

Henryk Kubzdela

**BADANIE CECH SEGMENTÓW
FONETYCZNO-AKUSTYCZNYCH
W UPROSZCZONYCH REPREZENTACJACH
WYZNACZONYCH Z WIDM**

40/1991

P.269



WARSZAWA 1991

Praca wpłynęła do Redakcji dnia 31 grudnia 1991 r.



56759



N a p r a w a c h r ę k o p i s u

Instytut Podstawowych Problemów Techniki PAN
Nakład 100 egz. Ark.wyd. 2,0 Ark.druk. 2,25
Oddano do drukarni w marcu 1992 r.

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

Henryk Kubzdela
Zakład Fonetyki Akustycznej
IPPT PAN

Badanie cech segmentów fonetyczno-akustycznych w uproszczonych
reprezentacjach wyznaczonych z widm.

S t r e s z c z e n i e

Przedstawiono zasady kilku przekształceń widma akustycznego sygnału mowy prowadzących do wyekstrahowania niektórych cech akustycznych mogących różnicować pewne segmenty fonetyczno-akustyczne. Tymi cechami są: położenie, szerokość, ilość i skorygowany poziom segmentów widma o wypukłej obwiedni. Użyto po raz pierwszy przekształcenie widmowe znacznie redukujące najniższy fragment widma samogłosek /e/, /a/, /o/, /u/ i przez to dodatkowo różnicujące wzajemnie obrazy czasowych segmentów samogłoskowych i nosowych. Obrazy te ukazują położenie i rangę wypukłych segmentów widmowych. Rangę określa względny poziom segmentu wypukłego w obrębie widma skorygowanego przez odpowiednią preemfazę będącą funkcją energii akustycznej w pewnych pasmach częstotliwości. Stosując wymienione przekształcenia widmowe wyznaczono wspomniane wyżej obrazy dla wypowiedzi 30-tu jednosylabowych wyrazów i logatomów. Obrazy te opisano zwracając uwagę na cechy różnicujące niejednakowe segmenty fonetyczno-akustyczne. Wskazano cechy charakteryzujące wyraźnie głoski nosowe i niektóre samogłoski. Nadają się one do wykorzystania w systemie automatycznej identyfikacji połączeń głównie samogłosek /e/, /a/, /o/, /u/ ze spółgłoskami nosowymi i trącymi.

1. WSTĘP

Badania przedstawione w niniejszej pracy stanowiły kontynuację prób znalezienia takiego opisu segmentów fonetyczno-akustycznych, który uwzględniałby jedynie cechy widmowe wspólne dla segmentów zaliczanych do tej samej klasy i odróżniające wyraźnie segmenty należące do różnych klas. Obrazy widm dźwięków mowy fonetycznie tożsamy, czyli zaliczanych do tej samej klasy głosek, różnią się między sobą. Różnice znaczne występują, gdy segmenty te pochodzą z wypowiedzi różnymi głosami, są jednak także wyraźne wówczas, gdy obrazy

widm dotyczą fonetycznie tożsamyh segmentów w wypowiedziach tym samym głosem. Fakt ten przedstawia dużą trudność w identyfikacji fonetyczno-akustycznych segmentów mowy, będącej pierwszym krokiem w procesie automatycznego rozpoznawania mowy ciągłej. Dla przeprowadzenia tej identyfikacji konieczna jest znajomość właściwych cech dyskryminacyjnych. Poszukiwaniu ich poświęcona jest niniejsza praca.

2. ZAKRES POSTAWIONEGO ZADANIA BADAWCZEGO

Nakreślonej we wstępie problematyce poświęcono dotychczas wiele badań. Ich wyniki nie przyniosły rozwiązania odpowiadającego wymogom automatycznego rozpoznawania segmentów fonetycznych w mowie ciągłej, wykazały natomiast, że problem jest bardzo trudny i złożony i wymaga dalszych badań. Specjalistów zajmujących się automatyczną identyfikacją segmentów fonetycznych w mowie nurtuje pytanie o zbiór cech, dzięki którym pewne segmenty fonetyczne w mowie ciągłej uznawane są za tożsame dźwięki mowy bez względu na cechy głosu mówcy. Podstawowe informacje o cechach mających charakteryzować różne segmenty fonetyczno-akustyczne dostarczają liczne prace z dziedziny fonetyki i akustycznej teorii wytwarzania mowy [1],[2],[3],[4]. Cechy te powinny znajdować swój wyraz w widmie amplitudowym. Siła i przejrzystość tego wyrazu w widmach naturalnej mowy różnych mówców są rozmaite. W niektórych przypadkach jest wyraźnie widoczna spodziewana cecha w innych natomiast oznaki jej obecności są ledwo zauważalne. Pewne cechy przewidywane na podstawie akustycznej teorii wytwarzania mowy nie pojawiają się w niektórych wypowiedziach w ogóle. Zadanie postawione w tej pracy polega na zbadaniu pewnych przekształceń widmowych prowadzących do wyselekcjonowania z pełni informacji o widmie jedynie niektórych szczegółów uznanych z założenia za bardziej ważne od pozostałych i stąd mogących charakteryzować różne segmenty fonetyczno-akustyczne w mowie. Uwagę skierowano na zakresy częstotliwości, w których obwiednia odpowiednio wygładzonego widma amplitudowego jest wyraźnie wypukła. Postanowiono zbadać przydatność pewnej reguły przyznawania

rangi ważności poszczególnym segmentom wypukłym widma dla dyskryminacji różnych segmentów fonetyczno-akustycznych na określonego rodzaju obrazie mowy ciągłej.

Program zadania obejmował :

- a) Ustalenie zasady pomiaru wypukłości obwiedni widma oraz nadawania rangi ważności poszczególnym segmentom widmowym o wypukłej obwiedni.
- b) Pomiar i wyznaczenie rangi ważności poszczególnych segmentów wypukłych w kolejnych widmach sygnału mowy pochodzącego z wypowiedzi jednym głosem materiału fonetycznego wybranego do badań.
- c) Opis wyselekcjonowanych cech w badanych segmentach fonetyczno-akustycznych i ocena przydatności niektórych z tych cech w identyfikacji pewnych segmentów.

3. REGUŁA POMIARU WYPUKŁOŚCI OBWIEDNI I LOKALNYCH KOREKCJI POZIOMU WIDMA AMPLITUDOWEGO.

Przekształcenia badane w pracy dotyczyły widm wyznaczonych 128 punktową FFT zredukowanych następnie do 64 punktów. Widma wygładzano według rozszerzonej zasady podanej w pracy [5]. Obwiednię widma wygładzonego uznawano za wypukłą w punkcie i , jeśli alternatywnie spełniony był jeden z dwóch kolejno badanych następujących warunków :

$$I \quad p_i \geq (p_{i-5} + p_{i+5})/2 + |(p_{i-5} - p_{i+5})|/k_1, \quad (1)$$

$$II \quad p_i \geq (p_{i-3} + p_{i+3})/2 + |(p_{i-3} - p_{i+3})|/k_2, \quad (2)$$

Założono : $k_1 = 5$ oraz $k_2 = 3$.

Na podstawie warunku I badano wypukłość krzywej obwiedni widma w punkcie i względem cięciwy poprowadzonej przez dwa punkty tej krzywej oddalone od punktu i o 5 kolejnych punktów odpowiadających pięciu przyrostom częstotliwości w obu kierunkach osi częstotliwości. W ten sposób wykrywano tzw. rozległą wypukłość obwiedni widma. Stwierdzenie jej obecności wystarczało do uznania krzywej obwiedni widma za wypukłą w

punkcie i . W przypadku nie stwierdzenia wypukłości na podstawie kryterium pierwszego szukano w punkcie i wypukłości mniej rozległej na podstawie kryterium II. Wynik tego badania przesądzał ostatecznie, czy w punkcie i obwiednia widma jest wypukła. Użycie podwójnego alternatywnego kryterium wypukłości obwiedni w podanej wyżej kolejności służyć miało pominięciu w pomiarze wypukłości małych, usytuowanych na wypukłości rozległej. W opisany sposób następowało wyselekcjonowanie segmentów widmowych charakteryzujących się wypukłą obwiednią. Spodziewano się, że w położeniu niektórych z tych segmentów wzdłuż osi częstotliwości i ewentualnie w ich szerokości zawarta może być informacja przydatna do identyfikacji dźwięków mowy. Założono, że o znaczeniu wyselekcjonowanych segmentów dla identyfikacji dźwięków mowy decyduje ich poziom skorygowany zależnie od poziomu widma w innych zakresach częstotliwości. Przeprowadzono szereg doświadczeń w celu wyłonienia takiej zasady korekcji, która zapewniałaby automatyczne podniesienie poziomu w segmentach widma, którym przypisuje się na podstawie wiedzy fonetycznej największą rolę dyskryminacyjną w rozróżnianiu dźwięków mowy. W przypadku samogłosek segmenty te pokrywają się z dwoma najważniejszymi spośród zwykle kilku istniejących formantów. Ze względu na formanty samogłoskowe zastosowano korekcję widma według następującej zasady : Każda składowa s_i o częstotliwości f_i większej od pewnej progowej częstotliwości f_g nieznacznie przekraczającej najwyższą częstotliwość drugiego formantu samogłoski /i/ zostaje stłumiona proporcjonalnie do sumy składowych o częstotliwościach zawartych w paśmie od f_d do $f_i - f_{i-\Delta d}$. Regułę tę wyraża następujący wzór :

$$s_{s_{d,i}} = s_i - k \sum_{j=d}^{i-\Delta d} s_j \quad (3)$$

gdzie przez $s_{s_{d,i}}$ oznaczono i -tą składową widma po stłumieniu składowymi dolnymi.

