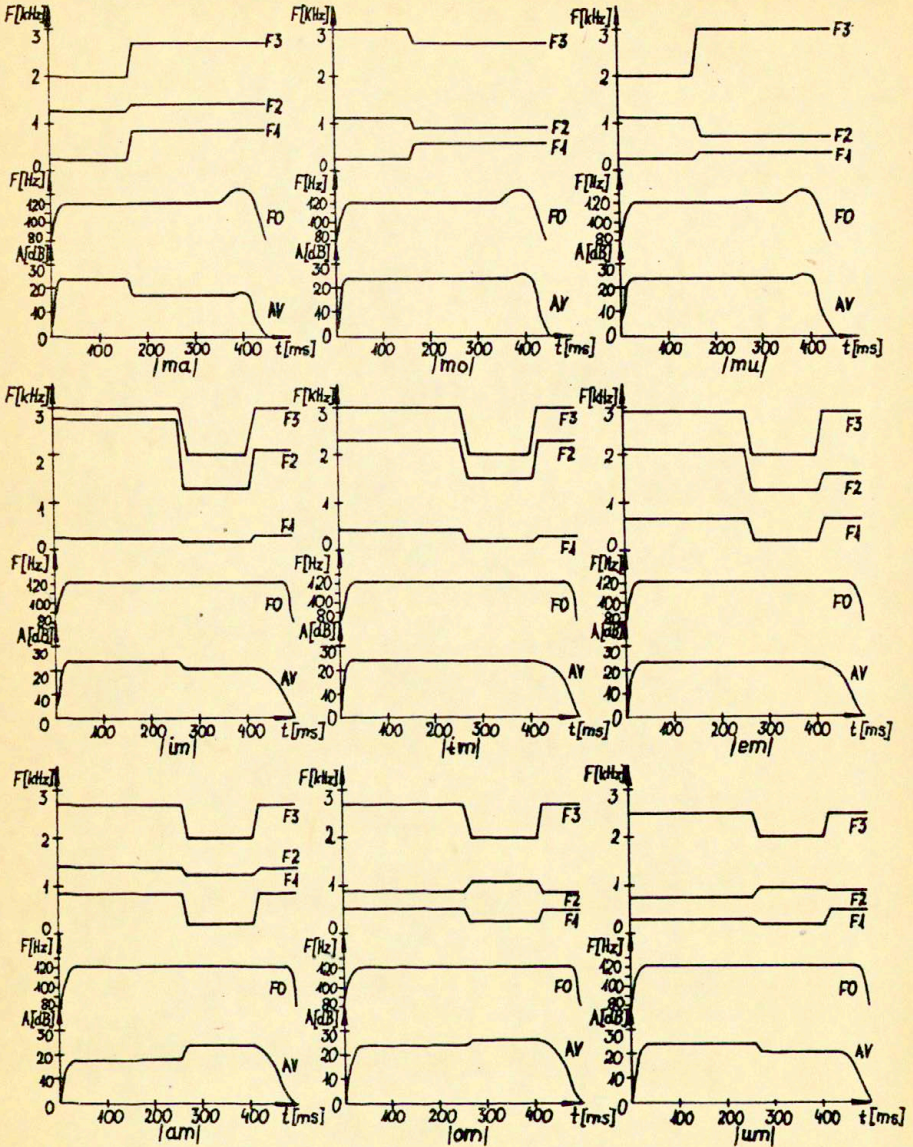
Rys. 2.1.b. Wykresy przebiegów parametrów akustycznych dźwięków syntetycznych.



