

~~c.d. 380 800 4 s. k.~~

k. inf.

2.13 — Analiza i synteza mowy

Grażyna Demenko

NORMALIZACJA PRZEBIĘGÓW
CZĘSTOTLIWOŚCI PODSTAWOWEJ

WARSZAWA 1931

ISSN 0208-5658

Praca wpłynęła do Redakcji dnia 13 listopada 1980 r.,
poprawiona wersja 7.VII.1981 r.

Zarejestrowana pod nr 26/1981



1085622 II



Na prawach rękopisu

Instytut Podstawowych Problemów Techniki PAN
Nakład 150 egz. Ark.wyd. 1. Ark. druk. 1,5 .
Oddane do drukarni w lipcu 1981 r.
Nr zamówienia 483/c

Bibl. UAM

81 EQ 598

Warszawska Drukarnia Naukowa, Warszawa,
ul. Śniadeckich 8

Grażyna Demenko
Pracownia Fonetyki Akustycznej
IPPT PAN

NORMALIZACJA PRZEBIEGÓW CZĘSTOTLIWOSCI PODSTAWOWEJ^{1/}

Streszczenie

W pracy przedstawiono metodę stosowaną do opisu przebiegów częstotliwości podstawowej. Dokonano wstępnego wygładzania danych pomiarowych F_0 otrzymanych w wyniku ekstrakcji tego parametru co 10 ms. Przyjęto pewien materiał językowy składający się z zapisanych na taśmie wzorców intonacyjnych oraz ich powtórzeń przez kilka osób. Wyniki normalizacji czasowej i częstotliwościowej porównano z rezultatami doświadczenia odśłuchowego, które miało za zadanie ustalić, czy powtarzane intonacje są podobne do wzorcowych. Metodę normalizacji zilustrowano szeregiem przykładów. Aproksymowano przebiegi intonacji funkcją typu $y_t = at^{b_e}ct$ metodą najmniejszych kwadratów. Podane przykłady aproksymacji danych doświadczalnych wskazują na właściwą postać przyjętej funkcji aproksymującej. Przedstawiono perspektywy dalszej kontynuacji pracy.

1. Wstęp.

Ze względu na szczególną zawartość informatywnych cech akustycznych i lingwistycznych częstotliwość podstawowa jest jednym z częściej analizowanych parametrów sygnału mowy. Przeprowadzono szereg badań statystycznych F_0 (por. np. [2], [3]) jak i również czasowej zmienności intonacji (por. np. [3], [5]). Brak jednak jest sprecyzowanego matematycznego opisu właściwości tego parametru.

Niniejsza praca jest kontynuacją jednej z wcześniejszych prac

IPPT [3] i ma na celu opracowanie matematycznego modelu normalizacji częstotliwości podstawowej. Znalazienie odpowiednich reguł, za pomocą których przebiegi F_0 - ekwiwalentne lingwistycznie choć z wielu przyczyn niejednakowe można będzie sprowadzić do inwariantnych postaci dla każdego wzorca intonacyjnego, jest konieczne dla automatycznej analizy tego parametru. Aby scharakteryzować kontury F_0 należy określić, które wśród wielu czynników oraz w jaki sposób wpływają na czasową zmienność częstotliwości podstawowej. Nie jest to zadanie proste, gdyż przebieg F_0 kształtuje szereg różnych przyczyn które zmieniają się w trakcie trwania wypowiedzi. Wypracowanie adekwatnego modelu opisu czasowych własności F_0 jest konieczne dla określenia stopnia przydatności tego parametru w szeregu podstawowych problemach fonetyki.

2. Przygotowanie analizowanego sygnału.

2.1. Materiał eksperymentalny.

Jedna osoba wymówiła frazę "do widzenia" sześciokrotnie tak aby wypowiedzi były różne intonacyjnie. Zostały one zapisane na taśmie magnetofonowej i traktowane jako wzorce intonacyjne. Wybrano losowo 7 osób, 5 głosów męskich i 2 żeńskie, które proszono o powtórzenie poszczególnych wzorcowych intonacji. Otrzymano nagranie typu : Głos 1 : Wzorzec 1 - odtworzenie 1, wzorzec 2 - odtworzenie 2, ..., Głos 7 : Wzorzec 1 - odtworzenie 1 ... itd. Łącznie z intonacjami wzorcowymi uzyskano 48 intonacji 8 osób wymawiających sześciokrotnie frazę "do widzenia".

Zespół odsłuchowy składał się z 10 osób, które miały za zadanie ocenić czy odtworzenie kolejnych intonacji wzorcowych przez poszczególne osoby jest takie samo, podobne lub niepodobne do wzorca. Wyniki doświadczenia ilustruje Tablica 1. W poszczególnych kolumnach tablicy podano łączną ilość odpowiedzi oznaczonych literkami T, P, N, udzielonych na pytanie czy słyszane intonacje były takie same, podobne lub niepodobne do intonacji wzorcowych.

Tablica 1

Wyniki doświadczenia odsłuchowego

Numer intonacji	Głos 1			Głos 2			Głos 3			Głos 4			Głos 5			Głos 6			Głos 7		
	T	P	N	T	P	N	T	P	N	T	P	N	T	P	N	T	P	N	T	P	N
1	3	7		4	5	1	8	2		5	3	2	2	2	6	3	6	1	5	4	1
2	5	3	2	3	6	1	9	1		6	3	1	2	7	1	6	4		5	4	1
3	6	3	1	5	5		8	2		7	3		3	5	2	6	4		8	2	
4	7	3		9	1		10			2	6	2	1	6	3	5	5		8	1	1
5	8	2		9	1		8	2		6	4		5	5		8	2		9	1	
6	2	6	2	4	6		10			7	3		3	7		5	5		8	2	

Najlepsze oceny otrzymał głos, który podawał wypowiedzi wzorcowe. Nagrane na taśmie wzorce intonacyjne oraz ich powtórzenia przez wszystkie osoby stanowiły materiał do dalszych analiz.

