

Na prawach rękopisu

INSTYTUT TELEKOMUNIKACJI I AKUSTYKI

POLITECHNIKI WROCŁAWSKIEJ

Dział A

Akustyka

Raport nr I28/PRE-009/82

FUNKCJA PRZENOSZENIA MODULACJI (MTF)

JAKO MIARA OCENY JAKOŚCI TRANSMISJI

MOWY POLSKIEJ

Stefan Brachmański

PRACA DOKTORSKA

Promotor: doc.dr inż. Wojciech Majewski

Słowa kluczowe: jakość transmisji mowy
zrozumiałość mowy
pomiarы telefonometryczne
sygnał testowy

WROCŁAW 1982

mgr inż. Stefan Brachmański

Instytut Telekomunikacji i Akustyki
Politechniki Wrocławskiej
Wrocław, ul. Wybrzeże Wyspiańskiego 27

Promotorowi

doc. dr inż. Wojciechowi Gajownikowi

za opiekę nad pracą oraz pomoc w zakresie

w realizacji jej realizacji

składam podziękowania

Raport wpłynął do redakcji

Stefan Brachmański

SPIS TREŚCI

	str.
1. WSTĘP	6
1.1. Wprowadzenie	6
1.2. Przegląd obiektywnych metod oceny jakości transmisji sygnału mowy	7
1.3. Teza pracy	13
1.4. Układ pracy	15
2. TEORETYCZNE PODSTAWY WYKORZYSTANIA FUNKCJI PRZESZCZEPNIA MODULACJI (MTF) DO OBIEKTYWNEJ OCENY JAKOŚCI TRANSMISJI SYGNAŁU MOWY	17
2.1. Wprowadzenie	17
2.2. Teoretyczne podstawy wkaźnika transmisji sygnału mowy	18
3. USTALENIE PARAMETRÓW DO REALIZACJI SYGNAŁU TESTOWEGO	25
3.1. Wprowadzenie	25
3.2. Gęstość widmowa mowy sygnału mowy polskiej	26
3.3. Gęstość prawdopodobieństwa wartości chwilowych sygnału mowy polskiej	27
3.4. Gęstość widmowa mowy obwiedni sygnału mowy polskiej	34
4. SYSTEM POMIAROWY MTF	43
4.1. Wprowadzenie	40
4.2. Realizacja blokowa sygnału testowego	42
4.3. Realizacja blokowa sygnału testowego	45
4.4. Realizacja blokowa sygnału testowego	46

Promotorowi

doc.dr inż. Wojciechowi Majewskiemu
za opiekę nad pracą oraz cenne wskazówki
w trakcie jej realizacji niniejszym
składam podziękowanie

Stefan Brachmański

SPIS TREŚCI

	str.
1. WSTĘP	6
1.1. Wprowadzenie	6
1.2. Przegląd obiektywnych metod oceny jakości transmisji sygnału mowy	7
1.3. Teza pracy	13
1.4. Układ pracy	15
2. TEORETYCZNE PODSTAWY WYKORZYSTANIA FUNKCJI PRZENOSZENIA MODULACJI (MTF) DO OBIEKTYWNEJ OCENY JAKOŚCI TRANSMISJI SYGNAŁU MOWY	17
2.1. Wprowadzenie	17
2.2. Teoretyczne podstawy wskaźnika transmisji sygnału mowy	18
3. USTALENIE PARAMETRÓW DO REALIZACJI SYGNAŁU TESTOWEGO.	25
3.1. Wprowadzenie	25
3.2. Gęstość widmowa mocy sygnału mowy polskiej	25
3.3. Gęstość prawdopodobieństwa wartości chwilowych sygnału mowy polskiej	27
3.4. Gęstość widmowa mocy obwiedni sygnału mowy polskiej	34
4. SYSTEM POMIAROWY MTF	40
4.1. Wprowadzenie	40
4.2. Realizacja bloku generującego sygnał testowy ..	42
4.3. Realizacja odbiornika	46
4.4. Realizacja bloku sterowania	48

5. SUBIEKTYWNY POMIAR WYRAZISTOŚCI	49
5.1. Wprowadzenie	49
5.2. Wybór punktów pomiarowych	49
5.3. Wyniki subiektywnych pomiarów wyrazistości ...	52
6. POMIAR WSKAŹNIKA TRANSMISJI SYGNAŁU MOWY METODĄ MTF	64
6.1. Metoda pomiaru	64
6.2. Obliczanie wskaźnika transmisji sygnału mowy .	66
6.2.1. Wprowadzenie	66
6.2.2. Wstępne ustalenie parametrów w metodzie MTF	66
6.2.3. Metoda redukcji wymiarów w metodzie MTF	67
6.2.4. Ustalenie liczby częstotliwości modulu- jących	69
6.2.5. Ustalenie optymalnych wartości współ- czynników wagowych pasm oktaowych	73
6.2.6. Ustalenie optymalnej wartości współ- czynnika maskowania	74
6.2.7. Parametry systemu pomiarowego MTF dla sygnału mowy polskiej	75
6.3. Ocena oryginalnej i zmodyfikowanej metody MTF.	76
7. PODSUMOWANIE	87
8. DODATKI	91
9. LITERATURA	110

1. WSTĘP

1.1. Wprowadzenie

Mowa jest najbardziej rozpowszechnionym środkiem przekazywania informacji pomiędzy ludźmi. Jakość przekazywania informacji zależy od obiektywnych fizycznych parametrów urządzeń wykorzystywanych do transmisji sygnału mowy, a także od czynników subiektywnych, związanych z nadawcą i odbiorcą przesyłanych informacji. Wszystkie rodzaje pomiarów jakości transmisji sygnału mowy powinny zatem, na pewnym etapie, uwzględniać czynniki subiektywne przez stosowanie subiektywnych metod pomiarowych lub przez oszacowanie wyników pomiarów obiektywnych w sposób uwzględniający czynniki subiektywne.

Subiektywne metody oceny jakości transmisji sygnału mowy, szczegółowo omówione w literaturze [1,3,17,20,29,45,52,54,56,75,76,91,92,93,96], opierają się na podstawowych cechach mowy jakimi są zrozumiałość i głośność, natomiast z reguły pomijają wymóg naturalności przesyłanej mowy jako nieistotny dla celów telekomunikacji porozumiewawczej. Należy zaznaczyć, że istnieje wzajemne powiązanie pomiędzy zrozumiałością i głośnością przesyłanego sygnału mowy, przy czym obydwie te cechy decydują o swobodnej wymianie myśli między rozmówcami, ponieważ dobra zrozumiałość przesyłanej mowy zapewnia ściśle przekazanie treści lingwistycznych, dostateczna głośność natomiast umożliwia odbiór przekazywanej treści bez wysiłku aparatu słuchowego i napięcia nerwowego.

Wszystkie metody oceny jakości transmisji sygnału mowy oparte o pomiary subiektywne są pracochłonne i kosztowne, niezależnie od tego czy przyjmują jako podstawowe kryterium

poziom głośności odbieranych sygnałów mowy (metoda tłumienności odniesienia [96]), czy też uwzględniają łącznie wpływ wszystkich czynników na obniżenie jakości transmisji sygnału mowy (metoda tłumienności równoważnej wyrazistości [96], pomiar wyrazistości [3,17,54,92,93], metoda częstości powtórzeń [96], itp).

Ponadto do wad metod subiektywnych należy zaliczyć również słabą powtarzalność wyników uzależnioną od wydolności słuchowej ekipy słuchaczy, stopnia ich wykształcenia oraz stanu psycho-fizycznego. Z tych względów metody subiektywne nie nadają się do powszechnego i szerokiego stosowania.

Istnieje zatem pilna konieczność opracowania obiektywnych metod pomiaru jakości transmisji sygnału mowy, które eliminując ekipę słuchaczy, dałyby wyniki zgodne z subiektywną opinią użytkowników urządzeń telekomunikacyjnych. Szereg interesujących propozycji w tej dziedzinie jest dostępnych w literaturze [28, 31,33,34,36,37,41,45,58,59,60,65,66,72,73,74].

Niniejsza praca zajmuje się niektórymi aspektami przedstawionej powyżej problematyki.

1.2. Przegląd obiektywnych metod oceny jakości transmisji sygnału mowy.

Obiektywizacja metod oceny jakości transmisji sygnału mowy jest zagadnieniem złożonym. Dokonanie krótkiego przeglądu dotychczas opracowanych obiektywnych metod oceny jakości transmisji sygnału mowy może ułatwić umiejscowienie niniejszej pracy w całości tej problematyki.

Poniżej omówiono niektóre, bardziej znane obiektywne meto-

dy oceny jakości transmisji sygnału mowy.

- Metoda różnicowa - (Pattern Correspondence Index) [64,66] - polega na pomiarze bezwzględnej różnicy wartości obwiedni czasowych nieznieskształconego sygnału wejściowego oraz sygnału wyjściowego znieskształconego w badanym torze transmisyjnym. Sygnałem testowym jest segment mowy naturalnej o czasie trwania około 15 s i rozkładzie amplitudowym oraz częstotliwościowym, odpowiadającym średnim statystycznym rozkładom mowy naturalnej. Stwierdzono [66], że metoda ta daje pozytywne wyniki w przypadku oddziaływania na sygnał testowy szumów jednostajnych, impulsowych, przesłuchów zrozumiałych oraz w przypadku występowania różnych znieskształceń sygnału testowego, takich jak np. znieskształcenia amplitudowo-częstotliwościowe (tłumieniowe) itp. Poważną wadą tej metody jest konieczność zapewnienia bardzo dokładnej synchronizacji między stroną nadawczą, a odbiorczą.

- Metoda korelacyjna [64,65,87] - polega na wyznaczeniu współczynnika korelacji wzajemnej obwiedni sygnału testowego na wejściu i wyjściu badanego układu. Wykazano [64], że metoda korelacyjna oraz metoda różnicowa są równoważne. Wymagania odnośnie stosowanych sygnałów testowych oraz wady i zalety są identyczne jak w metodzie różnicowej.

- Metoda LPC (Linear Predictive Coding) [28] - polega na wykorzystaniu techniki liniowego kodowania predykcyjnego do otrzymania miary odległości pomiędzy nieznieskształconym sygnałem wejściowym, a sygnałem wyjściowym znieskształconym w badanym torze transmisyjnym. Miarę odległości uzyskuje się przez porównanie różnic występujących we współczynnikach predykcyjnych w kolejnych odpowiadających sobie okienkach czasowych

sygnałów wejściowego i wyjściowego. Jako sygnał testowy stosuje się np. czytane przez mówcę 50 wyrazowe listy zrównoważone fonetycznie [28]. Metoda LPC może być stosowana w przypadku zakłóceń i zniekształceń w dziedzinie częstotliwości, natomiast nie uwzględnia zakłóceń i zniekształceń w dziedzinie czasu. Stanowi to istotną wadę metody. Do wad można zaliczyć również konieczność posiadania wyspecjalizowanego sprzętu cyfrowego.

- Metoda głośnościowa [26, 58, 59, 60, 96] polega na wyznaczeniu wskaźnika głośnościowego. Sygnałem testowym jest szum o gęstości widmowej mocy odpowiadającej średniej gęstości widmowej mocy mowy naturalnej [26] lub sygnał sinusoidalny o częstotliwości zmieniającej się od 200 Hz do 4 kHz i z powrotem do 200 Hz z prędkością chwilową (df/dt) będącą odpowiednio dobraną funkcją częstotliwości [96]. Szacowanie głośności nie jest odpowiednią miarą wierności przesyłania parametrów dystyngtywnych mowy. Za pomocą tej miary można wprawdzie oszacować pewne ograniczone parametry jakościowe kanału, należy jednak pamiętać, iż ta metoda przy szacowaniu głośności sygnału wyjściowego nie rozróżnia poziomu sygnału mowy od poziomu sygnału zakłócającego oraz nie wskazuje w jakim stopniu nastąpiło zniekształcenie sygnału testowego.

- Metoda wskaźnika wyrazistości [8, 23, 41, 44, 45] polega na określeniu wyrazistości mowy przesyłanej układem transmisyjnym na podstawie jego fizycznych parametrów. Procedurę obliczania wskaźnika wyrazistości można przedstawić w trzech krokach:

1. Obliczenie stosunku sygnał/szum (S/N) w pasmach częstotliwości (w wersji pierwotnej, French i Steinberg stosowali 20 pasm o jednakowej wyrazistości [23], natomiast w wersji uproszczonej - opracowanej przez Krytera - wykorzystano do

analizy pasma tercjowe lub oktawowe [41]). Dla każdego pasma określa się poziom sygnału mowy i poziom szumu uwzględniając szum maskujący docierający do uszu słuchacza.

2. Przeprowadzenie liniowej transformacji stosunku sygnał/szum uwzględniającej udział danego pasma częstotliwości we wskaźniku wyrazistości. Maksymalny udział danego pasma wynoszący 1 odpowiada wartości $S/N = +18$ dB, natomiast minimalny równy 0 odpowiada wartości $S/N = -12$ dB [8, 23].

3. Obliczenie średniej ważonej udziałów wszystkich pasm częstotliwości w celu określenia końcowej wartości wskaźnika wyrazistości i odpowiadającej jej wyrazistości mowy.

Wskaźnik wyrazistości może być stosowany w przypadkach zakłóceń i zniekształceń w dziedzinie częstotliwości, natomiast nie uwzględnia zakłóceń i zniekształceń w dziedzinie czasu, jak również zniekształceń nieliniarnych. Stanowi to istotną wadę metody.

- Metoda wskaźnika transmisji sygnału mowy (STI) [31,72,73]

opiera się o pomiar stopnia zachowania różnic poziomów sygnału mowy podczas transmisji. Sygnałem testowym jest szum o widmie gęstości mocy ukształtowanym zgodnie z obwiednią średniego widma mocy sygnału mowy naturalnej. Na wejście układu transmisyjnego podawane są dwa sygnały testowe o stałej, znanej różnicy poziomów. Po przejściu sygnałów przez badany układ transmisyjny mierzona jest różnica poziomów obu sygnałów w poszczególnych pasmach częstotliwości. Miarą oceny jakości transmisji sygnału mowy jest stopień zachowania tych różnic. Użyty w tej metodzie sygnał testowy może być otrzymany przy użyciu powszechnych i łatwo dostępnych urządzeń do obróbki sygnału. Nie jest również wymagana synchronizacja czasowa między wejściem, a wyjściem układu transmisyj-

nego. Omawiana metoda nie uwzględnia jednak efektu samomaskowania (dla bardzo silnych sygnałów mowy) oraz wpływu progu słyszalności (dla bardzo słabych sygnałów).

- Metoda określania jakości transmisji sygnału mowy na podstawie funkcji przenoszenia modulacji (Modulation Transfer Function - MTF) [33,34,53,74].

Metoda MTF polega na porównaniu głębokości modulacji sygnału wejściowego i wyjściowego. Sygnałem testowym jest szum o gęstości widmowej mocy ukształtowanej zgodnie z obwiednią średniej gęstości widmowej mocy mowy naturalnej, zmodulowany amplitudowo sygnałem harmonicznym w poszczególnych pasmach oktaowych. Dobór częstotliwości modulujących wynika z analizy widma obwiedni sygnału mowy naturalnej; natomiast dobór pasm oktaowych - z analizy uśrednionego widma mocy sygnału mowy naturalnej oraz zakresu pasma częstotliwości przenieszonego przez badany układ transmisyjny. Zmiany głębokości modulacji w poszczególnych pasmach, dla zbioru wartości częstotliwości modulujących, są podstawą do wyliczenia wskaźnika transmisji sygnału mowy W_{STI} stanowiącego obiektywny estymator jakości transmisji sygnału mowy.

Wyniki badań [33,34,74] wskazują, że metoda ta daje dobre rezultaty w przypadku zakłóceń szumowych, pogłosowych oraz w przypadku ograniczenia pasma i amplitudy. Użyty w tej metodzie sygnał testowy może być otrzymany przy użyciu powszechnie stosowanych urządzeń do obróbki sygnału. Nie jest również wymagana synchronizacja czasowa między wejściem a wyjściem układu transmisyjnego.

Podsumowując powyższy przegląd można stwierdzić, że dotychczas nie udało się opracować uniwersalnej, obiektywnej metody pozwalającej na otrzymanie, dla pełnego zbioru zakłóceń i zniekształceń sygnału mowy, oceny jakości transmisji sygnału mowy

zgodnej z wynikami subiektywnych pomiarów jakości transmisji. Uwzględniając zale^tży i wady omówionych metod obiektywnego pomiaru jakości transmisji sygnału mowy zdecydowano przyjąć jako przedmiot badań metodą MTF, gdyż w tej metodzie stosowany sygnał testowy posiada strukturę widmową i czasową zbliżoną do uśrednionej struktury sygnału mowy odzwierciedlając m.in. średnią gęstość widmową mocy sygnału mowy oraz średnią gęstość widmową obwiedni sygnału mowy [33,34,72,73,74]. Zapewnia to w trakcie badań warunki pracy kanału telefonicznego zbliżone do warunków rzeczywistych oraz co jest najistotniejsze, pozwala uwzględnić prawie wszystkie zniekształcenia i zakłócenia sygnału mowy wprowadzane przez kanał (w tym również zniekształcenia i zakłócenia w dziedzinie czasu).

Dalsze udoskonalenie metody winno iść zatem w kierunku jeszcze większego zbliżenia struktury sygnału testowego do struktury sygnału mowy naturalnej.

W wersji metody MTF podanej przez Houtgasta i Steenekena [34,74] sygnał testowy nie zapewnia zbieżności rozkładu gęstości prawdopodobieństwa wartości chwilowych sygnału testowego z rozkładem dla sygnału mowy naturalnej. Ten parametr może mieć istotne znaczenie, gdyż z badań przeprowadzonych przez Hlibowickiego [30] jednoznacznie wynika, że zniekształcenia gęstości prawdopodobieństwa wartości chwilowych sygnału akustycznego są łatwo postrzegane przez organ słuchu człowieka. Celowe jest zatem dokonanie modyfikacji metody MTF poprzez uzupełnienie generatora sygnału testowego układem kształtowania rozkładu gęstości prawdopodobieństwa wartości chwilowych sygnału.

1.3. Teza pracy

W oparciu o rozważania przedstawione w rozdziale 1.1 i 1.2 sformułowano następującą tezę pracy:

Istotnym czynnikiem w obiektywnych metodach oceny jakości transmisji sygnału mowy kanałem telefonicznym^{1/} jest zbieżność struktury sygnału testowego z fizyczną strukturą sygnału mowy. Wprowadzając zatem do metody MTF kształtowanie rozkładu gęstości prawdopodobieństwa wartości chwilowych sygnału testowego zgodnie z rozkładem dla sygnału mowy naturalnej dokonuje się zbliżenia struktury tego sygnału do struktury sygnału mowy, co poprawia dokładność metody MTF.

W niniejszej pracy przyjęto, że udowodnienie przedstawionej tezy zostanie przeprowadzone dla sygnału mowy polskiej, co powoduje, że prócz zadań badawczych wynikających z tezy, należy uprzednio przeprowadzić badania pozwalające na ustalenie parametrów sygnału testowego w którym powinny być uwzględnione wybrane parametry sygnału mowy polskiej.

Uwzględniając łącznie wymogi wynikające z tezy oraz powyższego założenia postawiono sobie następujące zadania badawcze:

1. Ustalenie - w oparciu o wyniki badań zawartych w literaturze oraz przeprowadzonych w ramach niniejszej pracy - parametrów opisujących fizyczną strukturę sygnału mowy polskiej (średnia gęstość widmowa mocy, średnia gęstość widmowa obwiedni amplitudowej, rozkład gęstości prawdopodobieństwa wartości chwilowych).

^{1/} W dalszej części pracy pod pojęciem "kanał telefoniczny" należy rozumieć drogę od wyjścia aparatu telefonicznego nadawczego do wejścia aparatu telefonicznego odbiorczego.

2. Opracowanie generatora sygnału testowego w metodzie MTF w wersji oryginalnej^{2/} oraz zmodyfikowanej^{3/}.
3. Przeprowadzenie cyklu badań subiektywnych dla różnych warunków transmisji sygnału mowy kanałem telefonicznym celem uzyskania miary odniesienia dla obiektywnych badań transmisji sygnału mowy metodą MTF w wersji oryginalnej i zmodyfikowanej.
4. Ustalenie parametrów metody MTF^{4/} (w wersji oryginalnej i zmodyfikowanej) dla transmisji sygnału mowy polskiej oraz ustalenie analitycznych związków pomiędzy obiektywnym wskaźnikiem uzyskanym w metodzie MTF a subiektywną miarą oceny jakości transmisji.
5. Dokonanie analizy porównawczej dokładności wyników uzyskiwanych metodą MTF oryginalną i zmodyfikowaną.

2/ W dalszej części pracy pod pojęciem oryginalnej wersji metody MTF przyjęto wersję podaną przez Houtgasta i Steenekena [33,74] z tym, że gęstości widmowe mocy sygnału testowego i obwiedni tego sygnału ukształtowano zgodnie z wynikami badań dla mowy polskiej.

3/ Pod pojęciem zmodyfikowanej wersji metody MTF przyjęto wersję oryginalną z modyfikacją polegającą na kształtowaniu rozkładu gęstości prawdopodobieństwa wartości chwilowych sygnału testowego zgodnie ze średnim rozkładem dla sygnału mowy polskiej.

4/ Pod pojęciem ustalenie parametrów metody MTF należy rozumieć ustalenie minimalnej liczby pasm częstotliwości oraz częstotliwości modulujących niezbędnych do uzyskania dokładnego wyniku oraz ustalenie wartości współczynników wagowych występujących w zależności wyrażającej syntetyczny wskaźnik jakości transmisji sygnału uzyskany w metodzie MTF.

1.4. Układ pracy

Praca składa się z siedmiu głównych rozdziałów, pięciu dodatków oraz wykazu literatury.

W rozdziale 1 omówiono stan badań nad obiektywizacją metod pomiaru jakości transmisji sygnału mowy. W rozdziale tym podano tezę pracy oraz podstawowe zadania badawcze.

W rozdziale 2 przedstawiono teoretyczne podstawy obiektywnej oceny jakości transmisji sygnału mowy w oparciu o funkcję przeniesienia modulacji (MTF).

W rozdziale 3 podano parametry sygnału mowy polskiej potrzebne do generacji sygnału testowego w metodzie MTF ustalone na podstawie wyników badań własnych oraz zamieszczonych w dostępnej literaturze przedmiotu.

W rozdziale 4 zaprezentowano stosowany w pracy model systemu pomiarowego MTF.

W rozdziale 5 omówiono warunki techniczne pomiarów subiektywnych oceny jakości transmisji sygnału mowy oraz podano ich wyniki.

W rozdziale 6 podano wyniki pomiarów obiektywnych oraz ustalone wartości parametrów systemu pomiarowego MTF dla sygnału mowy polskiej. Na podstawie analizy wyników pomiarów subiektywnych oraz obiektywnych, wyznaczono zależność umożliwiającą bezpośrednio obliczenie estymatora subiektywnej oceny jakości transmisji na podstawie wskaźnika transmisji sygnału mowy. W rozdziale tym dokonano również porównawczej analizy wyników uzyskanych dla oryginalnej i zmodyfikowanej wersji metody MTF i na tej podstawie wykazano, że metoda zmodyfikowana zwiększa dokładność pomiaru estymatora subiektywnej oceny jakości transmisji sygnału mowy.

Rozdział 7 stanowi podsumowanie całej pracy. Przedstawiono w nim najistotniejsze wnioski wynikające z pracy.

W celu uzyskania przejrzystości pracy, niektóre obliczenia oraz szczegółowe rozważania i ppisy zamieszczono w dodatkach A, B, C, D i E.

2.1. Opis metody

Zadaniem systemu telekomunikacyjnego jest przekazanie informacji na odległość w sposób niezawodny, w tym w szczególności zapewnienie bezpieczeństwa informacji. W tym celu należy wykonać szereg czynności, które gwarantują integralność i poufność danych. W niniejszym rozdziale opisano metodę kryptograficzną, która umożliwia bezpieczny przesył danych w kanałach publicznych. Metoda ta opiera się na wykorzystaniu funkcji matematycznych, które są trudne do odwrócenia bez posiadania odpowiedniej kluczy. W celu zapewnienia bezpieczeństwa danych, należy wykonać szereg czynności, które gwarantują integralność i poufność danych. W niniejszym rozdziale opisano metodę kryptograficzną, która umożliwia bezpieczny przesył danych w kanałach publicznych. Metoda ta opiera się na wykorzystaniu funkcji matematycznych, które są trudne do odwrócenia bez posiadania odpowiedniej kluczy.

Metoda ta opiera się na wykorzystaniu funkcji matematycznych, które są trudne do odwrócenia bez posiadania odpowiedniej kluczy. W celu zapewnienia bezpieczeństwa danych, należy wykonać szereg czynności, które gwarantują integralność i poufność danych. W niniejszym rozdziale opisano metodę kryptograficzną, która umożliwia bezpieczny przesył danych w kanałach publicznych. Metoda ta opiera się na wykorzystaniu funkcji matematycznych, które są trudne do odwrócenia bez posiadania odpowiedniej kluczy.

Metoda ta opiera się na wykorzystaniu funkcji matematycznych, które są trudne do odwrócenia bez posiadania odpowiedniej kluczy.

1. W celu zapewnienia bezpieczeństwa danych, należy wykonać szereg czynności, które gwarantują integralność i poufność danych. W niniejszym rozdziale opisano metodę kryptograficzną, która umożliwia bezpieczny przesył danych w kanałach publicznych. Metoda ta opiera się na wykorzystaniu funkcji matematycznych, które są trudne do odwrócenia bez posiadania odpowiedniej kluczy.

2. W celu zapewnienia bezpieczeństwa danych, należy wykonać szereg czynności, które gwarantują integralność i poufność danych. W niniejszym rozdziale opisano metodę kryptograficzną, która umożliwia bezpieczny przesył danych w kanałach publicznych. Metoda ta opiera się na wykorzystaniu funkcji matematycznych, które są trudne do odwrócenia bez posiadania odpowiedniej kluczy.

3. W celu zapewnienia bezpieczeństwa danych, należy wykonać szereg czynności, które gwarantują integralność i poufność danych. W niniejszym rozdziale opisano metodę kryptograficzną, która umożliwia bezpieczny przesył danych w kanałach publicznych. Metoda ta opiera się na wykorzystaniu funkcji matematycznych, które są trudne do odwrócenia bez posiadania odpowiedniej kluczy.

2.2. Opis metody

W niniejszym rozdziale opisano metodę kryptograficzną, która umożliwia bezpieczny przesył danych w kanałach publicznych. Metoda ta opiera się na wykorzystaniu funkcji matematycznych, które są trudne do odwrócenia bez posiadania odpowiedniej kluczy. W celu zapewnienia bezpieczeństwa danych, należy wykonać szereg czynności, które gwarantują integralność i poufność danych. W niniejszym rozdziale opisano metodę kryptograficzną, która umożliwia bezpieczny przesył danych w kanałach publicznych. Metoda ta opiera się na wykorzystaniu funkcji matematycznych, które są trudne do odwrócenia bez posiadania odpowiedniej kluczy.

2. TEORETYCZNE PODSTAWY WYKORZYSTANIA FUNKCJI PRZENOSZENIA MODULACJI (MTF) DO OBIEKTYWNEJ OCENY JAKOŚCI TRANSMISJI SYGNAŁU MOWY

2.1. Wprowadzenie

Zadaniem systemów telekomunikacyjnych jest możliwie wierne przesyłanie na odległość sygnału, w tym w szczególności sygnału mowy. Umożliwia to kanał transmisyjny łączący źródło sygnału z odbiornikiem. Wierna transmisja oznacza możliwie dokładne zachowanie czasowych zmian natężenia lub obwiedni sygnału mowy dla wszystkich częstotliwości słyszalnych [34,48].

Biorąc to pod uwagę Houtgast i Steeneken opracowali metodę oceny jakości transmisji sygnału mowy z wykorzystaniem funkcji przenoszenia modulacji.

Metodę tę można scharakteryzować w trzech punktach:

1. W czasie pomiaru kanał telefoniczny jest traktowany jak "czarna skrzynka", w której stosunek sygnał-szum (S/N) w pasmach oktawowych jest określony przez uśrednienie pomiarów wykonanych z wykorzystaniem specjalnego sygnału testowego.
2. Sygnał testowy odtwarza podstawowe statystyki widmowe i czasowe sygnału mowy naturalnej.
3. Metoda zawiera miarę pozwalającą na oszacowanie zniekształceń i zakłóceń w dziedzinie częstotliwości oraz czasu.

2.2. Teoretyczne podstawy wskaźnika transmisji sygnału mowy.

W metodzie MTF sygnałem testowym jest szum ukształtowany w dziedzinie częstotliwości zgodnie ze średnią gęstością widmową mocy mowy naturalnej (wersja oryginalna) oraz o rozkładzie gęstości

ci prawdopodobieństwa wartości chwilowych sygnału testowego zgodnym z rozkładem dla mowy naturalnej (wersja zmodyfikowana). Tak ukształtowany szum zostaje poddany, w pasmach oktaowych, modulacji amplitudowej sygnałem cosinusoidalnym o wartościach częstotliwości wynikających z analizy widma obwiedni sygnału mowy naturalnej. Dobór pasm oktaowych wynika z analizy częstotliwościowej charakterystyki średniej gęstości widmowej mocy sygnału mowy naturalnej oraz zaleceń CCITT co do szerokości pasma częstotliwości w telefonii naturalnej /300 - 3600 Hz/ [95].

Chwilową amplitudę testowego sygnału szumowego zmodulowanego amplitudowo, z głębokością modulacji $m_{n,1}=1$, w tym 1-tym pasmie oktaowym i dla n -tej częstotliwości modulującej można opisać zależnością:

$$x_{n,1}(t) = x_1 (1 + \cos 2\pi f_n t) \quad , \quad (2.1)$$

gdzie:

x_1 - amplituda sygnału modulowanego (szumu) w 1-tym pasmie oktaowym,

f_n - n -ta częstotliwość modulująca ,

$n=1, \dots, N$ N - liczba częstotliwości modulujących,

$l=1, \dots, L$, L - liczba pasm oktaowych.

Po przejściu sygnału testowego przez kanał telefoniczny, w którym na sygnał testowy oddziałują różnego rodzaju zakłócenia i zniekształcenia, chwilowa amplituda tego sygnału ma postać:

$$Y_{n,1}(t) = Y_1 [1 + m_{n,1} \cos 2\pi f_n (t + \xi)] \quad , \quad (2.2)$$

gdzie:

Y_1 - amplituda sygnału testowego w 1-tym pasmie oktawowym na wyjściu kanału telefonicznego,

$m_{n,1}$ - głębokość modulacji w 1-tym pasmie oktawowym i dla n -tej częstotliwości modulującej (można ją oznaczyć również jako $m_1(f_n)$, co podkreśla zależność głębokości modulacji na wyjściu kanału telefonicznego od częstotliwości modulującej),

ξ - przesunięcie fazowe wnoszone przez kanał telefoniczny.

Określenia głębokości modulacji, w przypadku nieznanego przesunięcia fazowego wnoszonego przez kanał telefoniczny, można dokonać obliczając korelację wzajemną między amplitudą $Y_{n,1}(t)$, a sygnałem sinusoidalnym oraz cosinusoidalnym. Podnosząc do kwadratu obliczone funkcje korelacji wzajemnej i dodając je stronami otrzymuje się wyrażenie na głębokość modulacji (szczegółowe obliczenia zamieszczono w dodatku A).

Jeżeli na sygnał testowy oddziałują tylko zakłócenia addytywne, to głębokość modulacji $m_{n,1}$ na wyjściu kanału telefonicznego nie zależy od częstotliwości modulującej i jest powiązana ze stosunkiem sygnał-szum (S/N).

W przypadku występowania również innych typów zakłóceń i zniekształceń, głębokość modulacji zależy od częstotliwości modulującej i jest powiązana nie tylko ze stosunku sygnał-szum ale i z odpowiedzią impulsową $h(t)$ kanału telefonicznego.

W dodatku B. przedstawiono za Houtgastem i Steenekenem [34] związek między głębokością modulacji $m_{n,1}$, a odpowiedzią impulsową $h(t)$ kanału telefonicznego oraz stosunkiem sygnał-szum $(S/N)_1$.

Zgodnie z istotą metody MTF, podstawową wielkością w tej metodzie jest pozorny stosunek sygnał-szum w pasmach oktawowych określony zależnością [33,34]:

$$(S/N)_{n,1}^{\text{poz}} = 20 \log \frac{m_{n,1}}{1 - m_{n,1}} \quad (2.3)$$

gdzie:

$(S/N)_{n,1}^{\text{poz}}$ - pozorny stosunek sygnał-szum w 1-tym pasmie oktawowym i dla n -tej częstotliwości modulującej.