Dla $i \leq \Delta d + d$ współczynnik k występujący w wyrażeniu określającym składnik tłumiący przyjmuje się zero.

Przekształcenie takie oznacza w przypadku samogłosek przytłumienie formantów wyższych od drugiego przez dwa formanty najniższe. Słabo oddziałuje ono na widmo spółgłosek trących, które za wyjątkiem głosek trących dźwięcznych posiadają znikomą energię w paśmie, z którego pochodzi składnik tłumiący.

Ze względu na samogłoski /e/, /a/, /o/ zastosowano dodatkowe przekształcenie, dzięki któremu w widmach tych samogłosek zostają stłumione najniższe składowe pod wpływem silnego formantu oczekiwanego w paśmie od f_n do f_m . Jeśli w tym paśmie składowe nie występują w ogóle lub są bardzo słabe, to pod wpływem tego przekształcenia stłumienie najniższych składowych nie zachodzi lub jest nieznaczne. Ma to miejsce w przypadku samogłosek /i/ oraz /i/ a także spółgłosek nosowych, oraz większości dźwięcznych trących i zwartych dźwięcznych. Pasma $f_n - f_m$ zostało tak dobrane, że w widmach tych głosek nie występuje w nim silny formant. Składnik tłumiący stanowiły sumy jedynie dodatnich różnic pomiędzy amplitudami składowych w zmiennym przedziale $f_m - f_{i+d}$ a pewnym progiem p ustalonym na poziomie o DF niższym od poziomu najsilniejszej składowej w paśmie $0 - f_m$. Zgodnie z powyższą zasadą składowa s_i w i -tym punkcie widma była tłumiona do wartości st_i obliczonej wzorem :

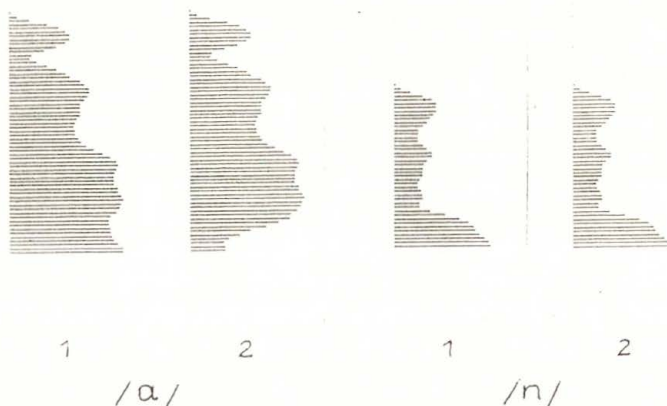
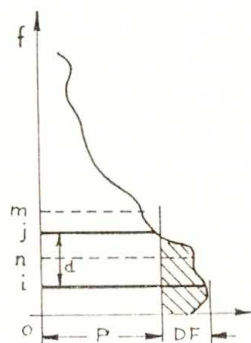
$$st_i = s_i - \sum_{j=m}^{i+d} (s_j - p) \quad (4)$$

dla $i = n, \dots, 1$ oraz $(s_j - p) > 0$, gdzie $m = n + d$ oraz j począwszy od wartości m jest malejące.

Na rys.1 zilustrowano zakres źródła składnika tłumiącego oraz punkt widma, w którym tłumienie od tego źródła wystąpi.

Na rys.2 przedstawiono przykłady selektywnego wpływu tego przekształcenia na poziom najniższych składowych. W przypadku przykładowo rozważonego widma spółgłoski /n/ powyższe przekształcenie nie wywołuje zauważalnego skutku natomiast w przykładzie widma samogłoski /a/ jego działanie jest wyraźnie widoczne.

Rys.1. Ilustracja zakresu źródła składnika tłumiącego oraz punktu widma, w którym tłumienie od tego źródła wystąpi. Oznaczenia zgodnie ze wzorem (4).



Rys.2. Przykłady selektywnego wpływu przekształcenia wyrażonego wzorem (4) na poziom najniższych składowych w widmach głosek /a/ oraz /n/. Numerem 1 oznaczono widma oryginalne a nr. 2 widma po warunkowym przekształceniu.

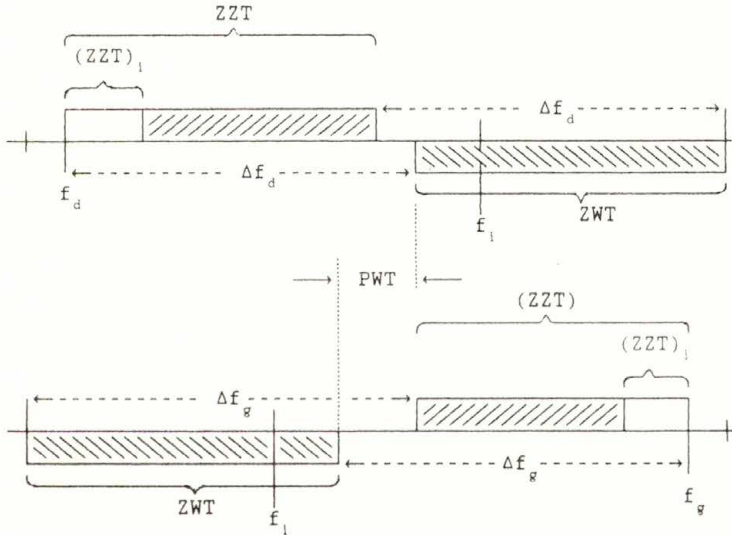
W celu osłabienia składowych widma w dolnym paśmie częstotliwości głosek trących zastosowano kolejną transformację widma amplitudowego. Składnik tłumiący utworzony zostaje ze składowych mieszczących się w paśmie najwyższym. Wielkość tego

tłumienia T_{gi} w i -tym punkcie widma jest wyrażona wzorem identycznym jak wzór (3) lecz z innym przedziałem sumowania a mianowicie :

$$T_{gi} = k_d \sum_{j=i+\Delta_g}^g s_j \quad (5)$$

Dla $i > g - \Delta_g$, T_{gi} przyjmuje się równe 0.

Na rys. 3 przedstawiono schematycznie zakresy częstotliwości ZZT, które mogą zawierać źródła tłumienia oraz zakresy częstotliwości ZWT, w których tłumienie od tych źródeł odpowiednio wystąpi. Δf_d oraz Δf_g oznaczają pasma częstotliwości, których szerokości wynoszą odpowiednio Δd i Δg .



Rys.3. Ilustracja zakresów częstotliwości ZZT, które mogą zawierać źródła tłumienia oraz zakresów częstotliwości ZWT, w których tłumienie od tych źródeł odpowiednio wystąpi.

Przyjęto $f_d + \Delta f_d > f_g - \Delta f_g$, wobec czego nie występuje w żadnym

punkcie widma równocześnie tłumienie od składowych niskich i wysokich. Jeżeli w danym punkcie widma ma miejsce tłumienie, to pochodzi ono z jednego z dwóch rozpatrywanych tutaj źródeł. W wyższych partiach widma tłumienie zachodzi, jeśli są obecne silne składowe w dolnym zakresie częstotliwości od f_d do $f_d + \Delta f_d$. Ewentualne tłumienie w dolnej partii widma pochodzi natomiast od silnych składowych z górnego zakresu częstotliwości od $f_g - \Delta f_g$ do f_g . Stworzono przedział PWT wolny od tłumienia.