2.2. Wstępne przetworzenie danych.

Dla wykonania pomiarów parametru F_0 wykorzystano tonometr TM 3. Metodę otrzymywania cyfrowych wyników ekstrakcji tego parametru opisano w pracy [2].

W kilku wątpliwych fragmentach otrzymane pomiary uzupełniono odczytami pierwszej harmonicznej ze spektrogramów.

W pewnych przypadkach na bardzo krótkich odcinkach czasowych pomiar F_0 okazał się niemożliwy ze względu na brak dźwięczności w środkowej części frazy "do widzenia". Zaznaczono to na wydrukach odpowiednimi czasowymi przerwami. Poszczególne intonacje wyprowadzono na oscyloskop a orientacyjne wykresy przebiegów F_0 na drukarkę wierszową.

Kierując się zasadą, że krzywa ilustrująca na wykresie związek funkcyjny między zmiennymi powinna być możliwie gładka oraz przebiegać w pobliżu wszystkich punktów, choć niekoniecznie przechodzić przez te punkty, otrzymane pomiary zostały wygładzone. Celem wygładzania danych powinno być nie tylko otrzymanie możliwie najbardziej gładkiej krzywej przedstawiającej te dane ale i najbardziej prawdopodobnego przebiegu. Zakładamy

tutaj, że prawdziwe wartości zmiennej tworzą gładki ciąg, natomiast w rzeczywistości zmierzonych wartościach występują nieregularności skutkiem przypadkowych błędów pomiaru. Przez wygładzanie rozumiemy proces znajdowania nowych wartości zmiennych zależnych $y'_1 \dots y'_n$ możliwie mało różniących się od wartości zmiennych zależnych pierwotnego układu $y_1 \dots y_n$ lecz takich dla których różnice tablicowe zmieniają się regularnie. Wielomian najbardziej odpowiedni w sensie aproksymacji średniokwadratowej ma następujące dwie własności :

1. Jego stopień jest na tyle wysoki, że dobrze przybliża prawdziwą funkcję
2. Nie jest zbyt ściśle dopasowany do danych empirycznych tak, by był możliwie nieczuły na zakłócenia lub niedokładności pomiarów.

Założono liniowość w kolejnych trzech punktach i zastosowano wygładzanie trzypunktowe według poniższych wzorów (por. [6], [7]) - dla punktu pierwszego

$$y_1(-1) = \frac{1}{6}(5f_{-1} + 2f_0 - f_1) \quad (1)$$

- dla punktu końcowego

$$y_1(1) = \frac{1}{6}(-f_{-1} + 2f_0 + 5f_1) \quad (2)$$

- dla punktów pozostałych

$$y_1(0) = \frac{1}{3}(f_{-1} + f_0 + f_{+1}) \quad (3)$$

Przez f_{-1} , f_0 , f_1 oznaczono wartości funkcji następujące kolejno po sobie.

Niekiedy celem poprawienia jakości wygładzenia stosujemy podane wzory nawet kilkakrotnie do danych już wygładzonych. Dla sprawdzenia jakości wygładzania wyznaczono pierwsze różnice tablicowe, a dla stwierdzenia zgodności obydwu układów danych obliczono różnice między wartościami danymi i wygładzonymi w kolejnych punktach czasowych. Pierwszą różnicę tablicową zdefiniowaną w punkcie $x_1 = x_0 + h$ (h jest interwałem argumentu tablicy) określamy wzorem :

$$\Delta y_i = y_{i+1} - y_i = f[x_0 + (i+1)h] - f(x_0 + ih) \quad (4)$$

Dla zilustrowania metody odczytane wartości F_0 oraz ich postać po wygładzeniu z pewnego fragmentu wydruku analizowanej intonacji zestawiono w Tabelicy 2.

Tablica 2

Wygładzanie wartości funkcji empirycznej

t_i	y_i	Δy_i	y'_i	$\Delta y'_i$	$y'_i - y_i$
170	197	-3	196	0	-1
180	194	+3	196	+2	+2
190	197	+5	198	+4	+1
200	202	+6	202	+5	0
210	208	+6	207	+7	-1
220	212	+9	214	+8	+2
230	221	+13	222	+10	+1
240	234	+6	232	+4	-2
250	240	-6	236	+1	-4
260	234	+3	237	-4	+3
270	237	-8	233	-2	-4
280	229	-1	231	-3	+2
290	228	0	228	+1	0
300	228	+4	229	+4	+1
310	232	+8	233	+5	+1
320	240	+34	238	+12	-2
330	243	+34	250	+9	+7
340	267	0	259	+20	-8
350	267	+36	279	+16	+12
360	303	+13	295	+18	-8
370	316	+3	313	+10	-3
380	319	+3	323	+6	+4
390	334	+15	329	+9	+5
400	334	0	329	+9	-5
410	334	0	338	+6	+4
420	347	+13	344	+6	-3
430	347	+3	348	+4	-2
440	350	-3	348	+5	-2
450	347	+3	349	+1	+2
460	350	+3	350	+1	0
470	353	+3	353	+2	0

Z podanego przykładu widać, że różnice tablicowe funkcji wygładzonej układają się bardziej regularnie niż danych doświadczalnych.

Wygładzone próbki tworzą wektor pomiarów F_0 .

$$F = \begin{pmatrix} F_{t1} \\ F_{t2} \\ \vdots \\ F_{tn} \end{pmatrix}$$

3. Normalizacja czasowa i częstotliwościowa konturów F_0 .

Aby istniała możliwość matematycznego porównania wyekstrahowanych przebiegów F_0 , konieczne jest dokonanie normalizacji czasowej oraz częstotliwościowej]. Długości badanych przebiegów były dość zróżnicowane zależnie od tempa mowy i wahały się w granicach 720-1020 ms, natomiast zróżnicowania w wartościach częstotliwości w określonych punktach czasowych w obrębie powtórzeń danej intonacji przez różne osoby dochodzą do stosunku 1:2.