Obliczenie pozornego stosunku sygnał-szum $(S/N)_{n,1}^{\text{poz}}$ wg zależności (2.3) obarczone jest pewnym błędem wynikającym z nieuwzględnienia występującego w organie słuchu maskującego oddziaływania niższych pasm częstotliwości na wyższe. To maskujące oddziaływanie można potraktować jako pozorne zwiększenia szumu w pasmie oktawowym 1-tym, a pochodzące od amplitudy dźwięku sąsiedniego 1-1 pasma oktawowego.

Amplitudę sygnału maskującego w 1-tym pasmie oktawowym można wyznaczyć na podstawie zależności:

$$Y_1 = Y_{1-1} \gamma \quad (2.4)$$

gdzie:

Y_1' - pozorna amplituda sygnału maskującego w 1-tym pasmie oktawowym,

Y_{1-1} - amplituda sygnału w pasmie oktawowym 1-1,

γ - współczynnik maskowania.

Zmniejszenie głębokości modulacji na skutek maskowania może być uwzględnione poprzez współczynnik korekcyjny β_1 :

$$\beta_1 = \frac{Y_1}{Y_1 + Y_1'} \quad (2.5)$$

Skorygowana głębokość modulacji opisana jest wówczas zależnością:

$$m'_{n,1} = \beta_1 \cdot m_{n,1} \quad (2.6)$$

Po uwzględnieniu skorygowanej głębokości modulacji zależności (2.3) przyjmuje postać :

$$(S/N)_{n,1}^{poz} = 20 \log \frac{m'_{n,1}}{1 - m'_{n,1}} \quad (2.7)$$

przy czym : $m'_{n,1} \in (0,1)$.

Z eksperymentów przeprowadzonych przez Houtgasta i Steenekena [74] wynika, że wartość pozornego stosunku sygnał/szum $((S/N)_{n,1}^{poz}$ (zal.2.3)) stosowana w obliczaniu syntetycznego wskaźnika W_{STI} można ograniczyć do przedziału $\langle -15 \text{ dB} + 15 \text{ dB} \rangle$,

przy czym

jeżeli $(S/N)_{n,1}^{poz} > 15 \text{ dB}$ to przyjmuje się $(S/N)_{n,1}^{poz} = 15 \text{ dB}$

natomiast

jeżeli $(S/N)_{n,1}^{poz} < -15 \text{ dB}$ to przyjmuje się $(S/N)_{n,1}^{poz} = -15 \text{ dB}$

Dla każdego pasma oktawowego obliczany jest wskaźnik funkcji przenoszenia modulacji, opisany zależnością :

$$W_{\text{mtf}}^l = \frac{1}{N} \sum_{n=1}^N \frac{(S/N)_{n,l}^{\text{poz}} + 15}{30} \quad (2.8)$$

gdzie:

$W_{\text{mtf}}^l \in \langle 0,1 \rangle$ - wskaźnik funkcji przenoszenia modulacji dla l-tego pasma oktawowego,

N - liczba częstotliwości modulujących.

Ostatecznie, z ważonej sumy wskaźników W_{mtf}^l obliczonych dla L pasm oktawowych, otrzymuje się wskaźnik transmisji sygnału mowy W_{STI} , opisany zależnością :

$$W_{\text{STI}} = \frac{1}{\mathcal{L}} \sum_{l=1}^L \mathcal{L}_l W_{\text{mtf}}^l = \quad (2.9)$$

$$= \frac{1}{\mathcal{L} N} \sum_{l=1}^L \sum_{n=1}^N \mathcal{L}_l \frac{(S/N)_{n,l}^{\text{poz}} + 15}{30}$$

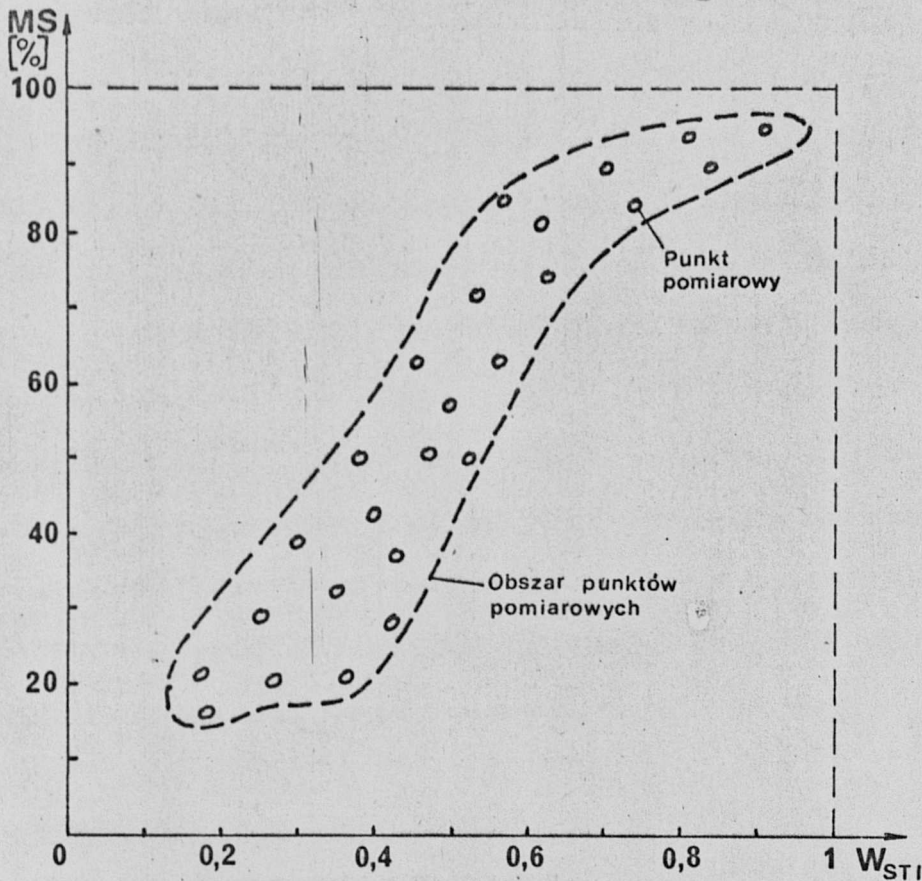
przy czym: $W_{\text{STI}} \in \langle 0,1 \rangle$ oraz $\mathcal{L}_l \in \langle 0,1 \rangle$,

$$\mathcal{L} = \sum_{l=1}^L \mathcal{L}_l$$

$$\mathcal{L} \leq 1$$

α_i - współczynnik wagowy 1-tego pasma oktawowego, określający udział danego pasma oktawowego w globalnym wskaźniku W_{STI} .

Wskaźnik transmisji sygnału mowy W_{STI} , zdefiniowany zgodnie z zależnością (2.9), jest pewną miarą jakości transmisji sygnału mowy kanałem telefonicznym.



Rys.2.1. Związek między miarą subiektywną - MS (np.wyrazistością), a miarą obiektywną - W_{STI} oceny jakości transmisji sygnału mowy.

W celu powiązania wskaźnika W_{STI} z subiektywną miarą MS oceny jakości transmisji sygnału mowy należy określić istniejący między nimi związek (rys.2.1), tzn. wyznaczyć zależność:

$$MS = f(W_{STI}), \quad (2.10)$$

Można tego dokonać np. aproksymując wielomianem n -tego rzędu punkty pomiarowe leżące na płaszczyźnie: miara subiektywna (np. wyrazistość) MS - wskaźnik transmisji sygnału mowy W_{STI} . Uzyskuje się w ten sposób zależność umożliwiającą obliczenie estymowanej miary subiektywnej oceny jakości transmisji sygnału mowy, na podstawie wyników pomiarów obiektywnych.

3. USTALENIE PARAMETRÓW DO REALIZACJI SYGNAŁU TESTOWEGO

3.1. Wprowadzenie

Sygnal testowy w metodzie MTF z uwzględnieniem przyjętej w pracy modyfikacji wymaga ustalenia następujących parametrów:

1. średniej gęstości widmowej mocy sygnału mowy,
2. rozkładu gęstości prawdopodobieństwa wartości chwilowych sygnału mowy,
3. średniej gęstości widmowej mocy obwiedni sygnału mowy.

W niniejszej pracy gęstość widmową mocy sygnału mowy polskiej przyjęto korzystając z dostępnych danych [25,43,89], pozostałe parametry natomiast ustalono w oparciu o badania własne [13,14]. Metody ustalania tych parametrów podano w następujących podrozdziałach.

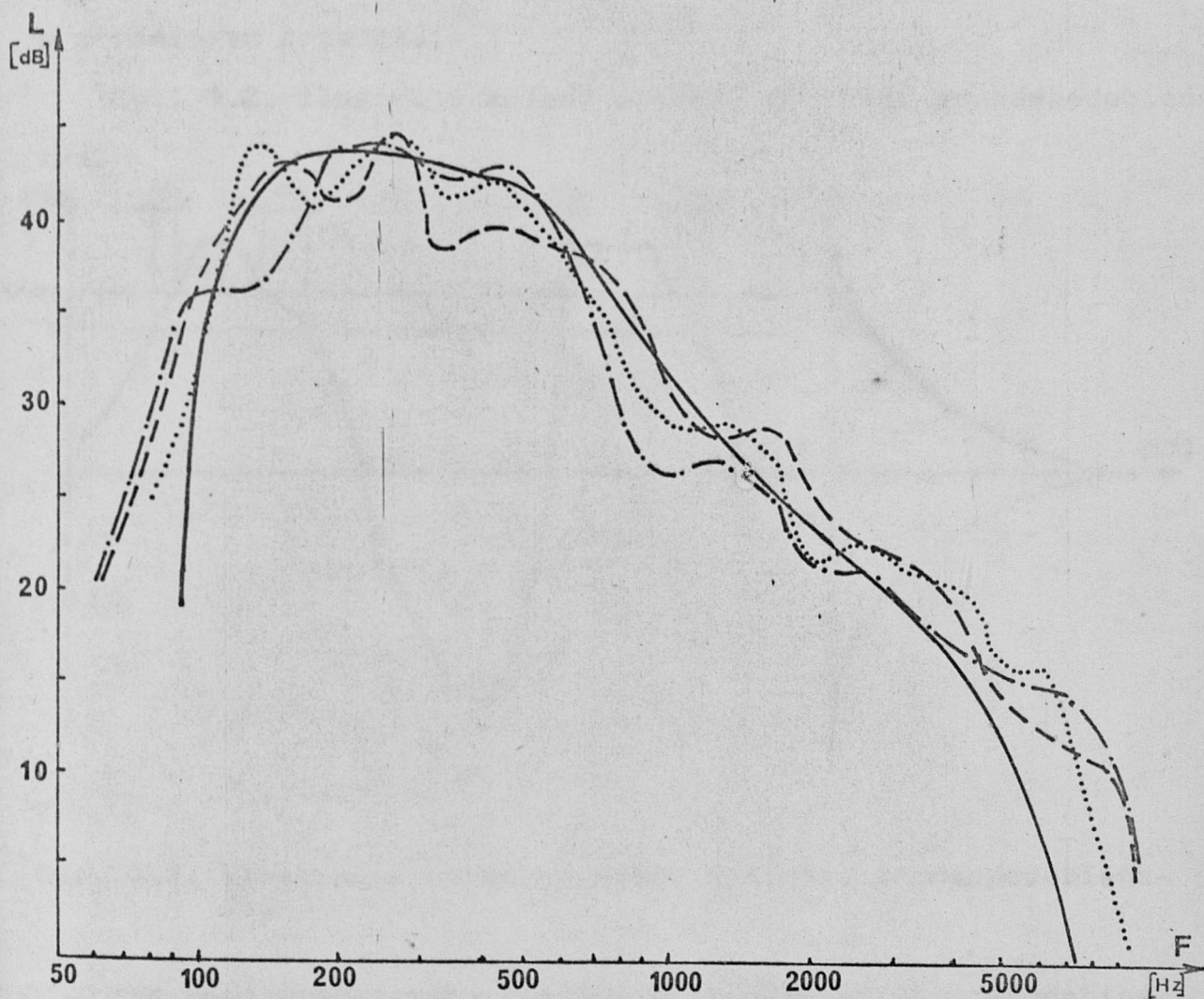
Przeprowadzając pomiar wybranych parametrów sygnału mowy przyjęto, że sygnał mowy ze względu na wyróżnione parametry statystyczne - można traktować jako realizację stacjonarnego, ergodycznego procesu losowego jeśli tylko rozpatruje się dostatecznie długi odcinek czasowy tego sygnału [5,10,14,85].

3.2. Gęstość widmowa mocy sygnału mowy polskiej

Szczegółowy opis metod otrzymywania gęstości widmowej mocy podano w pracach [6,50,51].

Wyniki prac eksperymentalnych nad pomiarem średniej gęstości widmowej mocy sygnału mowy podali m.in. Dunn i White [20], Guyonnaud [26], Harris i Waite [27], Jassem [35], Kosiel [40], Majewski, Rothman i Hollien [43], Majewski i Zalewski [89], Tarnoczy [77,78,79,80].

Widma mocy sygnału mowy polskiej podane w pracach [35,43, 89] przedstawiono na rys. 3.1. Stanowiły one podstawę do zaprojektowania korektora widma, w którym następowało kształtowanie widma mocy szumu białego zgodnie ze średnim widmem mocy mowy polskiej

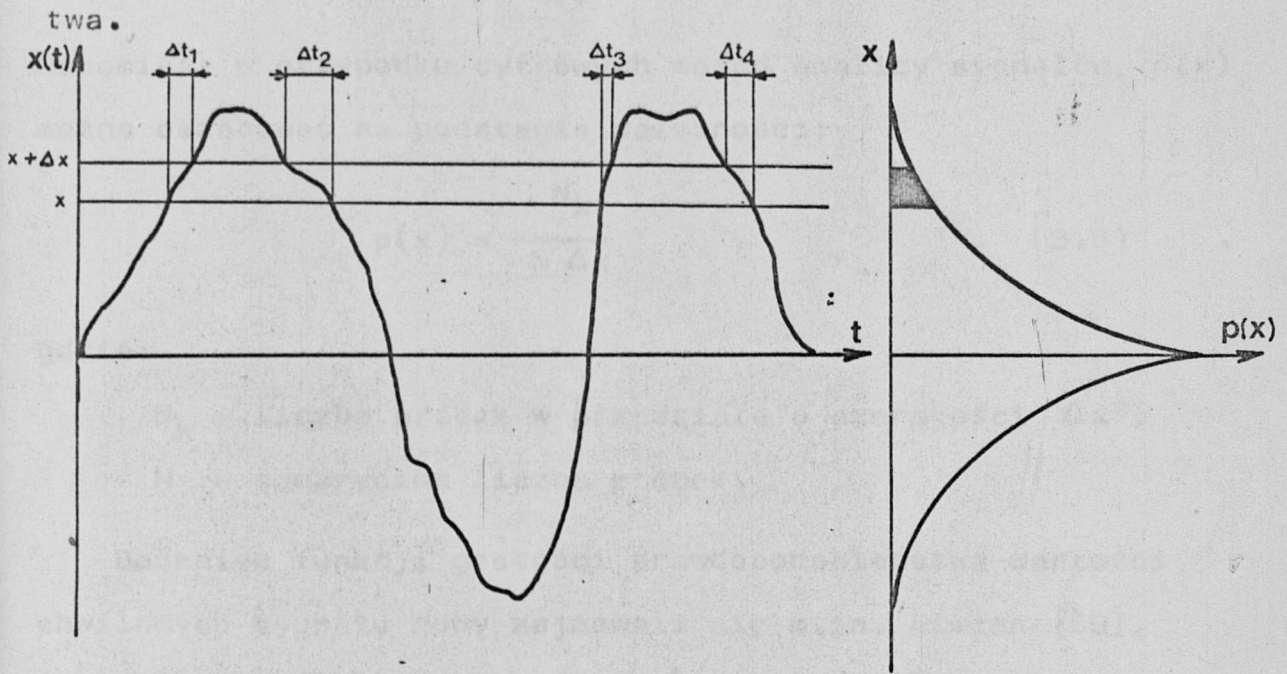


Rys.3.1. Gęstości widmowe mocy sygnału mowy polskiej
(- - - wg Jassema [35], -·-·- wg Zalewskiego i Majewskiego [89],
..... wg Majewskiego i in. [43]) oraz sygnału testowego (—).

3.3. Gęstość prawdopodobieństwa wartości chwilowych sygnału mowy polskiej

Gęstość prawdopodobieństwa $p(x)$ wartości chwilowych sygnału losowego $x(t)$ określa prawdopodobieństwo zdarzenia polegającego na tym, że wartości sygnału w dowolnej chwili są zawarte w określonym przedziale.

Rys. 3.2. ilustruje zasadę pomiaru gęstości prawdopodobieństwa.



Rys. 3.2. Ilustracja zasady pomiaru gęstości prawdopodobieństwa.

Dla zmiennej losowej ciągłej $x(t)$ gęstość prawdopodobieństwa $p(x)$ wartości chwilowych jest pochodną dystrybuanty zmiennej losowej $x(t)$ [7], co można zapisać w postaci:

$$p(x) = \frac{dF(x)}{dx} = \frac{P(x \leq x(t) \leq x+dx)}{dx} \quad (3.4)$$

W praktyce, w przypadku analogowych metod analizy sygnału, korzysta się z zależności :

$$p(x) = \frac{T_x}{T \Delta x}, \quad (3.5)$$

gdzie:

Δx - skończona szerokość przedziału zawierającego x ,

T_x - suma przedziałów czasu, w których sygnał $x(t)$ znajduje się wewnątrz przedziału wartości o szerokości Δx

T - czas trwania sygnału.

Natomiast w przypadku cyfrowych metod analizy sygnałów, $p(x)$ można oszacować na podstawie zależności:

$$p(x) = \frac{N_x}{N \Delta x}, \quad (3.6)$$

gdzie:

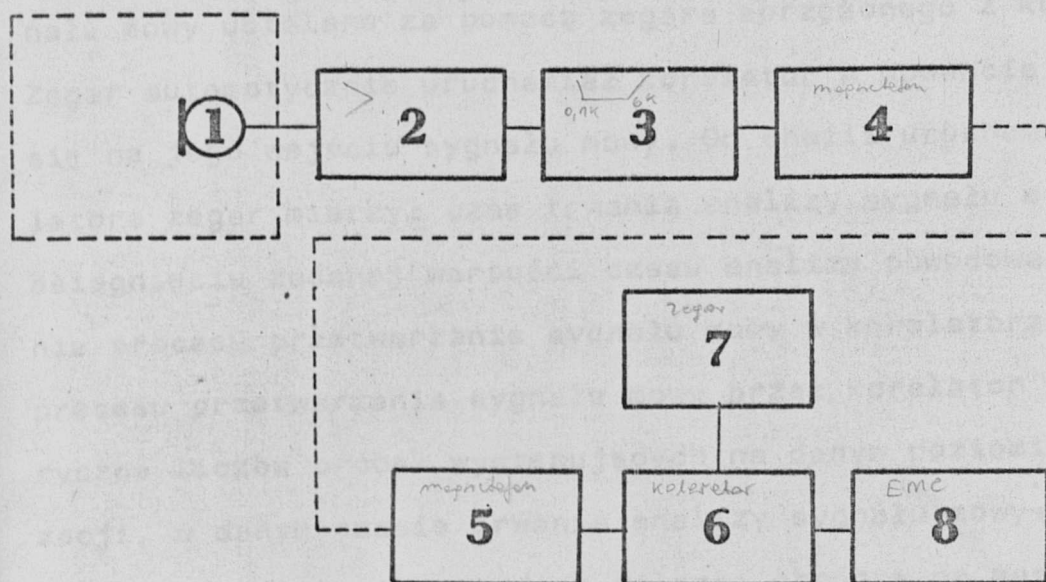
N_x - liczba próbek w przedziale o szerokości Δx ,

N - sumaryczna liczba próbek.

Badaniem funkcji gęstości prawdopodobieństwa wartości chwilowych sygnału mowy zajmowali się m.in. Sivian [69], Dunn i White [20], Davenport [18], Richards [57,68], Fersman [21], Rimski-Korsakow [61], Wieliczkin [83], Weiß [82], Wolf i Brehm [86], Wolf [84,85], Noll [49], Szitow i Bielkin (wyniki zamieszczone w pracy [39]), a dla mowy polskiej - Dobrogowski [19] oraz autor niniejszej rozprawy [14].

Poniżej zostaną omówione w skrócie badania przeprowadzone w celu otrzymania funkcji gęstości prawdopodobieństwa wartości chwilowych sygnału mowy polskiej oraz podane ich wyniki. Szerwsze omówienie wykonanych badań jest zawarte w pracy [14].

Schemat blokowy systemu do pomiaru gęstości prawdopodobieństwa wartości chwilowych sygnału mowy pokazano na rys. 3.3.



Rys. 3.3. Schemat blokowy systemu realizującego pomiar gęstości prawdopodobieństwa wartości chwilowych sygnału mowy
1 - mikrofon, 2 - wzmacniacz, 3 - filtr pasmowy (100-6000 Hz, 24 dB/okt), 4 - magnetofon, 5 - magnetofon, 6 - korelator, 7 - zegar, 8 - maszyna cyfrowa.

Pomiary oparto na materiale fonetycznym uzyskanym w wyniku odczytywania przez 10 lektorów (głosy męskie) jednakowego tekstu gazetowego, ze stałym średnim poziomem natężenia dźwięku. Czas trwania sygnału mowy wynosił 10 minut dla każdego mówcy. Nagrania wykonano w studio odsłuchowym ITA Politechniki Wrocławskiej. Do nagrań użyto mikrofonu pojemnościowego o charakterystyce kołowej firmy Neumann, typ UM 57, oraz magnetofonu stacyjnego MS 181. Nierównomierność charakterystyki skuteczności mikrofonu nie przekraczała wartości ± 2 dB w pasmie 30 - 15 000 Hz, natomiast nierównomierność charakterystyki częstotliwościowej toru zapis- odczyt magnetofonu nie przekraczała wartości ± 1 dB w pasmie 20 - 15 000 Hz. Mikrofon znajdował się

w pobliżu ust mówiącego (w polu bliskim).

Nagrany sygnał analizowano za pomocą korelatora firmy Hewlett-Packard, typ HP 3721 [97]. Czas trwania analizy sygnału mowy ustalano za pomocą zegara sprzężonego z korelatorem. Zegar automatycznie uruchamiał korelator w momencie pojawienia się na jego wejściu sygnału mowy. Od chwili uruchomienia korelatora zegar mierzył czas trwania analizy sygnału mowy i po osiągnięciu zadanej wartości czasu analizy powodował przerwanie procesu przetwarzania sygnału mowy w korelatorze. Wynikiem procesu przetwarzania sygnału mowy przez korelator była sumaryczna liczba próbek występujących na danym poziomie kwantyzacji, w danym czasie trwania analizy sygnału mowy. Otrzymane w tej postaci wyniki poddano dalszej obróbce na maszynie cyfrowej, w wyniku której otrzymano mierzone rozkłady w postaci znormalizowanej, w układzie współrzędnych $[p(z), z]$, gdzie:

$$z = \frac{x - \mu}{\sigma} \quad (3.7)$$

gdzie:

z - znormalizowana zmienna losowa o wartości średniej

$$\mu_z = 0 \text{ i wariacji } \sigma_z^2 = 1$$

μ - wartość średnia zmiennej losowej $x(t)$,

σ^2 - wariancja zmiennej losowej $x(t)$.

W wyniku pomiarów dokonanych przez Brachmańskiego i Majewskiego [14] ustalono, że w pomiarach rozkładu gęstości prawdopodobieństwa wartości chwilowych sygnału mowy należy przyjmować czas trwania sygnału nie mniejszy niż 60 s.

Porównanie rozkładów gęstości prawdopodobieństwa wartości chwilowych sygnału mowy polskiej dla różnych mówców przeprowa-

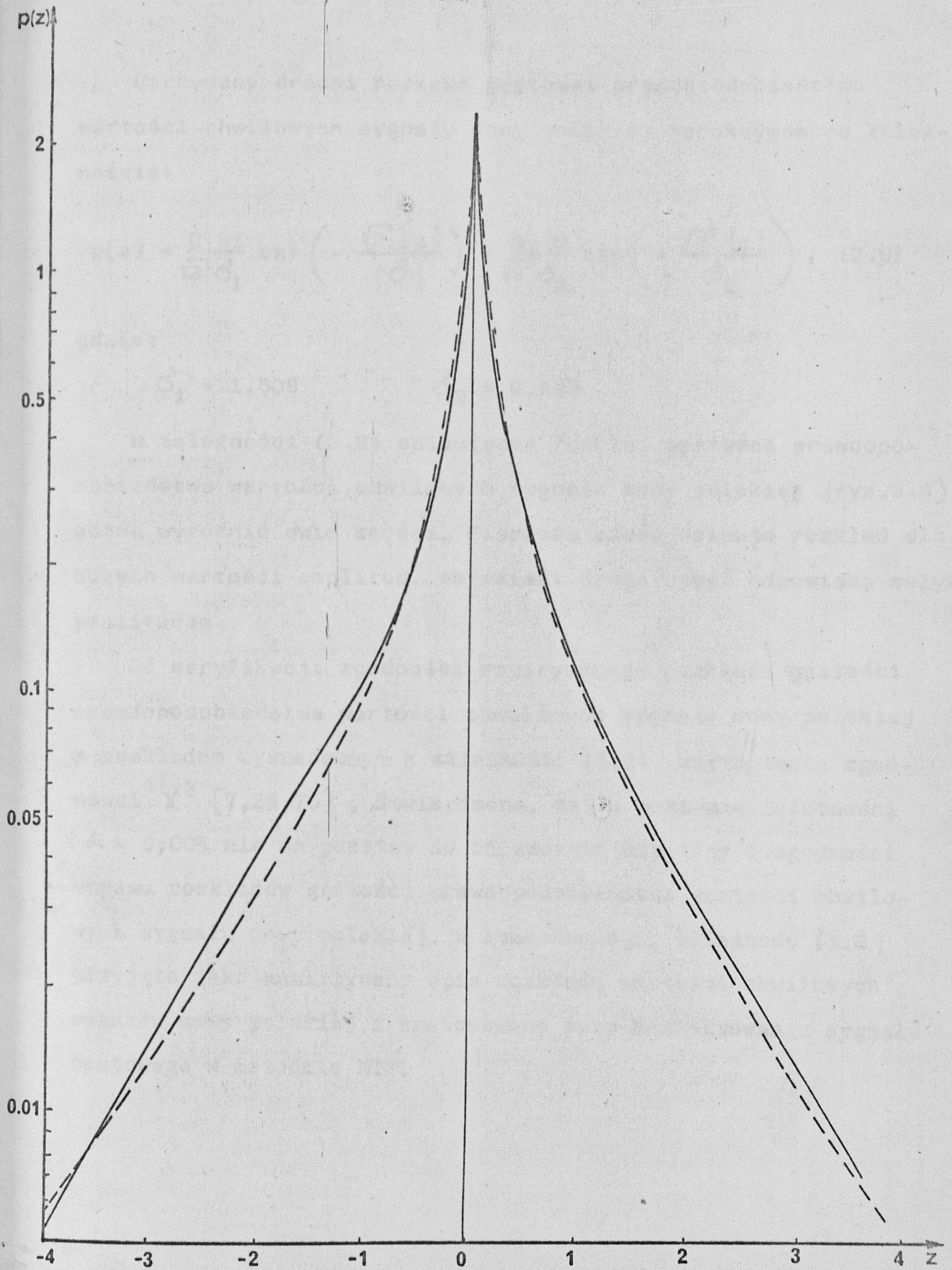
dzono za pomocą testu Smirnowa [25]. Stwierdzono, że na poziomie istotności $\alpha = 0,05$ wybrana populacja mówców stanowi populację jednorodną i nie ma podstaw do odrzucenia hipotezy o zgodności rozkładów.

W związku z tym otrzymane rozkłady posłużyły do wyznaczenia estymatora średniego rozkładu gęstości prawdopodobieństwa wartości chwilowych sygnału mowy polskiej. Rozkład ten jest wykreślony linią ciągłą na rys. 3.4.

Fizyczną interpretację określonego kształtu rozkładu gęstości prawdopodobieństwa wartości chwilowych sygnału mowy podał Davenport [18], wg którego zaostrzony wierzchołek krzywej rozkładu może wynikać z liczby pauz w mowie, w czasie których charakter sygnału jest określony pogłosem pomieszczenia, szumami pomieszczenia oraz szumami toru.

Inną interpretację podał Rimski - Korsakow [61], który przyjął, że przyczyną, dla której otrzymany rozkład różni się znacznie od normalnego jest fakt, iż mowa - jak również muzyka - jest złożonym procesem losowym. Oprócz fluktuacji faz i wartości chwilowych sygnału związanych z niekoherencją oddzielnych źródeł drgań i przejściami pomiędzy poszczególnymi dźwiękami (co powinno dać rozkład normalny), istnieją jeszcze stosunkowo wolne zmiany wariancji wynikające - w przypadku mowy - z prozodycznej modulacji siły głosu, akcentacji, wpływu procesu wydechowego itp.

Hipotezę Rimski - Korsakowa, że sygnał mowy ze względu na jego statystyczne własności należy rozpatrywać jako realizację procesu losowego z dyspersją modulowaną przez inny proces losowy, uzupełnił od strony analitycznej Wolf [16,85].



Rys.3.4. Rozkład gęstości prawdopodobieństwa wartości chwilowych sygnału mowy polskiej (—) oraz sygnału testowego (---).

Otrzymany średni rozkład gęstości prawdopodobieństwa wartości chwilowych sygnału mowy polskiej aproksymowano zależnością:

$$p(z) = \frac{0,51}{\sqrt{2} \sigma_1} \exp\left(-\frac{\sqrt{2} |z|}{\sigma_1}\right) + \frac{0,49}{\sqrt{2} \sigma_2} \exp\left(-\frac{\sqrt{2} |z|}{\sigma_2}\right), \quad (3.8)$$

gdzie:

$$\sigma_1 = 1,508 \quad , \quad \sigma_2 = 0,148$$

W zależności (3.8) opisującej rozkład gęstości prawdopodobieństwa wartości chwilowych sygnału mowy polskiej (rys.3.4) można wyróżnić dwie części. Pierwsza część opisuje rozkład dla dużych wartości amplitud, natomiast druga część odpowiada małym amplitudom.

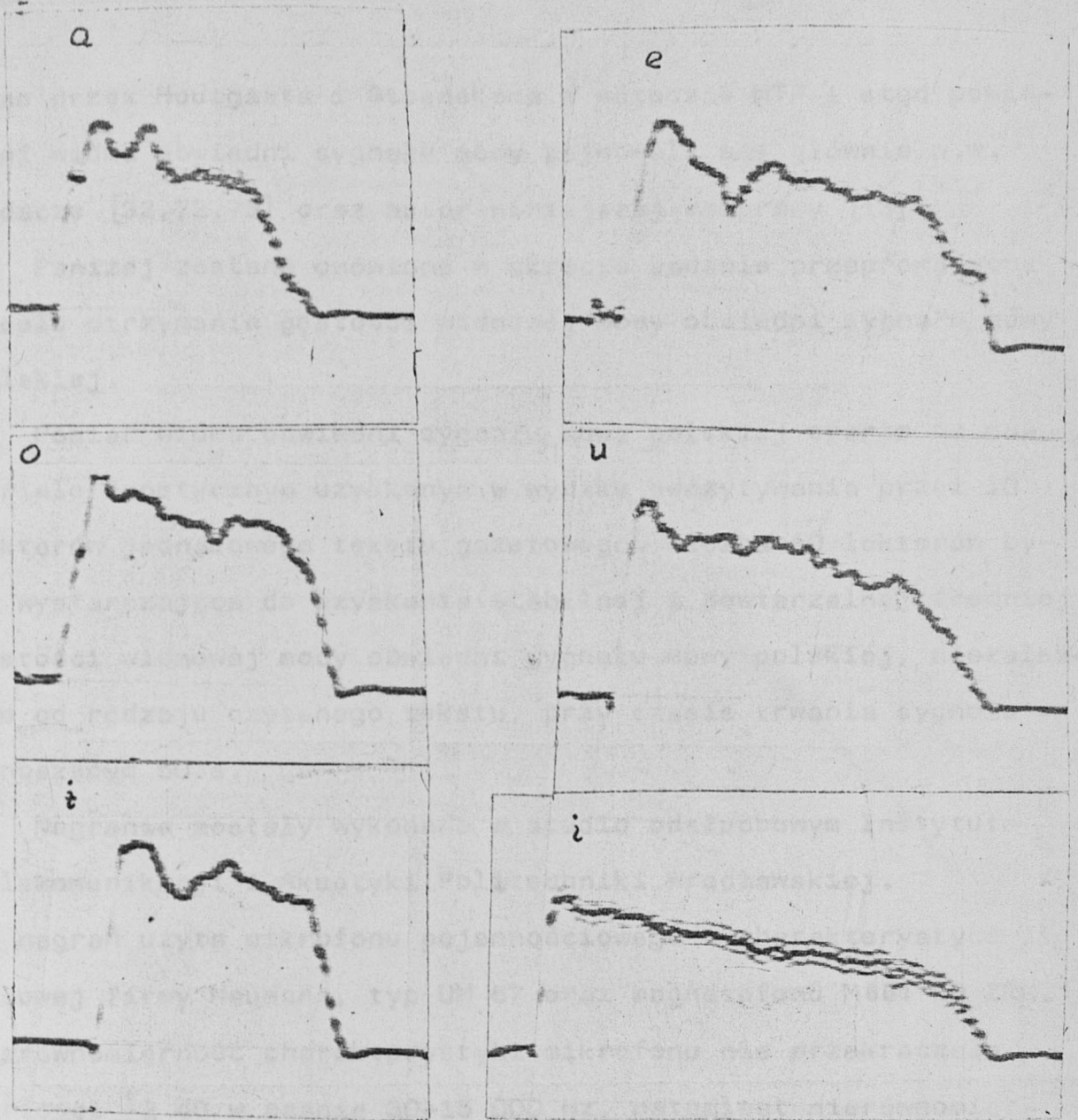
Do weryfikacji zgodności empirycznego rozkładu gęstości prawdopodobieństwa wartości chwilowych sygnału mowy polskiej z rozkładem wyznaczonym z zależności (3.8), użyto testu zgodności χ^2 [7,25,70]. Stwierdzono, że na poziomie istotności $\alpha = 0,005$ nie ma podstaw do odrzucenia hipotezy o zgodności obydwu rozkładów gęstości prawdopodobieństwa wartości chwilowych sygnału mowy polskiej. W konsekwencji, zależność (3.8) przyjęto jako analityczny opis rozkładu wartości chwilowych sygnału mowy polskiej i zastosowano przy kształtowaniu sygnału testowego w metodzie MTF.