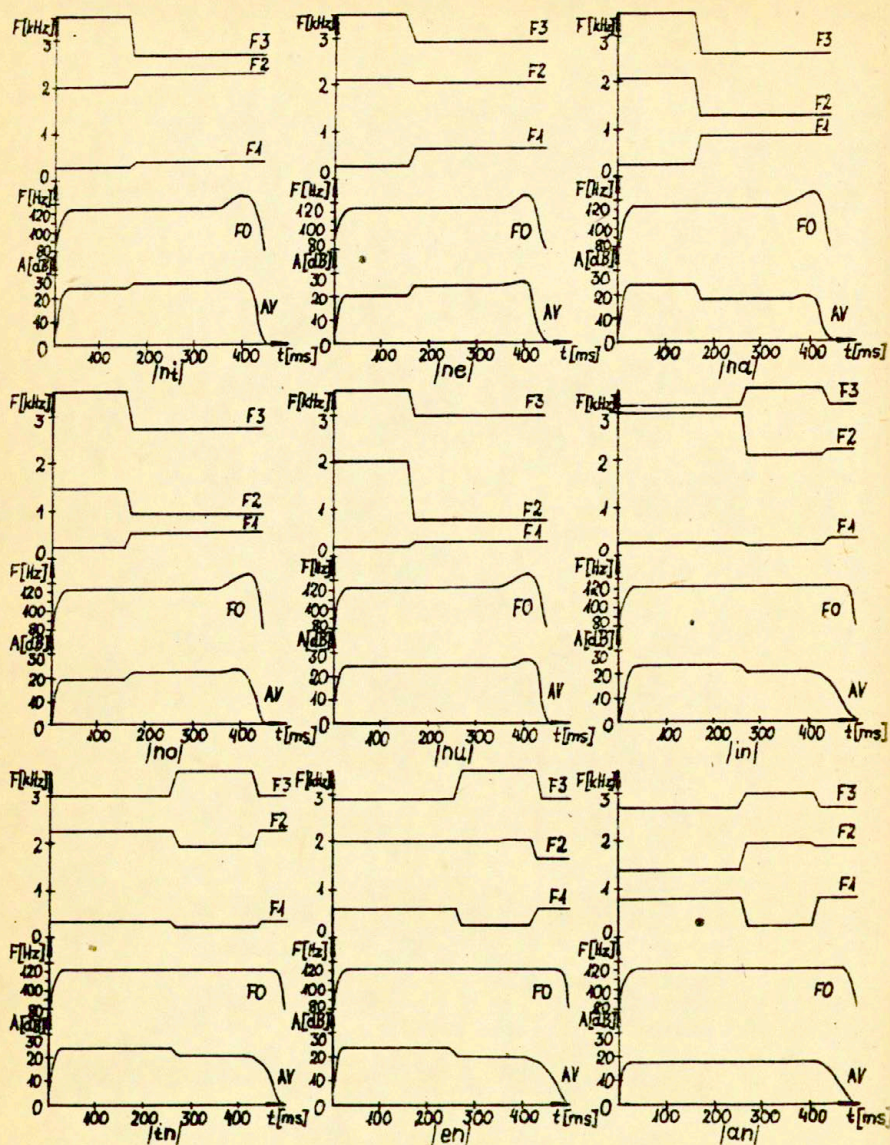
Rys. 2.1.c. Wykresy przebiegów parametrów akustycznych dźwięków syntetycznych.



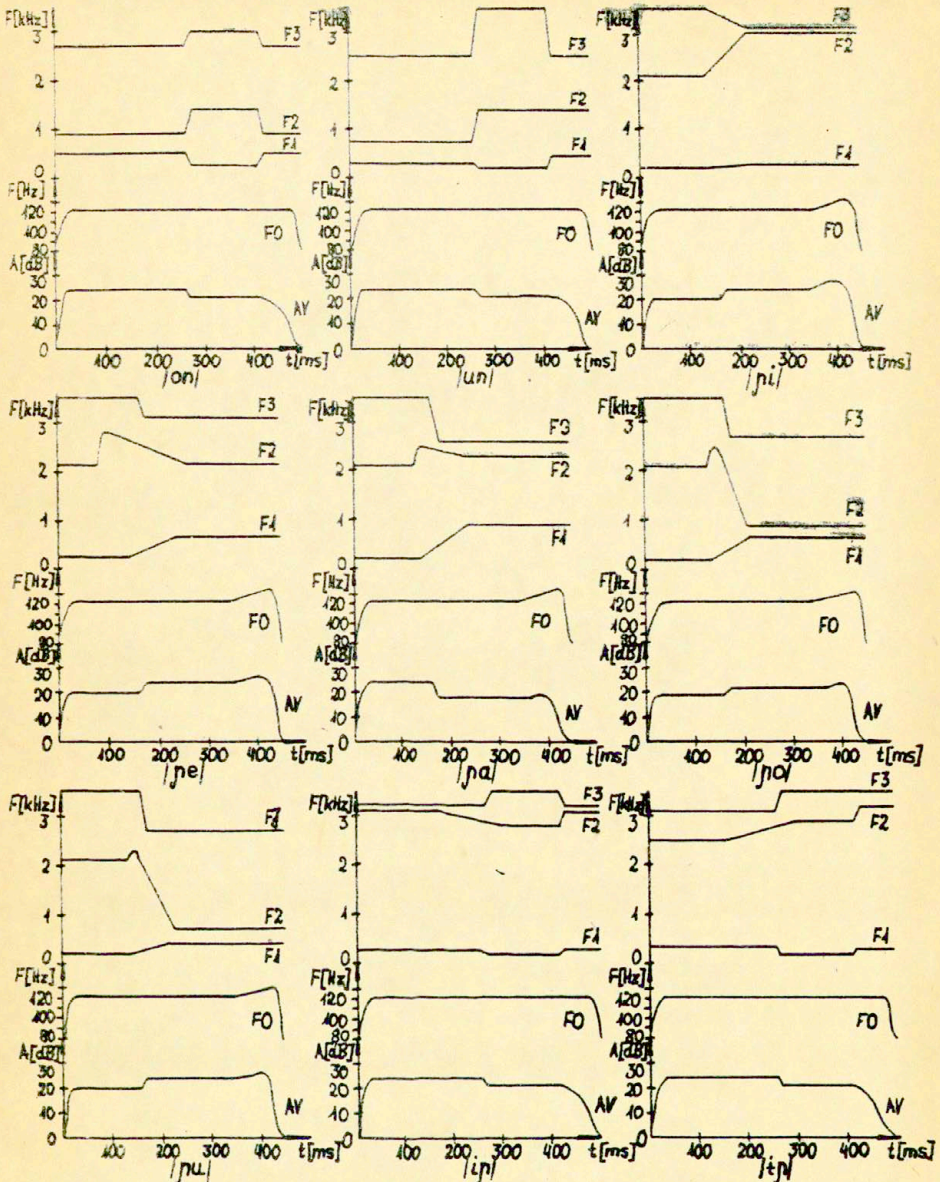