Normalizacja tyczyć będzie wyeliminowania wyżej wymienionych zróżnicowań. Traktując wartość średnią jako pewną stałą a stopień odchylenia od niej jako charakterystyczne cechy danego przebiegu nasuwa się sugestia, aby normalizację częstotliwościową przeprowadzić w oparciu o tę wartość.

Obliczono średnią częstotliwość sześciu intonacji wzorcowych, która została przyjęta jako punkt odniesienia dla pozostałych przebiegów. Określono częstotliwościowy współczynnik normalizujący zależnością :

$$K_{Fji} = \frac{\sum_{i=1}^N F_{ji}}{\bar{F}_w}$$

gdzie $\bar{F}_w = 102$ Hz (średnia częstotliwość intonacji wzorcowych)

j - numer intonacji $j = 1 \dots 6$

$i = 1 \dots N$ (N - ilość pomiarów).

Znormalizowane wartości częstotliwości obliczono na podstawie wzoru

$$F'_{ji} = \frac{F_{ji}}{K_{Fji}}$$

W ten sposób, nie zmieniając kształtu konturów F_0 , przez odpowiednie zwężenie lub poszerzenie dynamiki wykresu ustalona

została wspólna średnia wartość dla wszystkich przebiegów równa 102 Hz.

Normalizacja czasowa przebiegów F_0 jest problemem bardziej złożonym gdyż dotyczy zmian relacji czasowych w określonym układzie danych.

W niniejszej pracy zastosowano normalizację podobną do proponowanej w pracy [5]. Czasowa skala każdej intonacji została pomnożona lub podzielona przez odpowiedni współczynnik proporcjonalności tak, aby czas każdej intonacji równał się obliczonemu średniemu czasowi wszystkich wzorcowych intonacji wynoszącym 830 ms.

Określony został czasowy współczynnik normalizujący z zależności

$$K_{tj} = \frac{t_j}{\bar{t}}$$

gdzie $\bar{t} = 830$ ms

j = numer intonacji

Średnie czasy analizowanych intonacji oraz współczynniki normalizujące zestawiono w Tablicy 3.

W wyniku omówionych przekształceń otrzymano znormalizowane czasowo i częstotliwościowo przebiegi F_0 w układzie współrzędnych

$$t' = k_t t$$

$$F' = k_F F$$

Przykładowo na ryc. 1 przedstawiono efekt normalizacji częstotliwościowej jednego wzorca intonacyjnego dla wszystkich osób. Z ryciny widać, że intonacje powtarzane przez głosy 1 i 2 wizualnie najbardziej odbiegają od wzorcowej a intonacja powtarzana przez głos 3 jest najbardziej zgodna, co pokrywa się w pełni z wynikami doświadczenia odsłuchowego. Na ryc. 2,3,4,5 ukazano przebiegi F_0 znormalizowane czasowo i częstotliwościowo. Zasadnicze wnioski, które daje analiza wizualna wykresów są następujące :

1. W pewnych typach intonacji znormalizowane przebiegi wielokrotnie przecinają się, tak że zróżnicowania między głosami zostały usunięte (ryc. 3).

Tablica 3.
 Średnie czasy poszczególnych przebiegów F_0 .
 Czas podano w ms.

	Głos 1		Głos 2		Głos 3		Głos 4		Głos 5		Głos 6		Głos 7		Głos 8	
	\bar{t}_j	k_{tj}	\bar{t}_j	k_{tj}	\bar{t}_j	k_{tj}	\bar{t}_j	k_{tj}	\bar{t}_j	k_{tj}	\bar{t}_j	k_{tj}	\bar{t}_j	k_{tj}	\bar{t}_j	k_{tj}
1	1020	1,23	790	0,95	880	1,06	940	1,13	790	0,95	850	1,02	800	0,96	880	1,06
2	930	1,12	770	0,93	650	0,78	920	1,11	700	0,84	780	0,94	770	0,93	860	1,04
3	950	1,14	930	1,12	880	1,06	840	1,01	760	0,92	920	1,11	890	1,07	920	1,11
4	990	1,19	840	1,01	820	0,99	930	1,12	730	0,88	720	0,87	840	1,01	860	1,04
5	840	1,01	760	0,92	820	0,99	860	1,04	760	0,92	720	0,87	760	0,92	810	0,98
6	740	0,89	790	0,95	800	0,96	940	1,13	770	0,93	800	0,96	750	0,91	760	0,92