3.4. Gęstość widmowa mocy obwiedni sygnału mowy polskiej

Porównując przebiegi czasowe dwu różnych dźwięków mowy (samogłosek, spółgłosek, itp) można zauważyć, że jedną z zasadniczych informacji pozwalających na rozróżnienie tych dźwięków jest różnica w kształcie obwiedni (amplitudzie chwilowej) sygnału mowy [2,32,38]. Na rys. 3.5 przedstawiono przykładowo zapis obwiedni samogłosek a,e,i,o,u,y oraz zdania "Jutro będzie ładny dzień". Zapisu dokonano za pomocą sonografu 6061-B firmy Kay Electric Co. wyposażonego w przystawkę umożliwiającą rejestrację obwiedni sygnału mowy (Amplitude Display 6076C).

Przebieg obwiedni sygnału mowy stanowi odzwierciedlenie zarówno wydatku energetycznego źródła pobudzającego, jak i zmian parametrów transmisyjnych kanału głosowego. Zmiany te wynikają ze stanów przejściowych od ciszy (zerowy wydatek energetyczny źródła) do dużego (maksymalnego) natężenia dźwięku. Zmiany amplitudy obwiedni w czasie trwania wypowiedzi, odpowiadające zmianom zależności amplitudowo-czasowych w obrębie wypowiedzianego elementu, nie są tak gwałtowne jak na początku lub na końcu elementu wypowiedzi. Występowanie różnorodnych zmian obwiedni sygnału mowy związane jest z szybkością i sposobem wypowiedzania, treścią lingwistyczną i semantyczną oraz cechami osobniczymi głosu mówcy. Poziom mocy widma obwiedni dla wyższych częstotliwości jest znikomo mały w porównaniu z częstotliwościami małymi i w związku z tym można ograniczyć pasmo mierzonej obwiedni sygnału do 25 Hz [32].

Ponieważ obwiednia czasowa sygnału mowy ulega zmianom wskutek wpływu różnego typu zniekształceń i zakłóceń wnoszonych przez kanał telefoniczny, właściwość ta została wykorzysta-



E B/65 SONAGRAM © KAY ELEMETRICS CO. PINE BROOK, N. J.

j u t r o b ę d Ź e ł a d n i d Ź e p



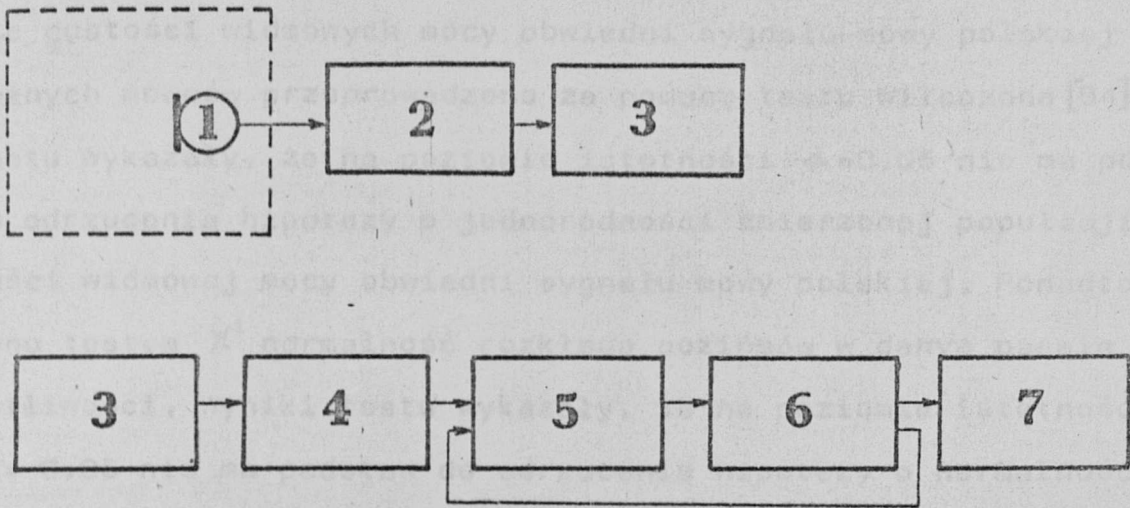
Rys. 3.5. Zapis obwiedni sygnału dla samogłosek : a,e,o,u, y (i), i oraz zdania "Jutro będzie ładny dzień" (fonetycznie - jutro będze ładni dżep).

tana przez Houtgasta i Steenekena w metodzie MTF i stąd pomiarami widma obwiedni sygnału mowy zajmowali się głównie w.w. badacze [32,72,73] oraz autor niniejszej rozprawy [13].

Poniżej zostaną omówione w skrócie badania przeprowadzone w celu otrzymania gęstości widmowej mocy obwiedni sygnału mowy polskiej.

Pomiar widma obwiedni sygnału mowy polskiej oparto na materiale fonetycznym uzyskanym w wyniku odczytywania przez 10 lektorów jednakowego tekstu gazetowego. Liczba 10 lektorów była wystarczająca do uzyskania stabilnej i powtarzalnej średniej gęstości widmowej mocy obwiedni sygnału mowy polskiej, niezależnie od rodzaju czytanego tekstu, przy czasie trwania sygnału wynoszącym 60 s.

Nagrania zostały wykonane w studio odsłuchowym Instytutu Telekomunikacji i Akustyki Politechniki Wrocławskiej. Do nagrań użyto mikrofonu pojemnościowego o charakterystyce kołowej firmy Neumann, typ UM 57 oraz magnetofonu M601 SD ZRK. Nierównomierność charakterystyki mikrofonu nie przekraczała wartości ± 2 dB w pasmie 30-15 000 Hz, natomiast nierównomierność charakterystyki toru zapis-odczyt magnetofonu nie przekraczała wartości ± 2 dB w pasmie 80-10 000 Hz. Mikrofon umieszczono w polu bliskim mówcy. Schemat blokowy systemu do pomiaru widma obwiedni sygnału mowy przedstawiono na rys. 3.6. Zasadniczym elementem tego systemu pomiarowego jest ekstraktor obwiedni, w którym zastosowano liniowy detektor prostownikowy oraz filtr dolnoprzepustowy o $f_{gr} = 25$ Hz i nachyleniu zbocza - 42 dB/okt /proces ekstrakcji omówiono w dodatku C/.



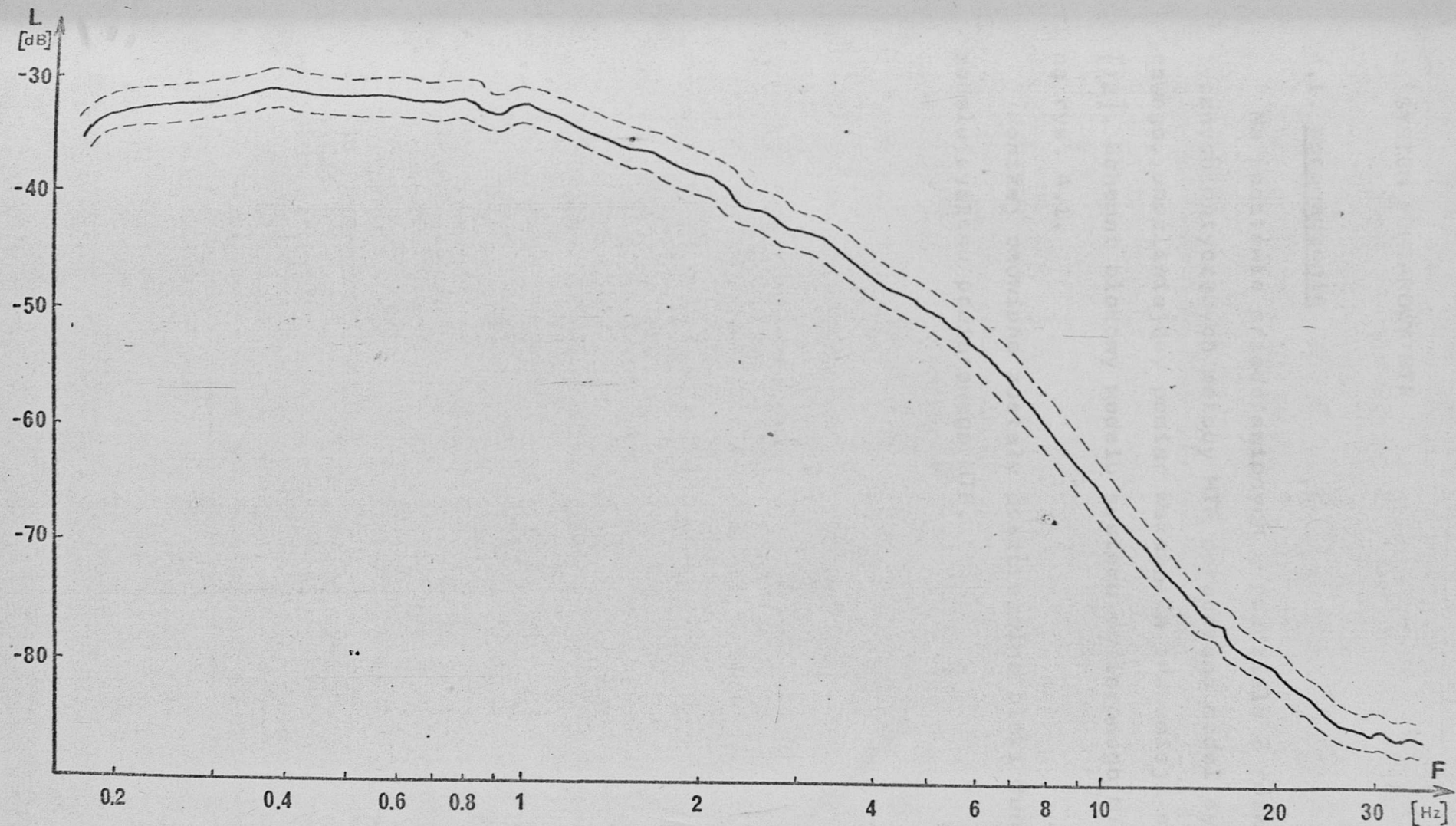
Rys.3.6. Schemat blokowy systemu realizującego pomiar obwiedni sygnału mowy

1 - mikrofon, 2 - wzmacniacz, 3 - magnetofon, 4 - ekstraktor obwiedni, 5 - cyfrowy rejestrator B&K 7502, 6 - Komputer Varian / B&K 7504, 7 - drukarka DZM 180

Obwiednią sygnału mowy, otrzymaną na wyjściu ekstraktora obwiedni, podawano na wejście cyfrowego rejestratora sygnałów losowych, pełniącego w układzie funkcję przetwornika analogowo-cyfrowego oraz pamięci zewnętrznej komputera. Po przetworzeniu na postać cyfrową sygnał obwiedni został wprowadzony do minikomputera Varian / B&K 7504. Gęstość widmową obwiedni sygnału mowy polskiej uzyskano w wyniku realizacji przekształcenia FFT wg. algorytmu zaczerpniętego z literatury [46,47]. Obliczona gęstość widmowa mocy obwiedni sygnału mowy polskiej złożona jest z 256 prążków widma rozmieszczonych równomiernie w przedziale częstotliwości od 0 do 25 Hz (rozdzielczość $\Delta f = 0,097654$ Hz), przy czym prążek pierwszy 0-0,097654 Hz został pominięty, gdyż odzwierciedla on składową stałą obwiedni sygnału mowy.

W wyniku przeprowadzonych pomiarów otrzymano 10 rozkładów gęstości widmowej mocy obwiedni sygnału mowy polskiej. Porównanie gęstości widmowych mocy obwiedni sygnału mowy polskiej dla różnych mówców przeprowadzono za pomocą testu Wilcozona [94]. Wyniki testu wykazały, że na poziomie istotności $\alpha=0,05$ nie ma podstaw do odrzucenia hipotezy o jednorodności zmierzonej populacji gęstości widmowej mocy obwiedni sygnału mowy polskiej. Ponadto zbadano testem χ^2 normalność rozkładu poziomów w danym pasmie częstotliwości. Wyniki testu wykazały, że na poziomie istotności $\alpha=0,05$ nie ma podstaw do odrzucenia hipotezy o normalności rozkładów poziomów we wszystkich pasmach częstotliwości. Stąd otrzymane gęstości widmowe posłużyły do wyznaczenia estymatora średniej gęstości widmowej mocy obwiedni sygnału mowy polskiej (rys 3.7). Stwierdzono, że poziom mocy widma obwiedni sygnału mowy polskiej dla częstotliwości wyższych od 25 Hz jest porównywalny z poziomem szumu systemu pomiarowego (stosunek sygnał/szum był równy 50 dB). W związku z tym, przedstawiono przebieg średniej gęstości widmowej obwiedni sygnału mowy polskiej w zakresie częstotliwości od 0,0976564 Hz do 30 Hz.

Przeprowadzając analizę błędów systemu pomiarowego obliczono błąd standardowy ϵ_0 oraz przedziały ufności estymatora gęstości widmowej. Teoretyczne podstawy analizy błędów systemu pomiarowego do wyznaczania widma obwiedni omówiono w dodatku D. W omawianych pomiarach, wartość błędów standardowych ϵ_0 estymatora średniej gęstości widmowej mocy obwiedni sygnału mowy polskiej wynosi $\epsilon_0 = 0,1291$. Obliczone dla poszczególnych pasm częstotliwości przedziały ufności wykreślono na rys. 3.7 linią przerywaną. Przedziały ufności obliczono przy prawdopodobieństwie $p=0,9$.



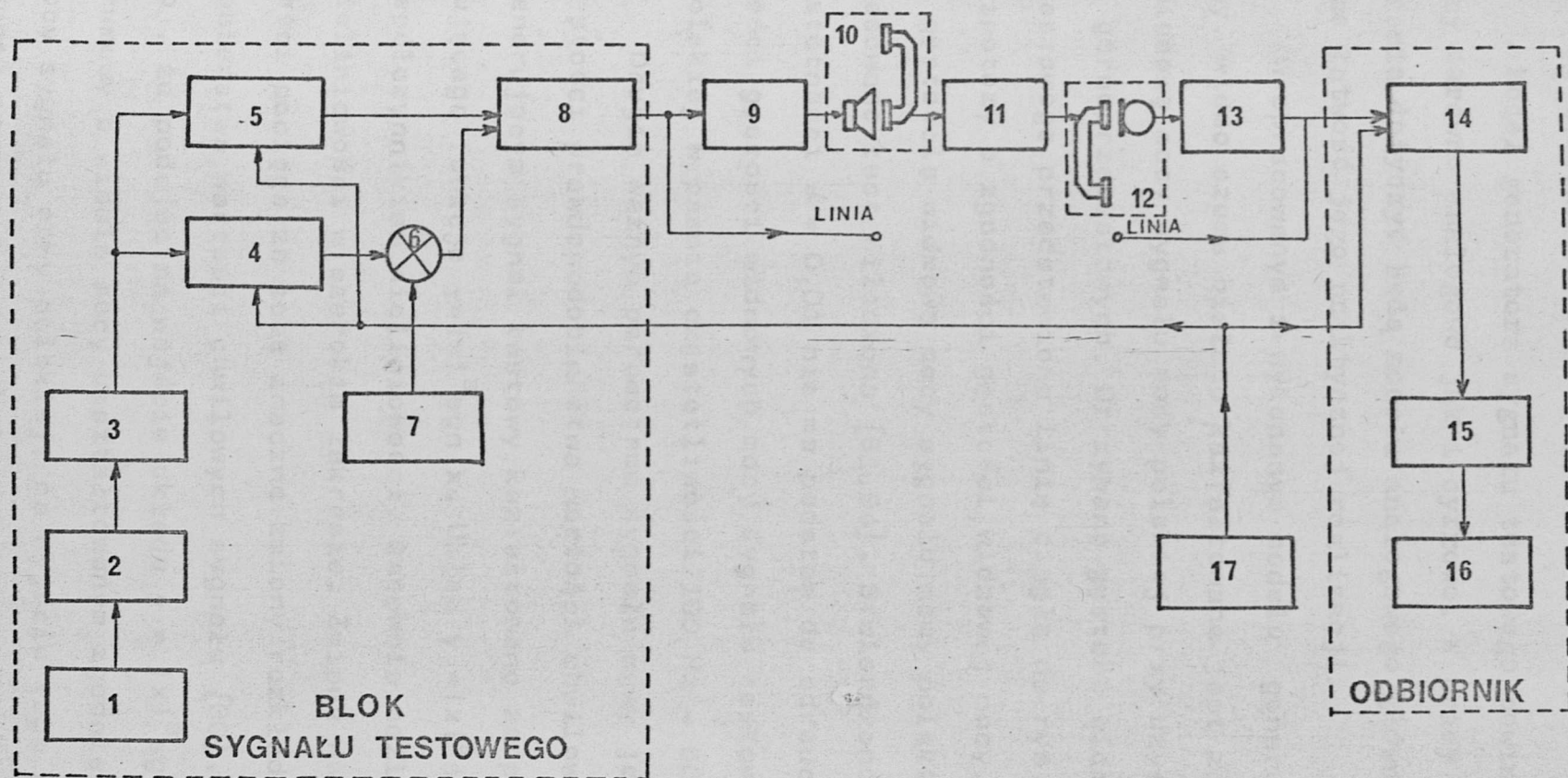
Rys.3.7. Gęstość widmowa obwiedni sygnału mowy polskiej analiza w pasmach o jednakowej szerokości .

4. SYSTEM POMIAROWY MTF

4.1. Wprowadzenie

Na podstawie przedstawionych w rozdziale 2 rozważań teoretycznych dotyczących metody MTF opracowano model systemu pomiarowego, umożliwiający pomiar wskaźnika transmisji sygnału mowy [12]. Schemat blokowy modelu systemu pomiarowego MTF pokazano na rys. 4.1.

Poniżej omówione zostały poszczególne bloki funkcjonalne modelu systemu pomiarowego MTF.



Rys.4.1. Schemat blokowy systemu pomiarowego MTF

1. Generator szumu białego; 2. Układ kształtowania widma; 3. Układ $y=|x|^a \text{sgn}x$; 4. Zestaw okta-
wowych filtrów pasmowo-przepustowych; 5. Zestaw okta-
wowych filtrów pasmowo-zaporowych; 6. Mo-
dulator amplitudy; 7. Generator sinusoidalny; 8. Sumator; 9. Wzmacniacz; 10. Sztuczne usta;
11. Łącze telefoniczne; 12. Sztuczne ucho; 13. Wzmacniacz mikrofonowy; 14. Zestaw okta-
wowych filtrów pasmowo-przepustowych; 15. Detektor synchroniczny; 16. Blok cyfrowy; 17. Blok ster-
owania

4.2. Realizacja bloku generującego sygnał testowy

Model generatora sygnału testowego powinien być realizowany zarówno analogowo jak i cyfrowo. W pracy przyjęto, że rozważania dotyczyć będą modelu analogowego głównie ze względu na łatwość jego praktycznej realizacji.

W opracowanym i wykonanym modelu generującym sygnał testowy, widmo szumu białego kształtowane jest zgodnie z gęstością widmową mocy sygnału mowy polskiej przy użyciu filtrów dolno- i górnoprzepustowych. Otrzymaną gęstość widmową mocy sygnału testowego przedstawiono linią ciągłą na rys. 3.1. Do weryfikacji hipotezy o zgodności gęstości widmowej mocy sygnału testowego z gęstością widmową mocy sygnału mowy polskiej (rys. 3.1) zastosowano test Wilcoxon [81,94]. Stwierdzono, że na poziomie istotności $\alpha = 0,01$ nie ma podstaw do odrzucenia hipotezy o zgodności gęstości widmowych mocy sygnału testowego i sygnału mowy polskiej w pasmie częstotliwości 100 Hz - 6000 Hz.

Drugim ważnym parametrem sygnału mowy jest średni rozkład gęstości prawdopodobieństwa wartości chwilowych, który w modelu generującym sygnał testowy kształtowano za pomocą układu realizującego funkcję $y = |x|^a \operatorname{sgn} x$. Układ $y = |x|^a \operatorname{sgn} x$, gdzie a jest współczynnikiem nieliniowości, zapewnia możliwość zmian stopnia nieliniowości w szerokim zakresie. Zmiana współczynnika nieliniowości pociąga za sobą znaczne zmiany rozkładu gęstości prawdopodobieństwa wartości chwilowych sygnału [30]. Empirycznie stwierdzono, że podając na wejście układu $y = |x|^a \operatorname{sgn} x$ ($a \approx 1.2$) sygnał szumowy o widmie mocy ukształtowanym zgodnie ze średnim widmem mocy sygnału mowy polskiej, na wyjściu tego układu uzyskuje się sygnał szumowy o pomijalnie małych zmianach w widmie mocy i

o rozkładzie gęstości prawdopodobieństwa wartości chwilowych sygnału przedstawionym na rys. 3.4 linią przerywaną. Do weryfikacji zgodności rozkładu gęstości prawdopodobieństwa wartości chwilowych sygnału testowego z rozkładem odnoszącym się do sygnału mowy polskiej (rys. 3.4 - linia ciągła) zastosowano test zgodności χ^2 [7,62, 90]. Stwierdzono, że na poziomie istotności $\alpha = 0,01$ nie ma podstaw do odrzucenia hipotezy o zgodności rozkładów gęstości prawdopodobieństwa wartości chwilowych sygnału testowego i sygnału mowy polskiej.

Tak ukształtowany szum podawany jest na zestaw oktaowych filtrów pasmowo-przepustowych oraz pasmowo-zaporowych (nachylenie charakterystyki częstotliwościowej wynosi 48 dB, częstotliwości środkowe 250 Hz, 500 Hz, 1000 Hz, 2000 Hz i 4000 Hz). Pasma częstotliwości filtru pasmowo-przepustowego są skorelowane z pasmami częstotliwości filtru pasmowo-zaporowego, tzn. jeżeli w danej chwili włączone jest oktaowe pasmo przepustowe o częstotliwości F_1 filtru pasmowo-przepustowego, to w tym samym momencie włączone jest zaporowe pasmo oktaowe o częstotliwości środkowej F_1 filtru pasmowo-zaporowego.

Stosując standardowe filtry oktaowe można podzielić pasmo częstotliwości łącza telefonicznego (w przypadku telefonii naturalnej - 300-3600-Hz [95,96]) na pięć pasm oktaowych o częstotliwościach środkowych 250 Hz, 500 Hz, 1000 Hz, 2000 Hz i 4000 Hz.

Powyższy podział pasma telefonicznego na pasma oktaowe został uwzględniony w procesie opracowywania i konstrukcji modelu systemu pomiarowego MTF. W każdym pasmie oktaowym sygnał szumowy jest poddany procesowi modulacji amplitudowej sygnałem sinusoidalnym o częstotliwościach wynikających z analizy widma obwiedni

sygnału mowy polskiej. Zbiór częstotliwości modulujących jest zbiorem 21 dyskretnych wartości częstotliwości odpowiadających częstotliwościom środkowym pasm tercjowych z zakresu od 0,25 Hz do 25 Hz. Wartości częstotliwości modulujących podano w tabeli 4.1.

Czas trwania sygnału modulującego (w pasmie oktawowym) dobierany jest tak, aby dla każdej częstotliwości modulującej wygenerowanych zostało 11 okresów sygnału modulującego, przy czym w procesie analizy sygnału testowego na wyjściu badanego kanału telefonicznego uwzględnionych jest 10 okresów sygnału testowego. Empirycznie stwierdzono bowiem, że dla danej częstotliwości próbkowania f_p sygnału testowego na wyjściu kanału telefonicznego do obliczeń głębokości modulacji metodą korelacyjną (patrz dodatek A) czas T_n trwania sygnału modulującego powinien być równy co najmniej jego 10 okresom .

Przez jedenasty okres sygnału modulującego trwa sygnał zmiany filtru oktawowego. Cykl 11 okresów powtarza się dla pozostałych filtrów pasm oktawowych . Po ostatnim piątym sygnale zmiany filtru generowany jest sygnał zmiany częstotliwości sygnału modulującego, który trwa przez jeden okres sygnału zmiany filtru. Cykl ten powtarzany jest dla wszystkich częstotliwości sygnału modulującego.

Czas trwania T_n sygnału modulującego w jednym pasmie oktawowym dla poszczególnych częstotliwości modulujących zamieszczono w tabeli 4.1.

Całkowity czas generowania sygnału modulującego dla $L=5$ i $N=21$ wynosi 1236,36 s.

Tabela 4.1. Wartości częstotliwości modulujących oraz odpowiadające im czasy trwania T_n sygnału modulującego

n	1	2	3	4	5	6	7	8	9	10	11
f_n [Hz]	0,25	0,315	0,4	0,5	0,63	0,8	1	1,25	1,6	2	2,5
T_n [s]	44	34,9	27,5	22	17,5	13,8	11	8,8	6,9	5,5	4,4
n	12	13	14	15	16	17	18	19	20	21	
f_n [Hz]	3,15	4	5	6,3	8	10	12,5	16	20	25	
T_n [s]	3,5	2,7	2,2	1,8	1,4	1,1	0,88	0,69	0,55	0,44	

Sygnal szumowy zmodulowany amplitudowo w pasmach oktaowych sygnałem sinusoidalnym zostaje połączony w sumatorze z sygnałem przepuszczonym przez filtr pasmowo-zaporowy. Na wejście badanego kanału telefonicznego podawany jest więc sygnał testowy reprezentujący omówione w rozdziale 3 podstawowe statystyki języka polskiego i spełniający zależność (2.1).

Dane techniczne sygnału testowego w wersji zmodyfikowanej

- szum biały o gęstości widmowej mocy ukształtowanej, w pasmie 100 Hz - 6000 Hz, zgodnie ze średnią gęstością widmową mocy sygnału mowy polskiej (rys. 3.1).
- rozkład gęstości prawdopodobieństwa wartości chwilowych opisuje zależność (3.8) aproksymująca rozkład odnoszący się do mowy polskiej
- wartości częstotliwości sygnału modulującego:

$$\{ 0,25 ; 0,315 ; 0,4 ; 0,5 ; 0,63 ; 0,8 ; 1,0 ; 1,25 ; 1,6 ; 2,0 ; 2,5 ; 3,15 ; 4,0 ; 5,0 ; 6,3 ; 8,0 ; 10,0 ; 12,5 ; 16,0 ; 20,0 ; 25,0 \} \text{ Hz}$$
- wartości częstotliwości środkowych pasm oktaowych, w których sygnał szumowy jest modulowany amplitudowo sygnałem sinusoidalnym :

$$\{ 250 ; 500 ; 1000 ; 2000 ; 4000 \} \text{ Hz}$$

5. głębokość modulacji - 100%
6. poziom napięcia sygnału testowego podawanego na wejście modelu kanału telefonicznego - 0 dB
7. czas trwania sygnału testowego - 1236,36 s.

4.3. Realizacja odbiornika

Zadaniem odbiornika jest wydzielenie obwiedni $Y_{n,1}(t)$ (zależność 2.2) zniekształconego sygnału modulującego, obliczenie głębokości modulacji, wskaźnika W_{STI} transmisji sygnału mowy oraz estymowanej wyrazistości logatomowej. Analiza sygnału na wyjściu kanału telefonicznego jest przeprowadzana w pasmach oktaowych. W ten sposób uwzględniony zostaje wpływ, wnoszonych przez kanał telefoniczny zniekształceń i zakłóceń będących funkcją częstotliwości. Zestaw oktaowych filtrów pasmowo-przepustowych jest analogiczny do znajdującego się w bloku sygnału testowego.

Estymacja obwiedni $Y_{n,1}(t)$ sygnału wyjściowego przeprowadzana jest za pomocą detektora synchronicznego. Na wybór detektora synchronicznego wpłynął fakt, że przy detekcji synchronicznej (iloczynowej) nie występuje tzw. efekt progowy, który polega na tym, że poniżej pewnej wartości progowej stosunku sygnał/szum (S/N) na wejściu detektora, stosunek ten na wyjściu detektora maleje znacznie szybciej niż na jego wejściu [38, 67]. W zrealizowa-

wanym systemie pomiarowym MTF synchroniczna detekcja obwiedni $y_{n,1}(t)$ zrealizowana jest w oparciu o detektor prostownikowy oraz filtr dolnoprzepustowy o częstotliwości granicznej 25 Hz i nachyleniu charakterystyki częstotliwościowej ok. 36 dB/okt.

Otrzymana obwiednia $y_{n,1}(t)$ zostaje przetworzona, za pomocą przetwornika analogowo-cyfrowego (A/C), na postać cyfrową i poprzez interfejs wprowadzona do maszyny cyfrowej. Działanie przetwornika A/C opiera się na zasadzie przetwarzania kompensacyjnego o częstotliwości próbkowania f_p będącej wielokrotnością częstotliwości modulującej, przy czym $f_p = 5 f_n$. Powyższa procedura zapewnia uzyskanie dla wszystkich częstotliwości modulujących jednakowej dokładności przetwarzania obwiedni $y_{n,1}(t)$ na postać cyfrową. Zakres przetwarzanych napięć wynosi 0 - +5V, przy czym maksymalną wartość napięcia (+5V) można zmniejszyć w sposób ciągły do wartości ok. +2V.

Obwiednia $y_{n,1}(t)$ zostaje skwantowana na 256 poziomów (8 bitów). Postać cyfrowa obwiedni $y_{n,1}(t)$ zostaje wprowadzona poprzez interfejs do maszyny cyfrowej. W maszynie cyfrowej następuje obliczenie: głębokości modulacji metodą korelacyjną (Dodatek A), wskaźnika transmisji sygnału mowy (zależność 2.9) oraz estymowanej subiektywnej oceny jakości transmisji sygnału mowy, którą maszyna cyfrowa wyprowadza za pomocą drukarki. Do celów badawczych (optymalizacyjnych) drukowane były również wyniki pośrednie, tzn. głębokość modulacji i wskaźnik W_{STI} transmisji sygnału mowy.

W wykonanym systemie pomiarowym MTF przewidziano możliwość wyprowadzania postaci cyfrowej obwiedni $y_{n,1}(t)$ na taśmie papierowej za pomocą perforatora DT-101.

4.4. Realizacja bloku sterowania

W zrealizowanym modelu systemu pomiarowego przewidziano możliwość ręcznego lub automatycznego pomiaru wskaźnika transmisji sygnału mowy. Zmiana częstotliwości f_n generatora sinusoidalnego oraz przełączanie filtrów pasmowo-przepustowych i pasmowo-zaporowych dokonywane jest za pośrednictwem zestyków kontaktowych oraz układów przełączających znajdujących się w bloku sterowania.

Zadaniem bloku sterowania jest wysyłanie do pozostałych bloków układu, sygnałów sterujących zapisem rejestrów i ich zerowaniem, co zachodzi synchronicznie ze zmieniającą się częstotliwością generatora sygnału sinusoidalnego. Prócz tego w bloku sterowania wytwarzane są sygnały: zerujący przerzutniki po włączeniu zasilania układu, inicjujący pomiar i blokujący pomiar podczas przełączania filtrów oktaowych.