4. OPIS OBRAZÓW UZYSKANYCH W WYNIKU ZASTOSOWANYCH PRZEKSZTAŁCEN WIDMOWYCH DLA WYBRANYCH CIĄGÓW GŁOSEK

Przekształcenia widmowe będące przedmiotem przeprowadzonych w tej pracy badań służyć mają hierarchicznej selekcji zakresów częstotliwości wypukłych segmentów widma. Jako podstawę tego rodzaju selekcji przyjęto średni poziom widma w segmentach z wypukłą obwiednią, skorygowanych przed uśrednieniem według podanych w poprzednim rozdziale zasad. Kolejne widma amplitudowe badanych dźwięków mowy poddawano następującym operacjom :

1. Wygładzenie.
2. Warunkowa korekcja poziomu w dolnym zakresie częstotliwości.
3. Wyznaczenie krzywej korekcji poziomu.
4. Wyznaczenie przedziałów częstotliwości, w których obwiednia widma wygładzonego wykazuje odpowiednio wydatną wypukłość.
5. Skorygowanie poziomu w przedziałach częstotliwości wypukłych segmentów zgodnie z wcześniej wyznaczoną krzywą korekcji oraz wyliczenie średniego poziomu oddzielnie w każdym przedziale.
6. Nadanie zakresom poszczególnych segmentów wypukłych odpowiednich wartości z pięcio-stopniowej skali wyrażających ich miejsca kolejne ze względu na średni poziom. Przedział, w którym średni poziom był najwyższy, otrzymywał największą liczbę ze skali, a przedziały, w

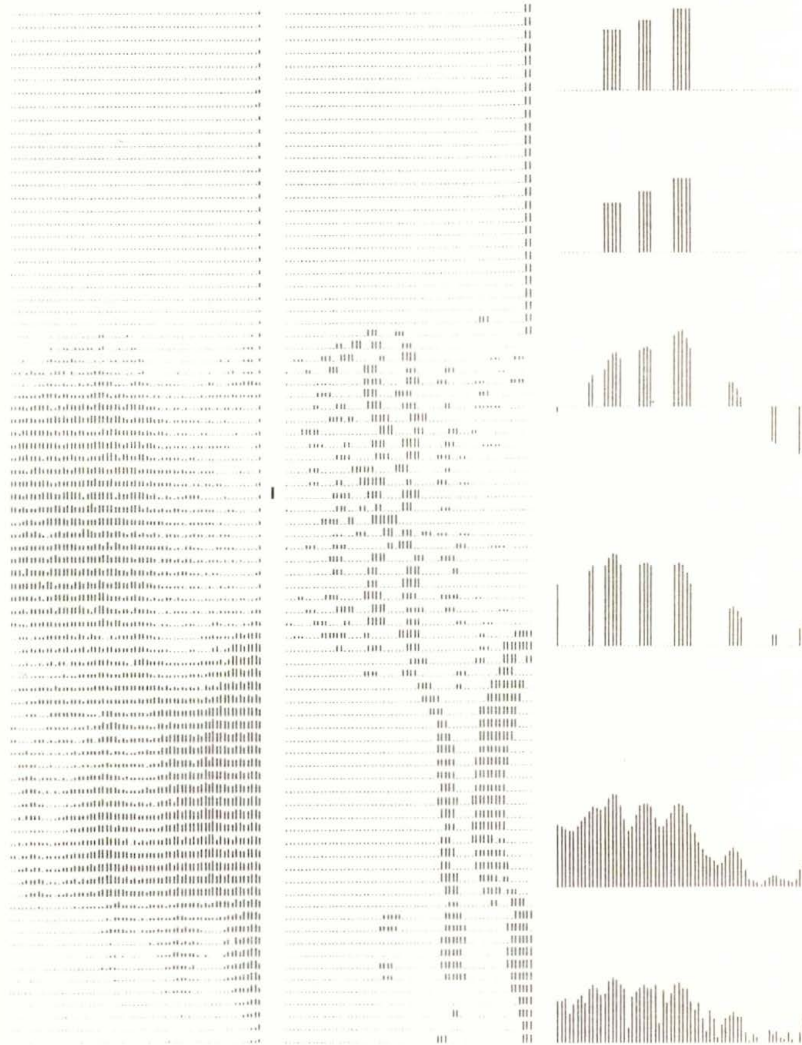
w których średni poziom był stopniowo niższy, otrzymywały liczby kolejno mniejsze.

Stosując wobec widm badanych wypowiedzi wyżej opisane przekształcenia uzyskano obrazy zawierające hierarchiczne wskazania zakresów częstotliwości tzw. wypukłych segmentów widma. Oznacza to, że oddzielnie dla każdego widma zaznaczona jest na tych obrazach ranga poszczególnych segmentów wypukłych wynikająca z wzajemnych relacji ich średniego poziomu. Przykład tego rodzaju obrazu dla wypowiedzi logatomu /maʃ/ pokazano na rys.4 zamieszczonym na następnej stronie. Na rys.4 poniżej obrazu wynikowego zamieszczono też ilustrację faz jego powstawania dla zaznaczonego punktu czasowego wypowiedzi.

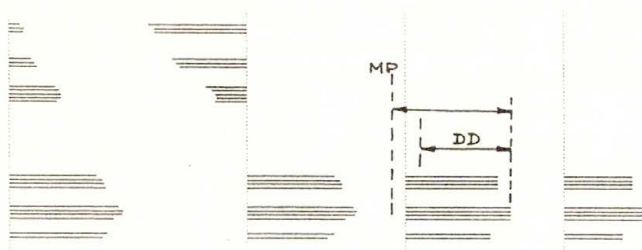
Zakres wypukłego segmentu widma o najwyższym średnim poziomie po korekcji zaznaczony jest najszerszej. Zakresy segmentów wypukłych o kolejno niższych średnich poziomach zaznaczone są stopniowo wężej. Spośród stwierdzonych wypukłych segmentów widma uwzględniano ostatecznie jedynie te, w których średni poziom przekraczał pewien próg MP oraz przypadał w przedziale rozciągającym się o DD jednostek poziomu w dół od średniego skorygowanego poziomu jednego z dwóch najsilniejszych segmentów wypukłych widma. Oba te ograniczenia przedstawiono schematycznie na rys.5.

Podpis do rys.4. zamieszczonego na str. 12. :

Uzyskany w wyniku wyżej opisanych przekształceń widmowych obraz wypowiedzi logatomu /maʃ/ oraz ilustracja faz jego powstawania dla przykładowo wybranego punktu czasowego. W dolnej części rysunku w kierunku poziomym przedstawione są kolejno : widmo pochodzące z FFT, widmo po wygładzeniu, segmenty wypukłe widma, segmenty wypukłe widma po skorygowaniu poziomu, segmenty wypukłe widma z poziomem skorygowanym i uśrednionym, prezentacja rangi poszczególnych segmentów wypukłych widma.



Rys. 4.



Rys.5. Ilustracja wyboru segmentów wypukłych widma ze względu na kryterium poziomu.

Przewidywano, iż w wyniku opisanych wyżej przekształceń widmowych uzyska się obrazy wyrażające na pierwszym miejscu zakresy występowania głównych formantów zaś na dalszych miejscach informacje o mniej ważnych dalszych formantach lub różnych lokalnych wzniesieniach widmowych. Przypuszczano także, iż na obrazach tych wyraźniej niż na klasycznych spektrogramach wystąpią znamiona granic między segmentami fonetyczno-akustycznymi. Dla sprawdzenia słuszności tych przewidywań dokonano przeglądu obrazów wypowiedzi jednym głosem męskim 18-tu logatomów i wyrazów typu CVC, wyznaczone drogą opisanych wyżej przekształceń. Listę tych ciągów głoskowych zamieszczono poniżej.

Lista ciągów głoskowych użytych w badaniach

kar	kaɸ	kam	dor	dom	dotʃ	mur	mun	mutʃ
vek	vem	veɸ	nik	niɸ	nim	lik	lin	litʃ

Poniżej omówione zostaną kolejno dla badanych ciągów głoskowych cechy obrazów uzyskanych w wyniku opisanych wyżej przekształceń widmowych. Określenia : niski i wysoki użyte w tych omówieniach wobec segmentów wypukłych widma odnoszą