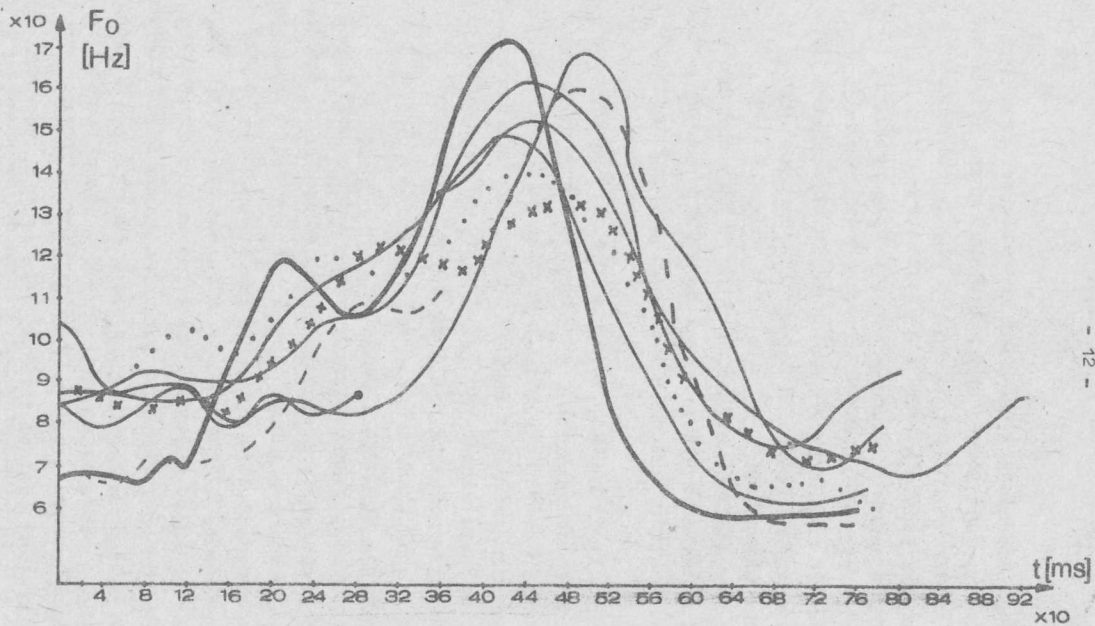
2. Intonacje posiadające bardziej zróżnicowany przebieg zostały znacznie gorzej znormalizowane (ryc. 5). Widać tutaj znaczenie relacji czasowych między charakterystycznymi punktami badanych przebiegów.
3. Porównując znormalizowane przebiegi F_0 z wynikami doświadczenia odsłuchowego stwierdza się zgodność wyników. Te intonacje, które uzyskały najgorsze i najlepsze oceny odsłuchowe również tak samo zostały ocenione wizualnie.

Charakterystyczne są dwa głosy : No. 3, który zarówno w odsłuchach i w analizie wykresów dał najkorzystniejszy wynik oraz głos No. 5, który konsekwentnie odbiegał najbardziej od wzorca.

Z podanych przykładów wynika, że normalizacja czasowa tego typu w niektórych przypadkach dość dobrze spełnia zadanie, w innych gorzej. Nasuwa się przypuszczenie, że normalizację czasową należy przeprowadzić nieliniowo w oparciu o krótsze segmenty np. sylaby. Biorąc pod uwagę fakt, że prędkość sygnału na całej wypowiedzi nie jest stała, rozwiązanie takie wydaje się słuszne.

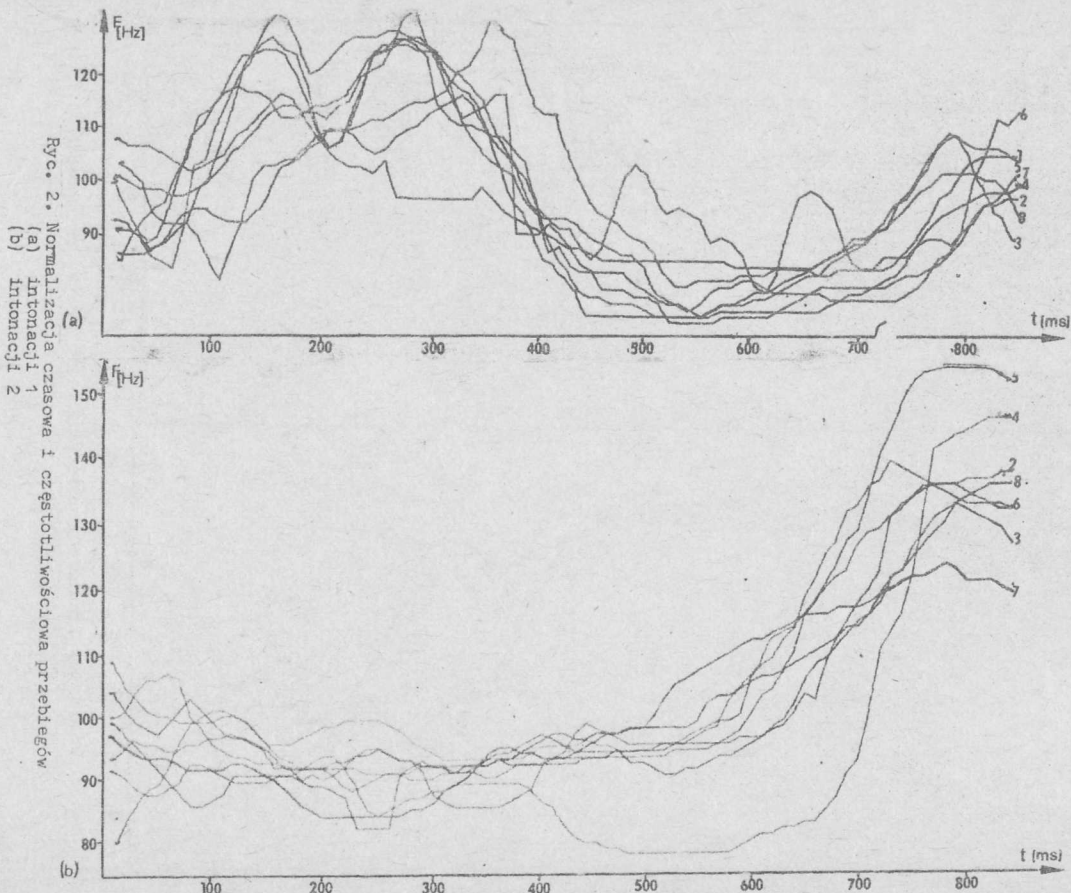
4. Aproksymacja badanych przebiegów.