Do realizacji modelu systemu pomiarowego MTF znaczny wkład wniósł autor niniejszej rozprawy. Opracował i wykonał w całości model generujący sygnał testowy oraz część odbiornika obejmującą następujące układy: zestaw oktaowych filtrów pasmowo-przepustowych oraz detektor synchroniczny.

Oprogramowanie maszyny cyfrowej (serii Odra 1300) oraz układy obejmujące przetwornik analogowo-cyfrowy, interface i blok sterowania zostały opracowane i wykonane przez innych członków zespołu.

5. SUBIEKTYWNY POMIAR WYRAZISTOŚCI

5.1. Wprowadzenie

Opracowując lub modyfikując obiektywną metodę oceny jakości transmisji sygnału mowy należy ją zweryfikować za pomocą miary odniesienia uzyskanej na drodze subiektywnych pomiarów jakości transmisji. Wyniki pomiarów subiektywnych powinny w maksymalnym stopniu zależeć od parametrów fizycznych badanego kanału telefonicznego, a nie od struktury testu językowego. Eliminację informacji na poziomie semantycznym zapewniają listy logatomowe, na podstawie których określa się wyrazistość logatomową lub fonemową. Dlatego też jako miarę odniesienia jakości kanału telefonicznego, spełniającą rolę weryfikatora skuteczności metody MTF, przyjęto w niniejszej pracy wyrazistość logatomową uśrednioną w zbiorze słuchaczy.

5.2. Wybór punktów pomiarowych

Wybór punktów pomiarowych ^{5/} powinien zapewnić uzyskanie możliwie dużego zakresu zmienności wartości wyrazistości logatomowej, przy jednoczesnym uwzględnieniu różnorodnych warunków transmisji sygnału mowy występujących w rzeczywistych kanałach telefonicznych.

W przeprowadzonych badaniach korzystano ze specjalnie do tego celu opracowanego i wykonanego w Instytucie Telekomunikacji i Akustyki Politechniki Wrocławskiej modelu kanału telefonicznego.

^{5/} Przez "punkt pomiarowy" należy rozumieć określone warunki transmisji sygnału mowy, którym odpowiada odpowiedni zestaw parametrów fizycznych opisujących dany kanał telefoniczny i panujące w nim warunki transmisji sygnału.

MKT-1 [4]. W modelu tym jest możliwe wprowadzenie następujących zakłóceń i zniekształceń do przesyłanego sygnału mowy:

1^o zakłócenia addytywne:

- a/ - szum biały o poziomie regulowanym w sposób ciągły,
- b/ - szum różowy o poziomie regulowanym w sposób ciągły,
- c/ - zakłócenia impulsowe o czterech różnych wartościach średnich rozkładu losowego generującego odstęp czasowy pomiędzy kolejnymi impulsami i o amplitudzie regulowanej w sposób ciągły,

2^o ograniczenie pasma częstotliwości - 78 kombinacji ustawianych skokowo w pasmie 100 - 4500 Hz oraz pozycja bez ograniczenia pasma,

3^o zniekształcenia liniowe charakterystyki amplitudowej (9 niezależnie włączonych korektorów liniowych).

4^o sygnały zakłócające zewnętrzne (np. przesłuch zrozumiały, przesłuch niezrozumiały itp.); w celu wprowadzenia tego typu zakłóceń w modelu kanału przewidziano odrębne wejście na sumator.

Wymienione możliwości modelu kanału pozwalają na symulację dużej liczby różnorodnych warunków transmisji sygnału mowy kanałem telefonicznym.

Z uwagi na ograniczone możliwości wykonania pomiarów subiektywnych, liczbę punktów pomiarowych czyli eksperymentów praktycznych przeprowadzonych w ramach niniejszej pracy musiano jednak ograniczyć do ok. 100 (jest to liczba eksperymentów porównywalna z wykonanymi przez Houtgasta i Steenekena [74]). Szczególną uwagę zwrócono na właściwy dobór warunków transmisji w symulowanym kanale telefonicznym, aby uzyskać zarówno różnorodne warunki

transmisji jak i równomierne rozłożenie uzyskiwanych wartości wyrazistości na skali wyrazistości. Przy wyborze warunków pomiaru przyjęto następujące założenia:

- 1^o Decydujący wpływ na wyrazistość w kanałach telefonicznych ma znany a priori ograniczony zakres zmian częstotliwości pasma przenoszenia i szumu addytywnego
- 2^o Ze względu na ograniczoną liczbę punktów pomiarowych, należy w większości przypadków badać łączny wpływ kilku czynników,
- 3^o Należy utrzymywać stały poziom sygnału mowy wynoszący 80 dB, tzn. poziom optymalny dla odbioru mowy polskiej w warunkach odbioru jednousznego [9].
- 4^o Czynniki decydujące (szum, przesłuch, ograniczenie pasma) należy badać dla kilku wartości,
- 5^o Maksymalny rozważany poziom szumu powinien być nie większy niż +12 dB w stosunku do poziomu sygnału mowy [23].
- 6^o Minimalny poziom szumu powinien wynosić -18 dB w stosunku do sygnału mowy [23],
- 7^o Przy ustalaniu rozważanych wariantów ograniczenia pasma przenoszenia należy uwzględnić przedział tłumienności optymalnych dla łącz telefonicznych określony zaleceniami G151 CCITT [95].
Ustalając konkretne wartości parametrów zakłóceń i zniekształceń korzystano z wyników badań nad zrozumiałością (wyrazistością) mowy zawartych w pracach [15,31,33,44,64,88]. Wyniki te należy potraktować jako bardzo przybliżone oszacowanie wynikowej wyrazistości, gdyż badania te przeprowadzono dla różnych warunków realizacji odsłuchów subiektywnych oraz różnego rodzaju materiału testowego. W związku z tym, w pierwszym etapie badań zbadano związek między wyrazistością logatomową a pojedynczym czynnikiem

zakłócającym, którym był szum biały, szum różowy, przydźwięk oraz przesłuch zrozumiały. Otrzymane wyniki przedstawiono na rys. 5.1.

Uwzględniając podane powyżej założenia, wyniki wstępnych badań własnych i cytowanych prac oraz możliwości ustawienia wartości parametrów stosowanego modelu kanału telefonicznego, przyjęto następujące wartości parametrów:

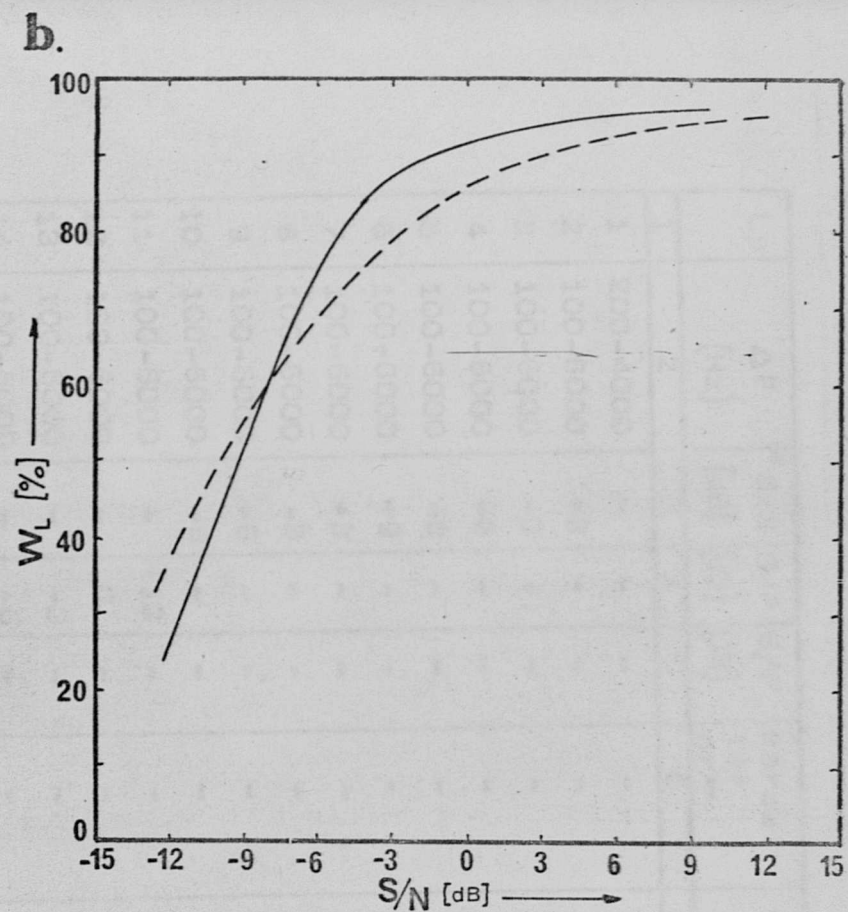
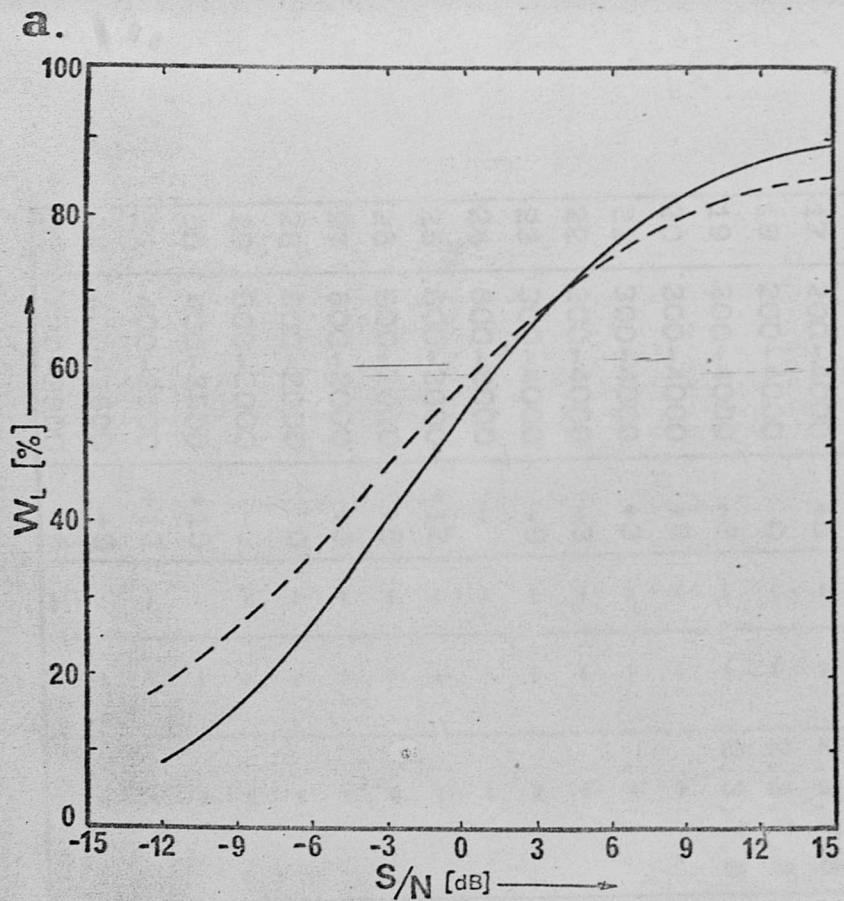
- 1^o szum biały, szum różowy, przesłuch zrozumiały oraz przydźwięk o poziomach względem poziomu sygnału mowy: +15 dB, +12 dB, +9 dB, +6 dB, +3 dB, 0 dB, -3 dB, -6 dB, -9 dB, -12 dB,
- 2^o szerokość pasma: 100 - 6000 Hz, 200 - 4000 Hz, 300 - 4000 Hz, 600 - 2000 Hz, 400 - 2500 Hz,
- 3^o dla niektórych wariantów ograniczenia pasma stosowano korektory, celem symulacji pewnego rodzaju nierównomierności charakterystyk amplitudowych występujących w rzeczywistych kanałach telefonicznych.

Wybrany zestaw punktów pomiarowych przedstawiono w tabeli 5.1.

5.3. Wyniki subiektywnych pomiarów wyrazistości

W przeprowadzonych pomiarach subiektywnych istotne jest, aby odbywały się w warunkach maksymalnie zbliżonych do rzeczywistych warunków transmisji sygnału, przy czym czynniki zakłócające i zniekształcające przesyłany sygnał powinny być znane i możliwe do wielokrotnego odtwarzania celem uzyskania możliwości porównania obiektywnej miary wyrazistości z miarą subiektywną.

Uwzględniając powyższe uwagi, w zrealizowanym eksperymencie zdecydowano się przeprowadzić subiektywne pomiary wyrazistości metodą odsłuchu jednousznego, stosując słuchawki nagłowne SN-60 firmy Tonsil (charakterystyki skuteczności słuchawek SN-60 zamiesz-



Rys. 5.1. Związek między wyrazistością logatomową a stosunkiem: poziom natężenia sygnału mowy - poziom natężenia pojedynczego czynnika zakłócającego (S/N), w pasmie 400-2500 Hz
 a/ — szum biały, --- szum różowy,
 b/ — przydźwięk, --- przesłuch zrozumiały.

Tabela 5.1. Parametry punktów pomiarowych oraz wyniki pomiarów wyrazistości $/W_L/$.

Lp	ΔF [Hz]	$^* S/N$ [dB]	S/P [dB]	S/P _r [dB]	korek tor nr**	Uwagi	W_L [%]
1	2	3	4	5	6	7	8
1	200-4000	-	-	-	-	-	96,3
2	100-6000	+3	-	-	-	sb	67.1
3	100-6000	-3	-	-	-	sb	31.2
4	100-6000	-6	-	-	-	sb	22.5
5	100-6000	-9	-	-	-	sb	9.2
6	100-6000	+9	-	-	-	sr	76.5
7	100-6000	+3	-	-	-	sr	52.6
8	100-6000	-3	-	-	-	sr	35.3
9	100-6000	-6	-	-	-	sr	19.5
10	100-6000	-9	-	-	-	sr	8.5
11	100-6000	-	+12	-	-	-	95.3
12	100-6000	-	0	-	-	-	89.0
13	100-6000	-	-3	-	-	-	75.6
14	100-6000	-	-6	-	-	-	56.3
15	200-4000	+12	-	-	2,3,7,8	sb	86.8
16	200-4000	+6	-	-	2,3,7,8	sb	78.2
17	200-4000	+3	-	-	2,3,7,8	sb	69.1
18	200-4000	0	-	-	2,3,7,8	sb	58,9
19	200-4000	-3	-	-	2,3,7,8	sb	32.4
20	300-4000	+9	-	-	-	sb	74.3
21	300-4000	+3	-	-	-	sb	61.0
22	300-4000	-3	-	-	-	sb	41.6
23	300-4000	-9	-	-	-	sb	29.2
24	600-2000	-	-	-	-	-	93.4
25	600-2000	+12	-	-	-	sb	83.7
26	600-2000	+6	-	-	-	sb	77.2
27	600-2000	+3	-	-	-	sb	59.9
28	600-2000	0	-	-	-	sb	58.5
29	600-2000	-3	-	-	-	sb	36.7
30	400-2500	+15	-	-	-	sb	89.3
31	400-2500	+12	-	-	-	sb	86.9
32	400-2500	+9	-	-	-	sb	83.2
33	400-2500	+6	-	-	-	sb	76.2

1	2	3	4	5	6	7	8
34	400-2500	+3	-	-	-	sb	66.6
35	400-2500	0	-	-	-	sb	53.8
36	400-2500	-2	-	-	-	sb	48.9
37	400-2500	-4	-	-	-	sb	42.4
38	400-2500	-6	-	-	-	sb	25.8
39	400-2500	-9	-	-	-	sb	16.4
40	400-2500	-12	-	-	-	sb	9.1
41	400-2500	+15	-	-	-	sr	84.9
42	400-2500	+9	-	-	-	sr	80.7
43	400-2500	+3	-	-	-	sr	67.9
44	400-2500	-3	-	-	-	sr	48.5
45	400-2500	-6	-	-	-	sr	34.2
46	400-2500	-9	-	-	-	sr	26.2
47	400-2500	-12	-	-	-	sr	19.2
48	400-2500	-	+6	-	-	-	94.9
49	400-2500	-	0	-	-	-	92.6
50	400-2500	-	-3	-	-	-	88.5
51	400-2500	-	-9	-	-	-	52.5
52	400-2500	-	-12	-	-	-	24.3
53	400-2500	-	+3	-	3,4,6,7	-	85.9
54	400-2500	-	-	+6	-	pzp	92.5
55	400-2500	-	-	0	-	pzp	85.4
56	400-2500	-	-	-3	-	pzp	78.6
57	400-2500	-	-	-6	-	pzp	68.1
58	400-2500	-	-	-9	-	pzp	56.2
59	400-2500	-	-	-12	-	pzp	37.4
60	400-2500	-	-	+12	-	pza	91.9
61	400-2500	-	-	+6	-	pza	87.4
62	400-2500	-	-	0	-	pza	80.8
63	400-2500	-	-	-3	-	pza	73.9
64	400-2500	-	-	-6	-	pza	66.2
65	400-2500	-	-	-9	-	pza	51.2
66	400-2500	-	-	-12	-	pza	33.1
67	400-2500	0	-	-	-	sbm/50Hz/	49.3
68	400-2500	-9	-	-	-	sbm/50Hz/	31.8
69	400-2500	-3	-	-	-	sbm/1kHz/	55.9
70	400-2500	-9	-	-	-	sbm/1kHz/	42.2

1	2	3	4	5	6	7	8
71	400-2500	+18	-	-3	-	sb, pzp	68.2
72	400-2500	+12	-	+18	3,4,6,7	sb, pzp	84.1
73	400-2500	+12	-	+12	3,4,6,7	sb, pzp	76.8
74	400-2500	+12	-	+12	3,4,6,7	sb, pn	73.0
75	400-2500	+12	-	+6	3,4,6,7	sb, pzp	74.6
76	400-2500	+12	-	+6	3,4,6,7	sb, pn	67.0
77	400-2500	+12	-	+3	3,4,6,7	sb, pzp	72.5
78	400-2500	+12	-	+3	3,4,6,7	sb, pn	72.0
79	400-2500	+12	+6	-	3,4,6,7	sb	71.8
80	400-2500	+12	+3	-	3,4,6,7	sb	68.7
81	400-2500	+6	+6	-	3,4,6,7	sb	59.0
82	400-2500	+6	+3	-	3,4,6,7	sb	62.0
83	400-2500	+3	+6	-	3,4,6,7	sb	68.8
84	400-2500	+3	+3	-	3,4,6,7	sb	55.2
85	400-2500	0	+6	-	3,4,6,7	sb	60.3
86	400-2500	0	-	0	3,4,6,7	sb, pzp	76.9
87	400-2500	0	-	0	3,4,6,7	sb, pn	68.0
88	400-2500	-4	-	-3	-	sb, pzp	10.4
89	400-2500	-4	-	-3	-	sb, pza	30.8
90	400-2500	-6	-	-	-	sb, sr	32.9
91	400-2500	-6	-	-6	-	sb, pzp	13.5
92	400-2500	-3	0	-	-	sbm/50Hz	25.6
93	400-2500	-2	-	-3	-	sr, pzp	38.8
94	400-2500	-3	-3	-	-	sr	20.9
95	400-2500	-6	0	-	-	sr	19.9

*S/N oznacza stosunek sygnał/szum, S/P - stosunek sygnał/przydźwięk, S/P_r - stosunek sygnał/przesłuch, ΔF oznacza pasmo częstotliwości, sb - szum biały, sr - szum różowy, sbm/50Hz/ - szum biały modulowany sygnałem sinusoidalnym o częstotliwości $f_{mod} = 50\text{Hz}$, sbm/1kHz/ - szum biały modulowany sygnałem sinusoidalnym o częstotliwości $f_{mod} = 1\text{kHz}$, pzp - przesłuch zrozumiały (język polski), pza - przesłuch zrozumiały (język angielski), pn - przesłuch niezrozumiały, W_L - wyrazistość logatomowa (subiektywna).

** - korektor nr 1 - częstotliwość środkowa $f_{sr} = 200\text{Hz}$, nr 2 - $f_{sr} = 250\text{Hz}$, nr 3 - $f_{sr} = 315\text{Hz}$, nr 4 - $f_{sr} = 400\text{Hz}$, nr 5 - $f_{sr} = 500\text{Hz}$, nr 6 - $f_{sr} = 2500\text{Hz}$, nr 7 - $f_{sr} = 3150\text{Hz}$, nr 8 - $f_{sr} = 4000\text{Hz}$.

czono w dodatku E).

Odsłuchu dokonywano w studio odsłuchowym Instytutu Telekomunikacji i Akustyki Politechniki Wrocławskiej przy optymalnym poziomie sygnału mowy równym 80 dB(A) [9].

Ważnym zagadnieniem w procedurze pomiarowej jest odstęp czasowy między sesjami odsłuchowymi, łączny czas trwania sesji odsłuchowej, tempo nadawania jednostek tworzących listy oraz forma rejestracji wyników odsłuchów. Ze względu na właściwości adaptacyjne słuchaczy odstęp czasowy między sesjami nie powinien być krótszy niż 1 doba i nie dłuższy niż kilka dni [92]. W omawianych badaniach przerwa między sesjami wynosiła 2 lub 3 dni.

Łączny czas trwania sesji nie przekraczał 2 godzin, co pozwalało na dokonanie pomiaru wyrazistości dla czterech punktów pomiarowych. Pomiędzy kolejnymi punktami pomiarowymi (czas trwania pomiaru wyrazistości dla jednego punktu pomiarowego wynosił ok. 20 minut) robiono 10 minutowe okresy odpoczynku. Odstęp pomiędzy poszczególnymi pozycjami list logatomowych powinien umożliwić słuchaczowi spokojną rejestrację odebranego bodźca.

Ustalono eksperymentalnie, że optymalnym tempem powtarzania wypowiedzi dla logatomów są wypowiedzi co 3 - 7 sekund zależnie od długości logatomów.

W przeprowadzonych eksperymentach, na każdy punkt pomiarowy przypadały cztery listy 100 logatomowe (400 logatomów) zrównoważone fonetycznie i strukturalnie [92].

Zgodnie z zaleceniami ISO [92] trening oraz testy odsłuchowe powinny być realizowane przynajmniej z jednym mówcą (kobieta lub mężczyzna). W niniejszych badaniach, subiektywne pomiary wyrazistości zrealizowano wykorzystując jednego mówcę (mężczyzna).

Listy logatomowe nagrano na taśmie magnetofonowej w studio Rozgłośni Wrocławskiej Polskiego Radia. Do nagrań list logatomo-

wych, czytanych przez spikera Rozgłośni Wrocławskiej Polskiego Radia, użyto mikrofonu pojemnościowego o charakterystyce kołowej - typ MCO 30, firmy Tonsil oraz magnetofonu stacyjnego MS181. Nierównomierność charakterystyki skuteczności mikrofonu nie przekracza wartości ± 3 dB w pasmie 20 - 20 000 Hz, natomiast nierównomierność charakterystyki częstotliwościowej toru zapis-odczyt magnetofonu nie przekraczała wartości ± 1 dB w pasmie 20 - 15 000 Hz. Mikrofon znajdował się w pobliżu ust lektora w polu bliskim. Stosunek sygnał/szum całego układu nagrywania był równy 50 dB.

Symulacja różnych warunków transmisji została przeprowadzona za pomocą omówionego modelu kanału telefonicznego MKT-1 [4]. Poziomy sygnału mowy oraz sygnałów zakłócających kontrolowano wzmacniaczem pomiarowym Brüel & Kjaer typ 2606, mierząc je na skali logarytmicznej według krzywej korelacyjnej A. Wartości skuteczne napięcia sygnału mowy oraz sygnału zakłócającego, odpowiadające założonym wartościom poziomu natężenia dźwięku przy uchu słuchacza zmierzono eksperymentalnie.

Ekipę odsłuchową dobrano spośród studentów Politechniki Wrocławskiej (wiek 20-24 lat). Liczebność ekipy wynosiła 12 osób. Kwalifikacja odbywała się w oparciu o badania audiometryczne przeprowadzone zgodnie z zaleceniami ISO [92]. Przed przystąpieniem do właściwych pomiarów wyrazistości, ekipę odsłuchową poddano 8 godzinnemu treningowi [92].

Jako miarę jakości transmisji sygnału mowy dla danego punktu pomiarowego przyjęto średnią wyrazistość logatomową zdefiniowaną jako stosunek liczby poprawnie odebrnych logatomów do liczby logatomów nadanych (w procentach). Jeśli k oznacza numer 100 logatomowej listy, a s numer słuchacza ekipy odsłuchowej, to wyrazistość średnią W_L^p określającą jakość transmisji sygnału mowy dla p -tego punktu pomiarowego można określić na podstawie zależności:

$$W_L^p = \frac{1}{SK} \sum_{k=1}^K \sum_{s=1}^S W_{s,k}^p \quad (5.1)$$

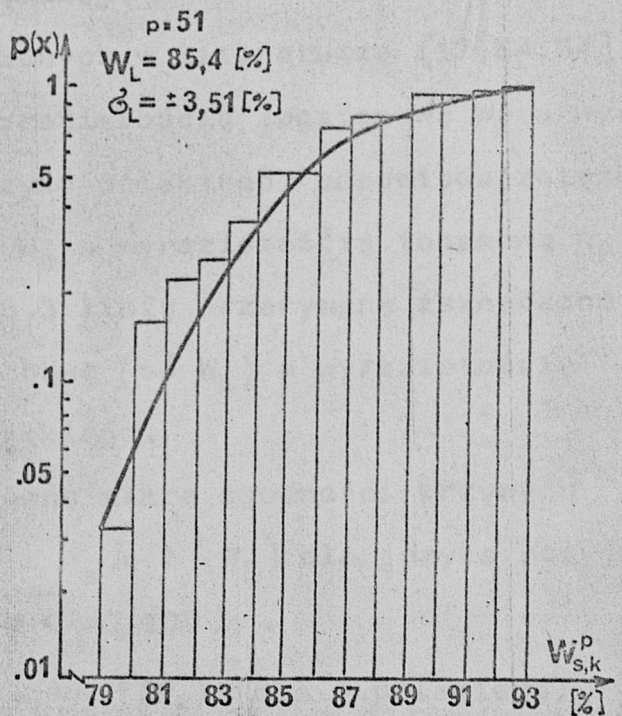
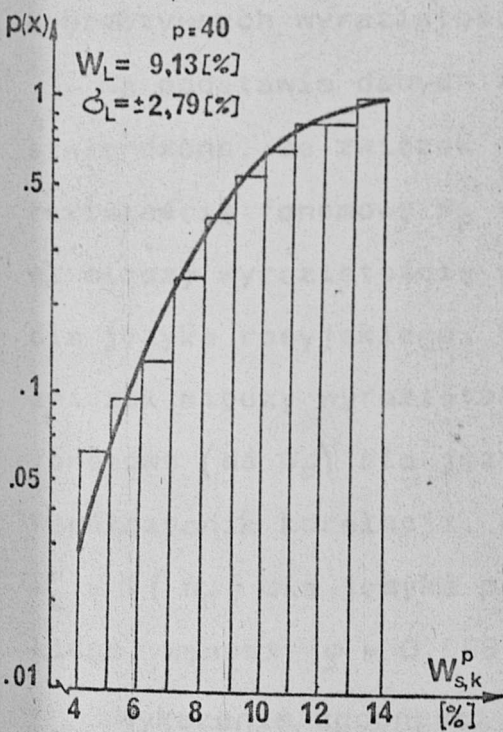
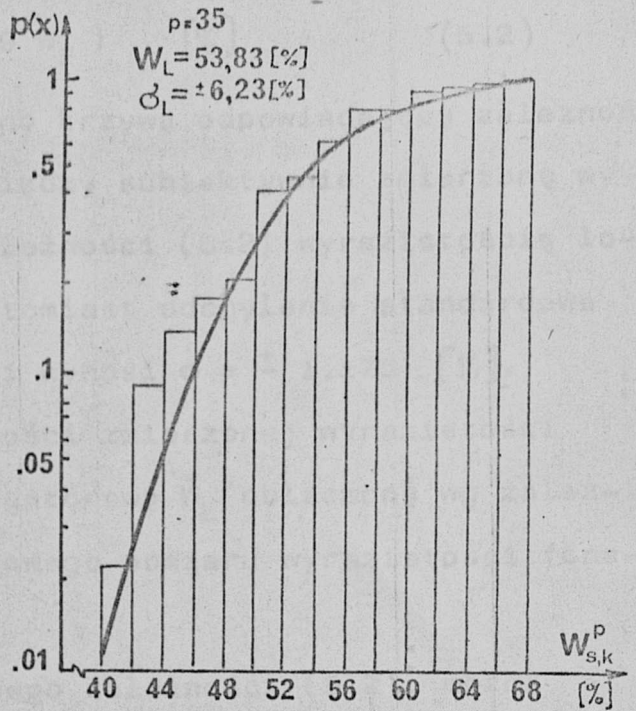
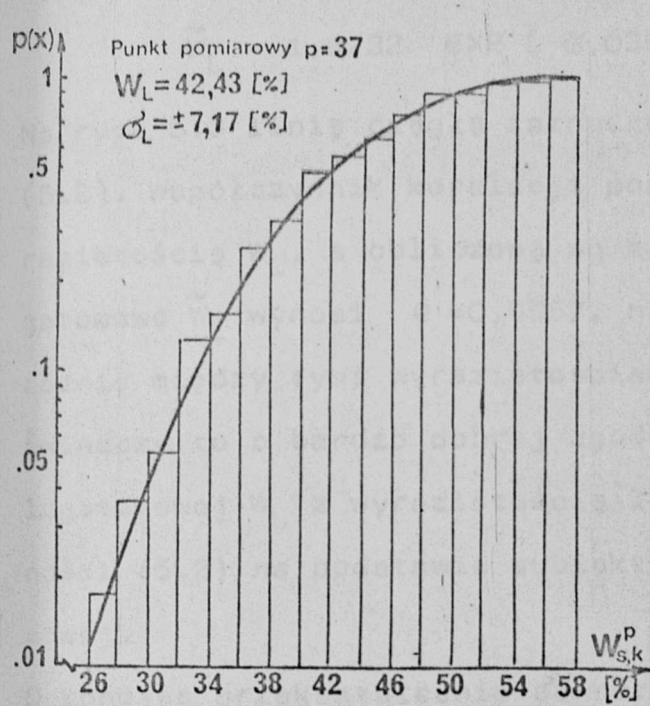
gdzie:

$W_{s,k}^p$ - wyrazistość logatomowa w [%], będąca miarą jakości transmisji sygnału mowy dla p-tego punktu pomiarowego, uzyskana przez odsłuch k-tej listy logatomowej przez s-tego słuchacza.

Obliczone wartości wyrazistości logatomowej, określające jakość transmisji sygnału mowy dla wszystkich 95 punktów pomiarowych, zestawiono w kolumnie 8 tabeli 5.1.

Dla kilku (4) wybranych losowo punktów pomiarowych porównano rozkład $W_{s,k}^p$ z rozkładem normalnym. Zgodność otrzymanego rozkładu $W_{s,k}^p$ (wyrazistości logatomowej po słuchaczach i listach logatomowych) z rozkładem normalnym sprawdzono za pomocą testu Kołmogorowa-Smirnowa [25]. Stwierdzono, że na poziomie istotności $\alpha = 0,05$ nie ma podstaw do odrzucenia hipotezy o zgodności rozkładów. Uzasadnione jest więc stosowanie wartości średniej wyrazistości logatomowej W_L^p jako estymatora wyrazistości logatomowej dla p-tego punktu pomiarowego. Na rys. 5.2 przedstawiono, dla czterech punktów pomiarowych, krzywą schodkową rozkładu $W_{s,k}^p$; linią ciągłą zaznaczono dystrybuantę teoretyczną rozkładu normalnego (wartość średnia $\mu = W_L^p$, odchylenie standardowe $\sigma = \sigma_L^p$).

W wielu praktycznych zastosowaniach do określania jakości kanału telefonicznego korzysta się z wyrazistości fonemowej, zdefiniowanej jako stosunek liczby poprawnie odebranych fonemów do liczby fonemów nadanych (w procentach). W związku z tym, dla wszystkich punktów pomiarowych określono również wyrazistość fonemową. Punkty leżące na płaszczyźnie wyrazistość logatomowa W_L -



Rys. 5.2. Krzywa schodkowa rozkładu $W_{s,k}^P$ oraz dystrybuanta teoretyczna rozkładu normalnego.

wyrazistość fonemowa W_F (rys. 5.3) aproksymowano funkcję eksponen-
cjalną, uzyskując zależność:

$$\tilde{W}_L = 1,8732 \text{ EXP} (0,0399 \cdot W_F) \quad [\%] \quad (5.2)$$

Na rys. 5.3 linię ciągłą zaznaczono krzywą odpowiadającą zależności
(5.2). Współczynnik korelacji pomiędzy subiektywnie zmierzoną wy-
razistością W_L , a obliczoną wg zależności (5.2) wyrazistością lo-
gatomową \tilde{W}_L wynosi $\rho = 0,9957$, natomiast odchylenie standardowe
różnic między tymi wyrazistościami wynosi $\sigma = \pm 1,175$ [%].