Rys. 2.1.d. Wykresy przebiegów parametrów akustycznych diał syntetycznych.



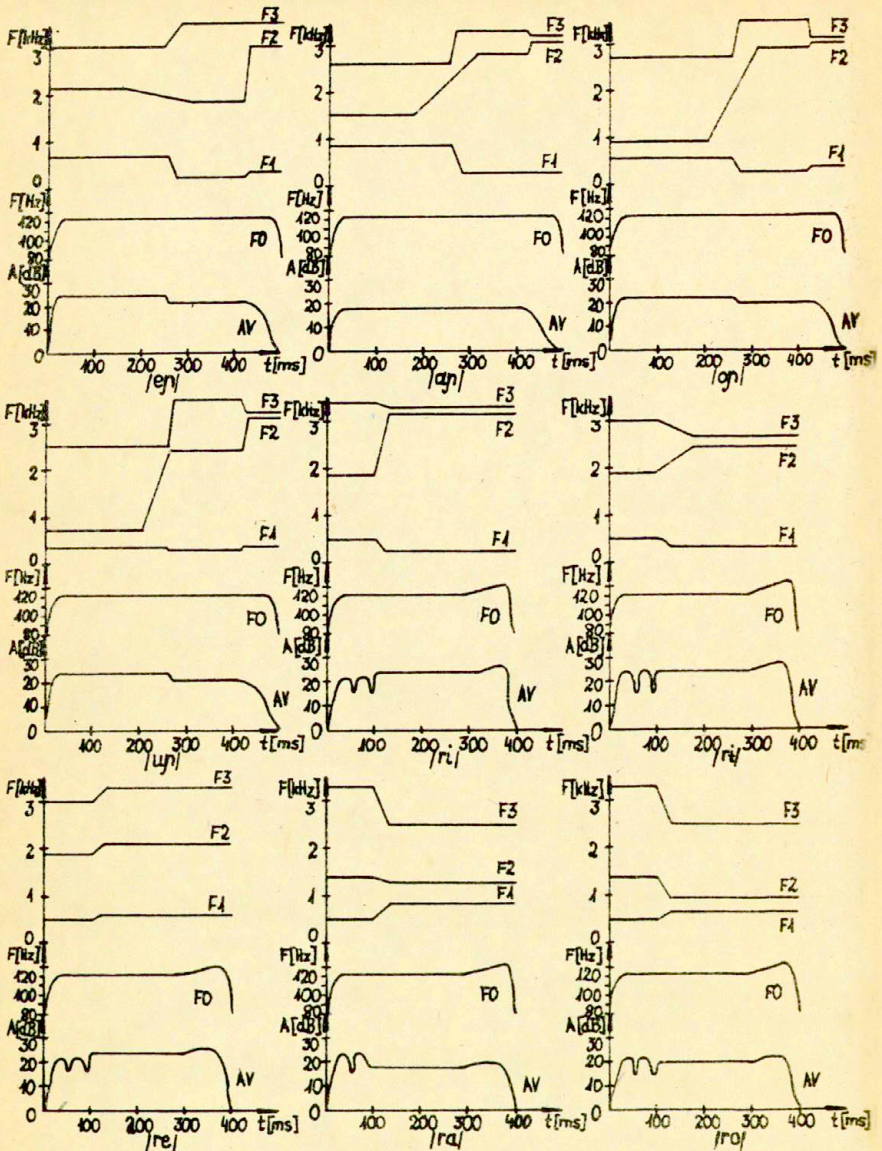
Rys. 2.1.e. Wykresy przebiegów parametrów