Chcąc wykorzystać jakąkolwiek sytuację musimy znać prawa rządzące jej rozwojem. W prostych sytuacjach prawa te wyrażają się wzorami znanymi z algebry. Przy doborze postaci równania empirycznego przedstawiającego dane eksperymentalne należy kierować się zasadą, że powinno ono możliwie najlepiej przedstawiać zależność między wartościami zmiennych wynikającymi z pomiarów oraz zawierać jak najmniejszą ilość stałych. Aproksymacja polega na zastąpieniu danej funkcji $y = f(x)$ inną funkcją $y = W(x)$ odtwarzającą jej przebieg niekoniecznie dokładnie, lecz wygodniejszą do matematycznych analiz. Założenie właściwej postaci funkcji jest centralnym problemem aproksymacji. W odniesieniu do częstotliwości podstawowej Levitt i Rabiner proponują (por. [5]) aproksymację F_0 wielomianami ortogonalnymi w założonym oknie czasowym. W kilku pracach proponuje się odcinkową aproksymację przebiegów F_0 liniami prostymi. W niniejszej pracy zastosowano aproksymację innego typu.

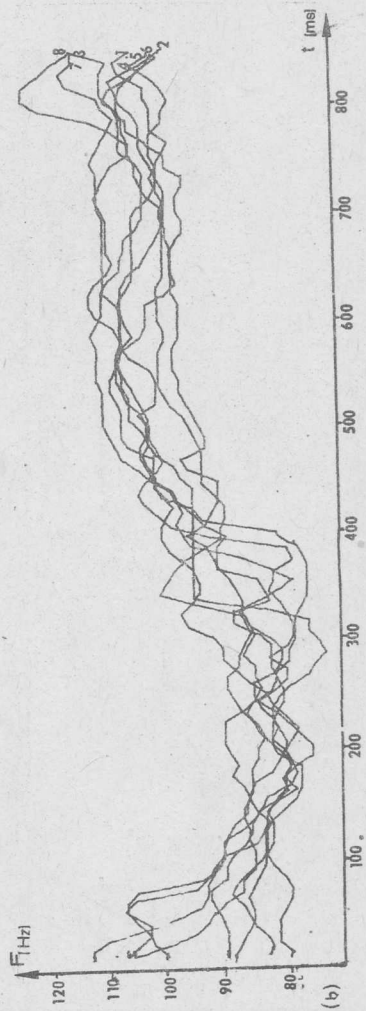
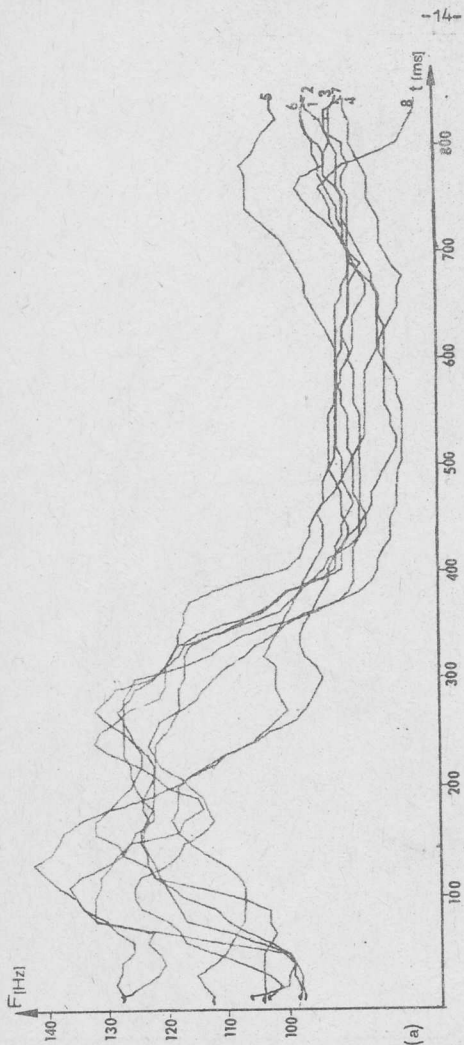


Ryc. 1. Znormalizowane częstotliwościowe przebiegi wszystkich intonacji. Przyjęto oznaczenia.