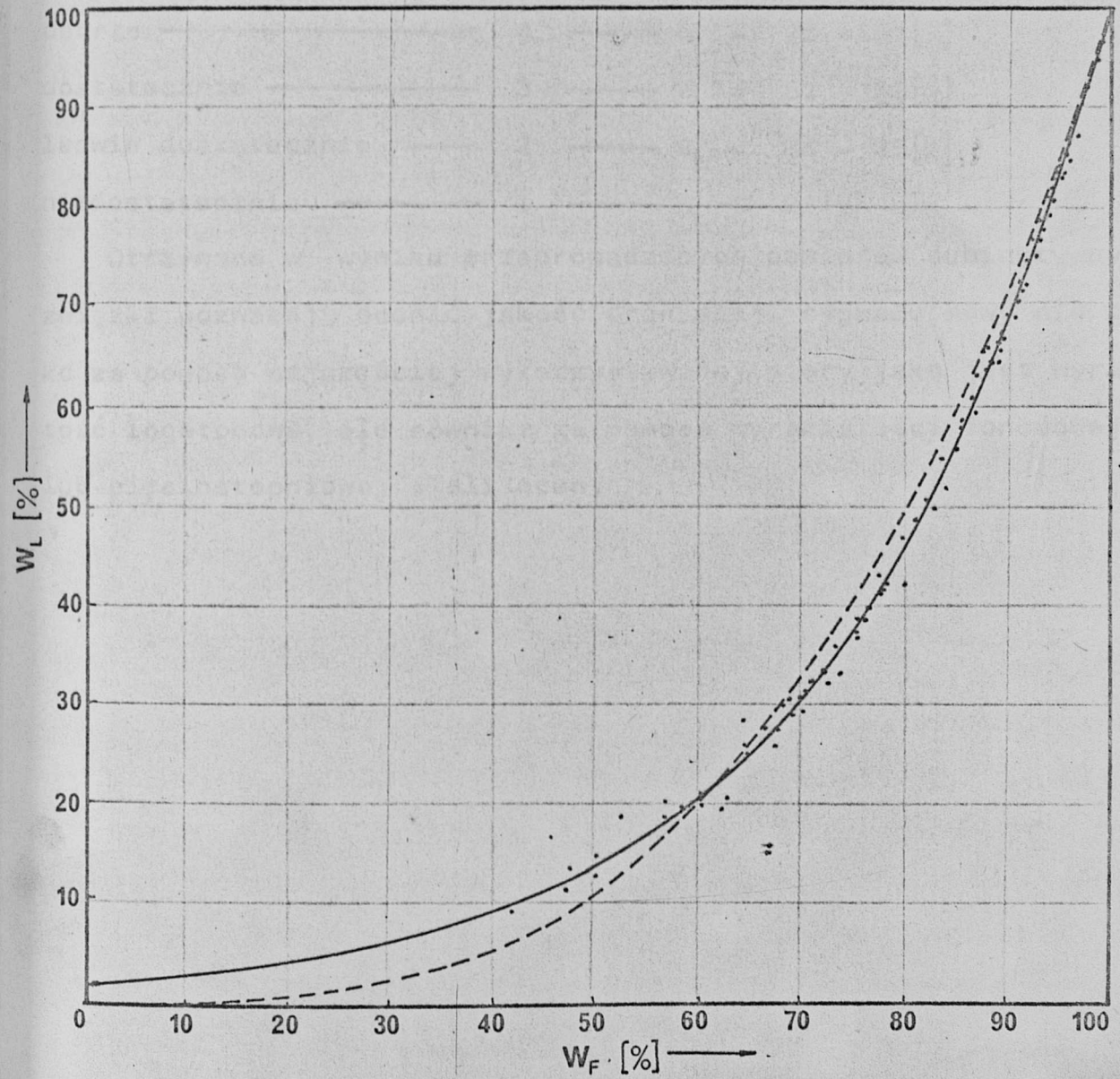
Świadczy to o bardzo dobrej zgodności zmierzonej wyrazistości
logatomowej W_L z wyrazistością logatomową \tilde{W}_L obliczoną wg zależ-
ności (5.2) na podstawie subiektywnego pomiaru wyrazistości fone-
mowej.

Dokonując przekształcenia odwrotnego zależności (5.2), można
obliczyć wyrazistość fonemową na podstawie otrzymanej w pomiarach
subiektywnych wyrazistości logatomowej.

Na podstawie danych zamieszczonych w literaturze [17,54,64]
stwierdzono, że związek między wyrazistością logatomową W_L a wy-
razistością fonemową W_F - dla języka polskiego, odpowiada związko-
wi między wyrazistością sylabową W_S a wyrazistością fonemową W_F
dla języka rosyjskiego. Na rys. 5.3 linią przerywaną zaznaczono
związek między wyrazistością sylabową (oś W_L) a wyrazistością
fonemową (oś W_F) dla języka rosyjskiego.

Współczynnik korelacji, będący pewną miarą zgodności krzywych
 $W_L = f(W_F)$ dla języka polskiego i $W_S = f(W_F)$ dla języka rosyjs-
kiego, wynosi $\rho = 0,9987$ dla $W_F \in \langle 40,100 \rangle$.

Wykazanie zgodności krzywej $W_L = f(W_F)$ (dla j. polskiego)
z krzywą $W_S = f(W_F)$ (dla j. rosyjskiego) pozwala na wykorzysta-
nie, wprowadzonej przez Bykova [2] dla języka rosyjskiego pię-



Rys.5.3 Związek między wyrazistością logatomową a wyrazistością fonemową dla języka polskiego (linia ciągła) oraz między wyrazistością sylabową a wyrazistością fonemową dla języka rosyjskiego (linia przerywana).

ciostopniowej skali ocen jakości transmisji sygnału mowy do oceny jakościowej sygnału mowy polskiej przesyłanego kanałem telefonicznym.

Korzystając z podanych w literaturze [17,54,64] danych dla języka rosyjskiego, otrzymano następujące związki między pięciostopniową skalą oceny jakości transmisji sygnału mowy polskiej, a odpowiednimi przedziałami wyrazistości logatomowej W_L :

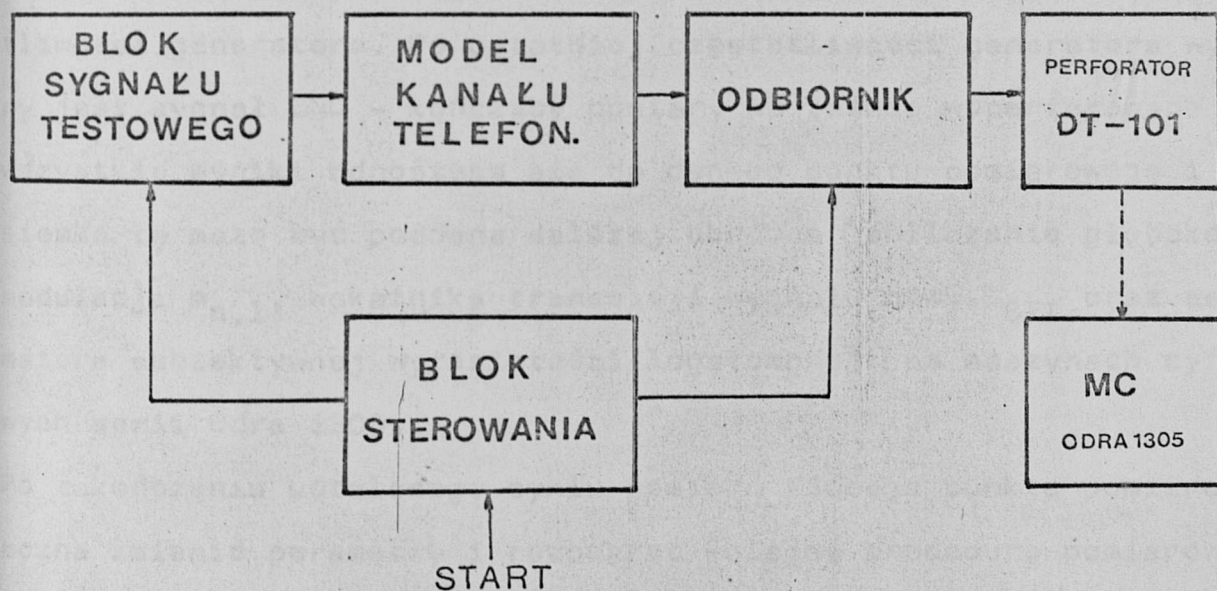
bardzo dobrze	—————	5	—————	$W_L = 100 - 73$ [%]
dobrze	—————	4	—————	$W_L = 73 - 51$ [%]
dostatecznie	—————	3	—————	$W_L = 51 - 38$ [%]
ledwie dostatecznie	—————	2	—————	$W_L = 38 - 25$ [%]
niedostatecznie	—————	1	—————	$W_L < 25$ [%]

Otrzymane w wyniku przeprowadzonych pomiarów subiektywnych związki pozwalają ocenić jakość transmisji sygnału mowy nie tylko za pomocą najczęściej wykorzystywanej miary jaką jest wyrazistość logatomowa, ale również za pomocą wyrazistości fonemowej lub pięciostopniowej skali ocen.

6. POMIAR WSKAŹNIKA TRANSMISJI SYGNAŁU MOWY METODĄ MTF

6.1. Metoda pomiaru

Pomiary wskaźnika transmisji sygnału mowy zostały wykonane za pomocą systemu pomiarowego MTF omówionego w rozdziale 4. W zrealizowanych obiektywnych pomiarach wskaźnika transmisji sygnału mowy, wyniki pośrednie (postać cyfrowa obwiedni $y_{n,1}(t)$ (zal. 2.2) sygnału testowego na wyjściu badanego kanału telefonicznego) wyprowadzono na taśmie papierowej wykorzystując perforator DT-101. Powyższa procedura wyprowadzania wyników pośrednich umożliwia przeprowadzenie redukcji wymiarów w metodzie MTF. Pomiary wskaźnika transmisji sygnału mowy wykonano w układzie przedstawionym na rys. 6.1.



6.1. Schemat blokowy systemu realizującego pomiar wskaźnika transmisji sygnału mowy.

Pomiar rozpoczyna się od chwili naciśnięcia przycisku "START". Naciśnięcie przycisku "START" powoduje wysłanie sygnału uruchamiającego silnik perforatora, włączającego pierwszy filtr pasmowy oraz pierwszą częstotliwość generatora sinusoidalnego.

W tym momencie sygnał testowy wygenerowany z bloku sygnału testowego zostaje podany na model kanału telefonicznego.

W odbiorniku włączony jest również pierwszy filtr pasmowy i sygnał transmitowany przez badany kanał telefoniczny zostaje analizowany w tym pasmie. Po dziesięciu okresach sygnału modulującego uruchamiany jest cykl perforowania taśmy papierowej.

Przez jedenasty okres sygnału modulującego trwa sygnał zmiany filtru. Cykl 11 okresów powtarza się dla pozostałych filtrów.

Po ostatnim sygnale zmiany filtru trwa sygnał zmiany częstotliwości generatora. Sygnał ten trwa przez jeden okres sygnału zmiany filtru. Cykl zmiany filtrów powtarzany jest dla kolejnych częstotliwości generatora. Po ostatniej częstotliwości generatora wysłany jest sygnał END - kończący pomiar. Na taśmie wyperforowane są wszystkie wyniki odnoszące się do danego punktu pomiarowego i taśmka ta może być poddana dalszej obróbce (obliczanie głębokości modulacji $m_{n,1}$, wskaźnika transmisji sygnału mowy W_{STI} oraz estymatora subiektywnej wyrazistości logatomowej) na maszynach cyfrowych serii Odra 1300.

Po zakończeniu ustalonego cyklu pomiaru jednego punktu pomiarowego można zmienić parametry i rozpocząć kolejną procedurę pomiarową.

6.2. Obliczanie wskaźnika transmisji sygnału mowy

6.2.1. Wprowadzenie.

Do obliczenia wskaźnika transmisji sygnału mowy zgodnie z zależnością (2.9) wymagane jest dokonanie stu pięciu pomiarów głębokości modulacji. Zarówno analiza fizycznej strony pomiaru metodą MTF jak i doświadczenia twórców metody - Houtgasta i Steenekena [33,74] wskazują na konieczność dokonania redukcji tej części wymiarów przestrzeni pomiarowej MTF, które nie są skorelowane lub są słabo skorelowane ze zjawiskami wpływającymi na jakość transmisji sygnału mowy. Redukcja wymiaru jest korzystna również ze względów technicznych (uproszczenie aparatury, skrócenie czasu pomiaru). Houtgast i Steeneken [74] w ostatecznej standardowej wersji MTF zredukowali liczbę częstotliwości modulujących do 14 a liczbę pasm oktawowych do 7; nie podali jednak metody i kryteriów redukcji. W związku z powyższym oraz kierując się założeniem o dostosowaniu metody MTF do sygnału mowy polskiej zdecydowano w niniejszej pracy przeprowadzić procedurę redukcji wymiarów metody MTF w oparciu o badania przeprowadzone dla mowy polskiej i na tej podstawie ustalić liczbę pasm częstotliwości i częstotliwości modulujących wraz z podaniem ich wartości oraz wartości współczynników wagowych występujących w zależności (2.9).

6.2.2. Wstępne ustalenie parametrów w metodzie MTF.

W oryginalnej i niezredukowanej wersji metody MTF Hougast i Steeneken stosowali 7 pasm oktawowych i 21 częstotliwości modulujących wybranych w oparciu o nierównomierną dyskretyzację uśrednionego widma obwiedni [33,74]. Bezwymiarowe współczynniki wagowe α pasm oktawowych dobrano na podstawie wkładu danego pasma w sumaryczną wyrazistość zgodnie z teorią i badaniami praktycznymi

nad wyrazistością w pasmach częstotliwości [23,74]. Bezwymiarowy współczynnik wagowy γ , uwzględniający maskujący wpływ sąsiednich pasm częstotliwości, dobrano w wersji oryginalnej w oparciu o badania Frencha i Steinberga [23,74] (szczegółowe omówienie fizycznej interpretacji współczynnika wagowego γ zawarto w rozdziale 2.2).

W niniejszej pracy, przy badaniu oryginalnej i zmodyfikowanej wersji metody MTF, współczynniki wagowe α dobrano w oparciu o badania Majewskiego i Żalewskiego [44] nad wskaźnikiem wyrazistości mowy polskiej w pasmach częstotliwości, natomiast współczynnik γ przyjęto identyczny jak w metodzie Houtgasta i Steenekena [74]. Wartości wstępne wymienionych współczynników α przedstawiono w tabeli 6.1 natomiast współczynnika γ w rozdziale 6.2.6.

6.2.3. Metoda redukcji wymiarów w metodzie MTF

Do redukcji wymiarów zastosowano znany z teorii optymalizacji program dynamiczny, sekwencyjny, pojedynczy [24,55].

Jako funkcję celu Z przyjęto maksymalizację współczynnika korelacji rang Spearmana r_s pomiędzy wskaźnikiem transmisji sygnału mowy W_{STI} a wyrazistością logatomową W_L otrzymaną z pomiarów subiektywnych (patrz rozdz. 5.3).

Niech h oznacza funkcję przekształcającą P -elementowy zbiór $\{W_{STI}\}$ na zbiór $\{h(W_{STI})\}$ w ten sposób, że $h(W_{STI}^i) < h(W_{STI}^j)$ gdy $W_{STI}^i > W_{STI}^j$ ($i, j = 1, 2, \dots, P$). Niech g oznacza funkcję przekształcającą P -elementowy zbiór $\{W_L\}$ na zbiór $\{g(W_L)\}$ w ten sposób, że $g(W_L^i) < g(W_L^j)$ gdy $(W_L^i) > (W_L^j)$. Inaczej mówiąc numeruje się elementy każdego zbioru według wielkości od najwyższej do najniższej (można również szeregować odwrotnie).

Współczynnik korelacji rang Spearmana jest określony wtedy zależnością:

$$r_s = 1 - \frac{6 \sum_{i=1}^P d_i^2}{P (P^2 - 1)} \quad (5.1)$$

przy czym :

$d_i = f(W_{STI}^i) - g(W_L^i)$ - i-ta różnica pomiędzy i-tą rangą
cech W_{STI} i W_L ,

P - liczba elementów w zbiorze $\{W_{STI}\}$ oraz $\{W_L\}$.

Następnie testuje się hipotezę H_0 , że korelacja między dwoma cechami jest nieistotna. W przypadku, gdy $P \geq 10$ hipotezę H_0 można weryfikować za pomocą statystyki t - Studenta [63,81].

W przeprowadzonych badaniach nad redukcją wymiarów w metodzie MTF wykorzystywano wyniki pomiarów subiektywnych (rozdz. 5.3) oraz wyniki obiektywnych pomiarów wskaźnika transmisji sygnału mowy wykonanych z wykorzystaniem omówionego w rozdziale 4 systemu pomiarowego MTF w wersji oryginalnej i zmodyfikowanej.

Zgodnie z zaleceniami CCITT [95,96], łącze telefoniczne telefonii naturalnej powinno posiadać pasmo 300-3600 Hz.

Podzielono więc pasmo telefoniczne na $L=5$ pasm oktaowych o częstotliwościach środkowych 250 Hz, 500 Hz, 1 kHz, 2 kHz i 4 kHz.

W każdym pasmie oktaowym przeprowadzono pomiar głębokości modulacji dla $N=21$ częstotliwości sygnału modulującego (wartości podano w tabeli 4.1). Uzyskano w ten sposób macierz M o wymiarach $N \times L$, tzn. 21×5 , której elementami były wartości głębokości modulacji $m_{n,1}$. Na podstawie wartości elementów macierzy M , obliczono - wg zależności (2.9) - wskaźnik transmisji sygnału mowy W_{STI} dla poszczególnych punktów pomiarowych.

6.2.4. Ustalenie liczby częstotliwości modulujących.

Kryterium minimalizacji liczby częstotliwości modulujących N_1 można sformułować następująco:

$$Z : \{ r_s = r_{s \max} \longrightarrow N_1 = N_{1 \min} \} , \quad (6.2)$$

gdzie:

N_1 - liczba częstotliwości modulujących f_n w l -tym pasmie

Dodatkowym warunkiem jest, aby:

$$\bigwedge_l f_{n,l} = \text{const} , \quad (6.3)$$

gdzie:

l - indeks w zbiorze pasm oktawowych,

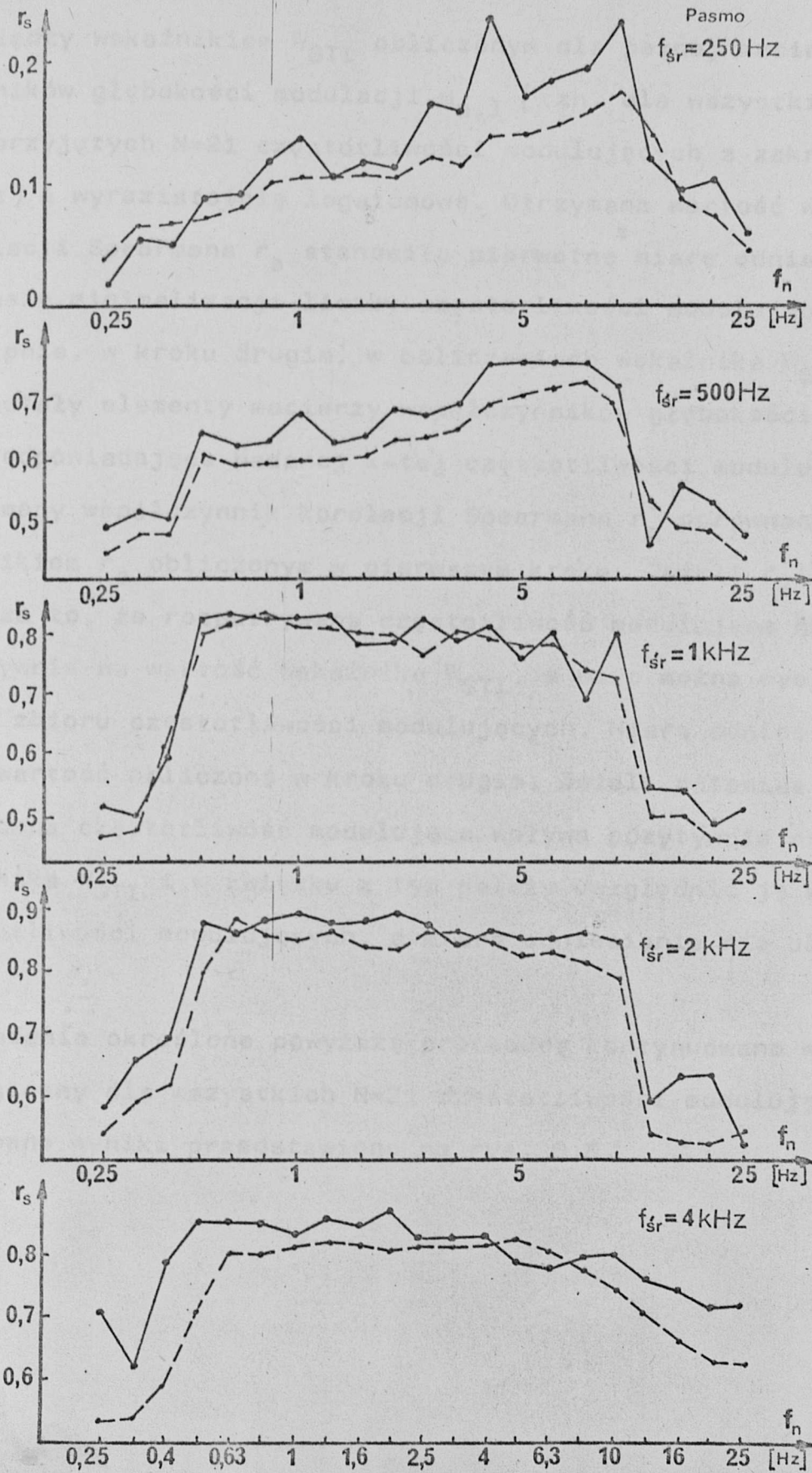
n - indeks w zbiorze częstotliwości modulujących.

Dla każdej częstotliwości modulującej i dla każdego pasma oktawowego przeprowadzono, za pomocą współczynników korelacji rang Spearmana r_s , analizę stopnia skorelowania współczynnika głębokości modulacji z subiektywnie zmierzoną wyrazistością logatomową. Przebieg współczynnika korelacji rang Spearmana r_s dla każdego pasma w zależności od częstotliwości modulującej f_n przedstawiono na rys. 6.2.

Kolejnym krokiem było dokonanie obliczeń wskaźnika W_{STI} w wersjach różniących się liczbą elementów macierzy współczynnika głębokości modulacji $m_{n,l}$. Obliczeń dokonano przy założeniu równomiernego wkładu wszystkich pasm oktawowych do wskaźnika W_{STI} , tzn przy $\alpha_l = \text{const}$ ($l = 1, 2, \dots, 5$).

Zasadę programu minimalizacji częstotliwości modulującej przedstawiono poniżej.

W pierwszym kroku otrzymano współczynnik korelacji Spearmana



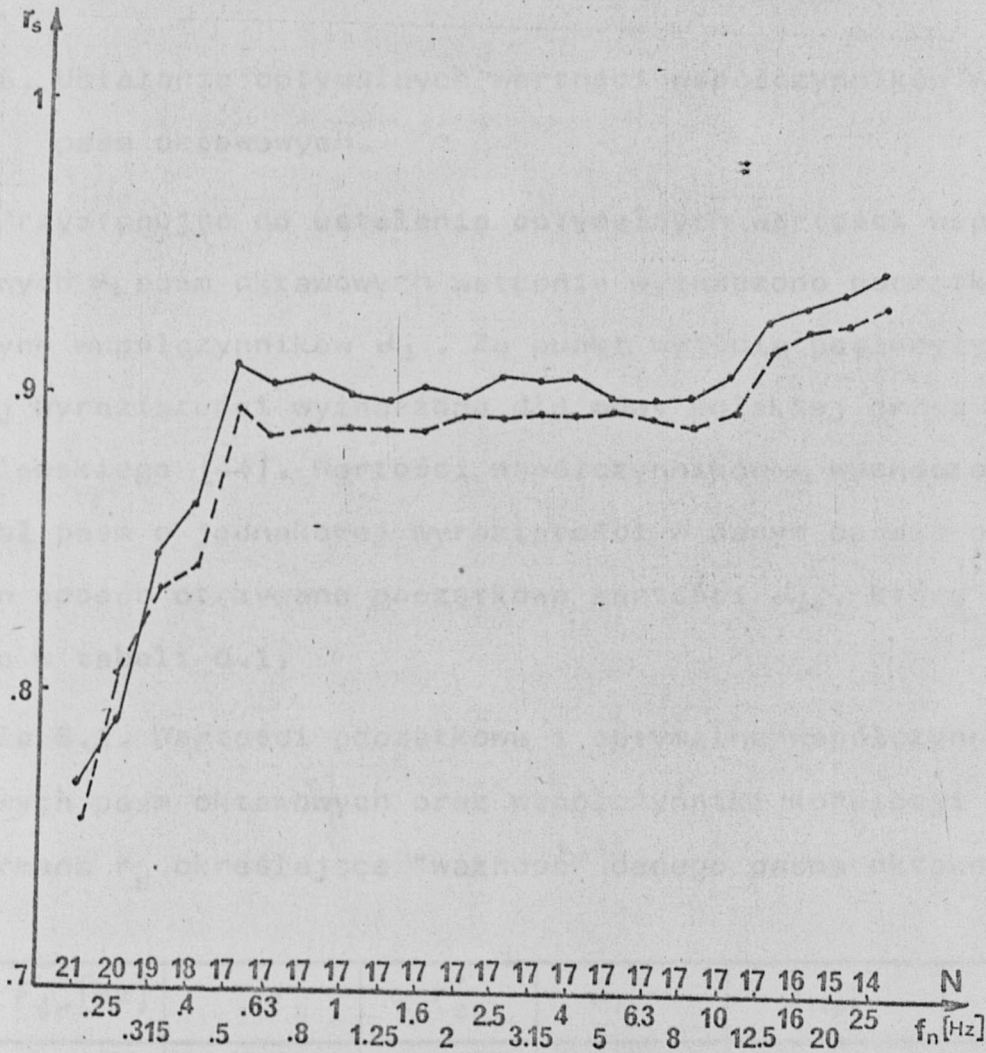
Rys. 6.2. Przebieg współczynnika r_s w funkcji częstotliwości modulującej f_n dla poszczególnych pasm oktaowych (--- wersja oryginalna, — wersja zmodyfikowana).

r_s między wskaźnikiem W_{STI} obliczonym dla pełnej macierzy współczynników głębokości modulacji $m_{n,1}$ (tzn. dla wszystkich wstępnie przyjętych $N=21$ częstotliwości modulujących z zakresu 0,25 Hz - 25 Hz) a wyrazistością logatomową. Otrzymana wartość współczynnika korelacji Spearmana r_s stanowiła pierwotną miarę odniesienia w procesie minimalizacji liczby częstotliwości modulujących.

Następnie, w kroku drugim, w obliczeniach wskaźnika W_{STI} pominięte zostały elementy macierzy współczynników głębokości modulacji $m_{n,1}$ odpowiadające badanej k -tej częstotliwości modulującej. Otrzymany współczynnik korelacji Spearmana r'_s porównano ze współczynnikiem r_s obliczonym w pierwszym kroku. Jeżeli $r'_s \geq r_s$, oznacza to, że rozpatrywana częstotliwość modulująca nie wpływa pozytywnie na wartość wskaźnika W_{STI} , a więc można wyeliminować ją ze zbioru częstotliwości modulujących. Miara odniesienia przyjmuje wartość obliczoną w kroku drugim. Jeżeli natomiast $r'_s < r_s$, to badana częstotliwość modulująca wpływa pozytywnie na wartość wskaźnika W_{STI} i w związku z tym należy uwzględnić ją w zbiorze częstotliwości modulujących, a miara odniesienia nie ulega zmianie.

Badania określone powyższą procedurą kontynuowano w sposób analogiczny dla wszystkich $N=21$ częstotliwości modulujących.

Otrzymane wyniki przedstawiono na rys. 6.3.



Rys.6.3. Przebieg współczynnika r_s w funkcji eliminowanej częstotliwości modulującej f_n oraz liczby częstotliwości modulujących N , przy równomiernym wkładzie każdego pasma oktawowego do wskaźnika W_{STI} (--- wersja oryginalna, — wersja zmodyfikowana).

Analiza wyników przedstawionych graficznie na rys. 6.3 wykazała, że w obydwu rozważanych przypadkach jest możliwe ograniczenie liczby częstotliwości modulujących do $N=14$. Wartości częstotliwości modulujących odpowiadają wówczas częstotliwościom środkowym pasm tercjowych z zakresu 0,5 Hz - 10 Hz. Po przeprowadzonej w ten sposób minimalizacji liczby częstotliwości modulujących przystąpiono do ustalenia optymalnych wartości współczynników

wagowych α_1 pasm oktawowych, przy czym obliczenia wykonano dla ostatecznie ustalonej liczby $N=14$ częstotliwości modulujących.

6.2.5. Ustalanie optymalnych wartości współczynników wagowych pasm oktawowych.

Przystępując do ustalania optymalnych wartości współczynników wagowych α_1 pasm oktawowych wstępnie wyznaczono początkowe wartości tych współczynników α_1 . Za punkt wyjścia posłużyły pasma jednakowej wyrazistości wyznaczone dla mowy polskiej przez Majewskiego i Zalewskiego [44]. Wartości współczynników α_1 wyznaczono badając udział pasm o jednakowej wyrazistości w danym pasmie oktawowym. W ten sposób otrzymano początkowe wartości α_{10} , które zamieszczono w tabeli 6.1.

Tabela 6.1. Wartości początkowe i optymalne współczynników wagowych pasm oktawowych oraz współczynniki korelacji rang Spearmana r_s określające "ważność" danego pasma oktawowego

Lp	F_{sr} [Hz]	r_s	r'_s	α_{10}	α_{1opt}	α'_{1opt}
1	250	0,1135	0,1501	0,050	0	0
2	500	0,5992	0,6181	0,150	0,32	0,080
3	1000	0,8014	0,8314	0,200	0,30	0,200
4	2000	0,8780	0,8926	0,325	0,15	0,480
5	4000	0,8103	0,8690	0,125	0,20	0,125

F_{sr} - częstotliwość środkowa pasma oktawowego,

α_{10} - początkowa wartość współczynnika wagowego pasm oktawowych,

α_{1opt} - optymalna wartość współczynnika wagowego pasm oktawowych (oryginalna wersja MTF),

α'_{1opt} - optymalna wartość współczynnika wagowego pasm oktawowych (zmodyfikowana wersja MTF),

r_s - współczynnik korelacji rang Spearmana (oryginalna wersja MTF),

r'_s - współczynnik korelacji rang Spearmana (zmodyfikowana wersja MTF).

W następnym kroku dokonano hierarchizacji pasm, którą przeprowadzono w oparciu o współczynnik korelacji rang Spearmana r_s między subiektywną wyrazistością logatomową, a wskaźnikiem W_{mtf} obliczonym dla danego pasma oktawowego według zależności (2.8). Wartości współczynników rang Spearmana r_s otrzymane dla poszczególnych pasm oktawowych zamieszczono w tabeli 6.1.

Z otrzymanych rezultatów wynika, że w obydwu wersjach pomiarowych kolejność "ważności" pasm oktawowych jest następująca: 2000 Hz, 4000 Hz, 1000 Hz, 500 Hz i 250 Hz.

Po wyznaczeniu początkowych wartości współczynników wagowych α_1 i przeprowadzeniu hierarchizacji pasm oktawowych przystąpiono do wyznaczenia optymalnych wartości współczynników wagowych α_1 .

W wyniku przeprowadzonych obliczeń z wykorzystaniem optymalizacyjnego programu dynamicznego, sekwencyjnego, pojedynczego [11, 24, 55] uzyskano wartości współczynników wagowych α_1 pasm oktawowych, które zamieszczono w tabeli 6.1. Po wyznaczeniu wartości współczynników wagowych α_1 , przystąpiono do ustalenia wartości współczynnika maskowania γ .

