

P. Łobacz, G. Demenko

ZALEŻNOŚĆ PERCEPCJI
SEGMENTALNYCH CECH
SAMOGŁOSEK POLSKICH
OD STRUKTURY DŁUGOTRWAŁEJ PAMIĘCI
LEKSYKALNO-FONEMATYCZNEJ

40/1983

P. 269



WARSZAWA 1983

<http://rcin.org.pl>

Praca wpłynęła do Redakcji dnia 22 listopada 1982 r.



56997



N a p r a w a c h r ę k o p i s u

Instytut Podstawowych Problemów Techniki PAN
Nakład 160 egz. Ark.wyd. 1,6. Ark. druk.2,25
Oddano do drukarni w październiku 1983 r.
Nr zamówienia 790/83

Warszawska Drukarnia Naukowa, Warszawa,
ul.Śniadeckich 8

Piotra Łobacz
Instytut Językoznawstwa UAM
Grażyna Demenko
Pracownia Fonetyki Akustycznej
IPPT PAN

ZALEŻNOŚĆ PERCEPCJI SEGMENTALNYCH CECH SAMOGŁOSEK
POLSKICH OD STRUKTURY DŁUGOTRWAŁEJ PAMIĘCI
LEKSYKALNO-FONEMATYCZNEJ¹

Streszczenie

W pracy podjęto próbę ustalenia interakcji pomiędzy przetwarzaniem fonetycznym i leksykalnym w percepcji mowy. Przedstawiono dwa doświadczenia psycholingwistyczne, z których pierwsze dotyczyło psychologicznej klasyfikacji polskich samogłosek, w drugim zbadano mechanizm korekty błędów fonematycznych w leksykalnym bloku przetwarzania sygnału mowy. Stwierdzono, że w większości przypadków błędnie usłyszanych wyrazów proces ich poprawnego dekodowania wykorzystuje dwie informacje zmagazynowane w pamięci długotrwałej: formę gramatyczną wyrazu oraz jego subiektywną częstość występowania w języku.

1. Fonetyczno-leksykalne przetwarzanie mowy.

1.1. Percepcja mowy na poziomie wyrazu.

W systemie przetwarzania informacji językowej przez człowieka wyrazowi jako jednostce lingwistycznej przypisuje się rolę szczególnie ważną. Wyraz jest tym elementem komunikatu, w obrębie którego następuje wzajemne przenikanie wstępującej informacji akustyczno-fonetycznej z zstępującą informacją syntaktyczno-semantyczną oraz całkowitą lingwistyczną, uświadamianą bądź nie uświadamianą, wiedzą słuchacza ([3,6,7,20,31,32,34]). Dlatego też

¹ Praca wykonana w ramach problemu węzłowego 06.9.

spośród wielu wstępnych modeli percepcji mowy najbardziej uzasadnione są układy leksykalnego dostępu, czyli takie, w których istnieje bezpośrednia możliwość wydobycia określonej informacji leksykalnej, zmagazynowanej w pamięci długotrwałej.

Opracowano dotychczas kilka modeli leksykalnego dostępu, wszystkie zostały przynajmniej częściowo potwierdzone eksperymentalnie. Poszczególne systemy różnią się sposobem przetwarzania wyrazu oraz rodzajem dostępu do pamięci długotrwałej.

System leksykalnego dostępu, w którym wyraz (bądź morfem leksykalny) jest dekodowany jako określona jednostka podstawowa opracowany został przez Mortona [29,30] w jego logogenowym układzie pamięci. Wyraz jako jednostka postaciowa (tzw. "gestalt"), określona zarówno przez swoje znaczenie jak i całkowity kształt akustyczny zaproponowana została w systemie Ahlgrena [1]². Akustyczna postać wyrazu porównywana jest z wewnętrznym leksykonem w terminach możliwych realizacji fonetycznych oraz możliwych znaczeń. W pamięci długotrwałej zawarta jest również informacja gramatyczna, a w niej m.in. fonologiczna interpretacja wyrazu.

Teza, że rozpoznawanie wyrazu jest pierwotne względem dekodowania jednostek niższego rzędu postawiona została także przez innych badaczy. Nooteboom stwierdza: "I will argue in this paper that linguistic processing of speech, and particularly word recognition, is not mediated by phonemes, but that rather phoneme perception as studied in phoneme identification tasks is mediated by word recognition" ([32] s. 143). Według Reppa [39] świadomość segmentów fonetycznych następuje po leksykalnym rozpoznaniu oraz na podstawie świadomej analizy wstępnego rozpoznawania. Zdaniem Fossa i Swinney'a [11] wyrazy pośredniczą w rozpoznawaniu fonemów, ponieważ w procesie normalnego słuchania mowy odbiorca nie monitoruje fonemów, a tylko znaczenia wyższych jednostek. Wykazano eksperymentalnie, że wyraz występujący w naturalnym kontekście identyfikowany jest przy małym wykorzystaniu wstępującej informacji akustycznej (audytorycznej) [14], przy czym istotniejsza dla rozpoznawanej jednostki leksykalnej

² Wpływ semantycznej płaszczyzny wyrazów na percepcję ich dźwiękowego kształtu badali także Maruszewski i Nowakowska [28].

jest informacja związana z początkowymi fragmentami wyrazu. Dla Cole'a, Jakimik i Coopera [8] najważniejsza jest pierwsza sylaba każdego wyrazu, stanowiąca dostęp do wewnętrznego leksykonu w pamięci odbiorcy, dla Marlsena-Wilsona i Welsha [27] dostęp tworzy kilka początkowych fonemów (najczęściej pierwsze trzy). Zarówno przetwarzanie sylab jak i sekwencyjna analiza początkowych fonemów są pierwotne w stosunku do dekodowania wyrazów. W modelu Marlsena-Wilsona i Welsha, w pierwszym etapie rozpoznawania "uaktywnione" zostają w pamięci zbiory jednostek leksykalnych ("kohorty") o wspólnych elementach początkowych. W etapie drugim wybór konkretnego wyrazu dokonany zostaje dzięki informacji zstępującej i ogólnej lingwistycznej wiedzy odbiorcy. Ostatni fonem, w linearnym ciągu elementów początkowych, który odróżnia dany wyraz od wszystkich pozostałych wyrazów kohorty autorzy nazwali punktem rozpoznania wyrazu. Na podstawie analizy błędów percepcyjnych typu przesłyszzeń Browman [5] także stwierdziła sekwencyjne przetwarzanie fonematyczne w obrębie wyrazu, przy czym określone elementy segmentalne różnią się siłą informacyjną. Punkty promienności wyrazowej według autorki stanowią: inicjalna spółgłoska, odcinek VC sylaby wygłosowej oraz sylaba akcentowana. W doświadczeniach nad wywoływaniem wyrazów z leksykonu wewnętrznego na podstawie dostępnej, tylko fragmentarycznej informacji (początkowej, środkowej bądź końcowej) Nootboom [33] ustalił rolę poszczególnych fragmentów w dekodowaniu jednostek leksykalnych. Mimo, że inicjalne fragmenty są znacznie bardziej efektywne jako wskazówki wywołujące określony wyraz z pamięci, to jednak dostęp do leksykonu wyłącznie w postaci końcowego fragmentu wyrazu może także spowodować prawidłową reakcję. Badania ilości informacji akustyczno-fonetycznej wykorzystywanej przez odbiorcę od początku wyrazu do punktu, w którym zostaje on wydzielony spośród wszystkich jednostek przeprowadzone przez Grosjeana [14] wykazały, że chociaż punkt izolacji występuje w wyrazie odpowiednio wcześniej (podobnie jak punkt rozpoznania u Marlsena-Wilsona), to słuchacz w dalszym ciągu kontroluje nadchodzącą informację fonetyczną, aż do momentu akceptacji swej wcześniejszej decyzji leksykalnej. Pełna akceptacja danego wyrazu może nastąpić znacznie później, nawet po odebraniu następnego wyrazu.