akustycznych diad syntetycznych.



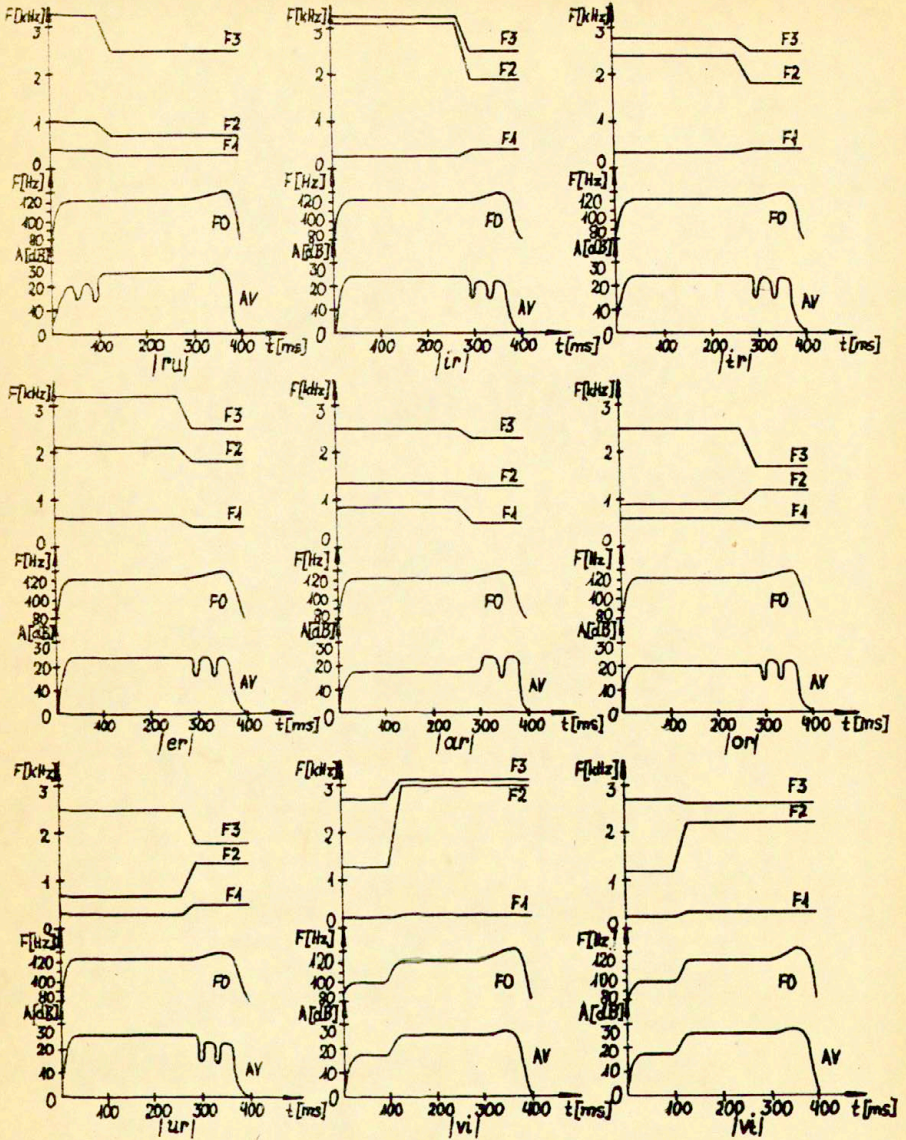
Rys. 2.1.2. Wykresy przebiegów parametrów akustycznych dla dźwięków syntetycznych.