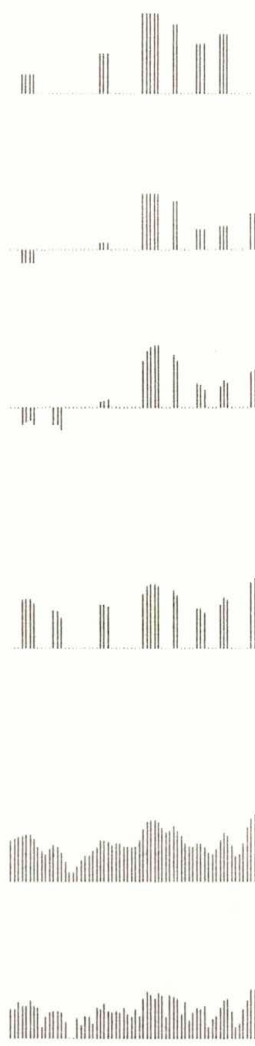
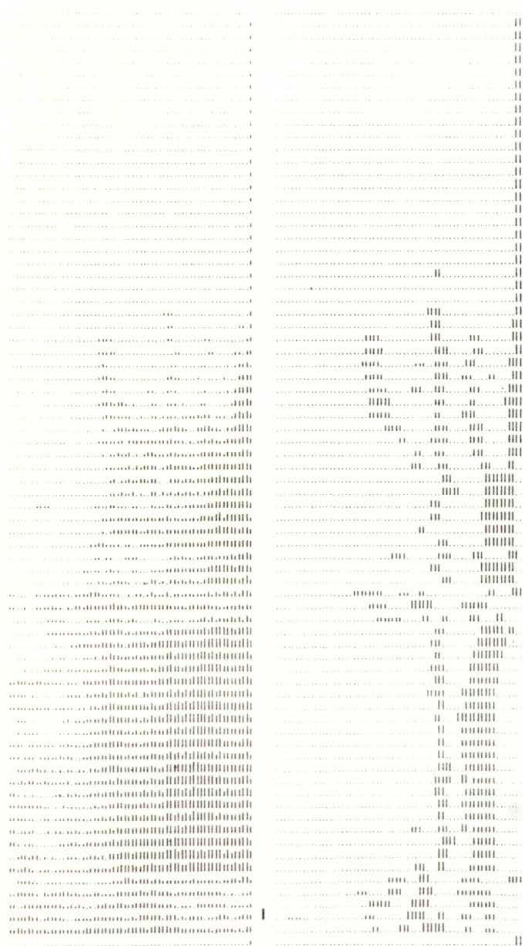
się do położenia tych segmentów na osi częstotliwości nie zaś do ich poziomu. Na rys.6 zamieszczono dla przykładu rozpatrywane w pracy obrazy wypowiedzi ciągów głoskowych /kar/, /kam/, /kaŃ/ , /kaŃ/ oraz wyniki kolejnych przekształceń widmowych prowadzących do uzyskania trzech fragmentów każdego z tych obrazów w zaznaczonych punktach czasowych.

Na stronach : 15, 16, 17 zamieszczono 3 - częściowy rys.6, którego dotyczy następujący wspólny podpis :

Rys.6. Obrazy wypowiedzi ciągów głoskowych /kar/, /kam/, /kaŃ/ oraz u dołu wyniki kolejnych przekształceń widmowych prowadzących do uzyskania trzech fragmentów każdego z tych obrazów w zaznaczonych punktach czasowych.

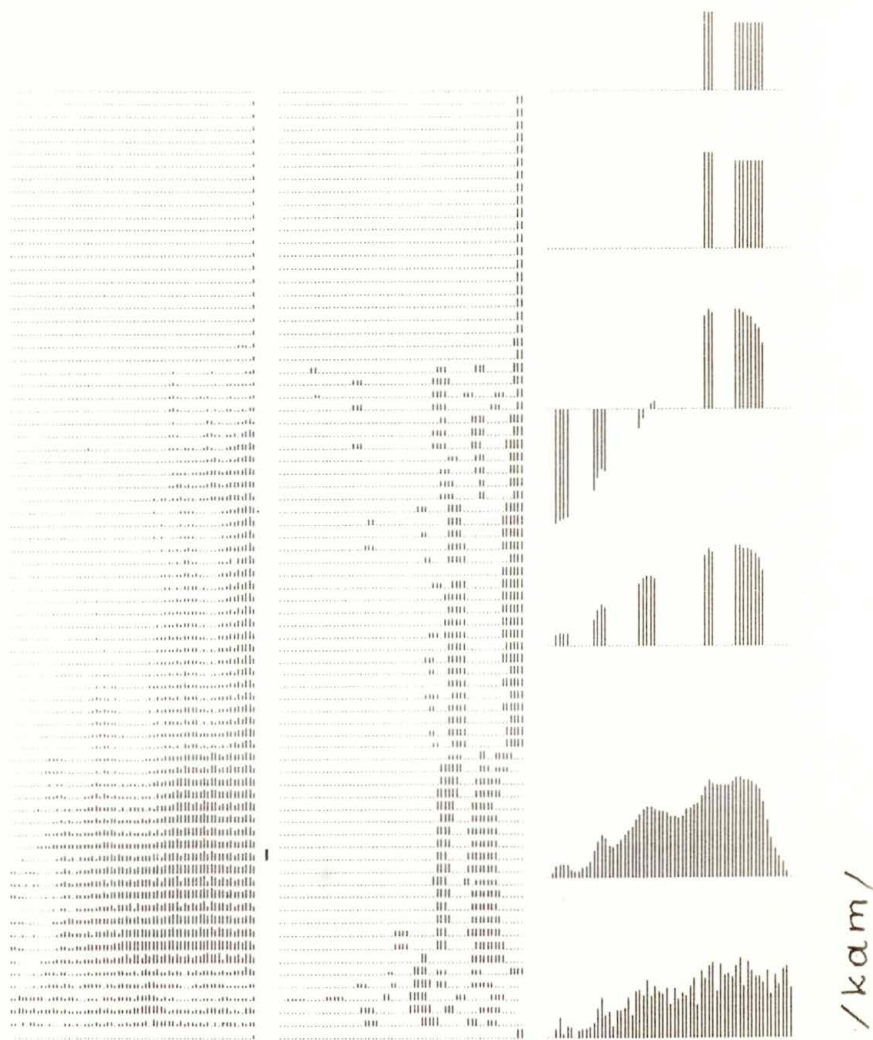
logatom /kar/

Obraz tego logatomu wykazuje pewną stałą regularność jedynie w obrębie dużej części samogłoski /a/. Regularność ta wyraża się obecnością przeważnie jedynie dwóch segmentów wypukłych widma mających w miarę stałe położenie. Segment wyżej położony ma stałą pozycję na całej długości samogłoski /a/. W takiej samej pozycji z nieznacznymi odchyleniami występuje jeden z kilku segmentów wypukłych w obrębie dużej części spółgłoski /r/ będąc tam jednak przeważnie zaledwie jednym z czterech segmentów wypukłych. Tak więc stwierdzenie wypukłości obwiedni widma w tym zakresie wskazuje zarówno na głoskę /a/ jak i /r/ w połączeniu /ar/. Niższy segment wypukły widma w obu tych głoskach ma po części różne położenie i szerokość. W samogłosce /a/ segment ten jest stale dość szeroki wykazując od początku tej samogłoski przez długi czas jej trwania w miarę niezmiennie położenie obniżające się dopiero tuż przed spółgłoską /r/ i to do pozycji, która po krótkiej niestabilności na początku /r/ będzie później na pewien czas trwania tej głoski położeniem jej najsilniejszego i najniższego segmentu wypukłego widma. W tym czasie segment ten zachowuje podobną szerokość, jaką miał niższy z segmentów wypukłych w widmie samogłoski /a/. W pozostałej fazie czasu trwania głoski /r/ ten szeroki niski segment rozdziela się na dwa wąskie, z

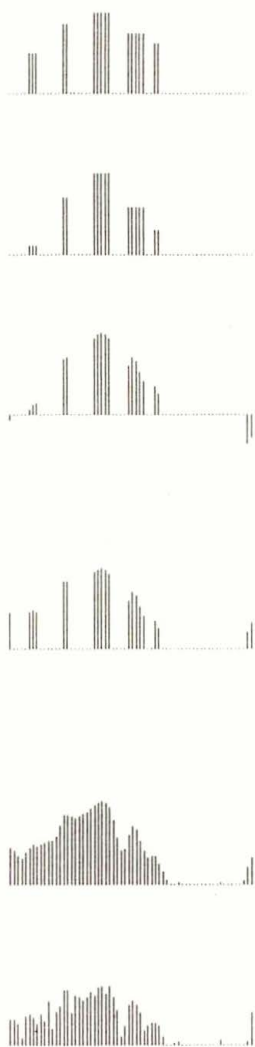
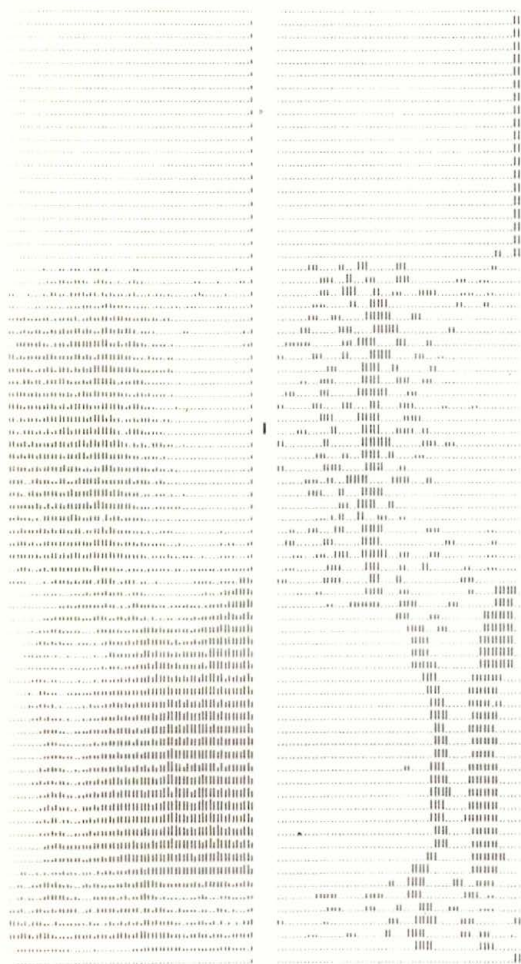


/kar/

Rys. 6.1



Rys. 6.2.