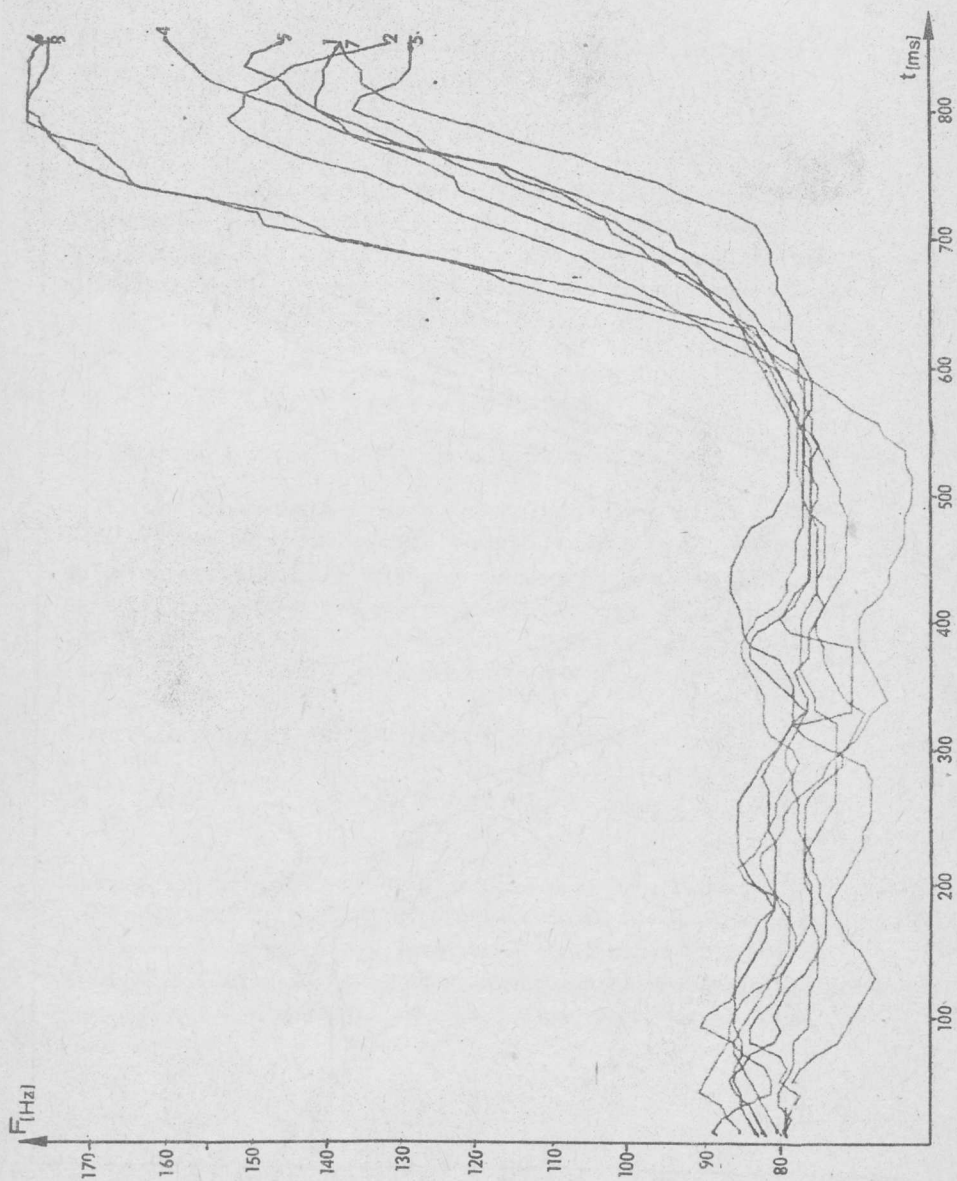
- Głos podający intonację wzorcową
- - - Głos 3
- ... Głos 2
- xxx Głos 1
- pozostałe intonacje



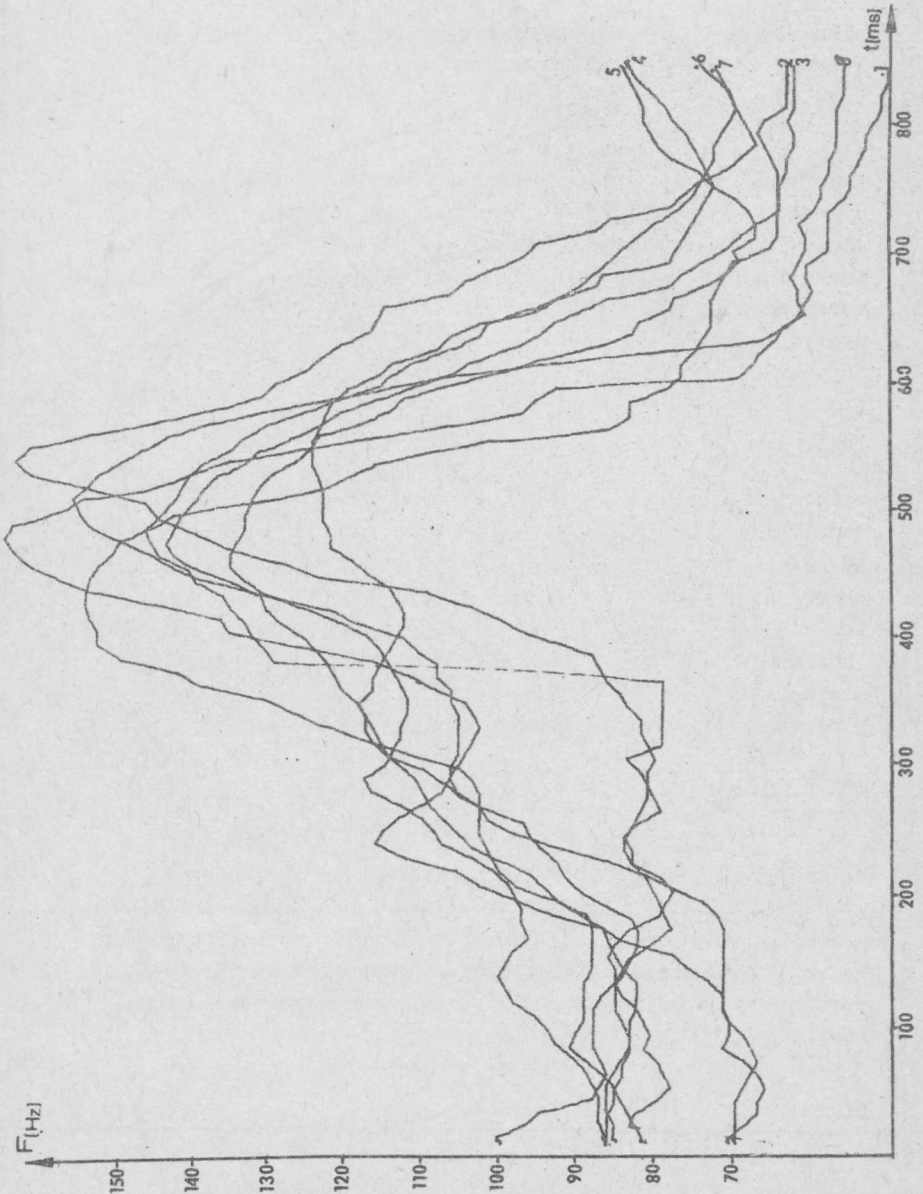
Ryc. 2. Normalizacja czasowa i częstotliwościowa przebiegów
(a) intonacji 1
(b) intonacji 2



Ryc. 3. Normalizacja czasowa i częstotliwościowa przebiegów
 (a) intonacji 3
 (b) intonacji 4



Ryc. 4. Normalizacja czasowa i częstotliwościowa przebiegów intonacji 5.