6.2.6. Ustalenie optymalnej wartości współczynnika maskowania

W procesie optymalizacji wartości współczynników wagowych α_1 pasm oktawowych oraz liczby N częstotliwości modulujących, do obliczenia wskaźnika W_{STI} transmisji sygnału mowy przyjęto wartość $\gamma = 0,0003$. Odpowiada to wartości współczynnika maskowania wyznaczonej przez Houtgasta i Steenekena w przypadku wykorzystywania metody MTF do oceny jakości transmisji sygnału mowy w pomieszczeniach [74].

Kryterium optymalizacji było osiągnięcie maksymalnej wartości współczynnika korelacji rang Spearmana r_s między wyrazistością

logatomową a wskaźnikiem W_{STI} obliczonym z uwzględnieniem badanej wartości γ . Wyjściową miarą odniesienia była w wersji zmodyfikowanej wartość współczynnika korelacji rang Spearmana $r_s = 0,9895$ między wyrazistością logatomową a wskaźnikiem W_{STI} obliczonym w warunkach optymalnej liczby częstotliwości modulujących, optymalnej wartości współczynników wagowych pasm oktaowych d_1 oraz $\gamma = 0,0003$, natomiast w wersji oryginalnej $r_s = 0,9630$.

Współczynnik korelacji r_s^i obliczony dla danego γ^i porównywano z miarą odniesienia, tzn. r_s^0 . Jeżeli $r_s^i \leq r_s^0$, to wpływ danej wartości γ^i na wskaźnik W_{STI} jest negatywny w porównaniu z wpływem wartości γ^0 i miara odniesienia nie ulega zmianie. Jeżeli natomiast $r_s^i > r_s^0$, to miara odniesienia przyjmuje wartość r_s^i , natomiast współczynnik maskowania przyjmuje wartość γ^i . Obliczenia wykonano dla $\gamma \in \langle 0,1, 0,00001 \rangle$ ze skokiem $\Delta\gamma = 0,00001$. W wersji oryginalnej dla $\gamma \in \langle 0,0002, 0,00038 \rangle$ wartość współczynnika korelacji rang Spearmana nie ulegała zmianie i wynosiła $r_s = 0,9630$, natomiast dla pozostałych γ współczynnik r_s był mniejszy od miary odniesienia.

W wersji zmodyfikowanej natomiast dla $\gamma \in \langle 0,00028, 0,00035 \rangle$ wartość współczynnika korelacji rang Spearmana nie ulegała zmianie i wynosiła $r_s = 0,9895$, natomiast dla pozostałych γ współczynnik r_s był mniejszy od miary odniesienia.

W związku z tym uznano, że wyznaczona przez Houtgasta i Steenekena dla języka holenderskiego wartość współczynnika maskowania $\gamma = 0,0003$ jest wartością optymalną również dla transmisji sygnału mowy polskiej kanałem telefonicznym.

6.2.7. Parametry systemu pomiarowego MTF dla sygnału mowy polskiej

W wyniku przeprowadzonych badań nad redukcją wymiarów oraz ustaleniem optymalnych wartości współczynników d i γ w zależ-

ności na W_{STI} (2.9) uzyskano następujące podstawowe parametry systemu pomiarowego MTF w wersji oryginalnej i zmodyfikowanej:

1. wartości częstotliwości sygnału modulującego:

{0,5; 0,63; 0,8; 1; 1,15; 1,6; 2; 2,5; 3,15; 4; 5; 6,3; 8; 10} Hz

2. wartości częstotliwości środkowych pasm oktaowych, w których sygnał szumowy jest modulowany amplitudowo sygnałem sinusoidalnym oraz w których jest analizowany sygnał testowy na wyjściu badanego kanału telefonicznego:

{500, 1000, 2000, 4000} Hz

3. głębokość modulacji sygnału testowego na wejściu kanału telefonicznego - 100 %

4. poziom napięcia sygnału testowego podawanego na wejście modelu kanału telefonicznego - 0 dB,

5. czas trwania sygnału testowego - 513 s

6. wartość współczynników wagowych pasm oktaowych:

6.1. dla systemu pomiarowego MTF w wersji oryginalnej

$\alpha_1 = 0,32$; $\alpha_2 = 0,3$; $\alpha_3 = 0,15$; $\alpha_4 = 0,2$

6.2. dla systemu pomiarowego MTF w wersji zmodyfikowanej

$\alpha_1 = 0,08$; $\alpha_2 = 0,2$; $\alpha_3 = 0,48$; $\alpha_4 = 0,125$

7. wartość współczynnika maskowania $\gamma = 0,0003$.

8. czas pomiaru jednego punktu pomiarowego w obu wersjach wynosi ok. 12 minut /nie licząc czasu potrzebnego do przygotowania systemu do pracy i obliczeń na maszynie cyfrowej/.

6.3. Ocena oryginalnej i zmodyfikowanej metody MTF

Otrzymany zbiór odpowiadających sobie wartości subiektywnie pomierzonych wyrazistości logatomowych W_L i zmierzonych w sposób obiektywny wskaźników transmisji sygnału mowy W_{STI} dla wszystkich

punktów pomiarowych przedstawiono w tabeli 6.2 oraz graficznie na rys. 6.4 (wersja oryginalna) i rys. 6.5 (wersja zmodyfikowana).

Punkty leżące na płaszczyźnie: wyrazistość logatomowa W_L - wskaźnik transmisji sygnału mowy W_{STI} aproksymowano, stosując metodę najmniejszych kwadratów [62], wielomianem czwartego rzędu uzyskując zależność umożliwiającą bezpośrednie obliczenie estymatora wyrazistości logatomowej \hat{W}_L na podstawie pomiaru W_{STI} :

1. dla wersji oryginalnej MTF

$$\hat{W}_L = -901,29 W_{STI}^4 + 2784,65 W_{STI}^3 - 3280,66 W_{STI}^2 + 1841,62 W_{STI} - 347,43 \quad [\%] \quad (6.4)$$

2. dla wersji zmodyfikowanej MTF

$$\hat{W}_L = -1447 W_{STI}^4 + 4221,74 W_{STI}^3 - 4609,47 W_{STI}^2 + 2308,57 W_{STI} - 378,68 \quad [\%] \quad (6.5)$$

Zależność (6.4) jest słuszna dla $W_{STI} \in \langle 0,36; 1 \rangle$, co odpowiada wyrazistości logatomowej $\hat{W}_L \in \langle 5,2; 96,9 \rangle$ [%]. Zależność (6.5) natomiast jest słuszna dla $W_{STI} \in \langle 0,31; 1 \rangle$, co odpowiada wyrazistości logatomowej $\hat{W}_L \in \langle 6,4; 95,2 \rangle$ [%]. Dla wyrazistości logatomowej $W_L < 8,5$ % stwierdzono, na podstawie pomiarów subiektywnych, przerwę w łączności.

Zależność (6.4) otrzymaną dla metody MTF w wersji oryginalnej wykreślono linią ciągłą na rys. 6.4, a obliczone wartości \hat{W}_L podano w kolumnie 4 tabeli 6.2. Współczynnik korelacji rang Speermana między estymowaną wyrazistością logatomową \hat{W}_L a subiektywnie zmierzoną wyrazistością logatomową W_L wynosi $r_s = 0,9630$, natomiast statystyka t wynosi $t = 9,3369$.

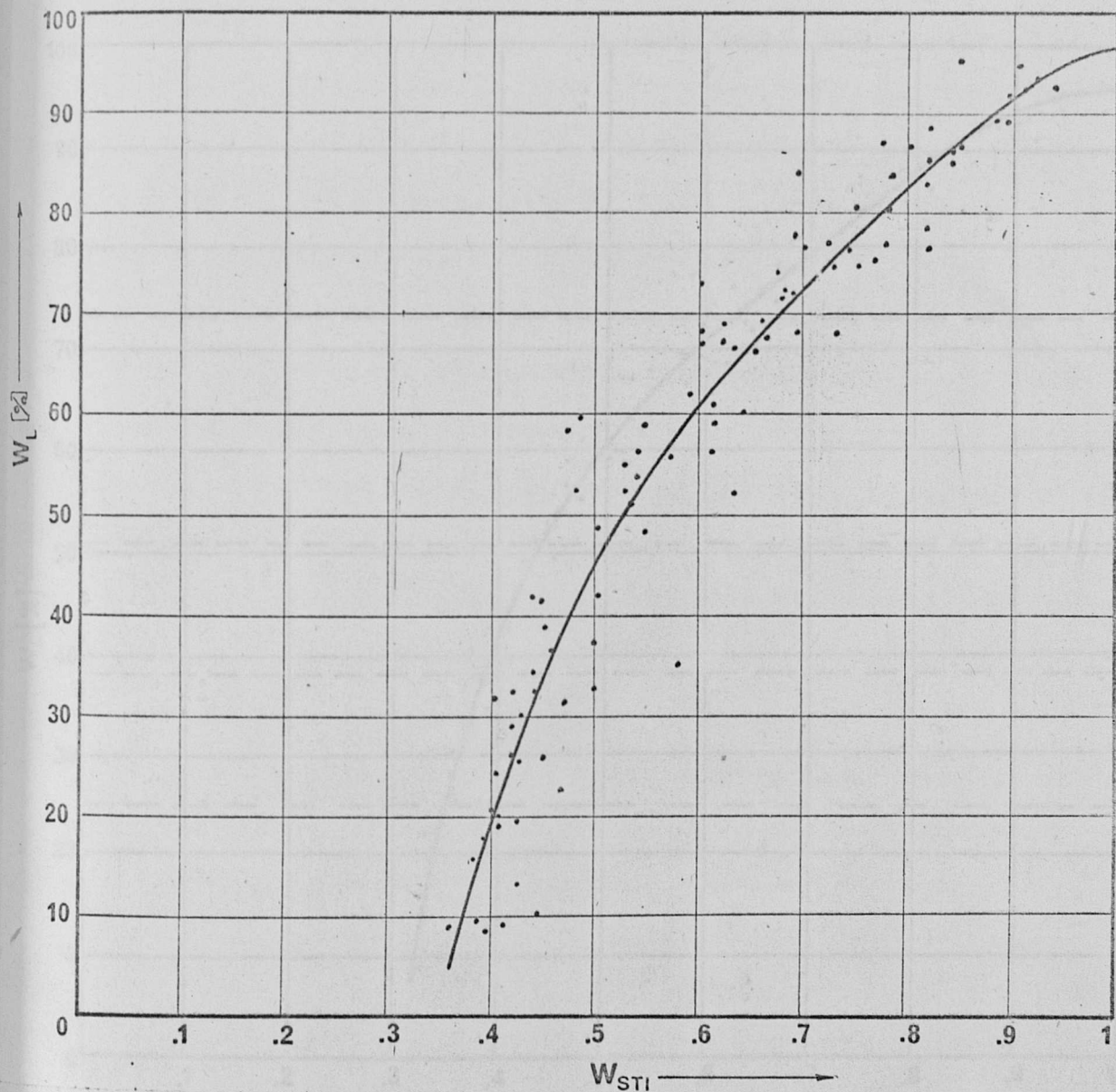
Z tablic statystycznych [63] przy γ stopniach swobody ($\gamma = 93$) odczytana wartość $t_{\alpha=0,001} = 3,50785$, a więc $t > t_{\text{tabl}}$. Hipotezę H_0 o braku korelacji między estymowaną wyrazistością logatomową \hat{W}_L a zmierzoną subiektywnie wyrazistością logatomową W_L można było odrzucić na poziomie istotności $\alpha = 0,001$. Odchylenie standardowe różnic między pomierzoną subiektywnie wyrazistością

Tabela 6.2. Wyniki pomiarów i oszacowań wyrazistości (W_L , \hat{W}_L) i wskaźnika W_{STI} .

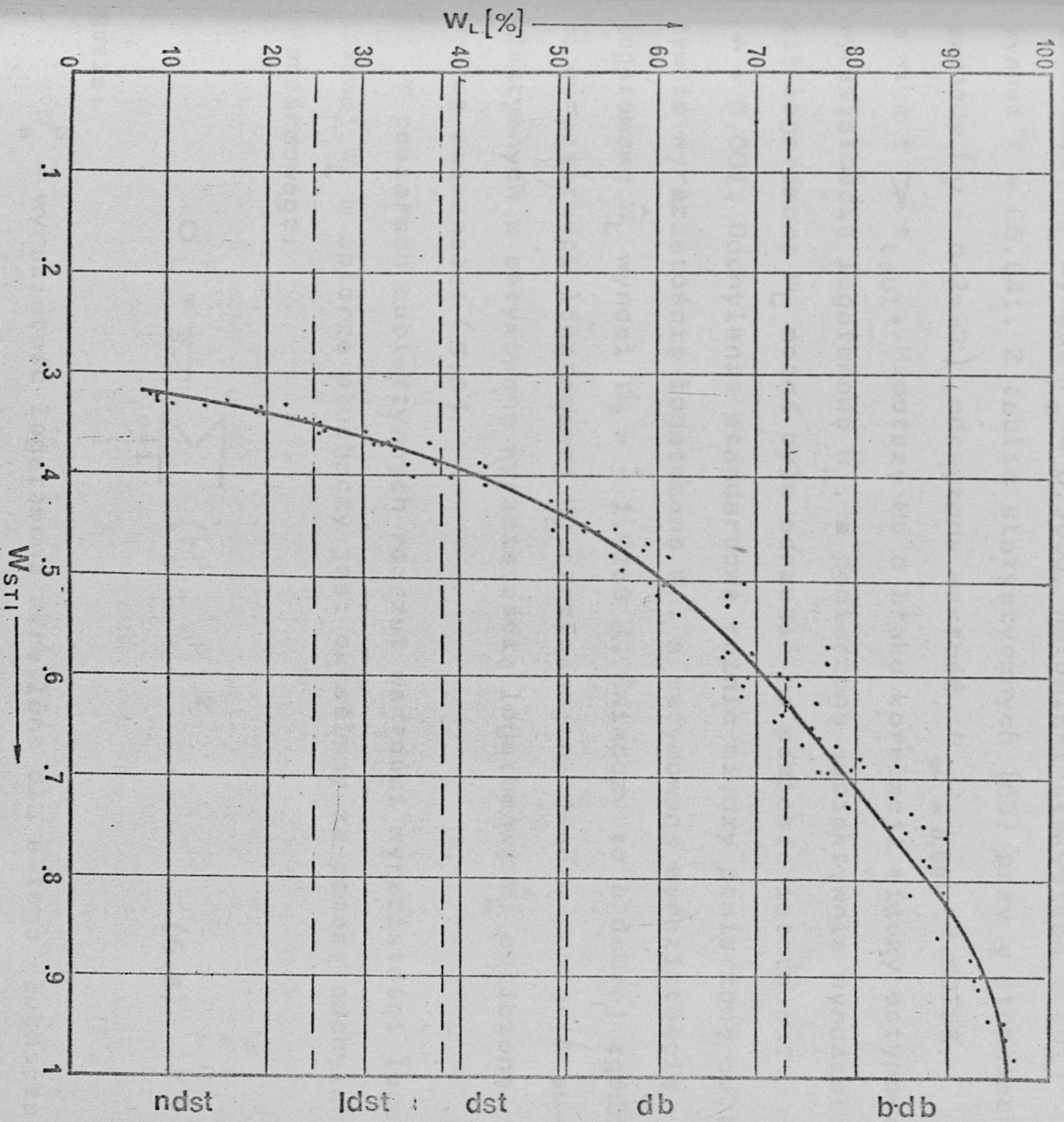
Lp	W_L [%]	wersja oryginalna		wersja zmodyfikowana	
		W_{STI}	\hat{W}_L [%]	W_{STI}	\hat{W}_L [%]
1	2	3	4	5	6
1	96,3	0,99965	96,88	0,98015	95,58
2	67,1	0,61462	63,09	0,52521	63,83
3	31,2	0,47030	38,63	0,37067	31,41
4	22,5	0,46781	38,06	0,33177	16,50
5	9,2	0,41144	23,05	0,33309	17,07
6	76,5	0,81982	84,64	0,66483	76,63
7	52,6	0,63374	65,46	0,45117	52,36
8	35,3	0,58191	58,69	0,40407	41,49
9	19,5	0,42661	27,51	0,33986	19,92
10	8,5	0,39545	17,97	0,32515	13,58
11	95,3	0,84531	86,91	0,96241	95,60
12	89,0	0,89101	90,68	0,81418	88,02
13	75,6	0,76417	79,40	0,65454	75,84
14	56,3	0,53832	51,97	0,46821	55,51
15	86,8	0,84751	87,10	0,75009	83,13
16	78,2	0,69083	71,94	0,67294	77,25
17	69,1	0,62234	64,06	0,57847	69,48
18	58,9	0,54333	52,80	0,47315	56,36
19	32,4	0,42419	26,82	0,371330	31,63
20	74,3	0,72736	75,75	0,61251	72,48
21	61,0	0,61345	62,94	0,48524	58,32
22	41,6	0,44662	32,90	0,38819	36,98
23	29,2	0,42334	26,57	0,36215	28,45
24	93,4	0,91955	92,77	0,94343	95,29
25	83,7	0,78341	81,25	0,72247	81,02
26	77,2	0,72242	75,24	0,57178	68,85
27	59,9	0,48332	41,52	0,52662	64,00
28	58,5	0,47271	39,18	0,47726	57,04
29	36,7	0,46492	37,39	0,37206	31,87
30	89,3	0,88469	90,19	0,76310	84,13
31	86,9	0,79748	82,58	0,78121	85,52
32	83,2	0,81542	84,24	0,75435	83,46
33	76,2	0,73971	76,99	0,62756	73,71

1	2	3	4	5	6
34	66,6	0,63309	65,39	0,53702	65,22
35	53,8	0,53961	52,18	0,45829	53,72
36	48,9	0,46800	38,10	0,42681	47,17
37	42,4	0,49470	43,90	0,41275	43,76
38	25,8	0,42881	28,13	0,35132	24,46
39	16,4	0,38640	14,91	0,33059	15,99
40	9,1	0,35969	5,04	0,32574	13,84
41	84,9	0,83844	86,31	0,75453	83,47
42	80,7	0,77973	80,90	0,68282	78,00
43	67,9	0,65811	68,33	0,54974	66,62
44	48,5	0,54581	53,21	0,43307	48,59
45	34,2	0,43860	30,80	0,38853	37,08
46	26,2	0,41541	24,25	0,35892	27,29
47	19,2	0,40789	21,96	0,34051	20,18
48	94,9	0,90316	91,60	0,94432	95,32
49	92,6	0,94145	94,19	0,90966	94,12
50	88,5	0,81566	84,26	0,86396	91,52
51	52,5	0,52712	50,04	0,45673	53,43
52	24,3	0,40455	20,91	0,34725	22,88
53	85,9	0,84031	86,47	0,82431	88,77
54	92,5	0,91153	92,20	0,90278	93,77
55	85,4	0,81864	84,53	0,73416	81,91
56	78,6	0,81871	84,54	0,72045	80,87
57	68,1	0,60019	61,22	0,61204	72,44
58	56,2	0,61103	62,63	0,49681	60,06
59	37,4	0,49689	44,34	0,38921	37,28
60	91,9	0,89200	90,76	0,88277	92,75
61	87,4	0,77521	80,47	0,78990	86,19
62	80,8	0,74864	77,88	0,69010	78,56
63	73,9	0,67305	70,01	0,65862	76,12
64	66,2	0,65332	67,78	0,58245	69,85
65	51,2	0,53249	50,98	0,43728	49,51
66	33,1	0,45038	33,85	0,37639	33,30
67	49,3	0,48573	40,96	0,45513	53,12
68	31,8	0,40106	19,80	0,37019	31,24
69	55,9	0,57095	57,11	0,45415	52,94
70	42,2	0,43768	30,56	0,39382	38,63

71	68,2	0,68822	71,77	0,62573	73,56
72	84,1	0,72803	70,55	0,69220	78,72
73	76,8	0,70221	76,24	0,58817	70,37
74	73,0	0,59871	61,01	0,60321	71,69
75	74,6	0,75107	78,12	0,67015	77,04
76	67,0	0,59904	61,05	0,60319	71,69
77	72,5	0,68162	70,95	0,64019	74,72
78	72,0	0,69474	72,36	0,60320	71,67
79	71,8	0,67704	70,45	0,65001	75,49
80	68,7	0,65856	68,38	0,60324	71,69
81	59,0	0,61511	63,15	0,50811	61,64
82	62,0	0,58714	59,43	0,54110	65,68
83	68,8	0,63407	65,50	0,51078	62,00
84	55,2	0,52644	49,92	0,49322	59,53
85	60,3	0,64375	66,66	0,51307	62,30
86	76,9	0,77525	80,47	0,69847	79,19
87	68,0	0,73007	76,08	0,57617	69,27
88	10,4	0,44978	33,70	0,33127	16,29
89	30,8	0,43242	29,13	0,35941	27,47
90	32,9	0,49874	44,71	0,36685	30,10
91	13,5	0,42706	27,63	0,33535	18,04
92	25,6	0,44918	33,55	0,36413	29,15
93	38,8	0,45009	33,78	0,40441	41,58
94	20,9	0,39995	19,44	0,34453	21,81
95	19,9	0,40637	21,48	0,34107	20,41



Rys.6.4. Związek między wyrazistością logatomową W_L a wskaźnikiem transmisji sygnału mowy W_{STI} / wersja oryginalna /.



Rys.6.5. Związek między wyrazistością logatową W_L a wskaźnikiem transmisji sygnału mowy W_{STI} / wersja zmodyfikowana /.

logatomową a estymowaną wyrazistością logatomową \hat{W}_L wynosi $\sigma_1 = \pm 5,1003 \%$.

Zależność (6.5) otrzymaną dla zmodyfikowanej wersji metody MTF wykreślono linią ciągłą na rys. 6.5, a obliczone wartości \hat{W}_L podano w kolumnie 6 tabeli 6.2.

Współczynnik korelacji rang Spearmana r_s między estymowaną wyrazistością logatomową \hat{W}_L a subiektywnie zmierzoną wyrazistością logatomową W_L wynosi $r_s = 0,9895$, natomiast statystyka t [31] wynosi $t = 65,541$. Z tablic statystycznych [63] przy γ stopniach swobody ($\gamma = n-2=93$) odczytana wartość $t_{\alpha} = 0,001 = 3,50785$, a więc $t > t_{\text{tabl}}$. Hipotezę H_0 o braku korelacji między estymowaną wyrazistością logatomową \hat{W}_L , a pomierzoną subiektywnie wyrazistością logatomową W_L można było odrzucić na poziomie istotności $\alpha = 0,001$. Odchylenie standardowe różnic między pomierzoną subiektywnie wyrazistością logatomową W_L , a estymowaną wyrazistością logatomową \hat{W}_L wynosi $\sigma_2 = \pm 1,9158 \%$. Świadczy to o dobrej zgodności wyrazistości logatomowej W_L otrzymanej na drodze pomiarów subiektywnych z estymowaną wyrazistością logatomową \hat{W}_L obliczoną według zależności (6.5).

W pomiarach subiektywnych rozrzut wartości wyrazistości logatomowej W_L w zbiorze słuchaczy jest określony za pomocą odchylenia standardowego:

$$\sigma_L^P = \frac{1}{S-1} \sum_{s=1}^S (W_s^P - W_L^P)^2, \quad (6.6)$$

gdzie:

W_s^P - wyrazistość logatomowa określona dla s -tego słuchacza i p -tego punktu pomiarowego.

W przeprowadzonych pomiarach subiektywnych $\sigma_L^P \in \langle 0,8; 7,6 \rangle$, a więc dla obydwu wersji pomiarowych odchylenie standardowe σ

mieści się w przedziale określonym przez rozrzut wyrazistości logatomowej uzyskanej w zbiorze słuchaczy. Wynika stąd, że różnice między estymowaną wyrazistością logatomową \hat{W}_L , a pomierzoną subiektywnie wyrazistością logatomową W_L są tego samego rzędu, co różnice wynikające z rozrzutu wyrazistości logatomowej otrzymanej w zbiorze słuchaczy. Można więc przyjąć, że wskaźnik transmisji sygnału mowy W_{STI} stanowi dobrą podstawę do obliczania zgodnie z zależnościami (6.4) lub (6.5) estymowanej wyrazistości logatomowej \hat{W}_L .

Dokładność obliczania estymowanej wyrazistości logatomowej \hat{W}_L w metodzie oryginalnej (zależność 6.4) i zmodyfikowanej (zależność 6.5) oceniono porównując odchylenia standardowe różnic między pomierzoną subiektywnie wyrazistością logatomową W_L a estymowaną wyrazistością logatomową \hat{W}_L . Stwierdzono, że odchylenie standardowe σ_2 różnic między wyrazistością logatomową W_L a \hat{W}_L obliczoną wg zależności (6.5) jest mniejsze od odchylenia standardowego σ_1 różnic między wyrazistością logatomową W_L a estymowaną wyrazistością logatomową \hat{W}_L obliczoną wg zależności (6.4). Istotność różnicy między odchyleniami standardowymi σ_1 i σ_2 zbadano testem F [25,90]. Obliczono wartość statystyki F wynosi $F = 7,087485$. Z tablic statystycznych [90] przy γ stopniach swobody ($\gamma = 94$) odczytana wartość $F_{\alpha/2} = 0,001 = 1,627$, a więc $F > F_{tabl}$. Hipotezę H_0 o braku istotności różnic między porównywanymi odchyleniami standardowymi σ_1 i σ_2 ($H_0: \sigma_1 = \sigma_2$) można odrzucić na poziomie istotności $\alpha = 0,02$. Stwierdzono więc, że dokonanie modyfikacji w metodzie MTF polegającej na wprowadzeniu bloku $1/x^a$ sgn x kształtującego rozkład gęstości prawdopodobieństwa wartości chwilowych sygnału testowego zgodnie z rozkładem gęstości prawdopodobieństwa wartości chwilowych sygnału mowy polskiej w sposób istotny poprawiło zbieżność obiektywnego estymatora wyrazistości

logatomowej \hat{W}_L z subiektywnie zmierzoną wyrazistością logatomową W_L .

Celem porównania wartości odchyłeń standardowych różnic między wyrazistością logatomową W_L a estymowaną wyrazistością logatomową \hat{W}_L dla rozważanego w niniejszej pracy przypadku mowy polskiej a wartością podaną przez Houtgasta i Steenekena dla mowy holenderskiej ($\sigma_H = \pm 5,6\%$ dla 167 punktów pomiarowych) [74] przeprowadzono test F o równości wariancji.

Postawiono hipotezę $H_0 : \sigma_A = \sigma_H$ i testowano ją na poziomie istotności $\alpha = 0,02$. Obliczona wartość statystyki F wynosi $F = 1.205548$ z tablic statystycznych [90] przy ν_A i ν_H stopniach swobody ($\nu_A = 94$, $\nu_H = 166$) odczytana wartość $F_{\alpha/2} = 0,01 = 1,5$, a więc $F < F_{\text{tabl}}$

Nie ma więc podstaw do odrzucenia hipotezy H_0 o równości odchyłeń standardowych σ_A i σ_H . Zatem wersje: oryginalna (w przypadku mowy polskiej) oraz podana przez Houtgasta i Steenekena (w przypadku mowy holenderskiej) pozwalają na uzyskanie zbliżonej dokładności obliczania estymowanej wyrazistości logatomowej \hat{W}_L .

Porównanie wyników uzyskanych w wersji zmodyfikowanej dla mowy polskiej i holenderskiej jest niemożliwe, gdyż brak jest odpowiednich badań dla mowy holenderskiej. Jednakże uwzględniając wyniki testu F o równości odchyłeń standardowych $\sigma_A = \sigma_H$, zdecydowano się porównać odchylenie standardowe σ_2 uzyskane dla mowy polskiej w metodzie zmodyfikowanej z odchyleniem standardowym σ_H podanym, dla wersji oryginalnej w warunkach języka holenderskiego, przez Houtgasta i Steenekena [74]. Istotność różnicy między odchyleniami standardowymi σ_2 i σ_H zbadano testem F.

Wykazano, że hipotezę H_0 o braku istotności różnic między porównywanymi odchyleniami standardowymi σ_2 i σ_H ($H_0 : \sigma_2 = \sigma_H$) można odrzucić na poziomie istotności $\alpha = 0,02$.

Na tej podstawie można stwierdzić, że wprowadzona modyfikacja w sposób istotny poprawia zbieżność obiektywnego estymatora wyrazistości logatomowej \hat{W}_L z subiektywnie zmierzoną wyrazistością logatomową W_L w porównaniu z wynikami otrzymanymi przez Houtgasta i Steenekena dla mowy holenderskiej.

7. PODSUMOWANIE

Zadania badawcze postawione w pracy (rozdz.1.3) służyły realizacji trzech celów:

1. adaptacji metody MTF, podanej przez Houtgasta i Steenekena, do oceny jakości transmisji sygnału mowy polskiej w kanałach telefonicznych,
2. modyfikacji metody MTF polegającej na ukształtowaniu rozkładu gęstości prawdopodobieństwa wartości chwilowych sygnału testowego zgodnie z rozkładem gęstości prawdopodobieństwa wartości chwilowych sygnału mowy polskiej,
3. zbadaniu czy słuszna jest teza pracy w której stwierdzono, że modyfikacja metody MTF poprawia dokładność pomiaru wyrazistości mierzonej tą metodą.

W oparciu o materiał eksperymentalny oraz wyniki analiz teoretycznych zawartych w pracy można stwierdzić, że zadania badawcze związane z trzema wyżej wymienionymi celami zostały zrealizowane.

Adaptując metodę MTF do warunków transmisji sygnału mowy polskiej określono obwiednię gęstości widmowej mocy szumu w sygnale testowym (rozdz.3.2) oraz gęstość widmową mocy obwiedni sygnału mowy polskiej (rozdz.3.4). Następnie przeprowadzono cykl badań subiektywnych (dla 95 punktów pomiarowych) w którym zmierzono 95 wartości wyrazistości logatomowych dla różnych warunków transmisji sygnału mowy. Do symulacji warunków transmisji sygnału mowy posłużono się skonstruowanym w Instytucie Telekomunikacji i Akustyki Politechniki Wrocławskiej modelem kanału telefonicznego (rozdz.5.2) Dla tych samych warunków transmisji sygnału mowy zmierzono również metodą obiektywną 95 wartości estymowanej wyrazistości logatomowej \hat{W}_L - stosując opracowany i skonstruowany w ramach pracy system

miarowy MTF, w którym uwzględniono podane statystyki dla mowy polskiej (rozdz. 4 i 6). Dokonując analizy porównawczej pomiarów subiektywnych i obiektywnych zredukowano liczbę wymiarów służących do wyliczenia wskaźnika transmisji sygnału mowy W_{STI} w zaadaptowanej metodzie MTF (w pracy zwanej wersją oryginalną MTF), przy czym podano analityczną metodę redukcji tych wymiarów. Dla zredukowanej oryginalnej wersji metody MTF obliczono wartości współczynników wagowych α i γ zapewniające maksymalną zbieżność pomiędzy pomiarami subiektywnymi a obiektywnymi i wartości te nazwano optymalnymi (rozdz.6.2).

Realizując cel drugi dokonano modyfikacji w oryginalnej metodzie MTF (zaadoptowanej do warunków transmisji sygnału mowy polskiej) polegającej na wprowadzeniu w bloku generującym sygnał testowy układu nieliniowego $|x|^a \operatorname{sgn} x$ pozwalającego na kształtowanie rozkładu gęstości prawdopodobieństwa wartości chwilowych sygnału testowego (rozdz.4.2). Wartość współczynnika a układu $|x|^a \operatorname{sgn} x$ dobrano tak, aby rozkład gęstości prawdopodobieństwa wartości chwilowych sygnału testowego był zgodny ze zmierzonym w ramach pracy rozkładem gęstości prawdopodobieństwa wartości chwilowych sygnału mowy polskiej (rozdz.3.3). Dla zmodyfikowanej wersji metody MTF powtórzono cykl obliczeń redukujących liczbę wymiarów oraz wyznaczono optymalne wartości współczynników wagowych α i γ (rozdz.6.2).

W oparciu o uzyskane wyniki podano dla obydwu wersji metody MTF (tzn. oryginalnej i zmodyfikowanej) analityczny związek pomiędzy wynikami uzyskanymi na drodze obiektywnej oraz subiektywnej (zależności 6.4 i 6.5, rozdz.6.3) i oszacowano zbieżność wyników pomiarów subiektywnych z wynikami pomiarów obiektywnych uzyskanymi za pomocą oryginalnej i zmodyfikowanej wersji metody MTF (rozdz.6.3)

Stwierdzono, że niedokładność metody MTF w wersji oryginalnej po dokonaniu redukcji liczby częstotliwości modulujących i optyma-

lizacji współczynników wagowych wynosi $\sigma_1 = \pm 5,1003\%$ dla mowy polskiej oraz $\sigma_H = \pm 5,6\%$ dla mowy holenderskiej [74].