/ ka 6/

Rys. 6.3

których jeden obniża swoje położenie a drugi podwyższa. W tym samym czasie widmo spółgłoski /r/ ma jeszcze jeden wysoko położony segment wypukły. Jego położenie przesuwa się od pozycji niższej zaczynając do pozycji bliskiej położeniu najwyższego segmentu wypukłego widma zidentyfikowanego w początkowej fazie spółgłoski /r/ bezpośrednio po samogłosce /a/. W tej końcowej fazie spółgłoski /r/ najniższe wypukłe segmenty widma starają się wrócić do położenia, jaki miał najniższy wypukły segment widma samogłoski /a/ przez przeważający czas jej trwania.

Badane w pracy przekształcenia nie prowadzą do stworzenia przejrzystego obrazu spółgłoski zwartej /k/ przynajmniej w rozpatrywanym przykładzie logatomu /kar/. O ile w przypadku samogłosek i głosek nietraczących przekształcenia te sprzyjają uzyskaniu obrazów wykazujących jedynie cechy widmowe w najważniejszych dla tych głosek zakresach częstotliwości o tyle w przypadku głoski takiej jak /k/, którą charakteryzuje duża rozległość widma, dochodzi do niepożądanego osłabienia wagi a nawet pominięcia niedrugorzędnych cech widmowych w wyższym zakresie częstotliwości. Tak więc uzyskany obraz głoski /k/ w rozpatrywanym przykładzie nie wyraża wszystkich cech nadających się do scharakteryzowania tej głoski. Obraz głoski /k/ w logatomie /kar/ obarczony takim brakiem uwidacznia jedynie moment początkowy plozji oraz fazę afrykacji. W momencie początkowym najsilniejszy segment wypukły widma jest szeroki i usytuowany nieco powyżej wyższego z dwóch segmentów wypukłych w skorygowanym widmie samogłoski /a/. Po obu stronach tego segmentu widma rozmieszczone są symetrycznie słabsze segmenty wypukłe. W fazie afrykacji najsilniejszy segment wypukły obniża swoje położenie. Słabsze segmenty rozmieszczone są dość nieregularnie. Silniejsze spośród położonych powyżej najsilniejszych też przemieszczają się w dół mając różną szerokość. Najniżej położone segmenty wypukłe są też różnej szerokości a ich położenie oscyluje wokół pozycji, jaką mają najniższe segmenty wypukłe w skorygowanym widmie następującej po /k/ samogłoski /a/.

Logatom /kaɸ/

Rozpatrywany obraz tego logatomu wykazuje kilka cech zbliżonych do tych, jakie zaobserwowano w obrazie części poprzedniego logatomu mającego również na początku głoski /k/ oraz /a/. Samogłoskę /a/ charakteryzują jedynie dwa segmenty wypukłe widma mające w środkowym przedziale jej czasu trwania stabilne położenie. Segmenty oddalają się od siebie na początku i końcu samogłoski /a/ lecz ze zrozumiałych względów inaczej niż w logatomie /kar/. Na początku głoski /a/ w logatomie /kaɸ/ niżej położony segment wypukły widma ma taką samą pozycję jak w środku tej samogłoski a segment wypukły wyższy jest położony wyżej niż w środku samogłoski /a/. Ta ostatnia różnica łącznie z różnicą szerokości wyższych segmentów wypukłych w widmach samogłoski /a/ w logatomach /kar/ i /kaɸ/ może być spowodowana różnym kontekstem prawostronnym samogłoski /a/ w obu tych logatomach. Obraz fazy końcowej samogłoski /a/ w logatomie /kaɸ/ wskazuje, że wypukłe segmenty widma oddalają się nawzajem poprzez obniżanie się segmentu dolnego a podwyższanie się segmentu górnego. Segment wypukły dolny jest szeroki na całej długości samogłoski /a/ w logatomie /kaɸ/, podobnie jak w przypadku logatomu /kar/.

W fazie afrykacji głoski /k/ w logatomie /ka/ najsilniejszy segment wypukły widma oscyluje wokół pozycji nieznacznie wyższej od położenia takiegoż segmentu w środkowej części samogłoski /a/. Po obu stronach tego segmentu rozmieszczone są nieregularnie inne segmenty wypukłe, kilka po każdej stronie. Podobna sytuacja ma miejsce w badanym tutaj obrazie spółgłoski /ɸ/. Najsilniejszy po skorygowaniu segment wypukły widma oscyluje wokół pewnej wysokiej pozycji, znacznie wyższej niż w przypadku głoski /k/. Po obu stronach tego segmentu rozmieszczone są nieregularnie liczne słabsze segmenty wypukłe.

logatom /kam/

Badany obraz tego logatomu w części dotyczącej głosek /k/ i /a/ jest podobny do obrazu tychże głosek w logatomie /kar/. Niewielką różnicę stanowi obniżenie położenia wyższego segmentu wypukłego widma na końcu samogłoski /a/ do pozycji,

jaką posiada drugi z najsilniejszych segmentów wypukłych widma głoski /m/ przez znaczną część jej czasu trwania licząc od jej początku tuż po skończeniu się samogłoski /a/. W obrazie głoski /m/ wyróżniają się dwa segmenty czasowe. Pierwszy obejmujący znaczną część jej czasu trwania charakteryzuje się głównie dwoma stabilnie położonymi segmentami wypukłymi widma. Silniejszy z nich położony jest ekstremalnie nisko a słabszy, jak już wspomniano, nieco niżej od najwyższego z wypukłych segmentów widma głoski /a/ w prawie całym okresie jej trwania. Sporadycznie występuje jeszcze w sąsiedztwie drugiego segmentu wypukłego widma głoski /m/ segment trzeci położony nieco wyżej. Drugi segment czasowy głoski /m/ charakteryzuje się trzema w miarę regularnie położonymi segmentami wypukłymi, z których najniższy jest nieco węższy niż w pierwszym segmencie czasowym głoski /m/ , drugi położony jest niemal pośrodku pomiędzy pierwszym i trzecim i ma wzrastającą szerokość a trzeci zmienia swoje położenie przesuając się lekko wwyż do pozycji drugiego segmentu wypukłego widma samogłoski /a/ poprzedzającej głoskę /m/. Ten segment czasowy głoski /m/ ma więc mieszaną strukturę - samogłoski /a/ , ze względu na położenie dwóch wyższych segmentów wypukłych i spółgłoski /m/ ze względu na położenie najniższego segmentu wypukłego widma. Od obrazu samogłoski /a/ odróżnia się obraz tego segmentu czasowego głoski /m/ szerokością środkowego segmentu wypukłego. Rozpatrywane wyżej obrazy wypowiedzi ciągów głoskowych /kar/, /kaʃ/, /kam/ /nie/ wskazują na obecność ekstremalnie nisko położonego segmentu wypukłego w obrębie samogłoski /a/, natomiast sygnalizują jego obecność w spółgłosce /m/ oraz w części końcowej głoski /r/. Ta cecha może być wykorzystana do wskazania granicy pomiędzy /a/ i /m/ oraz segmentem czasowym centralnym i wygłosowym w spółgłosce /r/. W rozpatrywanych przykładach zwrócono uwagę, na ile zastosowane przekształcenia widmowe pozwalają uwypuklić różnice pomiędzy głoską trącą /ʃ/ a segmentem aspiracji w głosce /k/. Stwierdzono dotychczas, że w obrazie segmentu aspiracji głoski /k/ oraz w obrazie głoski /ʃ/ w badanych logatomach istnieje różnica w położeniu najsilniejszego i

najniższego segmentu wypukłego widma.

Ciągi głoskowe /dor/ , /dom/ , /dotj/ .

W rozpatrywanym obrazie logatomu /dor/ samogłoska odróżnia się od innych głosek dwoma segmentami wypukłymi widma mającymi w miarę stabilne położenie i szerokość niemal na całej długości czasowej samogłoski /o/. Jedyne na samym początku i końcu tej samogłoski wyżej położony segment wypukły przyjmuje pozycję nieznacznie wyższą a segment niżej położony na początku /o/ jest węższy. Na rozpatrywanym obrazie segment plosji głoski /d/ posiadający nieregularnie rozmieszczone segmenty wypukłe widma odgradza samogłoskę od segmentu zwarcia dźwięcznego, reprezentowanego przez jeden ekstremalnie nisko położony segment wypukły widma przez cały czas trwania zwarcia. Bliższe scharakteryzowanie obrazu segmentu plosji nie jest możliwe, gdyż w podanych przykładach nie występują cechy powtarzalne. W obrazie głoski /r/ widocznych jest kilka różnych fragmentów. Wskazują one na różną szerokość, liczbę i stabilność segmentów wypukłych widma. Niektóre z tych segmentów wykazują w miarę stabilne położenie przez większość czasu trwania głoski /r/. Dotyczy to segmentów wyższych. Znaczne zmiany zachodzą głównie w położeniu i szerokości niższych segmentów wypukłych widma zwłaszcza na początku głoski /r/.