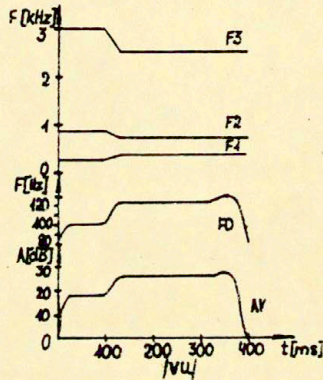
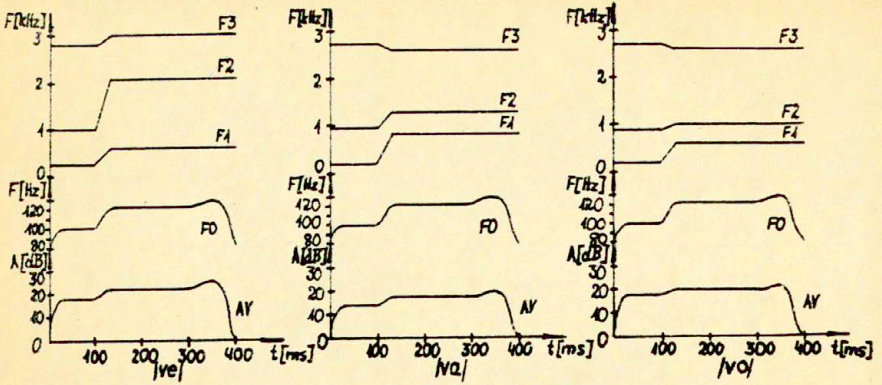
Rys. 2.1.g. Wykresy przebiegów parametrów akustycznych diał syntetycznych.



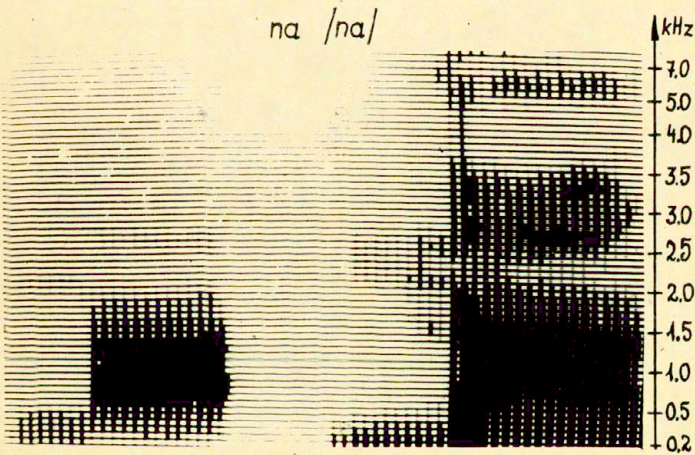
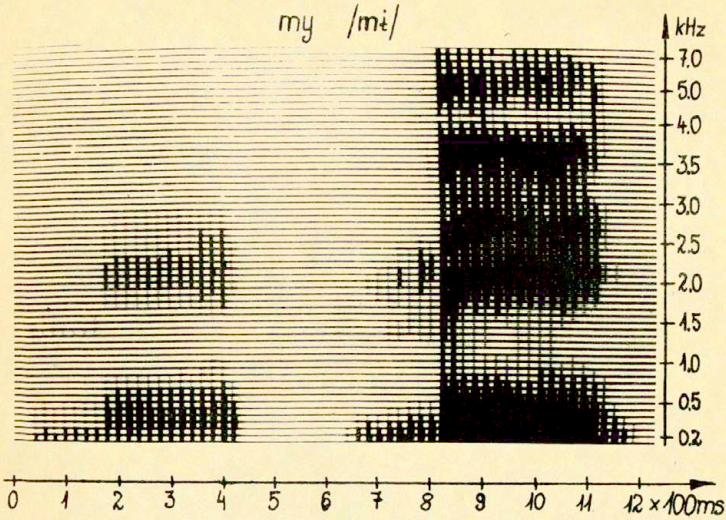
Fys. 2.1.h. Wykresy przebiegów parametrów akustycznych diad syntetycznych.