Spośród bardzo wielu systemów automatycznego rozpoznawania mowy wyróżnić można takie, które w swych założeniach teoretycznych wykorzystują w znacznym stopniu informacje o procesie percepcji naturalnej. Klatt [17, 18] opracował system LAFS (Lexical Access From Spectra), w którym oparł się na hipotezie, że w procesie przetwarzania sygnału mowy przez człowieka etap analizy fonetycznej może być stosowany tylko uzupełniająco i nie jest bezwzględnie konieczny. Według autora można zbudować adekwatny system przetwarzania informacji lingwistycznej stosując wyłącznie procedurę bottom-up, a dostęp do leksykonu jest w postaci psychoakustycznej. System analizy fonetycznej, istniejący poza LAFS - w układzie generowania mowy przez człowieka - bywa w procesie percepcji mowy wykorzystywany jako układ dodatkowy w przetwarzaniu nowych, niezmagazynowanych w leksykonie wyrazów.

Bezpośrednie odwzorowanie między prawie ciągłą informacją akustyczną a reprezentacją leksykalną w formie morfemów bądź wyrazów opracował także Marcus [24, 25] w swoich układach symulacyjnych: ERIS I i ERIS II. Podobnie jak Klatt autor wprowadza etap pośredni między informacją akustyczną a płaszczyzną leksykalną w wypadku rozpoznawania nowych wyrazów bądź jednostek nieznaczących o strukturze wyrazu. Marcus zakłada występowanie trzech następujących etapów: względnie pasywnej fazy wyrazowo-inicjalnej - ustalenie początku wyrazu na podstawie pozytywnej odpowiedzi z bodźcem, aktywnej fazy rozpoznawania - wykorzystującej negatywną i pozytywną informację dla bardziej prawdopodobnych kandydatów wyrazowych oraz etapu końcowego, w którym znaleziony zostaje koniec wyrazu i początek następnego. W fazie końcowej wykorzystana zostaje także informacja prozodyczna. Przegląd problematyki dotyczącej wzajemnych relacji pomiędzy automatycznym rozpoznawaniem mowy a percepcją naturalną przedstawiła Vaissiere [40].

1.2. Leksykon wewnętrzny.

Pod pojęciem leksykonu wewnętrznego rozumie się najczęściej nabyty w procesie akwizycji języka oraz w procesie uczenia się inwentarz wyrazów, zmagazynowany w pamięci długotrwałej, uporządkowany według różnych kryteriów lingwistycznych. Ponieważ informacja zawarta w leksykonie umysłowym potrzebna jest w procesie

mówienia, słuchania, czytania, pisania i dokonywania lingwistycznych ocen dla poszczególnych funkcji występują różne metody magazynowania informacji. Dla mówienia i słuchania sugeruje się nawet obecność dwóch odrębnych leksykonów (Klatt [18]). Reprezentacja wyrazu w leksykonie winna być abstrakcyjna w takim stopniu, by dla każdej konkretnej reprezentacji wyrazowej stanowić mogła adekwatny odpowiednik i zarazem winna być na tyle konkretna, by można ją było porównać z pojawiającym się sygnałem akustycznym na wejściu [40]. Najbardziej prostą formą leksykonu stanowi nieuporządkowany zbiór wszystkich wyrazów, z których każdy ma własny detektor (ze Forsterem [9]). Wickelgren [41] stwierdza, że wyrazy występują w pamięci jako nieuporządkowane zbiory allofonów z informacją o ich bezpośrednim obustronnym kontekście fonetycznym. Najczęściej przyjmuje się co najmniej dwie reprezentacje jednostek w magazynie wewnętrznym: według kryteriów semantycznych oraz według kryteriów fonologicznych³. W leksykonie zawarta jest także informacja gramatyczna (syntaktyczna). W "przeszukującym" modelu Forstera [9,10] wyodrębnione zostały następujące kartoteki dostępu bezpośredniego: ortograficzna, fonologiczna i semantyczno-syntaktyczna. We wszystkich kartotekach, składających się na wspólną kartotekę główną, jednostki uporządkowane są według częstości ich występowania w języku. Te same wyrazy mogą znajdować się w kilku różnych kartotekach. Według Fromkin [12,13] magazyn jest odpowiednikiem "tezaurusu" o drabinkowej strukturze, który zawiera cechy i adresy. Układ semantyczny występuje według kryteriów syntaktycznych (np. rzeczowniki, czasowniki itp. są grupowane osobno). Leksykon ponadto zawiera listę rdzeni, afiksów, idiomów, złożeń itp. Wyrazy podobnie brzmiące mogą być gromadzone w jednym podzbiore, podobnie jak osobno oddzielone są w pamięci wyrazy krótkie od długich (wielosylabowych) [2]. Zgodnie z teorią Marlsona-Wilsona i Welsha [27] oraz Cole'a, Jakimik i Coopera [8] jednostki w leksykonie pogrupowane są w podzbiory (kohorty) o takich samych fragmentach początkowych wyrazu.

³ W opracowaniu MIT [4] założono, że wyrazy są reprezentowane za pomocą morfemów, które z kolei występują w terminach fonemów.

1.3. Analiza fonetyczno-fonologiczna.

Bez względu na przedstawione w paragrafie 1.1. hipotezy dotyczące wzajemnych relacji pomiędzy przetwarzaniem wyrazów i fonemów, znaczna większość badaczy przyjmuje reprezentację fonetyczną (częściej fonologiczną) jednostek leksykalnych w pamięci długotrwałej człowieka. W metodologii badań nad leksykalnym dostępem stosuje się najczęściej trzy procedury testowe: naśladowanie (shadowing) określonych fragmentów tekstu, monitorowanie w prezentowanym materiale lingwistycznym wybranych z góry elementów fonologicznych oraz analizę błędów typu przesłyszeń (slip of the ear). W dwóch ostatnich analiza fonologiczna wysuwa się na plan pierwszy. Badacze w swoich doświadczeniach odwołują się do nieświadomej, bądź tylko częściowo uświadamianej wiedzy słuchacza o jednostkach segmentalnych mowy. Liczne eksperymenty psychofonetyczne wykazały, że wiedza ta zgromadzona w pamięci długotrwałej odnosi się do: prawidłowego rozpoznawania inwentarza fonemów danego języka, umiejętności klasyfikowania dźwięków mowy na podstawie subiektywnych (zdobytych w procesie akwizycji języka) oraz obiektywnych - lingwistycznych (przyswojonych w procesie dydaktycznym) kryteriów podobieństwa i odrębności pomiędzy dźwiękami mowy, a także relacji fonotaktycznych w linearnych ciągach wypowiedzi określonego języka.

Stawia się również hipotezę, że niezależnie od procesu dydaktycznego, użytkownik języka posiada "wiedzę" o odrębności pomiędzy spółgłoskami a samogłoskami, ponieważ w procesie percepcji mowy rozróżnienie spółgłoska - samogłoska ma znaczenie zasadnicze. Odrębność przetwarzania obu klas dźwięków występuje już w analizie peryferyjnej i na poziomie wstępnego przetwarzania mowy. Badania nad funkcjonalną asymetrią kory mózgowej wykazały znaczną przewagę lewej półkuli w przyjmowaniu informacji lingwistycznej nad półkulą prawą. Stwierdzono ponadto, że samogłoski zostają zdekodowane zanim napływający sygnał osiągnie dominującą dla zjawisk mowy półkulę, a informacja o nich zostaje przechowywana na przeciąg 200 - 300 ms w prekategoriálním magazynie słuchowym. W przeciwieństwie do spółgłosek, w naturalnej sytuacji odbioru komunikatu językowego, percepcja samogłosek jest znacznie słabiej związana z lingwistycznymi kategoriami przetwarzania fonetycznego.