Obrazy plosji głoski /d/ są w ciągach głoskowych /dom/ i /dotj/ podobne jak w przypadku logatomu /dor/. Zbliżone są też pozycje i szerokość niżej położonego segmentu wypukłego w widmie samogłoski /o/ we wszystkich trzech badanych ciągach głoskowych. Odmienne natomiast usytuowany jest w każdym z badanych trzech przykładów wyższy segment wypukły widma samogłoski /o/ , co niewątpliwie spowodowane jest różnym kontekstem prawostronnym tej samogłoski. We wszystkich tych trzech przykładach położenie tego segmentu w widmie samogłoski /o/ obniża się w miarę oddalania się od początku samogłoski, po czym w wyrazie /dom/ przyjmuje stałą pozycję, podobnie jak w logatomie /dor/ i podnosi się nieznacznie w logatomie /dotj/. Na samym końcu samogłoski /o/ w logatomie /dotj/ zwiększa się liczba segmentów wypukłych. Dwa najsilniejsze z

nich położone są na trajektorii jedynych dwóch poprzednio reprezentujących tę samogłoskę. Położenie jednego z nowych segmentów wypukłych wyraża początek fazy wygłosu tej samogłoski przypadającego w ciszy zwarcia należącego do następnej głoski /tʃ/. Wygłos ten reprezentuje także trwający nadal niższy segment wypukły zakończonej samogłoski /o/ .

W obrazie głoski /m/ z wyrazu /dom/ widoczne są trzy segmenty czasowe. Cechą charakterystyczną pierwszego z nich jest stabilne położenie dwóch najsilniejszych segmentów wypukłych widma, z których dolny usytuowany jest ekstremalnie nisko a drugi na przedłużeniu trajektorii wyżej położonego segmentu wypukłego widma poprzedzającej ten fragment samogłoski /o/ . Drugi krótkotrwały segment czasowy charakteryzuje się tylko jednym nisko położonym segmentem wypukłym widma. Trzeci wygłosowy segment czasowy wykazuje trzy wypukłe segmenty widma. Położenie najniższego z nich obniża się nieznacznie do pozycji ekstremalnie niskiej. Położenie wyższego segmentu wypukłego przypada pomiędzy dwoma najsilniejszymi segmentami wypukłymi poprzedniego segmentu czasowego i w połowie odległości pomiędzy jego segmentem najniższym i najwyższym. Pozostaje jeszcze do scharakteryzowania obraz czasowego segmentu szumowego głoski zwarto-trącej w logatomie /dotʃ/ . Najbardziej zmiennymi cechami tego obrazu są wskazania po pierwsze dużej liczby rozległe rozmieszczonych segmentów wypukłych widma i po drugie wysokiej pozycji segmentu najsilniejszego. W fazie ploszji najsilniejsze dwa segmenty wypukłe widma położone są zdecydowanie niżej niż po tej fazie.

Wyraz /mur/ i logatomy /mun/, /mutʃ/.

W rozpatrywanym obrazie wypowiedzi wyrazu /mur/ spółgłoskę /m/ charakteryzuje położenie dwóch najsilniejszych segmentów wypukłych widma a mianowicie położenie ekstremalnie niskie silniejszego z tych segmentów oraz położenie drugiego z tych segmentów mogące prawdopodobnie uchodzić za typowe dla głoski /m/ . Dzięki zastosowaniu warunkowej, dolnej korekcji widma omówionej w rozdziale 3 dotychczas tylko w głoskach nosowych, dźwięcznych trących i częściowo drżących miało

miejsce ekstremalnie niskie położenie najsilniejszego lub jednego z najsilniejszych segmentów wypukłych widma. Nie zachodzi taka sytuacja w obrazie samogłoski /u/, mimo iż ma ona niskie dwa formanty. Na obrazach uzyskiwanych w wyniku stosowanych w tej pracy przekształceń widmowych w dotychczas rozpatrywanych przykładach samogłoski /a/ oraz /o/ charakteryzują się obecnością jedynie dwóch segmentów wypukłych w skorygowanym widmie. Ta reguła obejmuje samogłoskę /u/ tylko częściowo. W rozpatrywanym przykładzie /u/ w wyrazie /mur/ na początku samogłoski występują tylko pojedyncze segmenty wypukłe wąskie i szerokie. Pierwszy przypadek może być wynikiem tego, że drugi segment wypukły był na tyle słaby, że nie udało się go wykazać. W drugim przypadku mogło mieć miejsce znaczne zbliżenie dwóch segmentów wypukłych, w wyniku czego doszło do wykazania ich jako jeden szeroki segment wypukły. W miarę trwania samogłoski /u/ tworzą się w jej obrazie dwa oddzielne segmenty wypukłe, z których niżej położony i jednocześnie silniejszy ma stabilną pozycję a drugi przesuwają się ku górze.

Obraz głoski /r/ w wyrazie /mur/ charakteryzuje kilka segmentów czasowych. Określa je położenie najsilniejszych segmentów wypukłych widma oraz liczba i położenie pozostałych segmentów. Są segmenty czasowe głoski /r/, których obraz wykazuje 3 segmenty wypukłe widma i są takie, których obraz wskazuje na obecność w widmie czterech segmentów wypukłych. Są też segmenty czasowe z ekstremalnie nisko położonym segmentem wypukłym widma. Te ostatnie występują w drugiej połowie głoski /r/ oraz krótkotrwanie na jej początku.

Obrazy logatomów /mun/ i /mut/ w części dotyczącej głosek /m/ oraz /u/ są z niewielkimi różnicami zbliżone do obrazu tychże głosek w wyrazie /mur/. Uwagi, które odnosiły się do obrazu głosek /m/ i /u/ w wyrazie /mur/ są aktualne także w przypadku logatomów /mun/ i /mut/. Różnice występujące w obrazach samogłoski /u/ sprowadzają się do trzech szczegółów i są spowodowane przez odmienne konteksty prawostronne tej samogłoski we wszystkich trzech przykładach. W rozpatrywanych obrazach samogłoski /u/ drugi segment wypukły widma w

przypadku logatomu /num/ oddala się od segmentu pierwszego bardziej niż w przypadku logatomu /mut/. W tym ostatnim logatomie nie występuje na początku samogłoski /u/ pojedynczy segment wypukły widma. W logatomie /mun/ na samym końcu samogłoski /u/ występuje symptom zaczynającej się po niej głoski /n/. Wyraża się on przez obniżenie najsilniejszego segmentu wypukłego widma do położenia ekstremalnie niskiego oraz przez pojawienie się dodatkowego trzeciego segmentu wypukłego usytuowanego w pobliżu pozycji, jaką posiada jeden z kilku segmentów wypukłych w widmie spółgłoski /n/. Rozpatrywany obraz tej spółgłoski charakteryzuje się zgodnie z wyżej wyrażoną zasadą ekstremalnie nisko położonym najsilniejszym segmentem wypukłym widma oraz trzema wyżej położonymi segmentami wypukłymi, z których dwa wyższe mają stabilną pozycję i pojawiają się regularnie. Niżej położony segment pojawia się z pewną nieregularnością a jego położenie nie jest stabilne.

Rozpatrywany obraz głoski /tʃ/ w logatomie /mutʃ/ charakteryzuje się w fazie zwarcia śladem wygłosu poprzedniej samogłoski wyrażającym się obecnością dwóch kolejno ukazanych segmentów wypukłych, najpierw jednego położonego ekstremalnie nisko a potem drugiego usytuowanego na wysokości segmentu niższego wcześniejszej samogłoski. Obraz fazy szumu charakteryzuje położenie najsilniejszego segmentu wypukłego. W fazie plosji położenie najsilniejszego segmentu wypukłego jest niższe niż w fazie szumu. W obrazie głoski /tʃ/ najsilniejszy segment wypukły zajmuje pozycję centralną w ciągu segmentów wypukłych widma.

Logatomy /vek/, /vem/ i wyraz /veɸ/.

Rozpatrywany obraz spółgłoski /v/ we wszystkich trzech przykładach jest inny. Na jej początku w wyrazie /veɸ/ i logatomie /vem/ występuje faza nagłosu. Jej obraz podobny jest do obrazu wpierw segmentu czasowego pauzy w głoskach zwartych dźwięcznych a następnie głoski nosowej /m/. Cechą szczególną obrazu tej fazy jest ekstremalnie niska pozycja najsilniejszego segmentu wypukłego widma. Po niej następuje segment czasowy

głoski /v/ wykazujący w rozpatrywanym obrazie wysokie segmenty wypukłe wskazujące na głoskę trącą. Taki segment występuje także w rozpatrywanym obrazie głoski /v/ w logatomie vek. W tym ostatnim segmencie czasowym głoski /v/ najsilniejszy segment wypukły widma podwyższa stopniowo swoje położenie startując z pozycji ekstremalnie niskiej.