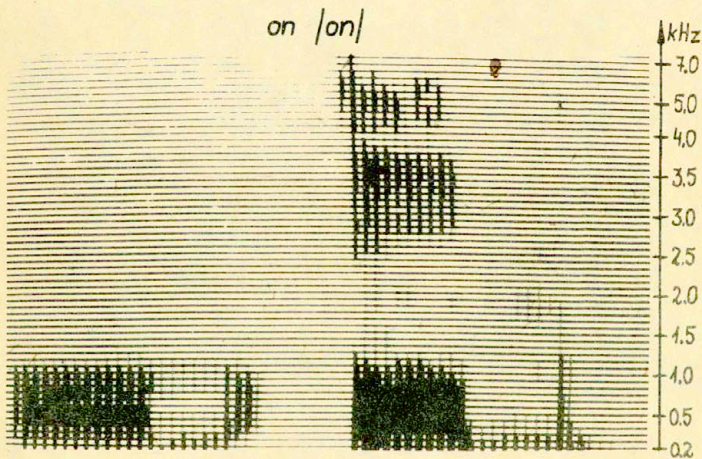
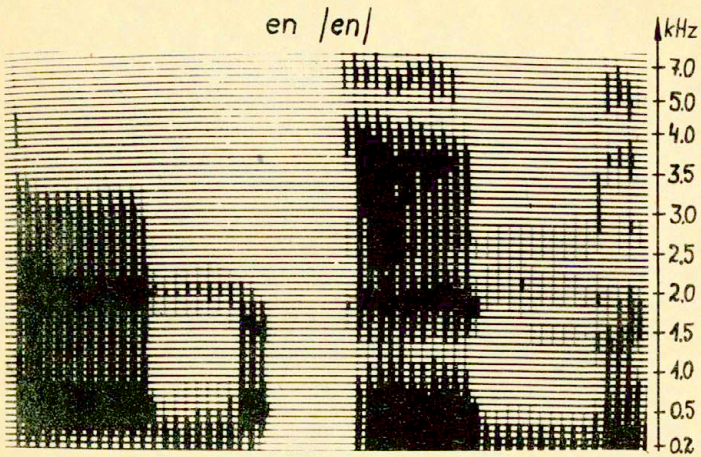
Rys. 2.1.1. Wykresy przebiegów parametrów akustycznych diał syntetycznych.



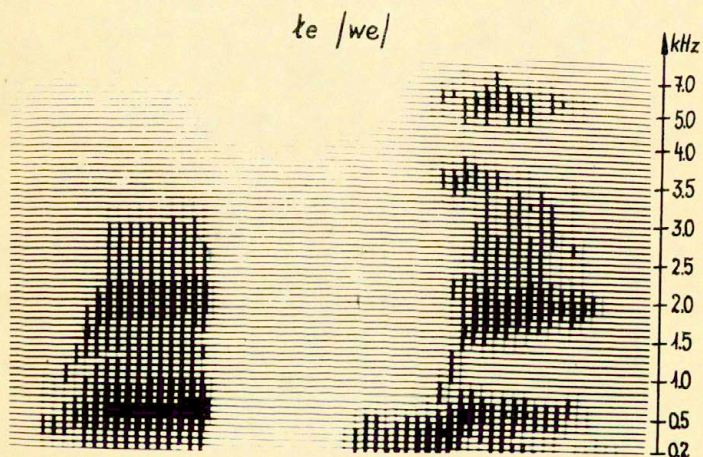
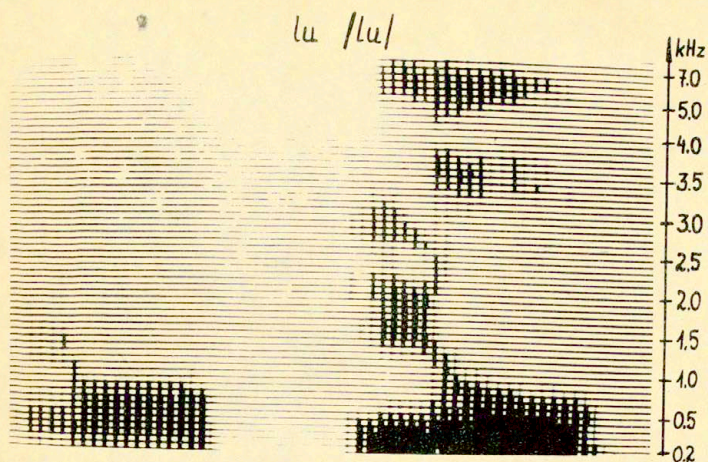
Rys. 2.1.3. Wykresy przebiegów parametrów akustycznych dźwięków syntetycznych.