Obszerniejszy przegląd wymienionych różnic zawarto w pracy [22]. W procesie porozumiewania się za pomocą języka naturalnego istnieją jednakże takie sytuacje, w których percepcja dźwięków wokalicznych nie zostaje ograniczona do "ciągłego" przetwarzania audytorycznego, podlega także bezwzględnej identyfikacji czyli przetwarzaniu kategorialnemu, ale na podstawie odmiennych cech klasyfikacyjnych niż ma to miejsce w wypadku spółgłosek.

2. Testy psycholingwistyczne.

W dalszym ciągu omówione zostaną dwa doświadczenia psycholingwistyczne, w których przedmiotem analizy będą interakcje pomiędzy przetwarzaniem fonetyczno-fonologicznym a leksykalnym. W obu testach zarówno elementy segmentalne mowy - polskie samogłoski, jak i jednostki leksykalne potraktowane zostały odmiennie. Doświadczenie pierwsze zostało tak skonstruowane by wyraz przetwarzany był jako "gestalt". Przedmiotem ocen były kształty dźwiękowe nadawanych wyrazów, różniących się wyłącznie elementem wokalicznym. Odwoływano się głównie do fonetycznej wiedzy użytkownika języka w zakresie jego zdolności kategorialnych ocen podobieństwa dźwięków mowy, a nie do semantyczno-syntaktycznej informacji o nadawanych sygnałach, zawartych w leksykonie wewnętrznym.

Eksperyment drugi był znacznie bardziej złożony. Jak wykazano w § 1.2. istnieje wiele hipotez dotyczących sposobu magazynowania informacji w leksykonie wewnętrznym. W niniejszej pracy zakłada się, że poszczególne podzbiory jednostek leksykalnych, np. według wspólnych fonemów początkowych, według podobieństwa fonetycznego całych wyrazów, według podobieństwa semantycznego, czy też według określonych kryteriów syntaktycznych mogą być tworzone ad hoc w zależności od konkretnej sytuacji komunikacyjnej na podstawie wieloaspektowej informacji zgromadzonej w leksykonie podstawowym, w którym wyrazy występują jako hasła wraz z wszelkimi możliwymi informacjami ilościowymi i jakościowymi. Uporządkowanie haseł następuje dopiero w momencie wyodrębnienia się określonego czynnika lub kilku czynników porządkujących, stanowiących drogi dostępu dla napływającej informacji komunikatu.

Przedmiotem testu było skojarzenie usłyszanego sygnału z jednym wyrazem polskiego języka. Wyrazy mogące stanowić potencjalną reakcję słowną na każdy bodziec tworzyły kilkuelementowe kohorty,

ponieważ różniły się tylko jedną samogłoską, a pozostałe elementy segmentalne posiadały takie same. Np.

warze, ważę	/vaʒe/
worze, wożę	/voʒe/
wirze	/viʒe/
wyrze, wyże	/vɨʒe/
węże	/vɛʒe/

Wyrazy w podzbiorach posiadały podobny kształt dźwiękowy, ale stanowiły odmienne typy syntaktyczne (występowały w dowolnych formach gramatycznych różnych części mowy), nie były uporządkowane według jakichkolwiek kryteriów semantycznych, co więcej jeden kształt dźwiękowy często był związany z dwoma znaczeniami.

We wstępnej części doświadczenia ilościowemu oszacowaniu poddano subiektywną częstość występowania wyrazów w każdej kohorcie.

Bodźcami wywołującymi reakcje słowne słuchaczy były wypowiedzi o rozciągłości wyrazu, w których wyeliminowany został zmienny element samogłoskowy tworzący kohortę (dla podanego powyżej podzbioru wyrazów bodziec był następujący : /vʒe/). Niepełna informacja fonetyczno-fonologiczna powodowała niejednoznaczność interpretację leksykalną. Przedmiotem analizy w niniejszej pracy było ustalenie roli niektórych czynników lingwistycznych na wywołanie wyrazu z leksykonu wewnętrznego. Zbadano wpływ informacji fonotaktycznej, subiektywnej częstości występowania wyrazów w podzbiorze oraz formy gramatycznej wyrazu.

2.1. Test subiektywnego podobieństwa samogłosek.

Głównym celem przeprowadzonego eksperymentu psychofonetycznego było sprawdzenie hipotezy, w jakim zakresie kategoryjalne oceny podobieństwa mają zastosowanie dla percepcyjnej klasyfikacji samogłosek.

Procedura testowa w znacznym stopniu jest, z założenia, powtórzeniem metodologii eksperymentalnej przeprowadzonej dla inwentarza spółgłoskowych fonemów języka podlaskiego [21]. Elementy samogłoskowe porównano ze sobą za pomocą metody triadowej ABC. Zadaniem słuchaczy było określenie, które dwa z trzech kolejno nadawanych sygnałów są do siebie bardziej podobne niż pozostałe. Spełniono warunek, by liczba porównań między każdą parą samogłosek była taka sama. Ponieważ przedmiotem badania był sześć-elementowy zbiór elementów wokalicznych zasada ta została spełniona przy zastosowaniu 20 potrójnych sygnałów. Słuchacze

dokonywali porównania jednosylabowych wyrazów znaczących z badanym elementem samogłoskowym w śródgłosie. Wyrazy były następujące : bit, b̄it, bet, bat, bot, but, a ich układ w materiale testowym przedstawiono w skrócie poniżej :

1. bat, bet, bot,
2. bet, b̄it, bat,

20. bit, bot, bet.

Porównanie kształtu dźwiękowego wymienionych wyrazów przeprowadzono przy zastosowaniu następującej instrukcji :

Usłyszycie bezpośrednio następujące po sobie trzy wyrazy, po których nastąpi krótka przerwa. Na formularzach ankietowych każdemu wyrazowi odpowiada jedna kratka. Jeśli waszym zdaniem są do siebie bardziej podobne wyrazy 1 i 2, połączcie kreską pierwsze dwie kratki, jeśli 1 i 3, kreskę należy przeprowadzić przez wszystkie trzy kratki, jeśli natomiast bardziej podobne są wyrazy 2 i 3, kreska powinna łączyć kratkę drugą i trzecią. Ocen podobieństwa należy zawsze dokonać tuż po usłyszeniu sygnału, bez zastanawiania się, ponieważ czas na decyzję jest bardzo krótki.

Ponieważ słuchacze nie posługiwali się ani transkrypcją fonetyczną ani formalną ortografią, wyeliminowano możliwość sugerowania się przy ocenie nawykami graficznymi.

W teście odsłuchowym wzięło udział 37 osób z normalnym słuchem (osoby dorosłe oraz licealiści starszych klas). Doświadczenie przeprowadzono dwukrotnie w odstępie kilku dni. W celu porównania zgodności ocen dla pierwszej i drugiej części testu zastosowano kryterium χ^2 . Słuchacze dokonywali nie przypadkowych ocen podobieństwa w 80 % dla pierwszej części testu, oraz w 83 % w drugiej części (poziom istotności $\alpha = 0.05$). Oceny przypadkowe dokonywane były dla następujących trójek wyrazów :

- b̄it, bet, but
- bet, bat, bot.

W dalszej interpretacji rezultatów wykorzystano wyniki drugiej części testu, które stanowiły punkt wyjścia do konstrukcji symetrycznej tablicy odległości samogłosek. Odległość między dwoma elementami samogłoskowymi rozumiana jest jako wielkość probabilistyczna, wyrażona wzorem :

$$d_{i,j} = 1 - \frac{f_{i,j}}{N}$$

gdzie : $f_{i,j}$ oznacza częstość reakcji preferujących skojarzenie i,j zaś N jest liczebnością wszystkich reakcji, w których możliwy był wybór i,j . Tak więc wielkość $\frac{f_{i,j}}{N}$ jest prawdopodobieństwem preferencji i,j czyli $P_{i,j}$. Przyjęta miara odległości stanowi zdarzenie przeciwne do $P_{i,j}$, czyli $d_{i,j} = 1 - P_{i,j}$ w sensie addytywności.