Na obrazach samogłoski /e/ w rozpatrywanych przykładach pojawiły się szczegóły, których nie było na obrazach wcześniej badanych samogłosek. W niektórych momentach trwania tej samogłoski, głównie w przykładach /vek/ i /veɕ/, stwierdzone zostały w jej widmie bardzo wąskie segmenty wypukłe położone nieznacznie poniżej silniejszych i szerokich segmentów wypukłych. Ponadto w kilku innych momentach pod koniec tej samogłoski stwierdzono obecność więcej niż dwóch segmentów wypukłych widma. Nie zdarzało się to w obrazach wcześniej rozpatrywanych samogłosek. Te szczegóły rozpatrywanych obrazów samogłoski /e/ nie mogą stanowić jej cechy charakterystycznej, gdyż pojawiają się sporadycznie. Cechą charakterystyczną rozpatrywanego obrazu samogłoski /e/ stanowią trajektorie wyznaczone w zdecydowanej większości przez położenie albo najsilniejszych albo najszerzych segmentów wypukłych widma. Zauważono, że spośród dwóch wzajemnie blisko położonych trajektorii wyznaczonych przez położenie dwóch wypukłych segmentów widma, z których jeden jest bardzo niski i wąski, rzędu 1 do 3 najmniejszych przyrostów częstotliwości (npc) widma, a drugi co najmniej dwukrotnie szerszy, przy odstępnie między sąsiednimi brzegami ich pasm wynoszącym 1 npc widma, trajektorię pierwszą można uważać jako drugorzędny szczegół obrazu

Cechą identyfikacyjną samogłoski na rozpatrywanych obrazach logatomów /vem/ i /vek/ oraz wyrazu /veɕ/ jest brak wskazania obecności ekstremalnie nisko położonego segmentu wypukłego widma. Obrazy samogłoski /e/ w jej części końcowej, w pobliżu sąsiadujących z nią spółgłosek są różne. Różnice zawierają się w obniżaniu lub podwyższaniu się położenia dwóch najsilniejszych segmentów wypukłych widma. Rozpatrywany obraz głoski /m/ w logatomie /vem/ ma cechy podobne jak w

logatomie /kam/ wykazując jedynie wyższe położenie drugiego i ewentualnie trzeciego od dołu segmentu wypukłego widma w przeważającej części czasu trwania tej głoski. Obraz końcowej części głoski /m/ w logatomie /vem/ w porównaniu z logatodem /kam/ charakteryzuje się na pewnym odcinku czasowym podwyższeniem pozycji najniższego segmentu wypukłego. Podwyższenie to upodabnia obraz tego fragmentu głoski /m/ do obrazu samogłoski /e/ zawierającej element nosowości wyrażający się obecnością dodatkowego niskiego segmentu wypukłego. Podobny przypadek zachodził w analogicznym segmencie czasowym głoski /m/ w logatomie /kam/.

W rozpatrywanym obrazie głoski /k/ w logatomie /vek/ w segmencie czasowym pauzy występują znamiona wygłosu zakończonej samogłoski /e/. Wyrażają się one przedłużeniem trajektorii położenia dwóch dominujących segmentów wypukłych w widmie samogłoski /e/. Przedłużenie trajektorii niżej położonej rozciąga się niemal na cały okres pauzy a nawet jeszcze, choć nieregularnie, na segment czasowy afrykacji. Przedłużenie trajektorii wyżej położonej urywa się na początku segmentu pauzy. Obraz segmentu plozji prawie nie odróżnia się od obrazu segmentu czasowego afrykacji, który z kolei charakteryzuje się wskazaniem dużej liczby segmentów wypukłych rozmieszczonych na ogół równomiernie wzdłuż całego rozpatrywanego zakresu częstotliwości. Segmenty wypukłe najsilniejsze sytuują się wokół środka tego zakresu tworząc jakby linię symetrii dla pozostałych segmentów wypukłych widma.

Obraz głoski /ɸ/ wskazuje także na dużą liczbę segmentów wypukłych widma rozmieszczonych w dwóch trzecich rozpatrywanego zakresu częstotliwości licząc od jego górnego brzegu w dół. Położenie najsilniejszego segmentu wypukłego oscyluje wokół środka tej części zakresu. Silniejsze z pozostałych segmentów wypukłych sytuują się częściej powyżej najsilniejszego segmentu wypukłego.

Logatomy /nik/, /niɸ/, /nim/.

Rozpatrywany obraz głoski /n/ we wszystkich tych logatomach składa się z dwóch segmentów czasowych. Obraz

pierwszego z nich wskazuje na obecność w widmie tylko jednego segmentu wypukłego. Położony on jest ekstremalnie nisko. Obraz drugiego segmentu czasowego charakteryzują 4 segmenty wypukłe. Najsilniejszy z nich położony jest ekstremalnie nisko a 3 pozostałe zajmują pozycje, które można uznać za charakterystyczne dla głoski nosowej. Zdumiewające jest, że obraz znacznej części głoski /m/ w logatomie /nim/ wskazuje na takie samo położenie trzech najniższych segmentów wypukłych widma. Jedynie najwyższy segment wypukły w obu tych głoskach ma nieco inne pozycje. W obrazie końca głoski /m/ w logatomie /nim/ wyróżnić można dwa segmenty czasowe. W obu są przeważnie jedynie 3 segmenty wypukłe. W pierwszym najsilniejszy segment wypukły jest usytuowany tak, jak w samogłosce poprzedzającej głoskę /m/ w logatomie /nim/. W drugim segmencie czasowym końca spółgłoski /m/ najsilniejszy segment wypukły jest ponownie w pozycji ekstremalnie niskiej. Kolejny drugi od dołu segment wypukły w pierwszym segmencie czasowym końcowej części głoski /m/ jest na wysokości trzeciego od dołu segmentu wypukłego w widmie początkowego fragmentu głoski /m/. W drugim segmencie czasowym końcowej części głoski /m/ kolejny drugi od dołu segment wypukły widma ma natomiast położenie nieco niższe niż drugi od dołu segment wypukły w początkowej części głoski /m/.

Rozpatrywane obrazy samogłoski /i/ w logatomach /nik/, /nim/, /niń/ wskazują na obecność w widmie więcej niż dwóch segmentów wypukłych. Położenia dwóch najsilniejszych z nich tworzą względnie regularne trajektorie. Segment wypukły najniższy nie zajmuje pozycji ekstremalnie niskiej, co również miało miejsce we wcześniej badanych obrazach innych samogłosek i może być uznane za cechę wspólną większości samogłosek na tego typu obrazach. Sporadycznie pojawiający się drugi niski segment wypukły w widmie samogłoski /i/ w logatomach /nik/ i /niń/ może być śladem nosowości w tej samogłosce.

Obraz głoski /k/ w logatomie /nik/ ma cechy podobne do tych jakie zauważono w obrazie tej samej głoski w logatomie /vek/. Także obraz głoski /β/ w logatomie /niń/ wykazuje podobne cechy do tych, jakie miał obraz tej samej głoski w wyrazie

/veɸ/.

Cechy obrazu głosek /k/ oraz /ɸ/ zostały bliżej przedstawione przy omawianiu obrazów logatomów /vek/ i/veɸ/.

Wyrazy /lik/, /lin/, /litʃ/.

W rozpatrywanych obrazach wszystkich trzech przykładów uderza brak silnych cech różnicujących obrazy głosek /l/ oraz /i/. Słabymi cechami pozwalającymi ewentualnie odróżnić wzajemnie obrazy tych głosek może być nieznacznie niższe położenie najsilniejszego segmentu wypukłego w przypadku widma głoski /l/ niż w przypadku widma głoski /i/ oraz chaotyczne ułożenie silniejszych z pozostałych segmentów wypukłych widma w obrazie głoski /l/. W obrazie głoski /i/ są co najmniej dwie w miarę regularnie przebiegające trajektorie położenia segmentów wypukłych widma. O ile stosowane tutaj przekształcenia widmowe w przypadku większości dotychczas badanych samogłosek dawały obrazy przeważnie posiadające dwie trajektorie położenia segmentów wypukłych widma, o tyle w przypadku samogłoski /i/ oraz częściowo /i/ uzyskano obrazy z większą liczbą takich trajektorii. Są one położone na różnych wysokościach. Te z dodatkowych trajektorii, które pochodzą od silniejszych segmentów wypukłych przeważnie położone są wyżej.