Wprowadzenie modyfikacji metody MTF, postulowanej w tezie pracy poprawia dokładność metody, gdyż uzyskiwana w tym przypadku niedokładność estymacji wyrazistości logatomowej \hat{W}_L wynosi $\sigma_2 = \pm 1.9158\%$ (rozdz.6.3), co zgodnie z wynikami testu podanymi w rozdziale 6.3 stanowi istotną poprawę dokładności metody MTF w porównaniu z wersją oryginalną stosowaną w warunkach mowy polskiej, jak również mowy holenderskiej. Na tej podstawie można stwierdzić, że teza pracy została potwierdzona.

Oceniając przeprowadzone badania oraz uwzględniając udowodnienie tezy pracy można obecnie wymienić elementy pracy stanowiące oryginalny wkład autora do rozwoju badań nad obiektywizacją pomiaru jakości transmisji sygnału mowy, a w szczególności nad udoskonaleniem metody MTF i jej przystosowaniem do mowy polskiej.

Elementy te są następujące:

1. adaptacja metody MTF do warunków transmisji sygnału mowy polskiej, poprzez podanie właściwych tej mowie parametrów statystycznych wykorzystywanych w metodzie MTF,
2. zmodyfikowanie metody MTF poprzez wprowadzenie kształtowania rozkładu gęstości prawdopodobieństwa wartości chwilowych sygnału testowego,
3. zmierzenie i podanie rozkładu gęstości prawdopodobieństwa wartości chwilowych sygnału mowy polskiej,
4. podanie analitycznej metody redukcji wymiarów służących do wyliczenia wskaźnika transmisji sygnału mowy W_{STI} ,
5. podanie analitycznej metody znajdowania wartości współczynników wagowych d i f występujących w zależności na W_{STI} , które zapewniają maksymalną zbieżność pomiarów obiektywnych z subiektywną miarą odniesienia,
6. wykazanie, że zmodyfikowana metoda MTF jest doskonalszą wersją

metody MTF podanej przez Houtgasta i Steenekena,

7. podanie dla zmodyfikowanej wersji metody MTF zaadaptowanej do warunków transmisji sygnału mowy polskiej liczby oraz wartości parametrów określających warunki pomiaru tą metodą, tzn
- liczby oraz wartości częstotliwości środkowych pasm oktawowych wraz z wartościami odpowiadających im współczynników wagowych α ,
 - liczby oraz wartości częstotliwości sygnału modulującego,
 - wartości współczynnika maskowania γ
- oraz podanie analitycznej zależności pomiędzy wskaźnikiem transmisji sygnału mowy W_{STI} a estymowaną wyrazistością logatomową \hat{W}_L wraz z oceną dokładności tej estymacji.

Dalsze badania związane z kontynuacją tematu pracy będą dotyczyły weryfikacji metody MTF przy uwzględnieniu innych typów zakłóceń i zniekształceń oraz rozszerzenie zakresu jej stosowania do oceny jakości transmisji sygnału mowy przesyłanego na drodze od ust mówcy do ucha słuchacza z uwzględnieniem przetworników elektroakustycznych aparatów telefonicznych, co wymagać będzie nadawania akustycznego sygnału testowego przez sztuczne usta oraz jego odbioru przez sztuczne ucho. Warto również podkreślić aspekt praktyczny wykonanych badań, posłużyły one bowiem do opracowania i wykonania modelu funkcjonalnego systemu pomiarowego MTF, który po wprowadzeniu automatyzacji obliczeń będzie mógł znaleźć szerokie zastosowanie w eksploatacji i produkcji sprzętu telekomunikacyjnego.

Wzrost i rozwój fizyczny dzieci w wieku 1-3 lat

Wzrost i rozwój fizyczny dzieci w wieku 1-3 lat jest procesem dynamicznym, który przebiega w sposób nierównomierny. Wzrost liniowy jest najszybszy w pierwszym roku życia, a następnie zwalnia. Wzrost przestaje się zwiększać w wieku 13-15 lat. Wzrost jest podległy wpływom hormonalnym, szczególnie hormonów tarczycy i przysadki przedniej. Wzrost jest również podległy wpływom czynników środowiskowych, takich jak odżywianie i warunki życia.

Wzrost i rozwój fizyczny dzieci w wieku 1-3 lat jest również podległy wpływom czynników genetycznych. Wzrost jest podległy wpływom genetycznym, szczególnie genetycznym. Wzrost jest podległy wpływom genetycznym, szczególnie genetycznym. Wzrost jest podległy wpływom genetycznym, szczególnie genetycznym.

8. DODATKI

$$y = a + b \cdot x \quad (8.1)$$

$$y = a + b \cdot x + c \cdot x^2 \quad (8.2)$$

$$y = a + b \cdot x + c \cdot x^2 + d \cdot x^3 \quad (8.3)$$

Wzrost i rozwój fizyczny dzieci w wieku 1-3 lat jest procesem dynamicznym, który przebiega w sposób nierównomierny. Wzrost liniowy jest najszybszy w pierwszym roku życia, a następnie zwalnia. Wzrost przestaje się zwiększać w wieku 13-15 lat. Wzrost jest podległy wpływom hormonalnym, szczególnie hormonów tarczycy i przysadki przedniej. Wzrost jest również podległy wpływom czynników środowiskowych, takich jak odżywianie i warunki życia.

$$y = a + b \cdot x + c \cdot x^2 + d \cdot x^3 + e \cdot x^4 \quad (8.4)$$

Dodatek A. Korelacyjna metoda obliczania głębokości modulacji

$$m_{n,1}$$

Rozważania teoretyczne odnośnie do wpływu zniekształceń fazowych na obwódnię sygnału zmodulowanego amplitudowo wykazały, że jeżeli sygnał nośny zostaje przesunięty w fazie, to obwódnia sygnału ulega zniekształceniom fazowym [71].

Metodą eliminującą wpływ błędnego obliczania głębokości modulacji, wywołany nieznaną charakterystyką fazową np. badanego kanału telefonicznego, jest metoda korelacyjna obliczania głębokości modulacji. Polega ona na obliczeniu korelacji wzajemnej obwódni sygnału testowego $y_{n,1}(t)$ na wyjściu badanego kanału telefonicznego z sygnałami harmonicznymi $V_n'(t)$ i $V_n''(t)$, przy czym:

$$Y_{n,1}(t) = Y_1 [1 + m_{n,1} \cos(2\pi f_n t + \xi)] \quad (A.1)$$

$$V_n'(t) = V_n \cos(2\pi f_n t), \quad (A.2)$$

$$V_n''(t) = V_n \sin(2\pi f_n t), \quad (A.3)$$

Obliczając korelację wzajemną między $Y_{n,1}(t)$ (zał.A.1) a $V_n'(t)$ (A.2) otrzymuje się zależność:

$$R_{yV}(\tau) = \frac{1}{T_n} \int_0^{T_n} Y_1 [1 + m_{n,1} \cos(2\pi f_n t + \xi)] V_n \cos 2\pi f_n (t + \tau) dt =$$

$$= \frac{V_n Y_1}{T_n} \left\{ \frac{2 \sin \frac{\omega_n T_n}{2} \cos \frac{\omega_n T_n + 2\tau}{2}}{\omega_n} + \frac{m_{n,1}}{2} [T_n \cos(\tau - \xi) + \right.$$

$$\left. \frac{2 \sin \omega_n T_n \cos(\omega_n T_n + \tau + \xi)}{\omega_n^2} \right\} \quad (A.4)$$

gdzie:

$$\omega_n = 2\pi f_n,$$

T_n - czas trwania sygnału o n -tej częstotliwości modulującej.

Korelację wzajemną między $Y_{n,1}(\tau)$ (A.1) a $V_n''(t)$ (A.2) opisuje zależność:

$$R_{YV}''(\tau) = \frac{1}{T_n} \int_0^{T_n} Y_1 \left[1 + m_{n,1} \cos(2\pi f_n t + \xi) \right] V_n \sin 2\pi f_n (t+\tau) dt =$$

$$= \frac{V_n Y_1}{T_n} \left\{ \frac{2 \sin \frac{\omega_n T_n}{2} \sin \frac{\omega_n T_n + 2\tau}{2}}{\omega_n} + \frac{m_{n,1}}{2} \left[T_n \sin(\tau - \xi) + \right. \right.$$

$$\left. \left. + \frac{2 \sin \omega_n T_n \sin(\omega_n T_n + \tau + \xi)}{\omega_n^2} \right] \right\}, \quad (A.5)$$

Podnosząc do kwadratu $R_{YV}''(\tau)$ (A.4) oraz $R_{YV}''(\tau)$ (A.5), a następnie dodając stronami otrzymuje się:

$$R_{n,1}^2 = R_{YV}''^2(\tau) + R_{YV}''^2(\tau) = V_n^2 Y_1^2 \left\{ \left(\frac{\sin \frac{\omega_n T_n}{2}}{\frac{\omega_n T_n}{2}} \right)^2 + m_{n,1} \frac{\sin \frac{\omega_n T_n}{2}}{\frac{\omega_n T_n}{2}} \cdot \right.$$

$$\cdot \cos \frac{\omega_n T_n + 2\xi}{2} \left[1 + \frac{\sin \omega_n T_n}{\omega_n} \right] + \frac{m_{n,1}^2}{4} \left[1 + 2 \frac{\sin \omega_n T_n}{\omega_n T_n} \cdot \right.$$

$$\left. \left. \cdot \cos(\omega_n T_n + 2\xi) + \left(\frac{\sin \omega_n T_n}{\omega_n T_n} \right)^2 \right] \right\}, \quad (A.6)$$

Jeżeli czas trwania analizy T_n sygnału o n -tej częstotliwości modulującej zostanie dobrany w ten sposób, że spełniony

jest warunek:

$$2 \pi f_n T_n = 2 \pi k \quad \text{gdzie } k=1,2,\dots$$

lub gdy

$$T_n \longrightarrow \infty$$

to

$$R_{n,1}^2 = \left(\frac{V_n Y_1}{2} \right)^2 m_{n,1}^2, \quad (\text{A.7})$$

Z przeprowadzonej analizy matematycznej wynika, że metoda korelacyjna może znaleźć zastosowanie do pomiaru głębokości modulacji sygnału zmodulowanego amplitudowo sygnałem harmonicznym na wyjściu dowolnego rzeczywistego kanału telekomunikacyjnego o nieznannej charakterystyce fazowej. W związku z tym, że w przypadku rzeczywistych kanałów telefonicznych nie jest znana ich charakterystyka fazowa, to metoda korelacyjna została zastosowana do obliczania głębokości modulacji sygnału testowego w metodzie MTF.

Dodatek B - Związek między głębokością modulacji $m_1(f_n)$, a odpowiedzią impulsową $h(t)$ kanału transmisyjnego i stosunkiem sygnał-szum (S/N)

Niech sygnał wejściowy będzie opisany zależnością:

$$x_{n,1}(t) = x_1(1 + \cos 2\pi f_n t) = x_1(1 + \operatorname{Re} \{ e^{j2\pi f_n t} \}), \quad (\text{B.1})$$

gdzie:

x_1 - amplituda sygnału w 1-tym pasmie oktawowym,

f_n - n-ta częstotliwość modulująca.

Sygnał na wyjściu układu liniowego opisany jest natomiast zależnością:

$$Y_{n,1}(t) = Y_1 [1 + m_{n,1} \cos 2\pi f_n (t + \xi)] = Y_1 [1 + m_1(f_n) \cos 2\pi f_n (t + \xi)], \quad (\text{B.2})$$

Z drugiej strony, sygnał wyjściowy może być otrzymany przez splot sygnału wejściowego i odpowiedzi impulsowej układu $h(t)$ [22,42,71] :

$$Y_{n,1}(t) = x_{n,1}(t) * h(t) = \int_{-\infty}^{\infty} x_{n,1}(t-t') h(t') dt', \quad (\text{B.3})$$

Podstawiając równanie (B.1) do (B.3) oraz zakładając, że $h(t) = 0$ dla $t < 0$ i że $h(t)$ jest odpowiedzią impulsową układu liniowego o wymiernej rzeczywistej i dodatniej funkcji transmitancji $H(s)$ dla której zachodzi $H(j\omega) = |H(\omega)| e^{j\varphi(\omega)}$ otrzymuje się:

$$\begin{aligned} Y_{n,1}(t) &= \int_{-\infty}^{\infty} x_1(1 + \operatorname{Re} \{ e^{j2\pi f_n (t-t')} \}) h(t') dt' = & (\text{B.4}) \\ &= x_1 \left[\int_{-\infty}^{\infty} h(t') dt' + \operatorname{Re} \left\{ e^{j2\pi f_n t} \int_{-\infty}^{\infty} e^{-j2\pi f_n t'} h(t') dt' \right\} \right] = \\ &= x_1 \int_{-\infty}^{\infty} h(t') dt' \left[1 + \frac{1}{\int_{-\infty}^{\infty} h(t') dt'} \operatorname{Re} \left\{ e^{j2\pi f_n t} \int_{-\infty}^{\infty} e^{-j2\pi f_n t'} h(t') dt' \right\} \right] \end{aligned}$$

Z porównania równań (B.2) i (B.4) wynika, że:

$$m_{n,1} = m_1(f_n) = \frac{\left| \int_0^{\infty} e^{j2\pi f_n t} h(t) dt \right|}{\int_0^{\infty} h(t) dt}, \quad (\text{B.5})$$

Jeżeli na sygnał przechodzący przez układ transmisyjny o funkcji transmitancji $H(s) = ke^{-st_0}$ (warunek niezniekształcającej transmisji [42]) oddziałują jedynie zakłócenia addytywne, to sygnał wyjściowy ma postać:

$$Y_{n,1}(t) = kx_1(1 + \cos 2\pi f_n t) + N_1 = (Y_1 + N_1) \left(1 + \frac{Y_1}{Y_1 + N_1} \cos 2\pi f_n t \right), \quad (\text{B.6})$$

Porównując równanie (B.2) i (B.6) otrzymuje się zależność na głębokość modulacji (niezależną od częstotliwości modulującej f_n):

$$m_1 = \frac{Y_1}{Y_1 + N_1} = \left(1 + 10^{(S/N)/20} \right)^{-1} \quad (\text{B.7})$$

gdzie:

$(S/N)_1$ - stosunek sygnał/szum w 1-tym pasmie oktawowym
wyrażony w dB wg zależności:

$$(S/N)_1 = 20 \log \frac{Y_1}{N_1}, \quad (\text{B.8})$$

W przypadku gdy na sygnał wejściowy, przechodzący przez układ transmisyjny o odpowiedzi impulsowej $h(t)$, oddziałują również zakłócenia addytywne o średniej wartości amplitudy N_1 w 1-tym pasmie częstotliwości (np. oktawowym), to sygnał wyjściowy można opisać zależnością:

$$\begin{aligned}
 Y_{n,1}(t) &= x_1 \int_0^{\infty} h(t') dt' \left(1 + \frac{1}{\int_0^{\infty} h(t') dt'} \operatorname{Re} \left\{ e^{j2\pi f_n t} \int_0^{\infty} e^{-j2\pi f_n t'} \cdot \right. \right. \\
 &\cdot h(t') dt' \left. \right\} \Big) + N_1 = x_1 \int_0^{\infty} h(t') dt' \left(1 + \frac{\left| \int_0^{\infty} e^{-j2\pi f_n t'} h(t') dt' \right|}{\int_0^{\infty} h(t') dt'} \operatorname{Re} \left\{ e^{j2\pi f_n t} \right\} + \right. \\
 &+ N_1 = \left[x_1 \int_0^{\infty} h(t') dt' + N_1 \right] \left(1 + \frac{\left| \int_0^{\infty} e^{-j2\pi f_n t'} h(t') dt' \right|}{\int_0^{\infty} h(t') dt'} \cdot \right. \\
 &\left. \frac{x_1 \int_0^{\infty} h(t') dt'}{x_1 \int_0^{\infty} h(t') dt' + N_1} \operatorname{Re} \left\{ e^{j2\pi f_n t} \right\} \right. \quad (B.9)
 \end{aligned}$$

Uwzględniając w równaniu (B.9) zależność (B.7) otrzymuje się:

$$\begin{aligned}
 Y_{n,1}(t) &= \left[x_1 \int_0^{\infty} h(t') dt' + N_1 \right] \left[1 + \frac{\left| \int_0^{\infty} e^{-j2\pi f_n t'} h(t') dt' \right|}{\int_0^{\infty} h(t') dt'} \cdot \right. \\
 &\cdot \left(1 + 10^{-(S/N)_{1/20}} \right)^{-1} \operatorname{Re} \left\{ e^{j2\pi f_n t} \right\} \quad (B.10)
 \end{aligned}$$

Porównując równania (B.2) i (B.10) otrzymuje się zależność na głębokość modulacji obowiązującą dla wszystkich warunków transmisji sygnału :

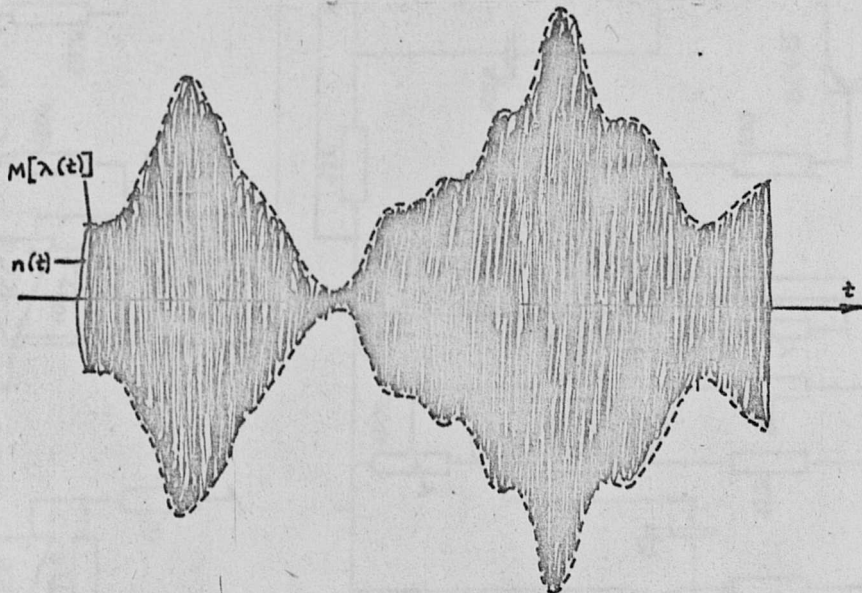
$$m_1(f_n) = \frac{\left| \int_0^{\infty} e^{j2\pi f_n t} h(t) dt \right|}{\int_0^{\infty} h(t) dt} \left(1 + 10^{-(S/N)_{1/20}} \right)^{-1} \quad (B.11)$$

Dodatek C - Zasady ekstrakcji obwiedni sygnału mowy.

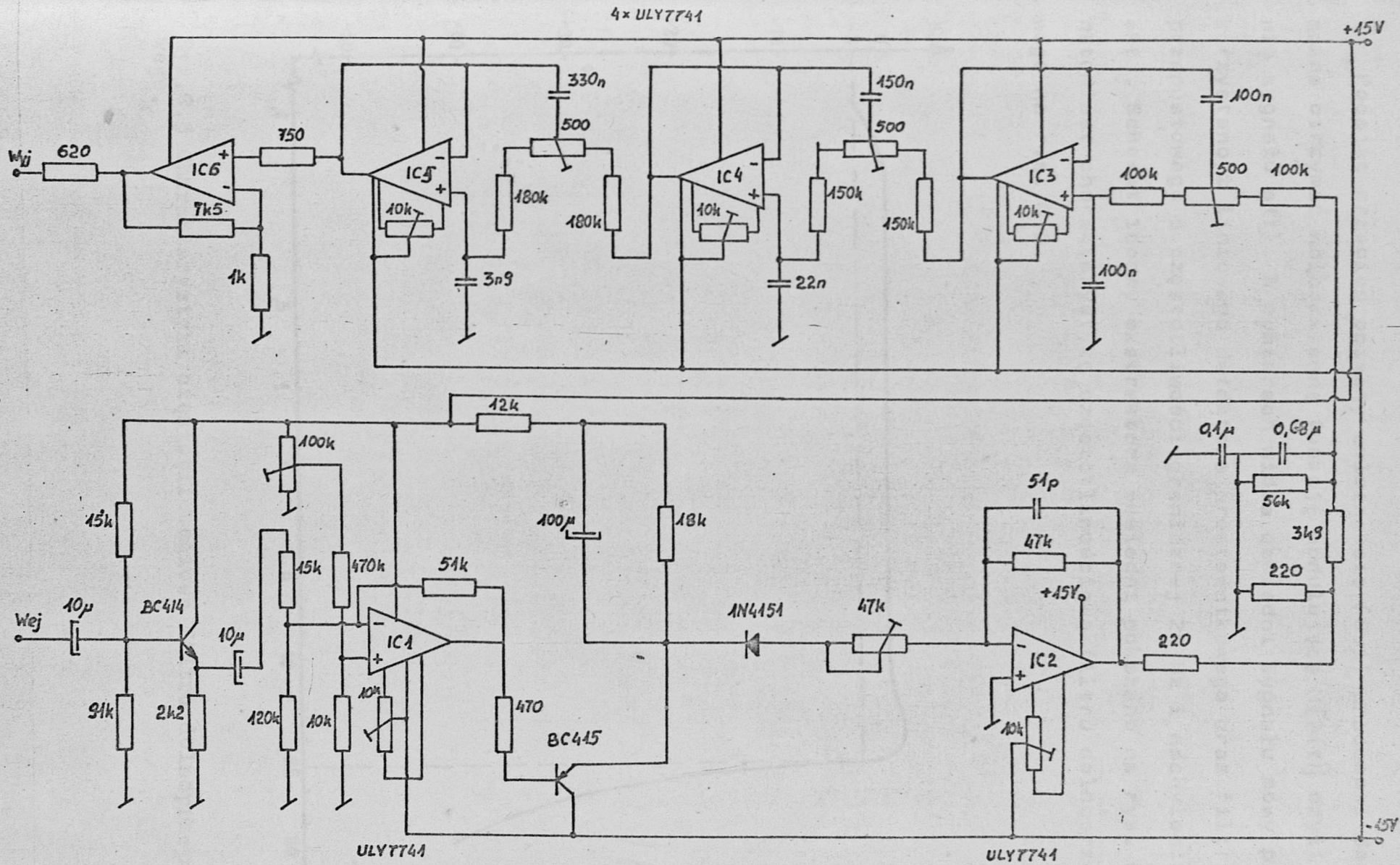
Niech sygnał mowy $s(t)$ będzie procesem losowym, otrzymanym przez zmodulowanie amplitudowe stacjonarnego procesu losowego $n(t)$ stacjonarnym procesem losowym $\lambda(t)$ [48]. Inaczej mówiąc, sygnał mowy $s(t)$ można potraktować jako funkcję czasu $N[n(t)]$ zmodulowaną przebiegiem modulującym $\lambda(t)$. Funkcję modulującą $M[\lambda(t)]$ można uważać jako obwiednię czasową sygnału (Rys. C.1). Matematyczny opis sygnału mowy przyjmuje wówczas postać:

$$s(t) = M[\lambda(t)] \cdot N[n(t)] \quad (C.1)$$

Zależność (C.1) jest słuszna przy założeniu rozdzielczości widm procesów losowych $\lambda(t)$ i $n(t)$. W przypadku sygnału mowy założenie to może być poczynione, gdyż z badań nad obwiednią sygnału mowy wynika, że widmo obwiedni nie zawiera składowych o $f > 25$ Hz [13,32,72,93], natomiast najniższa częstotliwość składowa sygnału mowy (ton krtaniowy) jest wyższa od 70 Hz [64].

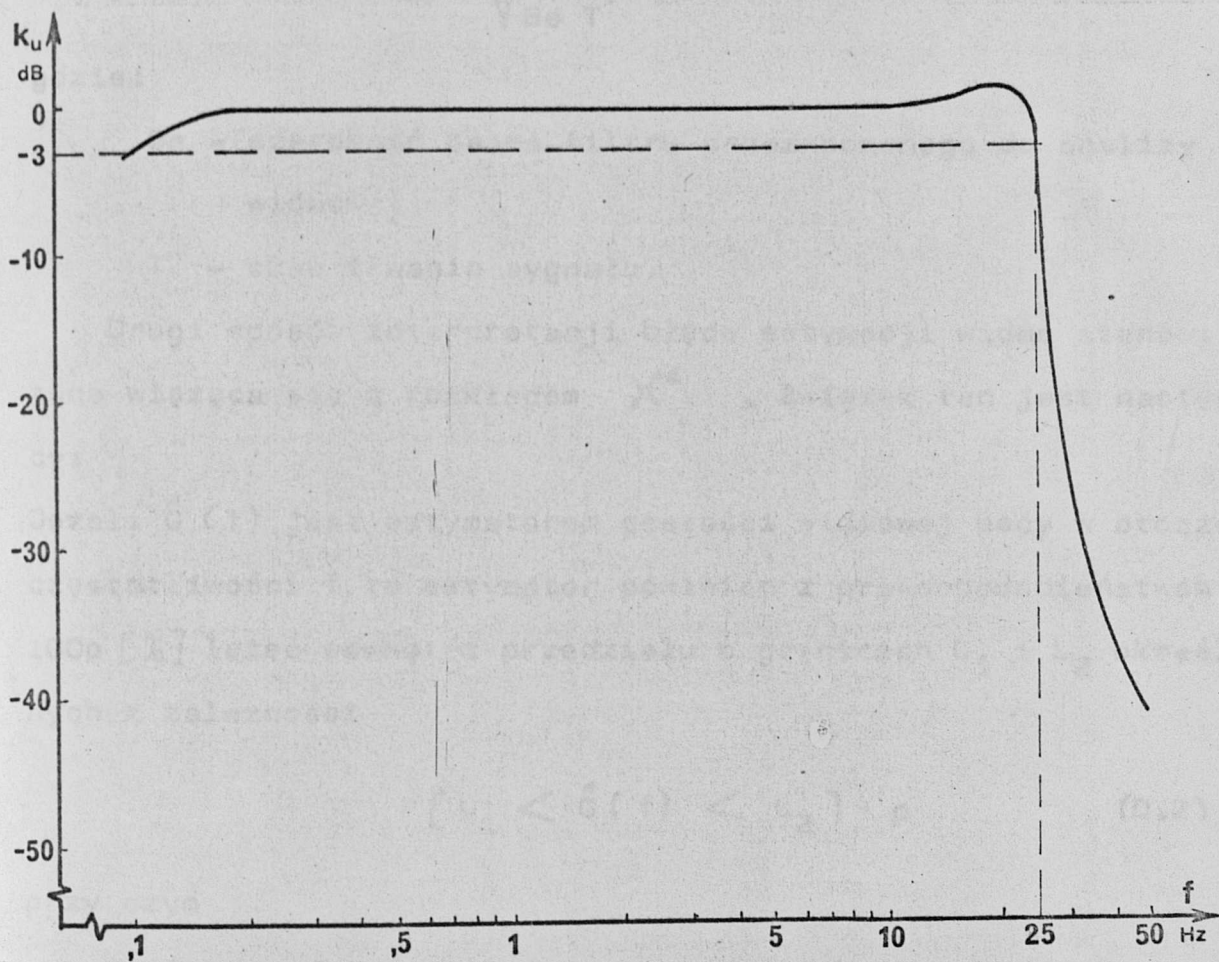


Rys. C.1 Przebieg sygnału $n(t)$ zmodulowanego amplitudowo przebiegiem $\lambda(t)$.



Rys.C.2 Schemat ideowy ekstraktora obwiedni.

Podając przebieg opisany zależnością (C.1) procesowi demodulacji można otrzymać wolnozmienną funkcję modulującą $M[\lambda(t)]$ czyli obwiednię sygnału $s(t)$. W pomiarach widma obwiedni sygnału mowy polskiej korzystano z liniowego detektora prostownikowego oraz filtra dolnoprzepustowego o częstotliwości granicznej 25 Hz i nachyleniu $-42\text{dB}/\text{okt}$. Schemat ideowy ekstraktora obwiedni pokazano na rys. C.2, natomiast charakterystykę częstotliwościową filtra dolnoprzepustowego na rys. C.3.



Rys.C.3 Charakterystyka częstotliwościowa filtra dolnoprzepustowego.

Dodatek D - Analiza błędu systemu pomiarowego

Pomiary gęstości widmowej mocy ze skończonej próby danych należy traktować jako estymację wprowadzającą pewne błędy w stosunku do rzeczywistych wielkości. Wyróżnia się dwa podstawowe sposoby analizy błędu oszacowania [51]. Pierwszy sposób polega na wyznaczeniu błędu standardowego ε_0 estymatora gęstości widmowej mocy $\hat{G}(f)$ z zależności [6.7, 51] :

$$\varepsilon_0 = \frac{1}{\sqrt{B_e T}} \quad (D.1)$$

gdzie:

B_e - szerokość pasma filtru przeznaczonego do analizy widmowej

T - czas trwania sygnału.

Drugi sposób interpretacji błędu estymacji widma stanowi metoda wiążąca się z rozkładem χ^2 . Związek ten jest następujący:

Jeżeli $\hat{G}(f)$ jest estymatorem gęstości widmowej mocy w otoczeniu częstotliwości f , to estymator powinien z prawdopodobieństwem 100p [%] leżeć wewnątrz przedziału o granicach L_1 i L_2 określonych z zależności

$$P[L_1 < \hat{G}(f) < L_2] = p \quad (D.2)$$

przy czym

$$L_1 = \frac{n}{\chi^2_{n, \alpha/2}} \hat{G}(f) \quad (D.3)$$

$$L_2 = \frac{n}{\chi^2_{n, 1-\alpha/2}} \hat{G}(f) \quad (D.4)$$

$$\alpha = 1-p \quad (D.5)$$

p - prawdopodobieństwo, że wartość gęstości widmowej leży w przedziale ufności (najczęściej $p=0,96; 0,9$ lub $0,8$)

n - liczba stopni swobody ($n = 2Be T$)

$\chi_{n,\alpha}^2$ - wartość zmiennej losowej χ_n^2 dla której

$$P [\chi_n^2 \geq \chi_{n,\alpha}^2] = \alpha$$

Dla wykonanych w ramach niniejszej pracy, pomiarów średniej gęstości widmowej obwiedni sygnału mowy polskiej został obliczony zarówno błąd standardowy ϵ_0 jak i przedział ufności (dla $p=0,9$) (rozd. 3.4).

Dodatek E. Pomiar charakterystyk słuchawek nagłownych SN-60

W subiektywnych pomiarach wyrazistości, stanowiących miarę odniesienia spełniającą rolę weryfikatora skuteczności metody MTF, przyjęto ze względów praktycznych słuchawki nagłowne typu SN 60 produkcji firmy Tonsil. Słuchawki te zapewniają stabilność położenia i wystarczającą izolacyjność od zewnętrznych zakłóceń akustycznych.

Zgodnie z danymi podanymi przez producenta charakterystyka amplitudowo-częstotliwościowa skuteczności tych słuchawek jest płaska w zakresie częstotliwości od 16 do 20 000 Hz, przy czym w zakresie częstotliwości od 16 do 2000 Hz zmierzono ją przy użyciu sztucznego ucha, natomiast w zakresie od 2 do 20 kHz w polu fali swobodnej, a więc w warunkach różnych od rzeczywistych warunków pracy.

Uwzględniając badania Richardsa [60], gdzie wykazano, że warunki pomiaru słuchawek w istotny sposób wpływają na uzyskiwane charakterystyki częstotliwościowe, celowe jest przed przystąpieniem do pomiarów subiektywnych zmierzenie charakterystyk stosowanych słuchawek w rzeczywistych warunkach pracy, tzn. na głowie słuchacza.

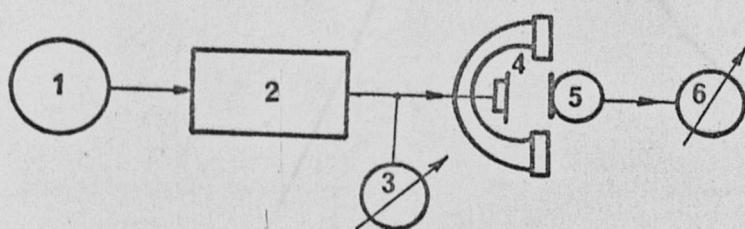
W ramach badań charakterystyk słuchawek wykonano pomiary charakterystyk amplitudowo-częstotliwościowych skuteczności słuchawek przy różnych obciążeniach ich strony akustycznej.

Eksperyment 1: pomiar charakterystyk słuchawek w polu zbliżonym do swobodnego.

Podawane przez producenta charakterystyki amplitudowo-częstotliwościowe skuteczności słuchawek nagłownych dotyczą samego przetwornika i nie uwzględniają wpływu poduszki wokółusznej. Celem

stwierdzenia wpływu warunków pracy słuchawki na jej charakterystykę skuteczności przeprowadzono pomiar tej charakterystyki dla kompletnego zestawu przetwornik-poduszka wokółuszną w warunkach różnych od rzeczywistych (w polu quasiswobodnym).