Symetryczna tablica odległości przedstawia się następująco :

	i	ɛ	e	a	o	u
i	-	0,284	0,750	0,872	0,872	0,818
ɛ		-	0,736	0,831	0,851	0,662
e			-	0,473	0,568	0,750
a				-	0,446	0,682
o					-	0,391
u						-

Na podstawie wyznaczonych odległości można wyjaśnić przypadkowość ocen dla dwóch triad wymienionych na str. 11. W pierwszej triadzie dwie odległości są prawie takie same : băt-bet (0,756), a bet-but (0,750). Również trzecia odległość jest znaczna i wynosi 0,662 dla pary : băt-but. Dla drugiej triady odległości nie są zróżnicowane, wynoszą odpowiednio : bet-bat (0,473), bet-bot (0,568) i bot-bat (0,446) i świadczą o większym podobieństwie porównywanych elementów.

Uzyskane w tablicy wartości dla poszczególnych par samogłosek porównano z odległościami Mahalanobisa między tymi samymi elementami, wyznaczonymi na podstawie kryteriów akustycznych (częstotliwości pierwszych czterech formantów, por. [15]). Największe odległości występują zgodnie w obu tablicach dla /i,a/ oraz /i,o/. Najmniejsza odległość percepcyjna dla /i,ɛ/ również potwierdzona została wynikami klasyfikacji akustycznej. Rozbieżności stwierdzono dla pary /ɛ,e/, w której z akustycznego punktu widzenia oba elementy są bardziej do siebie podobne, niż wynika to z przeprowadzonego testu odsłuchowego. Percepcyjnie samogłoska /ɛ/ wykazuje znacznie większe podobieństwo do /u/ niż do /e/. Samogłoska /a/ jest akustycznie najbardziej podobna do samogłoski /o/, podczas gdy klasyfikacja percepcyjna wykazała większe podobieństwo /o/ do /u/ niż do /a/.

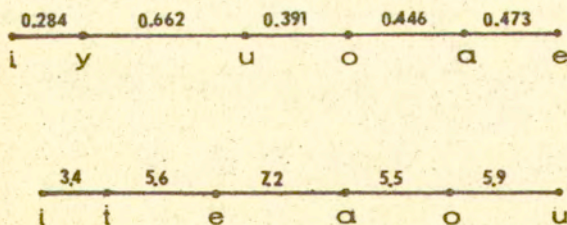
Rozbieżności w obu pracach wyjaśnić można następująco. Analizie akustycznej poddano dźwięki samogłoskowe wymówione w izolacji, zaś w porównaniu percepcyjnym elementy te umieszczone w jedno-

rodnym kontekście spółgłoskowym w wyrazach znaczących. Spowodowało to prawdopodobnie następujące typy decyzji badanych osób :

- 1/ według kryterium podobieństwa kształtów dźwiękowych nadawanych sygnałów (zgodnie z apriorycznym układem testu),
- 2/ według subiektywnej częstości występowania wyrazów w języku stanowiących sygnały testowe,
- 3/ według podobieństwa semantycznego bodźców.

Odbiciem kryterium pierwszego jest niewątpliwie najmniejsza odległość w tablicy pomiędzy /i/ oraz /ɨ/ (wyrazy : bit, bɨt). Mała odległość dla pary /ɨ,u/ wskazuje na wpływ częstości wyrazowej w podejmowaniu decyzji słuchaczy, ponieważ wyrazy but oraz bɨt otrzymały w teście ocen częstości wyrazowej rangi 1 i 2 (por. niżej 2.2.1). Podobieństwo semantyczne wyrazów but i bot wzmocniło prawdopodobnie decyzję opartą na kryterium brzmienia dla odległości /u,o/.

Symetryczna tablica odległości posłużyła do konstrukcji dendrytu taksonomicznego (por. ryc. 1a)⁴. Dendryt stanowi dogodne odwzorowanie na płaszczyźnie współzależności zachodzących między testowanymi elementami. Na ryc. 1b zamieszczono analogiczny dendryt zacytowany z pracy [15], oparty na odległościach wyznaczonych za pomocą kryteriów akustycznych.



Ryc. 1. Dendryty taksonomiczne samogłosek polskich
a/ odległości percepcyjne,
b/ odległości akustyczne.

W obu metodach otrzymano dendryt liniowy. Istotna różnica pomiędzy nimi dotyczy połączeń /ɨ,u/ (kryteria percepcyjne) oraz /ɨ,e/ (kryteria akustyczne).

⁴ Dendryt wykonano według metody opisanej przez Perkala [37].

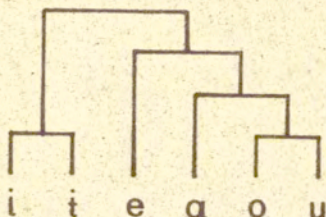
W dalszym ciągu przedstawiono sześć metod grupowania samogłosek na podstawie subiektywnych odległości określających miarę ich wzajemnego podobieństwa. W pracach z tego zakresu [19, 42] punktem wyjścia dla każdej z opisanych metod grupowania jest trójkątna macierz odległości euklidesowych. Dla celów niniejszej pracy program i algorytm opublikowany przez Karońskiego i Calińskiego [16] został przystosowany do odległości zdefiniowanych probabilistycznie.

Wspólną cechą zaprezentowanych niżej strategii skupień jest traktowanie, na początku analizy, wszystkich obiektów jako jednoelementowych grup. W każdym kolejnym kroku dokonuje się tylko jednego połączenia bądź dwu obiektów, bądź jednego obiektu z już poprzednio utworzoną grupą, bądź wreszcie dwu grup. We wszystkich metodach połączeniu w grupy podlegają te obiekty albo grupy poprzednio utworzone, dla których odległość jest najmniejsza. Metody różnią się wyłącznie statystycznymi definicjami odległości między grupami.

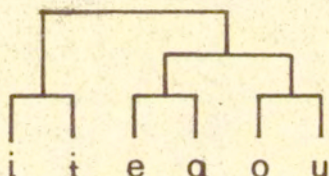
Definicje te są następujące :

1. W metodzie najbliższego sąsiedztwa - odległość między najbliższymi elementami (obiektami) pochodzącymi z dwu grup,
2. W metodzie najdalszego sąsiedztwa - odległość między parą najbardziej odległych elementów każdej z grup,
3. W metodzie mediany - odległość wyrażona dystansem pomiędzy medianami wyznaczonymi dla dwóch grup,
4. W metodzie środka ciężkości - dystans między środkami ciężkości dwu grup,
5. W metodzie średniej grupowej - odległość jako średnia wartość dystansów pomiędzy poszczególnymi elementami dwu grup,
6. W metodzie Warda - odległość między sumami kwadratów odchyleń wszystkich elementów w grupach od środków ciężkości grup, do których należą.

Metody : najbliższego sąsiedztwa, mediany oraz środka ciężkości dały w wyniku identyczną klasyfikację :



Pogrupowanie za pomocą pozostałych trzech metod : najdalszego sąsiedztwa, średniej grupowej oraz metody Warda także jest wspólne :



W obu przedstawionych powyżej klasyfikacjach odrębną grupę stanowią samogłoski /i/ oraz /ɨ/, które we wszystkich metodach wydzielone zostają w pierwszym kroku klasyfikacyjnej procedury. Pozostałe samogłoski są bądź łączone w pary : /o,u/ i /e,a/, bądź też tworzą sukcesywnie większe grupy. Bardziej zrównoważonemu podziałowi według metod najdalszego sąsiedztwa, średniej grupowej i Warda można przypisać następujące tradycyjne cechy fonetyczne :

- 1 - przednia/nieprzednia,
wysoka/niewysoka
- 2 - tylnia/nietylnia
- 3 - wysoka/niewysoka

Na poziomie pierwszego podziału działają równocześnie dwie cechy według tradycyjnego opisu fonetycznego.