Ryc. 5. Normalizacja czasowa i częstotliwościowa przebiegów intonacji 6.

Jako właściwie spełniającą wyżej wymienione założenia przyjęto funkcję o następującej strukturze :

$$y_t = a t^b e^{ct} \quad (5)$$

Tego typu funkcję sugeruje Khoker (por. [4]). Większość metod wyznaczenia stałych równania odnosi się do równań, które są liniowymi funkcjami stałych lub też mogą być sprowadzone do funkcji liniowej poprzez transformacje zmiennych. Logarytmując wyżej podaną zależność otrzymujemy funkcję zależną liniowo od stałych a, b, c .

$$\ln y_{(t)} = \ln a + b \ln t + ct \quad (6)$$

gdzie a, b, c - stałe, które należy wyznaczyć, t - czas.

W celu wyznaczenia stałych występujących w równaniu posłużono się metodą najmniejszych kwadratów, która pozwala nie tylko na osiągnięcie dość dużej dokładności ale również na określenie błędów wyznaczonych stałych. Przyjmujemy tutaj założenie, że wartości eksperymentalne są najlepiej przedstawione przez równanie, jeśli suma kwadratów różnic osiąga wartość minimalną.

W rozważanym przypadku otrzymujemy zależność

$$\sum_{i=1}^N \left[\ln y_i - (\ln a + b \ln t_i + c t_i) \right]^2 = \min_{a, b, c} \quad (7)$$

Warunek ten prowadzi do układu równań, z których możemy wyznaczyć stałe a, b, c . Jest to układ równań przedeterminowany. Nie może on być spełniony jednocześnie dla wszystkich równań. Przez minimalizowanie kwadratów różnic możemy wyznaczyć najlepsze wartości wektora niewiadomych. Oznaczając wartości wektora niewiadomych

$$x_1 = a, \quad x_2 = b, \quad x_3 = c$$

otrzymujemy

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Układ równań rozwiązano metodą największej wiarygodności. Największą ufnością obdarzamy ten zbiór parametrów, dla którego funkcja wiarygodności osiąga maksymalną wartość. Metodę rozwiązywania tego typu problemów szczegółowo przedstawiono w pracy [1]. Dokładność otrzymanych tutaj pomiarów nie jest jednakowa, waha się ona od 1 % do 2,5 %, lecz wstępnie dla uproszczenia metody przyjęto pomiary o równej dokładności i otrzymano układ równań

$$\begin{pmatrix} \ln x_1 \\ x_2 \\ x_3 \end{pmatrix} = - \begin{pmatrix} 1 \dots \dots \dots 1 \\ \ln a_{21} \dots \dots \ln a_{2i} \\ a_{31} \dots \dots \dots a_{3i} \end{pmatrix} \cdot \begin{pmatrix} 1 \ln a_{21} a_{31} \\ \vdots \\ \dots \dots \dots \\ \vdots \\ 1 \ln a_{2i} a_{3i} \end{pmatrix}^{-1} \cdot \begin{pmatrix} 1 \dots \dots \dots 1 \\ \ln a_{21} \dots \dots \ln a_{2i} \\ a_{31} \dots \dots \dots a_{3i} \end{pmatrix} \cdot \begin{pmatrix} \ln y_1 \\ \vdots \\ \vdots \\ \ln y_i \end{pmatrix} \quad (8)$$

Przyjęto oznaczenia analogiczne jak w pracy [1] przez wprowadzenie zmiennych

$a_i = t_i$ - czas
 $y_i = F_i$ - częstotliwość.

Układ równań został rozwiązany, po czym wprowadzono szereg niżej podanych oznaczeń dla otrzymania rozwiązania w postaci nadającej się do oprogramowania

$$z_i = \ln a_i$$

$$k_i = \ln y_i$$

$$B = \sum_{i=1}^N (z_i)^2$$

$$C = \sum_{i=1}^N (a_i)^2$$

$$D = \sum_{i=1}^N z_i$$

$$E = \sum_{i=1}^N a_i z_i$$

$$F = \sum_{i=1}^N a_i$$

$$G = \sum_{i=1}^N a_i k_i$$

$$H = \sum_{i=1}^N k_i$$

$$I = \sum_{i=1}^N z_i k_i$$

Rozwiązanie zapisano następująco +

$$M = N \cdot B \cdot C + 2 \cdot D \cdot E \cdot F - B \cdot F^2 - N \cdot E^2 - C \cdot D^2$$

$$\ln x_1 = \frac{B \cdot C - E^2 \cdot H + E \cdot F - C \cdot D \cdot I + D \cdot E - B \cdot F \cdot G}{M} \quad (9)$$

$$x_2 = \frac{F \cdot E - D \cdot C \cdot H + N \cdot C - F^2 \cdot I + D \cdot F - N \cdot E \cdot G}{M} \quad (10)$$

$$x_3 = \frac{E \cdot D - F \cdot B \cdot H + F \cdot D - N \cdot E \cdot I + N \cdot B - D^2 \cdot G}{M} \quad (11)$$

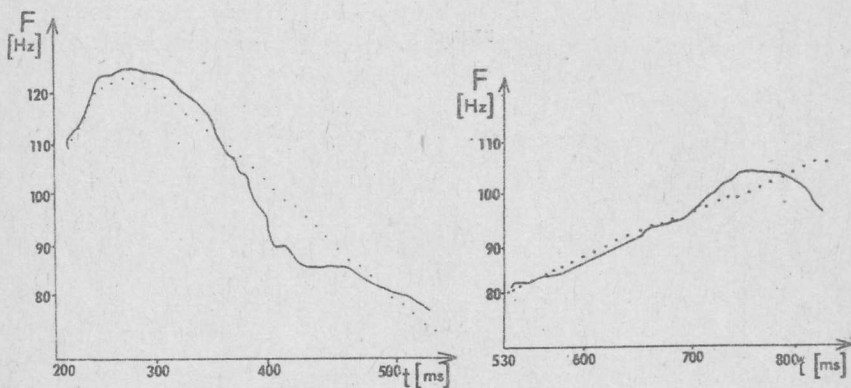
Stała N oznacza ilość aproksymowanych punktów.