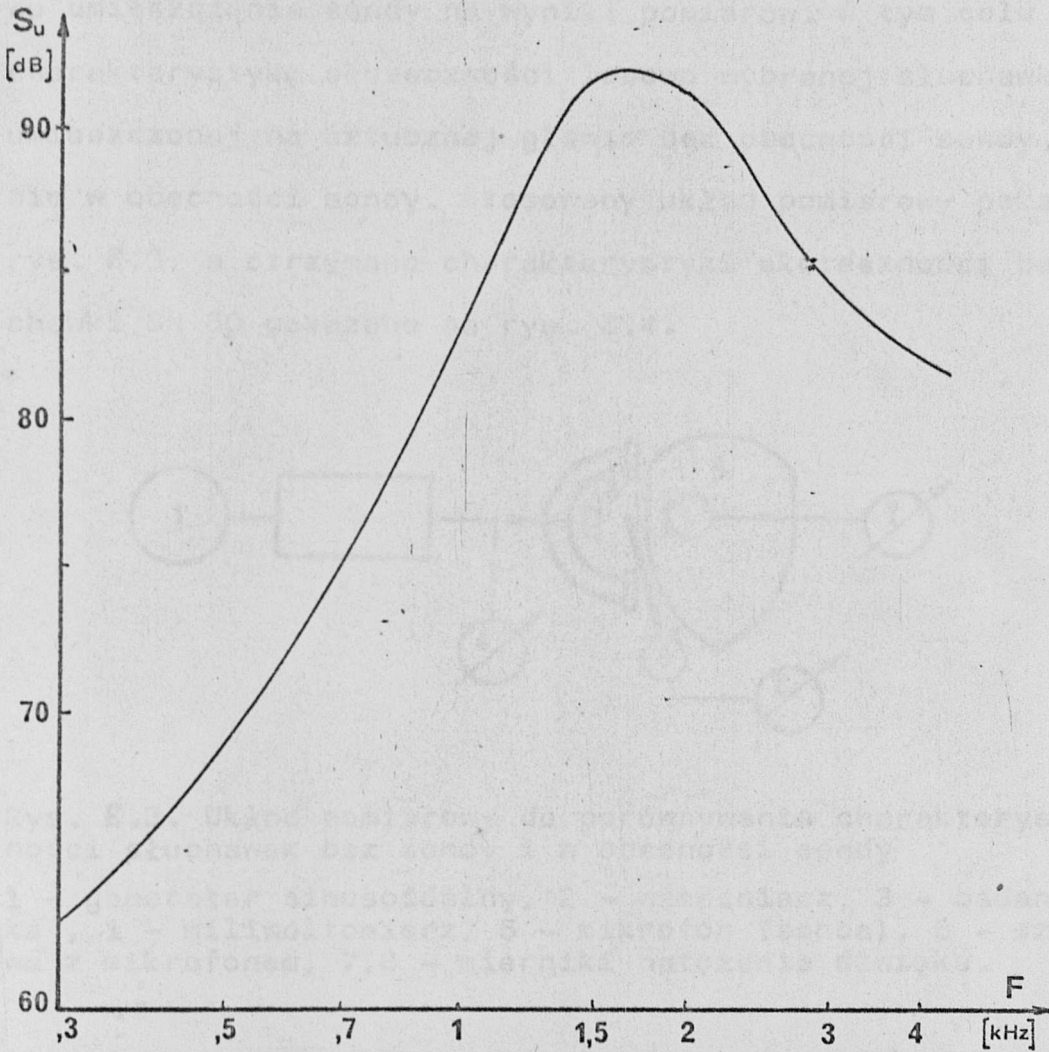
Pomiary przeprowadzono w układzie z rys. E.1.



Rys. E.1. Układ pomiarowy do zdejmowania charakterystyk słuchawek w polu quasiswobodnym

1 - generator sinusoidalny, 2 - wzmacniacz, 3 - miliwoltomierz, 4 - badana słuchawka, 5 - mikrofon, 6 - miernik natężenia dźwięku

Otrzymany zbiór 20 charakterystyk skuteczności słuchawek SN 60 poddano statystycznej analizie metodą testu Wilcoxon [94]. Wyniki testu wykazały, że na poziomie istotności $\alpha = 0,05$ nie ma podstaw do odrzucenia hipotezy o jednorodności zmierzonej populacji charakterystyk skuteczności. W konsekwencji przyjęto charakterystykę średnią (rys. E.2) jako reprezentatywną.

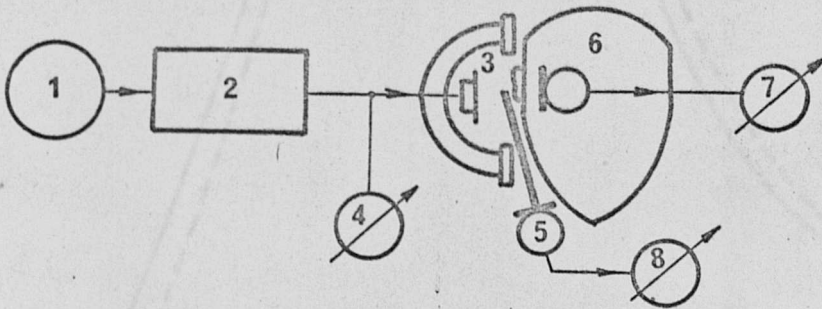


Rys. E.2. Średnia charakterystyka skuteczności słuchawek SN 60 otrzymana w polu quasiswobodnym

Eksperyment 2: pomiar charakterystyk skuteczności słuchawek w rzeczywistych warunkach pracy.

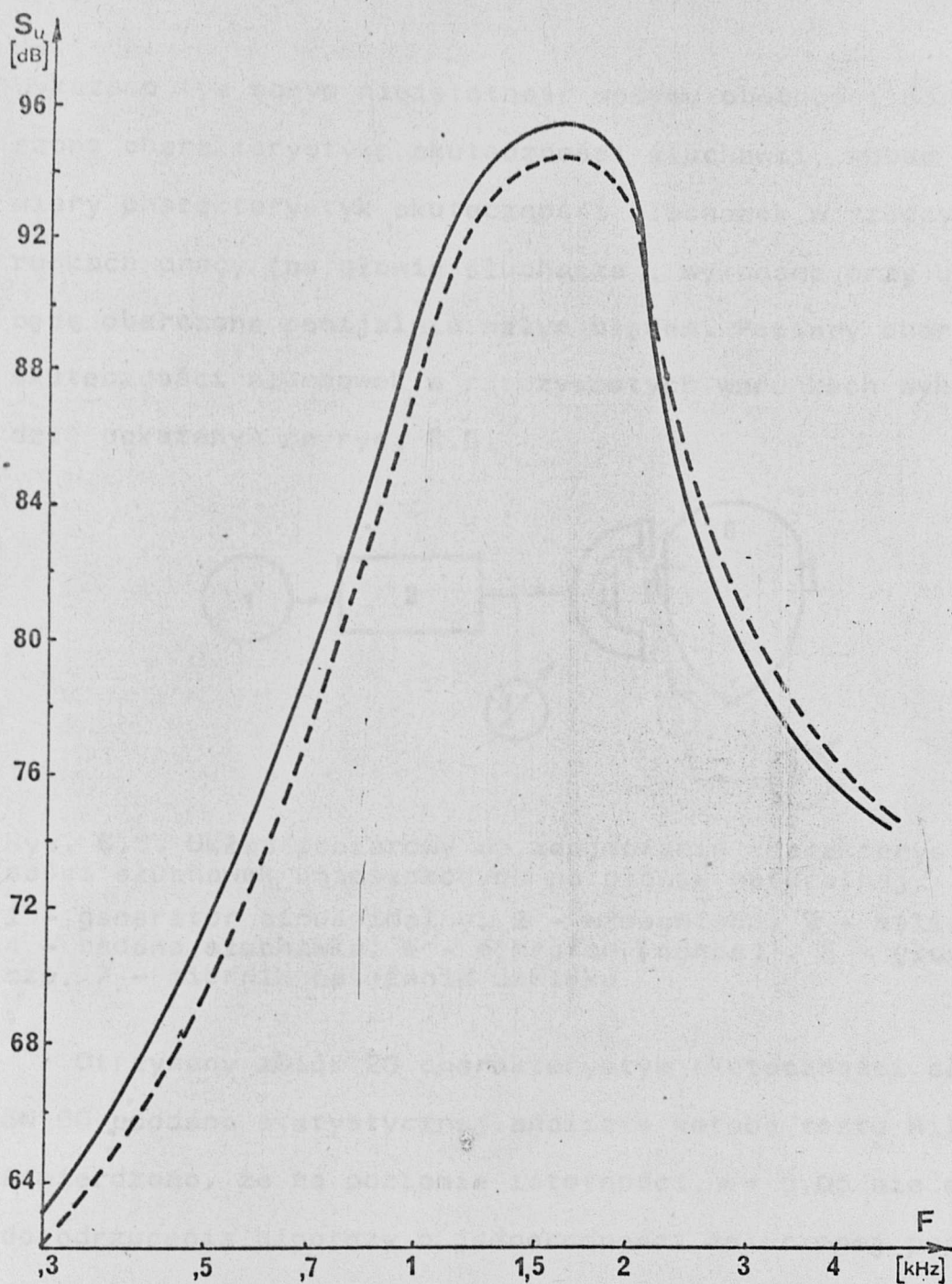
W pomiarach charakterystyk skuteczności słuchawek w rzeczywistych warunkach pracy tzn. na głowie słuchacza wykorzystywano sondę mikrofonu pomiarowego (1/4") wsuniętą od dołu do komory utworzonej przez poduszkę wokółuszną w ten sposób, aby jej

koniec znajdował się przy wlocie kanału usznego. Stosowana technika pomiaru wymaga uprzedniego eksperymentalnego zbadania wpływu umieszczenia sondy na wyniki pomiarów. W tym celu zmierzono charakterystykę skuteczności losowo wybranej słuchawki SN 60 umieszczonej na sztucznej głowie bez obecności sondy, a następnie w obecności sondy. Stosowany układ pomiarowy pokazano na rys. E.3, a otrzymane charakterystyki skuteczności badanej słuchawki SN 60 pokazano na rys. E.4.



Rys. E.3. Układ pomiarowy do porównywania charakterystyk skuteczności słuchawek bez sondy i w obecności sondy

1 - generator sinusoidalny, 2 - wzmacniacz, 3 - badana słuchawka, 4 - miliwoltomierz, 5 - mikrofon (sonda), 6 - sztuczna głowa z mikrofonem, 7, 8 - mierniki natężenia dźwięku.

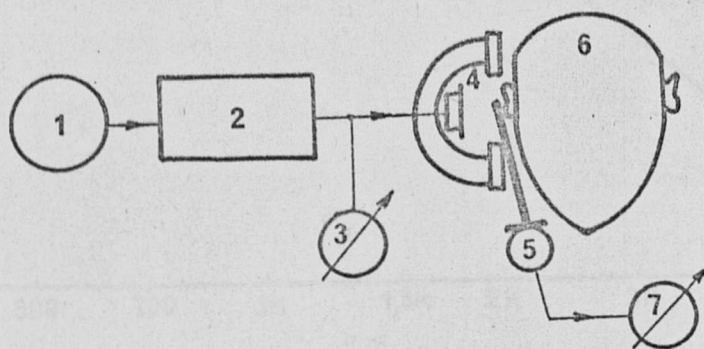


Rys. E.4. Charakterystyka skuteczności słuchawki SN 60 zmierzona na sztucznej głowie

(— bez obecności sondy, ---- w obecności sondy).

Dla charakterystyk z rys. E.4 stwierdzono metodą łącznego testu F i t , że nie ma podstaw do odrzucenia hipotezy o pochodzeniu obydwu charakterystyk z tej samej populacji.

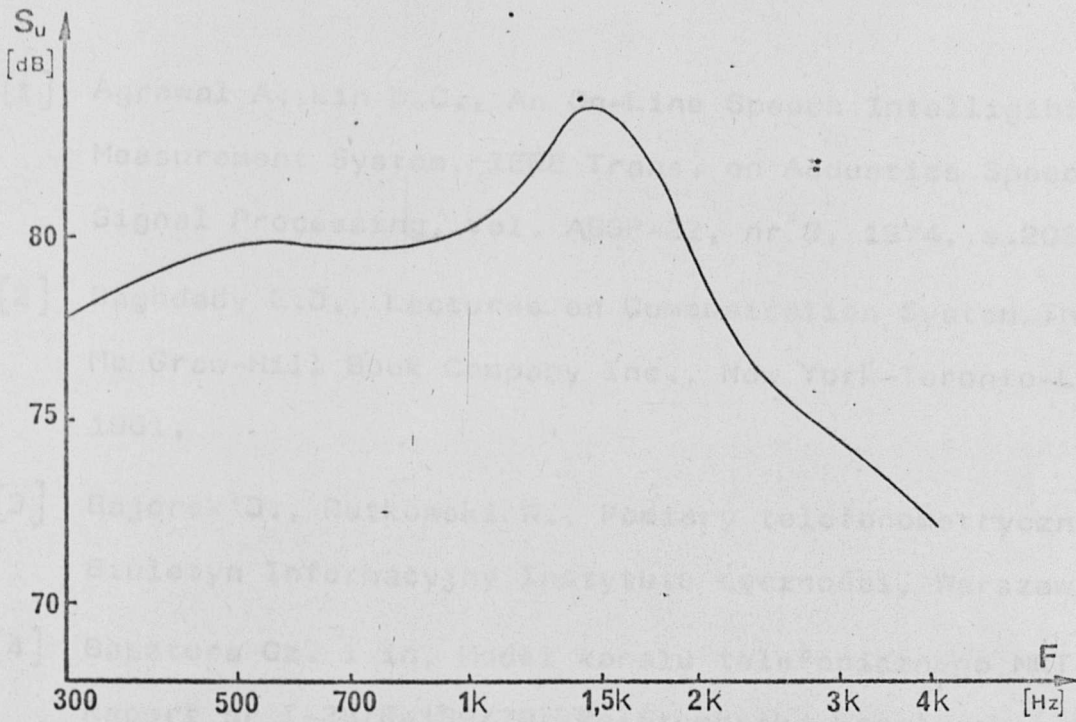
Wykazano tym samym nieistotność wpływu obecności sondy na mierzoną charakterystykę skuteczności słuchawki, wobec czego pomiary charakterystyk skuteczności słuchawek w rzeczywistych warunkach pracy (na głowie słuchacza), wykonane przy użyciu sondy, będą obarczone pomijalnie małym błędem. Pomiary charakterystyk skuteczności słuchawek w rzeczywistych warunkach wykonano w układzie pokazanym na rys. E.5.



Rys. E.5. Układ pomiarowy do zdejmowania charakterystyk skuteczności słuchawek umieszczonych na głowie naturalnej

1 - generator sinusoidalny, 2 - wzmacniacz, 3 - miliwoltomierz, 4 - badana słuchawka, 5 - mikrofon (sonda), 6 - głowa słuchacza, 7 - miernik natężenia dźwięku

Otrzymany zbiór 20 charakterystyk skuteczności słuchawek SN 60 poddano statystycznej analizie metodą testu Wilcoxon. Stwierdzono, że na poziomie istotności $\alpha = 0,05$ nie ma podstaw do odrzucenia hipotezy o jednorodności zmierzonej populacji charakterystyk skuteczności. Uzasadnione jest więc stosowanie charakterystyki średniej jako estymatora charakterystyk skuteczności słuchawek SN 60. Otrzymaną średnią charakterystykę skuteczności pokazano na rys. E.6.



Rys. E.6. Średnia charakterystyka skuteczności słuchawek SN 60 w rzeczywistych warunkach pracy.

9. LITERATURA

- [1] Agrawal A., Lin W.C., An On-Line Speech Intelligibility Measurement System, IEEE Trans. on Acoustics Speech and Signal Processing, Vol. ASSP-22, nr 3, 1974, s.203-206.
- [2] Baghdady E.J., Lectures on Communication System Theory, Mc Graw-Hill Book Company Inc., New York-Toronto-London, 1961.
- [3] Bajorek J., Rutkowski W., Pomiarы telefonometryczne, Biuletyn Informacyjny Instytutu łączności, Warszawa 1977. ✓
- [4] Basztura Cz. i in, Model kanału telefonicznego MKT-1, Raport nr I-28/R-139/78, Politechnika Wrocławska, 1978.
- [5] Basztura Cz., Majewski W., Zastosowanie długoterminowej analizy przejść przez zero sygnału mowy do automatycznego rozpoznawania mówców, Archiwum Akustyki, vol.13, nr 1, 1978, s. 3-15.
- [6] Beauchamp K.G., Przetwarzanie sygnałów metodami analogowymi i cyfrowymi, WNT, Warszawa 1978.
- [7] Bendat J.S., Piersol A.G., Metody analizy i pomiaru sygnałów losowych, PWN, Warszawa 1976.
- [8] Beranek L.L., The Design of Speech Communication Systems, Proceedings of the I.R.E, nr 35, 1947, s.880-890.
- [9] Bieńkowski P., Myślecki W., Zieliński J., Warunki techniczne pomiaru wyrazistości z zastosowaniem słuchawek nagłownych, Biuletyn Techniczny Ministerstwa łączności, nr 1, 1980, s. 20-24. ✓

- [10] Bieńkowski P., Reguły generacji sygnałów reprezentujących stytystyczne rozkłady parametrów sygnału mowy, Praca doktorska, Wrocław 1981.
- [11] Brachmański S.P., Optymalizacja parametrów w metodzie funkcji przenoszenia modulacji (MTF), Raport nr I-28/PRE/060/81, Politechnika Wrocławska 1981.
- [12] Brachmański S.P., Basztura Cz., System pomiarowy funkcji przenoszenia modulacji (MTF) do oceny wyrazistości mowy przesyłanej łączem telefonicznym, Biuletyn Techniczny Ministerstwa łączności, nr 1, 1980, s. 1-10.
- [13] Brachmański S.P., Gos A., Gęstość widmowa obwiedni czasowej sygnału mowy polskiej, Materiały XXVI Otwartego Seminarium z Akustyki, Wrocław-Oleśnica 1979, s.73-76.
- [14] Brachmański S.P., Majewski W., Gęstość prawdopodobieństwa wartości chwilowych sygnału mowy polskiej. Archiwum Akustyki, vol 14, nr 3, 1979, s.203-214.
- [15] Brachmański S.P., Jarycki J., Kozak M., Wskaźnik transmisji mowy jako predykcyjna miara oceny jakości transmisji mowy, Materiały XXIV Otwartego Seminarium z Akustyki, Gdańsk - Władysławowo 1977 s.64-67.
- [16] Brehm H., Jüngst E.W., Wolf D. Simulation von Sprachsignalen, AEU, vol. 28, nr 11, 1974, s.445-450.
- [17] Bykow J.S., Teoria razborczivosti rieczii i powyszenie efektiwnosti radiotielefonnoj swiazi, Gesenergoizdat, Moskwa 1959.
- [18] Davenport A.B., An Experimental Study of Speech-Wave Probability Distributions, JASA, vol.24, nr 4, 1952, s.390-399.

- [19] Dobrogowski A., Investigations into the Distribution of the Instantaneous Values of the Signal in the Case of Polish Speech at Quasi-constant Voice Effort, *Speech Analysis and Synthesis*, Tom 3, PWN, Warszawa 1973, s.11-30.
- [20] Dunn H.K., White G.D., Statistical Measurement on Conversational Speech, *JASA*, vol 11, 1940, s.278-288.
- [21] Fersman B.A., Eksperymentalnoje issledowanie statistycznych swojstw muzykalnych i rzeczowych radiowieszczatielnych signalow, *Akustyczny Żurnal*, t.3, nr 3, 1957, s.274-281.
- [22] Franks L.E., *Teoria sygnałów*, PWN, Warszawa 1966.
- [23] French N.R., Steinberg H., Factor Governing the Intelligibility of Speech Sound, *JASA*, vol 19, nr 1, 1947, s.90-119.
- [24] Goliński J., *Metody optymalizacyjne w projektowaniu technicznym*, WNT, Warszawa 1974.
- [25] Góralski A., *Metody opisu i wnioskowania statystycznego w psychologii*, PWN, Warszawa 1976.
- [26] Guyonnaud J., Progres recents dans la determination objective des equivalents relatifs des systemes d'abonnes: principes et appareillage de mesure utilises, *Note Technique NT/LAA/TSS/1*, Centre National d'Etudes des Télécommunications, Lannion 1980.
- [27] Harris C.M., Waite W.M., Measurement of Speech Spectra Recorded with a Cloze-Talking Microphone, *JASA*, vol.37, nr 5, 1965, s.926.
- [28] Hartman W.J., Objective Performance Measures for Voice System, 2 nd Symp.Tech.Exhibition on Elektromagn. Comp. IEEE, Montreux 1977.

- [29] Hecker M.H., Guttman N., Survey of Methods for Measuring Quality. Journ. of the Audio Engineering Society, vol 15, nr 4, 1967, s.400-403.
- [30] Hlibowicki S., Audibility Examinations of Change of Probability Density of Acoustical Signal Instantaneous Values, 59 th Convention AES, Hamburg 1978, Preprint No.1315(E).
- [31] Houtgast T., Steeneken H.J.M. Evaluation of Speech Transmission Channels by Using Artificial Signal, Acustica, vol 25, nr 6, 1971, s.355-360.
- [32] Houtgast T., Steeneken H.J.M. Envelope Spectrum and Intelligibility of Speech in Enclosures, IEEE-AFCRL Speech Conference, Newton 1972.
- [33] Houtgast T., Steeneken H.J.M. The Modulation Transfer Function in Rooms Acoustics as a Predictor of Speech Intelligibility, Acoustica, vol. 28, nr 1, 1973, s.66-73.
- [34] Houtgast T., Steeneken H.J.M., Plomp R., Predicting Speech Intelligibility in Rooms from the Modulation Transfer Function, I General Room Acoustics, Acoustica, vol. 46, nr 1, 1980, s.60-72.
- [35] Jassem W., Piela B., Steffen M., Average Spectra of Polish Speech, Proceedings of Vibration Problems, nr 2, Warszawa 1959, s.59-71.
- [36] Kett R.W., Design Principles of a New Telephone Apparatus Measuring System, Part 1, Telecom. Journal of Aust., vol. 27, nr 3, 1977, s.198-206.

- [37] Kett R.W., Design Principles of a New Telephone Apparatus Measuring System, Part 2., Telecom. Journal of Aust. vol 28, nr 1, 1978, s.56-63.
- [38] Knoch L., Ekiert T., Modulacja i detekcja, WKiŁ, Warszawa 1979.
- [39] Kopotow P.G., Maksimow J.P., Markow W.W., Nielinieinyje iskażenia rieczewowo signala, Elektroswiaz, nr 8, 1976, s. 29-34.
- [40] Kosiel U., Statistical Analysis of Speakerdependent differences in the Long-term Average Spectrum of Polish Speech, Speech Analysis and Synthesis, Tom 3., PWN, Warszawa 1973, s.181-208.
- [41] Kryter K.D. Methods for the Calculation and Use of the Articulation Index, JASA, vol.34, nr 11, 1962, s.1689 - 1697.
- [42] Lathi B.P., Teoria sygnałów i układów telekomunikacyjnych, PWN, Warszawa 1970.
- [43] Majewski W., Rothman H.B., Hollien H., Acoustic Comparisons of American English and Polish, Journal of Phonetics, nr 5, 1977, s.247-251.
- [44] Majewski W., Zalewski J., Wskaźnik wyrazistości mowy polskiej, Praca doktorska, Politechnika Wrocławska 1964.
- [45] Majewski W., Zalewski J., Wskaźnik wyrazistości mowy polskiej i jego związek z wyrazistością fonemową, Zeszyty Naukowe Politechniki Wrocławskiej, nr 101, 1965, s.59-70.
- [46] Monro D.M., Complex Discret Fast Fourier Transform. Algorithm AS 83, Applied Statistics J. Royal Stat. Soc.,

vol 24, nr 1, 1975 s.153-160.

- [47] Monro D.M., Real Discret Fast Fourier Transform, Algorithm AS97, Applied Statistics J. Royal Stat. Soc., vol 25, nr 2, 1976, s. 166-172.
- [48] Nazarow M.W., Prochorow H.J., Optimalnoje oceniwanie parametrow rieczowych signalow, Elektroswiaz, nr 10, 1975, s. 16-18.
- [49] Noll R., Redundanzreduzierende Sprachcodierung. Eine vergleichende Betrachtung verschiedener Verfahren, NTG - Fachtagung "Signalverarbeitung", Erlangen 1973, s. 354-361.
- [50] Oppenheim A.V.Schafer R.W., Cyfrowe przetwarzanie sygnałów, WKiŁ, Warszawa 1979.
- [51] Otnes R.K. Enochson L., Analiza numeryczna szeregów czasowych, WNT, Warszawa 1978.
- [52] Pirogow A.A., Wokodernaja telefonija, Izdatelstwo "Svjaz" Moskwa 1974 .
- [53] Plomp R., Steeneken H.J.M., Houtgast T., Predicting Speech Intelligibility in Rooms from the Modulation Transfer Function.II Mirror Image Computer Model Applied to Rectangular Rooms, Acustica, vol 46, nr 1, 1980, s.73-81.
- [54] Pokrowski N.B., Rascioł i izmierenie razborczivosti riecz, Swjazizdat, Moskwa 1962.
- [55] Polański Z., Współczesne metody badań doświadczalnych, WP, Warszawa 1978. ✓

- [56] Preusse J.W., Semiautomatic Speech Intelligibility Measurements, IEEE Trans. Audio Electroacoust. Vol. AU-15, 1967, s. 188-191.
- [57] Richards D.L., Statistical Properties of Speech Signals, Proc IEE, vol 111, 1964, s.941-949.
- [58] Richards D.L., Loudness Ratings of Telephone Speech Paths, Proc IEE, vol 118, nr 3/4, 1971, s.423-436.
- [59] Richards D.L., Calculation of Opinion Scores for Telephone Connections, Proc.IEE, vol 121, nr 5, 1974, s.313-323.
- [60] Richards D.L., Telecommunication by Speech. The Transmission Performance of Telephone Networks, London Butterworths, 1973.
- [61] Rimski-Korsakow A.W., Statisticzeskije swojstwa radio-wieszczatielnogo signala, Akusticzeskij žurnal, vol 6, nr 3, 1960, s.360-369.
- [62] Rumszyski L.Z., Matematyczne opracowanie wyników eksperymentu, WNT, Warszawa 1973. ✓
- [63] Sadowski W., Tablice statystyczne, PWN, Warszawa 1957.
- [64] Sapożkow M.A., Sygnał mowy w telekomunikacji i cybernetyce, WNT, Warszawa 1966.
- [65] Sapożkow M.A., Korielacjonnyj metod izmierienia koeficjenta iskażenii peredaczi, Akusticzeskij žurnal, vol 2, nr 3, 1956.
- [66] Schwarzlander H., Intelligibility Evaluation of Voice Communications, Electronics, vol 32, nr 22, 1959, s.88-91

- [67] Seidler J., Statystyczna teoria odbioru sygnałów, PWN, Warszawa - Wrocław 1963.
- [68] Shearme J.N., Richards D.L. The Measurement of Speech Level, P.O. Electr. Engrs. J., vol 47, 1954, s.159-161.
- [69] Sivian L.J., Speech Power and its Measurement, Bell.Syst. Techn.J., vol.8, 1929, s.646-661.
- [70] Smirnow N.W., Dunin-Borkowski I.W., Krótki kurs statystyki matematycznej dla zastosowań technicznych, PWN, Warszawa 1966.
- [71] Sobkowski J., Częstotliwościowa analiza sygnałów, WMON, Warszawa 1975.
- [72] Steeneken H.J.M. Houtgast T., Intelligibility in Telecommunication Derived from Physical Measurement, Proc FASE, Symposium on Speech Intelligibility. Liège 1973, s.316-324.
- [73] Steeneken H.J.M. Houtgast T., A fast Method for the Determination of the Intelligibility of Running Speech, AES 44 th Conv. Rotterdam 1973, s.1-8.
- [74] Steeneken H.J.M., Houtgast T. A Physical Method for Measuring Speech Transmission Quality. JASA, vol 61, nr 1, 1980, s.318-326.
- [75] Steinhardt K.H., Verbesserte Meßeinrichtungen zur Bestimmung der Sprachverständlichkeit. Frequenz, vol 31, nr 1, 1977, s.13-22.
- [76] Steinhardt K.H., Anwendungsmöglichkeiten und Einsatzfälle moderner Sprachverständlichkeitstests, Frequenz, vol 31, nr 1, 1977, s.45-55.

- [77] Tarnoczy T., Determination of the Speech Spectrum through Measurements of Superposed Samples, JASA, vol 28, 1955, s.1270-1274.
- [78] Tarnoczy T., Détermination du spectre de la parole avec une methode nouvel, Acustica, vol 8, 1958, s.392-395.
- [79] Tarnoczy T., Die Sprechchor-Methode, Acustica, vol 23, 1970, s.174-188.
- [80] Tarnoczy T., Über das Individuelle Sprach-Spectrum, Proc. Fourth.Int. Cong. Phonetic Sci, 1961, s.259-264.
- [81] Wawrzynek J., Wybrane testy nieparametryczne, Listy Biometryczne, nr 16-18, 1967, s.1-44.
- [82] Weiß W., Statistische Untersuchungen an Sprachschwingungen, NTZ, nr 7, 1964, s.339-348.
- [83] Wieliczkin A.I., Statisticheskoe issledowanie reczewo processa, Elektroswjaz, t.8, nr 8, 1961, s.3-10.
- [84] Wolf D., Entropy of Speech Signals and Rate Distortion Function, Proc. 1974 Internat.Zürich Seminar on Digital Communications, B8, s.1-3.
- [85] Wolf D., Analytische Beschreibung von Sprachsignalen, AEÜ, vol 31, nr 9, 1977, s.392-398.
- [86] Wolf D., Brehm H., Zwiedimensionale Verteilungsdichten der Amplituden von Sprachsignalen, NTG-Fachtagung Signalverarbeitung, Erlangen 1973, s.378-385.
- [87] Zajcew T.W., O Korrelacionnom metode opriedelenia naturalnosti zwuczania i razborczivosti rieczzi, pieriedawajemoj po traktom swjazi, Elektroswjaz, pr.10, 1958, s.38-46.

- [88] Zalewski J., Badanie wyrazistości fonemowej mowy polskiej i w funkcji częstotliwości, Zeszyty Naukowe Politechniki Wrocławskiej; nr 101, 1965, s.83-92.
- [89] Zalewski J., Majewski W., Polish Speech Spectrum Obtained from Superposed Samples and its Comparison with Spectra of Other Languages. Proc. of 7 th International Congress on Acoustic, Budapest 1971.
- [90] Zieliński R., Tablice statystyczne, PWN, Warszawa 1972.
- [91] Żyszkowski Z., Podstawy elektroakustyki, PWN, Warszawa 1966.
- [92] Acoustics - Recommended Methods for Measuring the Intelligibility of Speech., ISO/DP4870 Second Draft Proposal, August 1976. ✓
- [93] ANSI Standard, Method for Measurement of Monosyllabic Word Intelligibility, s.3.2-1960.
- [94] Commissariat a l'Energie Atomique, Statistique appliquée a l'exploitation des mesures, Tom I i II, Masson, Paris-New York-Barcelone-Milano 1978.
- [95] CCITT, Księga Zielona, Transmisja przewodowa, Tom III-1, WKiŁ, Warszawa 1976.
- [96] CCITT, Księga Zielona. Jakość transmisji telefonicznej, sieci miejscowe i aparaty telefoniczne. Tom V, WKiŁ, Warszawa 1976.
- [97] Hewlett-Packard, Service Manual Model 3721 Correlator, 1971.

Modulation Transfer Function (MTF) as a Measure
of Polish Speech Transmission Quality

The purpose of the dissertation was an adaptation of a method for speech transmission quality evaluation in auditoria by means of modulation transfer function (MTF) to the evaluation of Polish speech transmission quality in telephone channels.

A model generating a test signal of parameter distributions approximating chosen parameters of Polish speech is presented. Basic statistics of Polish speech signal: spectral power density, envelope spectrum and distribution of instantaneous values necessary to generate the test signal are established and presented.

As a result of an optimization of the method and the MTF measuring system the number of modulation frequencies, the values of octave bands weighting factors and the value of masking factor were established with regard to evaluation of Polish speech transmission quality in telephone channels.

Subjective measurements of speech intelligibility and objective measurements of speech transmission index carried out for chosen measurement points representing typical parameters of telephone channels are presented and compared. On this basis the estimation of the usefulness of the MTF method was carried out. It is concluded that the MTF method may be successfully used to evaluate the Polish speech transmission quality in telephone channels.