Wyniki tego testu potwierdzają słabość opozycji fonologicznej /i-ɨ/ ponieważ oba fonemy oddziela od pozostałych tylko jeden poziom.

Przedstawione powyżej strategie grupowania dały odmienne rezultaty od istniejących wcześniej klasyfikacji, ponieważ w analizie percepcyjnej podstawę stanowiły oceny subiektywne a nie parametry fizyczne. Odrębność wyników uwarunkowana została ponadto interakcją czynników pozafonetycznych wynikających z jednoczesnego przetwarzania sygnałów na poziomie leksykalnym.

2.2. Test kojarzenia wyrazów.

Zgodnie z ogólnymi założeniami testu zreferowanymi w § 1.2. zasadnicze doświadczenie psycholingwistyczne poprzedzone zostało specjalnym przygotowaniem materiału językowego, które obejmowało wyznaczenie subiektywnej częstości występowania wyrazów w zbiorach komutacyjnych (por. przykład, str.10) oraz procedurę wycinania segmentów samogłoskowych z wypowiedzi zawierających samogłoskę /e/ w miejscu wymiennego elementu wokalicznego.

2.2.1. Test subiektywnej częstości występowania wyrazów w szeregach komutacyjnych.

Przedmiotem ocen 50 osób była lista zawierająca 90 szeregów komutacyjnych wyrazów jedno i dwu sylabowych. Pełna ankieta obejmowała łącznie 406 jednostek leksykalnych. Liczba wyrazów w szeregach wahała się od 4 do 6 elementów.

224 wyrazy wystąpiły w 56 szeregach czteroelementowych, 110 wyrazów w 22 zbiorach pięcioelementowych oraz 72 wyrazy w 12 szeregach sześćelementowych.

Subiektywną częstość wyrazową oceniano dla każdego szeregu oddzielnie. Zadaniem osób ankietowanych było uporządkowanie wyrazów od najczęściej do najrzadziej używanego w języku. Przedmiotem oceny były wszystkie znaczenia z jakimi kojarzyła się określona postać wyrazowa (np. dla wyrazów : "warze" oraz "ważę" dokonywano łącznej oceny częstości ich użycia, ale tylko dla tej formy gramatycznej w jakiej wystąpiły w ankiecie).

W celu określenia częstości wyrazu w obrębie szeregu, jak również porównania odpowiednich rang w różnych szeregach zastosowano rangę ważoną ze zrównoważoną liczebnością szeregu, wyznaczoną według następującego wyrażenia :

$$\bar{R}_j = \frac{N \sum_{i=1}^N f_j R_i}{\sum_{i=1}^N \sum_{j=1}^N f_j R_i}$$

dla $j = 1, \dots, N$

gdzie :

N - liczba wyrazów w szeregu

R_i - ranga wyrazu według oceny uczestników testu $i=1, \dots, N$

f - częstość rangi R_i dla wyrazu j .

Każdy wyraz w szeregu charakteryzowany był za pomocą dwóch liczb :

- 1/ najczęstszej oceny dla wyrazu w grupie badanych osób oraz
- 2/ rangi ważonej.

Dla przykładu podanego na str. 10 wyznaczone wartości przedstawiają się następująco :

	(1)	(2)
vaze	1	0,59
voze	2	0,71
věze	3	1,15
vize	4	1,24
vāze	5	1,31

Wartości rangi ważonej wskazują na fakt nierównych odległości pomiędzy sąsiadującymi pozycjami w kolejności. W podanym przykładzie największa różnica występuje pomiędzy rangą drugą i trzecią (0,44). Zakresy zmienności dla poszczególnych rang wyrażone miarą średniej ważonej dla wszystkich testowanych szeregów nie są rozłączne, co ilustruje poniższe zestawienie :

1	0,33	-	0,88
2	0,55	-	1,16
3	0,65	-	1,31
4	0,95	-	1,58
5	1,26	-	1,62
6	1,39	-	1,69

Fakt częściowego nakrywania się rang dla wyrazów z różnych szeregów oraz nierównomiernych odstępów między kolejnymi pozycjami wyrazów w tym samym szeregu może w istotny sposób wpłynąć na reakcje słuchaczy w drugiej części doświadczenia.

2.2.2. Wycinanie samogłosek.

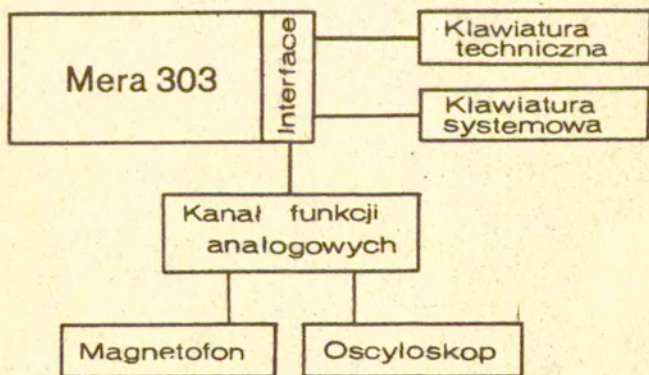
Każdemu szeregowi komutacyjnemu przyporządkowano jedną wypowiedź testową, przyjmując za sygnał testowy ten wyraz w szeregu, w którym wymienialnym elementem wokalicznym była samogłoska /e/

⁵ Zastosowana procedura obliczeniowa jest metodologicznym powtórzeniem pracy [21].

lub też tworząc logatom o zgodnych regułach fonotaktycznych z określonym szeregiem, który zawierał na wymiennej pozycji samogłoskę /e/. Dla cytowanego powyżej przykładu sygnałem testowym była wypowiedź /vege/. Wybór samogłoski /e/ podyktowany był najbardziej neutralną pozycją toru głosowego w czasie jej artykulacji, co w znacznym stopniu pozwoliło ograniczyć wpływ samogłoski na bezpośredni kontekst.

Bodźce testowe nagrane zostały przez trzy osoby. Do przetworzenia cyfrowego wykorzystano tylko jedną realizację każdej wypowiedzi.

Sygnał mowy po przetworzeniu analogowo-cyfrowym poddano filtrowaniu czasowemu w zadanej części pierwotnego przebiegu. Częstotliwość próbkowania wynosiła 15 kHz, a kwantyzacja amplitudy 8 bitów. Opracowana cyfrowo wypowiedź po konwersji cyfrowo-analogowej została zapisana na taśmie magnetofonowej. Schemat konfiguracyjny aparatury przedstawiono na ryc. 3.



Ryc. 3. Schemat blokowy układu do wycinania segmentów sygnału mowy.

Zestaw ten oprogramowano w języku wewnętrznym minikomputera. Program umieszczony w pamięci MERY był możliwie krótki i elastyczny, co pozwoliło na wyprowadzenie wypowiedzi długości pół sekundy. Pozwalał on operatorowi na elastyczną obsługę minikomputera po-

przez wciskanie i wyciskanie przydzielonych do określonych funkcji kluczy klawiatury technicznej oraz systemowej. Algorytmy programu wraz z opisem zamieszczono w dodatku.

Likwidowanie kolejnych próbek z sygnału było dobierane ręcznie w zależności od dostępnej informacji wzrokowej (na ekranie oscyloskopu) oraz słuchowej (w głośniku magnetofonu). Była to zatem segmentacja nie całkiem obiektywna, tym bardziej, że czynnikiem decydującym było wrażenie słuchowe⁶. Poniżej podano pięć przykładów wyeliminowania samogłoski /e/ z różnych kontekstów fonetycznych. Na ryc. 4a przedstawiono usunięcie samogłoski nagłosowej w wypowiedzi /erka/. Pozostawienie pierwszego okresu było zamierzonym sygnałem początku wypowiedzi. Ryc. 4b ilustruje usunięcie /e/ śródgłosowego z wypowiedzi /tek/. Ryc. 4c i 4d stanowią przykłady pozostawienia początkowych i końcowych fragmentów wokalicznych, dzięki którym otrzymano lepszą identyfikację sąsiednich spółgłosek. Ryc. 4e stanowi przykład całkowitego wycięcia pierwszej samogłoski w wyrazie /ferma/.