Rozpatrywany obraz głoski /k/ w wyrazie /lik/ wykazuje w segmencie czasowym zwarcia ślad wygłosu poprzedzającej samogłoski. Segment plozji charakteryzuje niskie położenie najsilniejszych segmentów wypukłych widma, które są zdecydowanie niżej usytuowane niż w segmencie czasowym afrykacji. W tym ostatnim segmencie czasowym najsilniejsze segmenty wypukłe widma przyjmują różne położenia. Trajektorja tych położenia jest wprawdzie opadająca a potem utrzymuje się na mniej więcej stałej wysokości, prawie na środku rozpatrywanego zakresu częstotliwości. Inne segmenty wypukłe zajmują miejsca po obu stronach tej trajektorii, przy czym silniejsze z nich przypadają przeważnie powyżej. Niektóre z położonych poniżej zajmują bardzo niskie pozycje.

W rozpatrywanym obrazie głoski /tʃ/ w wyrazie /litʃ/ w jej czasowym segmencie zwarcia występują wskazania bardzo wysoko

położonych segmentów wypukłych widma. Obraz plozji nie zawiera cech, które odróżniałyby go od obrazu segmentu czasowego afrykacji. Wskazuje on na położenie najsilniejszego segmentu wypukłego podobne do tego, jakie ma miejsce w wielu momentach fazy afrykacji. Wynika stąd, że samo położenie segmentów wypukłych w widmie nie jest wystarczającą cechą do zidentyfikowania głoski zwarto-trącej /tʃ/. Rozpatrywany obraz w segmencie czasowym afrykacji głoski /tʃ/ wskazuje na rozległe rozmieszczenie segmentów wypukłych widma wzdłuż całego zakresu częstotliwości. Silniejsze z nich są położone w górnej połowie zakresu. Możliwe jest wskazanie stosunkowo wąskiego zakresu częstotliwości, w którym położony jest przeważnie jeden z dwóch najsilniejszych segmentów wypukłych widma.

Rozpatrywany obraz głoski /n/ w wyrazie /lin/ jest niejednorodny. Można w nim wyróżnić wpierw dwa segmenty czasowe przedzielone krótkotrwałym ubogim dźwiękiem niskotonowym z dwoma wąskimi nisko położonymi segmentami wypukłymi widma. W pierwszym z tych dwóch segmentów czasowych dominują trzy w miarę regularne trajektorie położenia segmentów wypukłych widma. Najsilniejszy z nich jest szeroki i położony ekstremalnie nisko. Segment kolejno słabszy od niego jest bardzo wąski i zajmuje drugie miejsce wzdłuż osi częstotliwości. Trzeci z tych segmentów wypukłych jest przeważnie szeroki i przypada nieco powyżej drugiego z dwóch najsilniejszych segmentów wypukłych poprzedniej głoski /i/. Na krótki czas ten układ zostaje zakłócony. W miejsce jednego szerokiego i najwyżej położonego segmentu wypukłego pojawiają się dwa wąskie w bliskiej odległości od siebie. W dużej części omawianego segmentu czasowego głoski /n/, w połowie odległości pomiędzy dotychczas wymienionym drugim i trzecim segmentem wypukłym widma, istnieje jeszcze jeden segment wypukły. Brak go lub przyjmuje on pozycję niższą jedynie na początku tegoż segmentu czasowego. W drugim segmencie czasowym głoski /n/ w miejsce tegoż dodatkowego, jednego ze słabszych i wąskiego segmentu wypukłego pojawia się segment szeroki, który jest przeważnie drugim z dwóch najsilniejszych segmentów wypukłych. Najsilniejszy jest dalej najniżej położonym lecz częściej już

nie ekstremalnie nisko. Pozostałe segmenty wypukłe są słabsze i rozmieszczone nieregularnie w otoczeniu przedłużenia trajektorii położeń dwóch wyższych spośród trzech głównych segmentów wypukłych poprzedniego segmentu czasowego głoski /n/ w rozpatrywanym obrazie głoski /lin/ .

5. WYNIKI I WNIOSKI

W rozdziale 4-tym opisano rozpatrywane tutaj obrazy wypowiedzi wybranych ciągów głosek. W opisach tych wskazywano głównie cechy jakościowe widm. Nie brano natomiast pod uwagę wartości parametrów, które obrazy te wyrażały. Obrazy różnych segmentów fonetyczno-akustycznych charakteryzowano na podstawie wskazanych położeń segmentów wypukłych widma. Zwracano uwagę na : ilość wykazanych segmentów wypukłych w widmie, usytuowanie najsilniejszych z nich, regularność trajektorii ich położeń w kolejnych widmach, rozległość położeń segmentów wypukłych. Ważenie poszczególnych segmentów wypukłych w widmie i pomijanie niektórych z nich następowało w drodze selektywnej i warunkowej korekcji poziomu. Zastosowane przekształcenie przyjęto jako sposób na zróżnicowanie obrazów różnych segmentów fonetyczno-akustycznych. Możliwe i celowe jest dalsze rozbudowanie systemu tego rodzaju korekcji widma tak, aby uzyskać maksymalnie wyraźne zróżnicowanie obrazów różnych segmentów a nawet subsegmentów fonetyczno-akustycznych.

Na rozpatrywanych obrazach zauważono, że samogłoski /e/, /a/, /o/, /u/, są reprezentowane przeważnie przez jedynie dwie trajektorie wyznaczone przez położenie dwóch najsilniejszych segmentów wypukłych w ich kolejnych widmach. Powyżej wyższego z dwóch najsilniejszych segmentów wypukłych na obrazach tych samogłosek poza nielicznymi wyjątkami nie było wskazań innych jeszcze segmentów wypukłych widma. W obrazach wszystkich samogłosek nie było wskazań ekstremalnie nisko położonego segmentu wypukłego widma. Wskazanie takiego położenia segmentu wypukłego występowało natomiast na rozpatrywanych obrazach niemal regularnie w obrębie segmentów czasowych obejmujących głoski nosowe. Segment tak położony na obrazach głosek nosowych

był zawsze klasyfikowany jako najsilniejszy. Nie zauważono wyraźnych i konsekwentnie powtarzalnych różnic pomiędzy rozpatrywanymi obrazami głosek /m/ oraz /n/. Głoski nosowe na końcu rozpatrywanych ciągów głoskowych wymawiano dłużej niż w innych kontekstach. Ich obrazy były przeważnie niejednorodne. Składały się zwykle z kilku odmiennych segmentów czasowych. Scharakteryzowano je bliżej w opisie obrazów badanych ciągów głoskowych.

Na rozpatrywanych obrazach trudno było wskazać cechy jednoznacznie charakteryzujące głoski /l/ oraz /r/. Z poszczególnych wypowiedzi tych głosek otrzymano obrazy nie mające wiele cech wspólnych.

Głoski trące wykazują na rozpatrywanych obrazach pewne cechy charakterystyczne, nadające się do wykorzystania w ich automatycznej identyfikacji. Te cechy wyrażają się w położeniu najsilniejszych i najniższych segmentów wypukłych oraz w rozległości położen wszystkich segmentów wypukłych widma.

Na podstawie oceny reprezentatywności różnych głosek przez ich rozpatrywane w pracy obrazy stwierdzić należy, że zastosowaną przy wyznaczaniu tych obrazów 8-stopniową skalę wagi segmentu wypukłego widma wyrażającą kolejność tych segmentów ze względu na ich poziom należałoby zmodyfikować. Byłoby wskazane wagę segmentu uzależnić ściślej od różnicy jego poziomu względem poziomu segmentu najsilniejszego. Nasuwa się też wniosek, że rozpatrywane w pracy obrazy należałoby uzupełnić o informacje o położeniu i głębokości wydatnie wklęsłych segmentów widma wyraźnie odgradzających segmenty wypukłe.

Te wnioski i uwagi zostaną uwzględnione w kolejnym etapie poszukiwań sposobu identyfikacji segmentów i subsegmentów fonetyczno-akustycznych w mowie ciągłej.

Przedstawione opisy badanych obrazów wybranych segmentów fonetycznych mowy są przykładem charakteryzowania tychże segmentów. Cechy, na jakie w tych opisach wskazywano dają się łatwo zdyskryminować i zapamiętać. Reguły przedstawionego opisu nadają się do sformalizowania a sam opis po takiej formalizacji do zautomatyzowania.

B i b l i o g r a f i a

- [1]. FANT, G., *Acoustic theory of speech production*, The Hague, 1960.
- [2]. FLANAGAN, J.L., *Speech analysis synthesis and perception*, Kommunikation und Kybernetik in Einzeldarstellungen, t. 3, Springer-Verlag, Berlin, 1965.
- [3]. JASSEM, W., *Podstawy fonetyki akustycznej*, PWN, Warszawa, 1973.
- [4]. DUKIEWICZ, L., *Polskie głoski nosowe*, PWN, Warszawa, 1967.
- [5]. KUBZDELA, H., *Metoda globalnego rozpoznawania wyrazów na podstawie spektrogramów binarnych*, Prace IPPT 28/1986, Warszawa, 1986.



56759