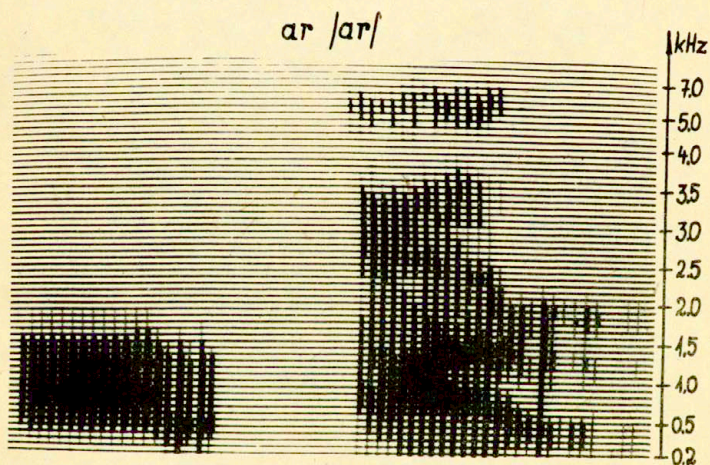
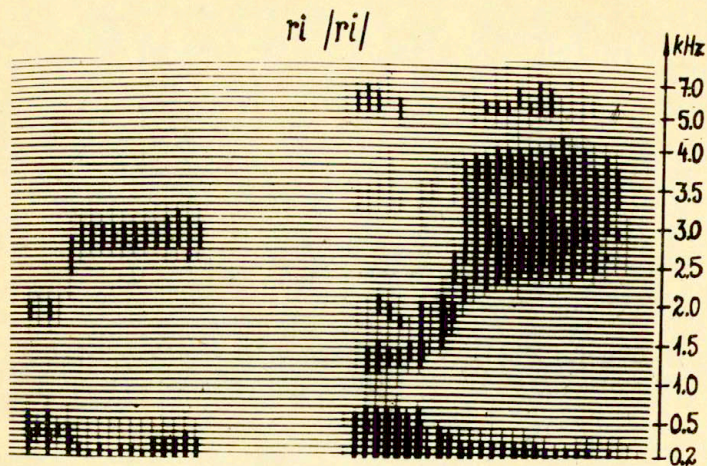
Rys. 2.2.a. Spektrogramy wybranych diad wykonane dla mowy naturalnej /po stronie prawej/ oraz dla produktu jej syntetycznej symulacji /po stronie lewej/.