2.2.3. Test odsłuchowy.

Sygnały testowe nagrane w przypadkowej kolejności zaprezentowano słuchaczom, zadaniem których było skojarzenie sygnału z dowolną formą gramatyczną konkretnego wyrazu języka polskiego. W teście wzięło udział 30 osób. W przeważającej większości byli to uczniowie maturalnej klasy liceum ogólnokształcącego. Słuchacze udzielali pisemnych odpowiedzi. W swych ocenach posługiwali się w czasie sesji formularzem zawierającym schematyczną reprezentację bodźców, uwzględniającą strukturę spółgłoskowo-samogłoskową nadawanych sygnałów. Formularz był szczególnie pomocny w tych wypadkach, w których usunięcie z wyrazu fragmentu o rozciągłości 100-150ms⁷ nie stwarzało dostatecznie wyraźnego wrażenia braku dźwięku mowy. Oto dwa przykłady graficznej reprezentacji sygnałów /kr# t#/ oraz /s# s/ :

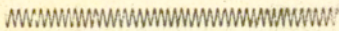


⁶ Na podobnych obiektywno-subiektywnych ocenach oparta była izolacja samogłosek z jednosylabowych wyrazów przeprowadzona przez Polsa [38].

⁷ Iloczas wszystkich wyeliminowanych z wypowiedzi samogłosek mieścił się w granicach : 100-150 ms.



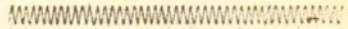
krak



Ryc. 4a



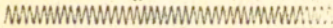
tk



Ryc. 4b



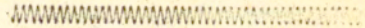
bt



Ryc. 4c



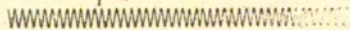
d3t



Ryc. 4d



frma



Ryc. 4e

Ryc. 4. Oscylogramy wypowiedzi z wyciętymi samogłoskami.

Liczba kratak odnosi się do liczby dźwięków. Kratka zakreskowana odpowiada miejscu wyciętej samogłoski, kratka z kółkiem informuje o pozycji w wyrazie pozostawionej w sygnale samogłoski. Sesję odsłuchową przeprowadzono dla grup pięcioosobowych, którym każdy sygnał prezentowano czterokrotnie co 5 sekund. Pierwsze odtworzenie wypowiedzi testowej oceniane było wyłącznie słuchem, drugie przeznaczone było na porównanie ze schematem na formularzu, trzecie na skojarzenie z wyrazem a czwarte na zapisanie odpowiedzi. Dla niektórych sygnałów odpowiedzi następowały już po dwukrotnym odtworzeniu, ale były i takie wypowiedzi, których czterokrotna prezentacja nie wywołała reakcji.

Przedstawione poniżej omówienie wyników testu ma charakter wstępny i opisowy. Uzyskane odpowiedzi wskazują na współwystępowanie co najmniej kilku zmiennych, o tak odrębnym charakterze, że wszelkie próby analizy statystycznej okazują się na obecnym etapie bezprzedmiotowe.

2.2.3.1. Poziom trudności testu.

Na 2700 możliwych odpowiedzi (30 osób x 90 sygnałów) uzyskano 2187 odpowiedzi spełniających wszystkie warunki testu, co stanowi 81 % całego zaprezentowanego słuchaczom materiału. W 123 przypadkach (niecałe 5 %) nie udzielono żadnej odpowiedzi. 390 odpowiedzi było błędnych (14 %). Analiza odpowiedzi nie spełniających wszystkich warunków testu wykazała, że zdecydowana większość błędów była wynikiem przesłyszenia. W 34 przypadkach uczestnicy doświadczenia nie potrafili dać odpowiedzi zgodnej ze schematem strukturalnym wyrazu i świadomie zapisywali inne najbliższe skojarzenie (np. dla sygnału /m#wa/ odpowiedź /mogwa/). W 288 wypowiedziach testowych błędnie została odebrana jedna spółgłoska, w 52 co najmniej 2 spółgłoski, a w 16 sygnałach obok błędów spółgłoskowych nastąpiło błędne zdekodowanie niewymiennej samogłoski. W grupie wyrazów z jedną niepoprawnie odebraną spółgłoską do najczęstszych błędów należały : substytucje /r - l/ 30 razy, /m - n/ 27 razy, /p - t/ 25 razy i /v - m/ 22 razy. Są to podstawienia spółgłosek najbardziej precepcyjnie podobnych (por. [21]). Mechanizm powstawania błędów w dekodowaniu wypowiedzi testowych potwierdza współwystępowanie dwóch czynników związanych z funkcjonowaniem fonologicznej pamięci długotrwałej :

1/ subiektywnej wiedzy o podobieństwie fonetycznym spółgłosek oraz

2/ informacji o częstości występowania ich w języku.

2.2.3.2. Wiedza fonetyczna.

W grupie uwarunkowań fonetycznych omawianego testu znajdują się czynniki kontekstowe i częstościowe wymieniających samogłosek. W konstrukcji szeregów komutacyjnych wykorzystano wszystkie fonemy samogłoskowe języka polskiego w większości kontekstów spółgłoskowych. Dokonano odstępstwa od przyjętych ustaleń w zakresie polskiej fonologii dotyczących tzw. samogłosek nosowych / \tilde{e} , \tilde{o} /. Nie tylko przyjęto ich występowanie jako jednostek fonologicznych przed spółgłoskami trącymi, jak np. w wyrazie / $v\tilde{e}ʒe$ / co jest zgodne z niektórymi opisami fonematycznymi, lecz także przed spółgłoskami zwartymi (np. w wyrazie "lęk"), zakładając wpływ wyuczonej wiedzy ortograficznej w procesie kojarzenia wyrazów na podstawie bodźców słuchowych.

Kontekst spółgłoskowy tylko w nielicznych wypadkach wprowadzał ograniczenia w liczbie komutujących elementów samogłoskowych. Po niektórych spółgłoskach nie mogła wystąpić samogłoska /i/, po niektórych samogłoska /ɛ/.

Częstość występowania poszczególnych samogłosek uwarunkowana była jedynie możliwością utworzenia określonych szeregów komutacyjnych i nie odzwierciedla ich występowania w mówionej formie języka polskiego. Poniżej podano zestawienie częstości występowania elementów wokalicznych w mówionych tekstach (kolumna 1), w słowniku testowym (kolumna 2) oraz w odpowiedziach słuchaczy (kolumna 3), wszystkie dane wyrażono w procentach :

	(1)	(2)	(3)
i	8,7	11,1	6,7
ɛ	9,9	12,6	11,2
e	27,3	12,8	8,3
a	24,8	20,5	30,2
o	20,5	16,5	20,3
u	7,3	18,3	20,6
ɛ	0,2	4,2	1,3
o	1,3	4,0	1,1

Wartości częstości występowania elementów w materiale słownikowym testu charakteryzują się mniejszą rozpiętością niż frakcje dla mowy ciągłej. Największe rozbieżności występują dla samogłosek : /e,u/. Częstość występowania /e/ została w materiale testowym znacznie ograniczona, natomiast /u/ uległa znacznemu podwyższeniu. W reakcjach słuchaczy zwraca uwagę częste użycie samogłoski /a/, co związane jest z nawykami ortograficznymi badanych osób (a jest pierwszą literą alfabetu), a także z występowaniem tej samogłoski w 28 wyrazach z rangą 1 (por. § 2.2.1.).