Na ryc. 6 podano przykłady aproksymacji fragmentów jednej z badanych intonacji. Linia cieńszą oznaczono funkcję aproksymującą.

Na ryc. 7 przedstawiono przykład aproksymowania całego analizowanego przebiegu. Krzywa (1) aproksymuje cały przebieg,

(2) aproksymuje krótszy fragment, (3) tylko początkowy fragment przebiegu.

Z podanych przykładów wynika, iż funkcja ta bardzo dobrze aproksymuje kombinacje rosnąco-opadających równych przebiegów. Gorsze wyniki uzyskuje się aproksymując całą i to w dodatku dość zróżnicowaną intonację.



Ryc. 6. Aproksymacja fragmentów intonacji No.

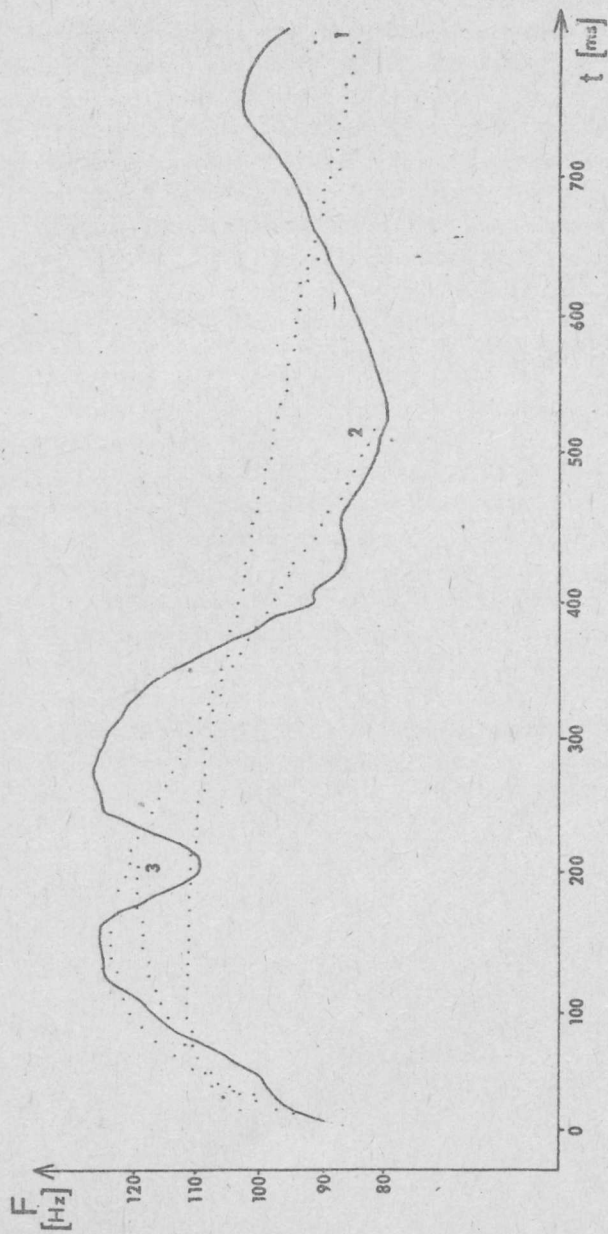


Рис. 7. Апроксимация przebiegu F_0 .

Wnioski.

Wyniki wykonanej wstępnie pracy nie mogą być jeszcze uogólnione, są jednak dość zachęcające dla dalszych badań. Problemy związane z normalizacją przebiegów F_0 nie są całkowicie rozwiązane. Również metoda aproksymacji nie jest jeszcze do końca opracowana. Brak jest analizy błędów aproksymacji oraz badania wpływu stałych a, b, c na przebieg funkcji. Wymienione zagadnienia będą przedmiotem dalszej pracy. Biorąc pod uwagę dużą użyteczność tego typu badań np. w nauce intonacji języków obcych, lub też nauce intonacji głuchych [9] czy też nawet w syntezie mowy należy zagadnienia te rozwiązać szczególnie starannie.

BIBLIOGRAFIA

- [1] BRANDT, S. : Metody statystyczne i obliczeniowe analizy danych, PWN, Warszawa, 1978.
- [2] DEMENKO, G. : Statystyczne własności rozkładów chwilowych wartości parametru F_0 w mowie ciągłej, Prace IPPT 31/80, Warszawa, 1980.
- [3] JASSEM, W., KUDELA-DOBROGOWSKA, K. : Inwarianty w przebiegach parametru F_0 , Prace IPPT 25/73, Warszawa, 1973.
- [4] KLOKER, D. : A technique for the automatic location and description of pitch contours, IEEE Conference on Acoustics, Speech and Signal Processing, 55-58, Philadelphia, 1976.
- [5] LEVITT, H., RABINER, L.R. : Analysis of Fundamental Frequency Contours in Speech, JASA, Number 2, 1971.
- [6] RALSTON, A. : Wstęp do analizy numerycznej, PWN, Warszawa, 1975.
- [7] STRZAŁKOWSKI, A., SŁIŻYŃSKI, A. : Matematyczne metody opracowywania wyników pomiarów, PWN, Warszawa, 1978.
- [8] STUART, L., MEYER : Data analysis for scientists and engineers, 1975.
- [9] WICHERN, P.U.N., BOVES, L. : Visual Feedback of F_0 curves as an aid in learning intonation contours, Proceedings 4, 1980.