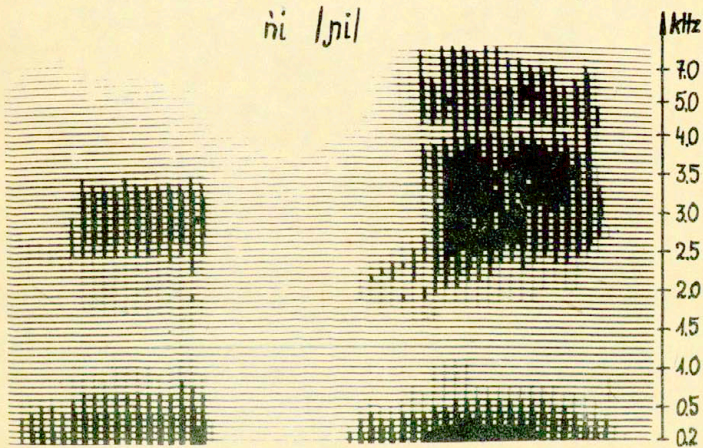
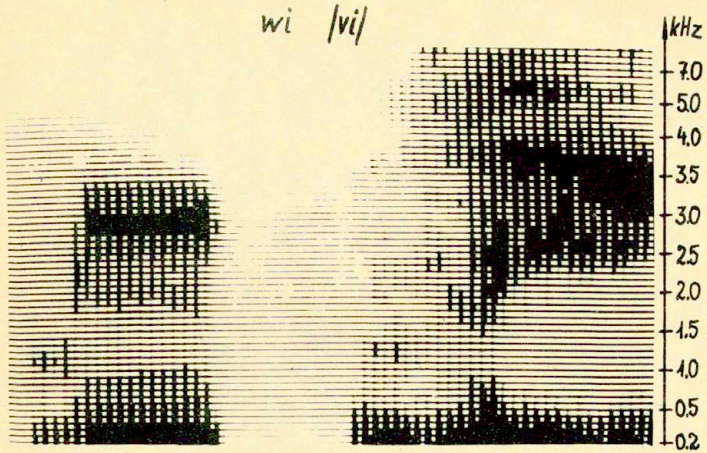
Rys. 2.2.b. Spektrogramy wybranych dźwięków wykonane dla mowy naturalnej /po stronie prawej/ oraz dla produktu jej syntetycznej symulacji /po stronie lewej/.



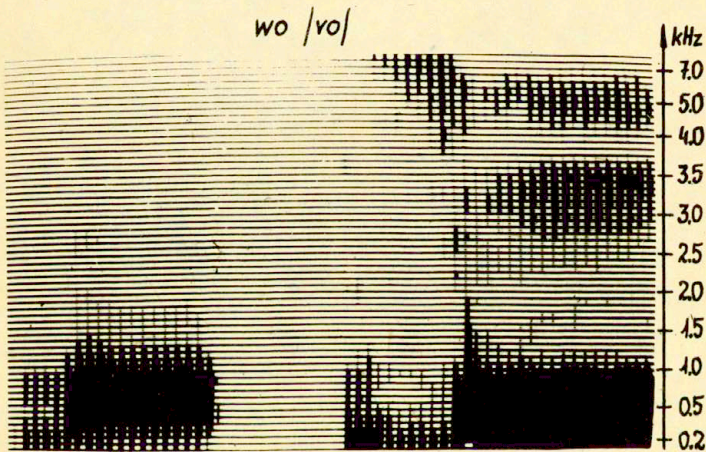
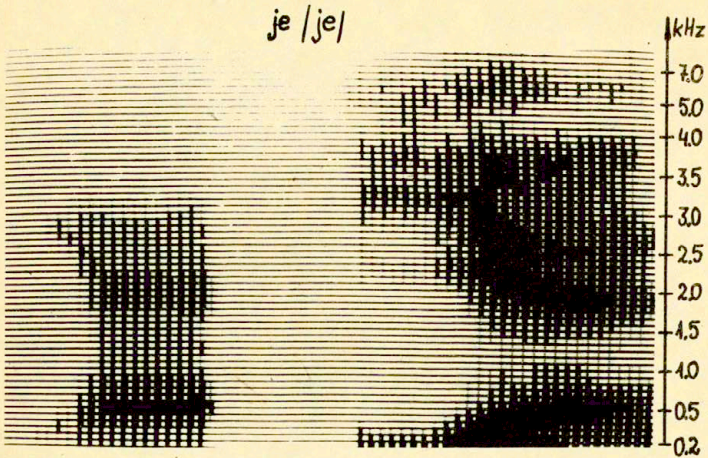
Rys. 2.2.c. Spektrogramy wybranych diad wykonane dla mowy naturalnej /po stronie prawej/ oraz dla produktu jej syntetycznej symulacji /po stronie lewej/.



Rys. 2.2.d. Spektrogramy wybranych diad wykonane dla mowy naturalnej /po stronie prawej/ oraz dla produktu jej syntetycznej symulacji /po stronie lewej/.



Rys. 2.2.e. Spektrogramy wybranych diad wykonane dla mowy naturalnej /po stronie prawej/ oraz dla produktu jej syntetycznej symulacji /po stronie lewej/.



Rys. 2.2.f. Spektrogramy wybranych dźwięków wykonane dla mowy naturalnej /po stronie prawej/ oraz dla produktu jej syntetycznej symulacji /po stronie lewej/.