Samogłoska /e/ wystąpiła tylko w 8 wyrazach z rangą 1. Fakt ten tłumaczy niską frekwencję dla tego elementu w realizacjach słuchaczy. Kryterium częstości wyrazowej nie znalazło odbicia we frakcjach testowych samogłosek nosowych. Dla / \tilde{o} / oraz / \tilde{e} / wystąpiło łącznie 5 wyrazów z rangą 1. W ankietach słuchaczy tylko wyraz / \tilde{v} osci/ otrzymał najwięcej odpowiedzi dla danego szeregu. Większość słuchaczy stosowała samogłoski nosowe jako elementy wymienne kierując się bardziej nawykami ortograficznymi niż faktyczną realizacją fonetyczną tych samogłosek w niektórych kontekstach. Podsumowując stwierdza się, że na odpowiedzi słuchaczy wpływ miały następujące czynniki związane bezpośrednio z realizacją określonej samogłoski :

- 1/ świadomość inwentarza samogłosek (szczególnie ustnych)- wszystkie osoby wykorzystywały pełny zbiór elementów,
- 2/ częstość występowania samogłosek skorelowana z częstością występowania wyrazów.

2.2.3.3. Czynniki gramatyczne.

W niniejszym paragrafie omówiony zostanie mechanizm korzystania przez słuchaczy z określonych form fleksyjnych wyrazów. Konstrukcja testu dopuszczała występowanie wyrazów w dowolnej formie gramatycznej. Można się było spodziewać, że w doświadczeniu polegającym na skojarzeniu sygnału dźwiękowego z aktualnym wyrazem języka, bez jakichkolwiek wskazówek kontekstu leksykalnego, bodźce wywoływać będą z magazynu pamięci długotrwałej poszczególne odpowiedzi w postaci hasłowo-słownikowej znacznie częściej niż w tzw. formach zależnych. Dla celów niniejszej pracy przyjęto, że tzw. formy podstawowe stanowią wszystkie rzeczowniki w mianowniku liczby pojedynczej i mnogiej, przymiotniki w mianowniku liczby pojedynczej oraz wszystkie czasowniki w czasie

teraźniejszym (najczęściej w 3 osobie l.p.). Poniżej podano strukturę gramatyczną słownika testowego (1) oraz odpowiedzi słuchaczy (2), dane w % :

	(1)	(2)
rzeczowniki (mian. l.p.)	35,3	46,9
rzeczowniki (mian. l.m.)	23,5	27,1
przymiotniki (mian. l.p.)	4,0	6,0
czasowniki (czas teraźn.)	7,1	6,4
pozostałe części mowy	30,1	13,7

Użycie poszczególnych form gramatycznych w odpowiedziach nie tylko wskazuje na hasłowy (uporządkowany w formach podstawowych) układ wyrazów w leksykonie wewnętrznym, ale i na łatwiejszy dostęp do rzeczowników, które w magazynie pamięciowym stanowią prawdopodobnie odrębny podzbiór⁸.

2.2.3.4. Wpływ częstości występowania wyrazów.

Na oddziaływanie częstości wyrazów zwrócono już uwagę w § 2.2.3.2. Pełniejszy opis tego zjawiska nastrocza trudności, ponieważ nie można go oddzielić od innych czynników. Podstawę porównania stanowią rangi wyznaczone w doświadczeniu wstępnym (2.2.1.) oraz liczba identycznych reakcji u poszczególnych słuchaczy w grupie, która także może być podstawą do wyznaczenia rangi. Opracowanie statystyczne istotności różnic między rangami nie jest możliwe, ponieważ szeregi komutacyjne są zbyt krótkie, a wyznaczone średnie rangi ważone wykazały nierówne odległości między wyrazami o kolejnych rangach porządkowych.

Spśród 90 wyrazów stanowiących najczęstsze odpowiedzi dla poszczególnych szeregów komutacyjnych otrzymano :

- a/ 46 wyrazów zgodnych z jednostkami o randze 1
- b/ 23 wyrazy zgodne z jednostkami o randze 2
- c/ 21 wyrazów, dla których liczba identycznych reakcji była zgodna z dalszymi rangami.

Dla wyrazów w grupie (a) oraz (b) wyznaczono średnie odległości

⁸ Dla prób 10000 wyrazów tekstu mówionego i pisanego Zarebina [43,44] ustaliła, że rzeczowniki stanowią 44 % wszystkich wyrazów w mowie i 45 % wyrazów w piśmie. Materiał testowy poddany ocenie słuchaczy łącznie zawierał ponad 60 % rzeczowników, a w odpowiedziach słuchaczy rzeczowniki stanowiły aż 74 %.

między średnimi rangami ważonymi dla pozycji pierwszej i drugiej w szeregu komutacyjnym. Odległości wynoszą : dla wyrazów z grupy (a) 0,37, a dla wyrazów z grupy (b) 0,21. Nasuwa się hipoteza, że świadomość zróżnicowania między częstością występowania wyrazów użytkowników języka nie jest zbyt ostra. Znaczna liczba wyrazów w grupie (c) jest tego dalszym dowodem.

Dominującymi czynnikami oddziaływującymi na reakcje skojarzeniowe słuchaczy były : leksykalna produktywność samogłoski /a/ oraz bezpośredni dostęp do kategorii rzeczowników w leksykonie wewnętrznym.

3. Uwagi końcowe.

Przedstawiono powyżej dwa eksperymenty psycholingwistyczne, dotyczące interakcji między przetwarzaniem na płaszczyźnie fonetycznej i leksykalnej. W doświadczeniu pierwszym w psychologicznej klasyfikacji elementów segmentalnych wykorzystano kształt dźwiękowy wyrazów, w doświadczeniu drugim badano mechanizm korekty błędów fonematycznych (brak samogłosek) w bloku wyrazowego przetwarzania informacji. W obu częściach pracy odwoływano się do struktury i zawartości pamięci długotrwałej (wiedzy lingwistycznej) użytkowników języka.

Wyniki obu testów potwierdziły hipotezę postawioną w § 1.3., że konkretna sytuacja komunikacyjna tworzy ad hoc dostęp do leksykonu wewnętrznego. W doświadczeniu pierwszym wyeliminowano wszystkie czynniki wiedzy lingwistycznej o wyrazach, poza kryterium ich semantycznego podobieństwa, którego udział w procesie subiektywnej klasyfikacji samogłosek można, na obecnym etapie badań, określić jako prawdopodobny. W drugim teście psycholingwistycznym połączono ze sobą oddziaływanie kilku czynników wiedzy językowej jak :

- 1/ świadomość inwentarza określonej klasy dźwięków mowy,
- 2/ znajomość struktury gramatycznej (syntaktycznej wyrazów),
- 3/ poczucie częstości występowania wyrazów w określonych podzbiorach.

Rola każdego czynnika z osobna potwierdzona została w wielu pracach innych autorów. Uwikłanie wszystkich tych czynników w jednym doświadczeniu stwarza bardziej naturalne warunki dla odbioru sygnału mowy, ale z drugiej strony stanowi istotną trudność w interpretacji uzyskanych rezultatów. W procesie

kojarzenia poszczególnych sygnałów testowych ze znaczącymi jednostkami języka, każda z osób posługiwała się odmiennymi wskazówkami otwierającymi dostęp do leksykonu wewnętrznego. Doświadczenie potwierdziło odwoływanie się słuchaczy do wszystkich wymienionych powyżej wskazówek. Na obecnym etapie badań - przy prezentacji nie związanych ze sobą semantycznie wyrazów - najczęściej wykorzystywana była informacja gramatyczna. Typ tej informacji (formy podstawowe wyrazów) sugeruje, że w omawianym doświadczeniu leksykon umysłowy był najczęściej porządkowany według kryteriów przyjętych dla większości słowników hasłowych. Osobną grupę w takim słowniku stanowiły rzeczowniki.

D O D A T E K

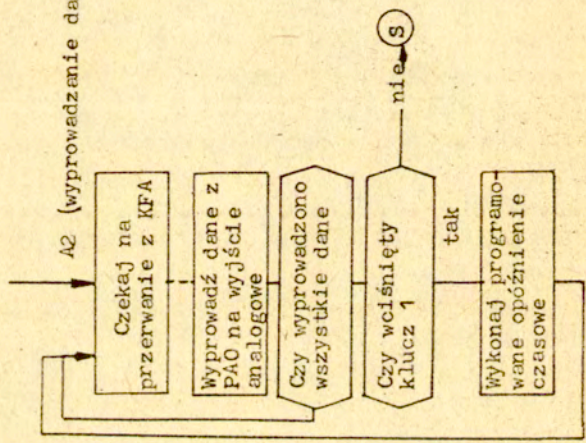
Algorytm programu wycinania segmentów sygnału mowy

Program sterujący zbudowany jest z kilku podprogramów realizujących następujące funkcje :

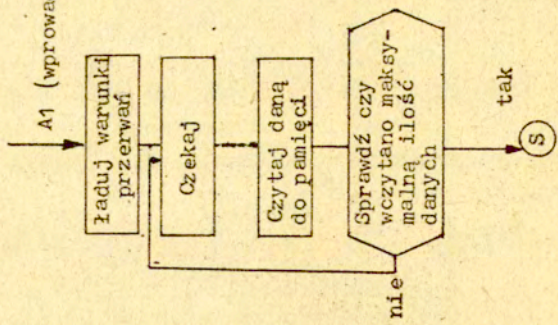
- a/ wprowadzanie do pamięci minikomputera sygnału mowy z magnetofonu,
- b/ wyprowadzanie sygnału z pamięci na wyjścia analogowe :
(oscyloskop, głośnik),
- c/ wycinanie wybranych segmentów sygnału mowy.

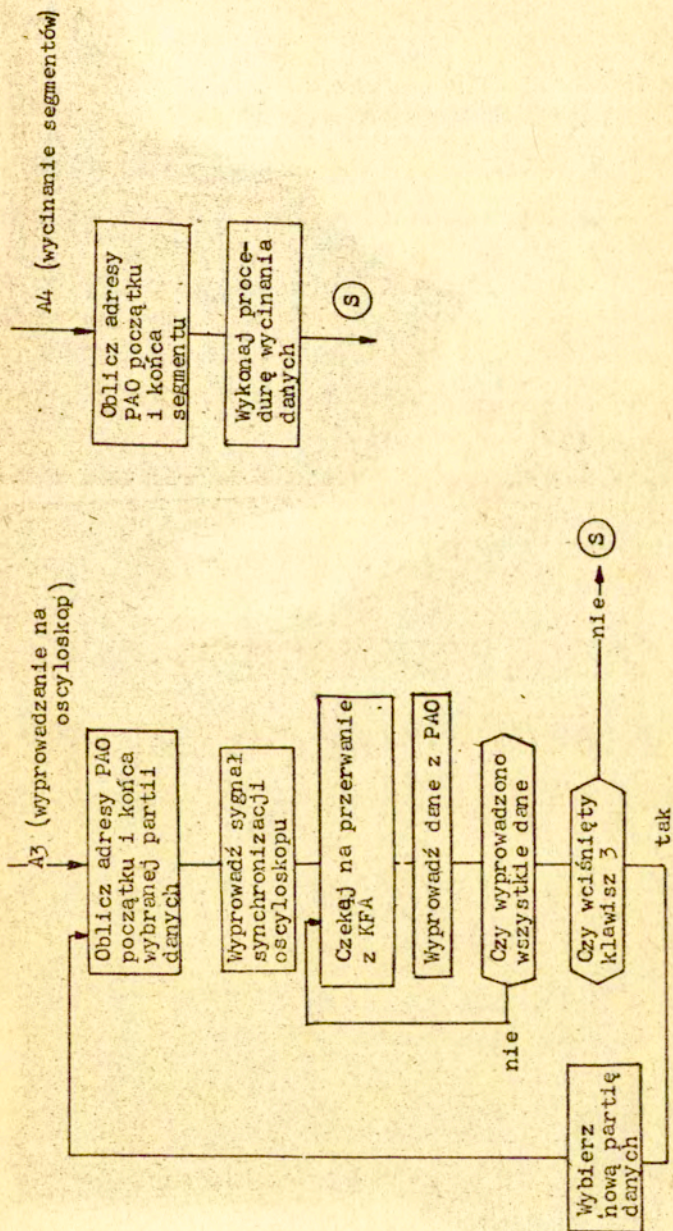
Wprowadzanie danych do pamięci następuje po przekroczeniu przez sygnał z góry zadanego progu tak, aby nie nastąpił start programu pod wpływem szumów. Program wyprowadzania na wyjście foniczne (magnetofon - odsłuch) wykonywany jest w dwóch kolejnych cyklicznie powtarzanych fazach, najpierw następuje wyprowadzenie sygnału na wyjście foniczne a później programowane czasowe opóźnienie mające na celu uzyskanie ciszy między kolejnymi wyprowadzeniami. Czas trwania ciszy jest nastawiany ręcznie przez wciśnięcie na klawiaturze technicznej odpowiedniej dla zadanego czasu wartości. Dla uzyskania stałej częstotliwości wyprowadzania zastosowano system przerwań z Kanału Funkcji Analogowych. Program wyprowadzania sygnału na oscyloskop umożliwia obejrzenie całego wprowadzanego uprzednio do pamięci sygnału lub też w zależności od układu wciśniętych klawiszy na klawiaturze technicznej jego kolejnych fragmentów. Długości wybieranych fragmentów sygnału mogą być różne w zależności od wciśniętej wartości na klawiaturze. Program wycinania umożliwia usunięcie części danych wpisanego do pamięci sygnału. Adresy w PAO początku oraz końca wycinanego segmentu ustawia się poprzez wpisywanie do programu uprzednio określonych wartości. Po procedurze wycinania danych wyprowadza się sygnał na wyjście foniczne lub oscyloskop. W przypadku nie uzyskania zadowalającego wyniku procedurę wycinania powtarza się modyfikując adresy początku oraz końca wycinanego segmentu.

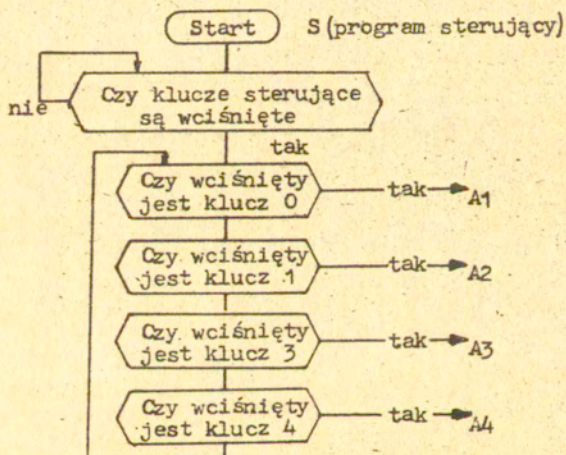
A2 (wyprowadzanie danych na wyjście foniczne)



A1 (wprowadzanie danych)







BIBLIOGRAFIA

- 1 AHLGREN, J., Interaction between form and content in linguistic perception, Monographs from the Institute of Linguistics, University of Stockholm, 1975.
- 2 BAKER, L., The lexicon : some psychological evidence, Working papers in Phonetics No. 26, U.C.L.A., 1974
- 3 BEVER, T.G., The influence of speech performance on linguistic structures, Advances in Psycholinguistics (D'Arcais G.B.F., Levelt W.J.M. red.), North-Holland, 4-30, 1970.
- 4 BLUMSTEIN, S.E., CONCIA, R., DELGUTE, B., GURLEKIAN, J.A., HALLE, M., HENKE, W.L., ISENBERG, D., KEYSER, S.J., KEATING, P.A., KLATT, D.M., OHDE, R.N., PERKELL, J.S., SENEFF, S., SHADLE, Ch.H., SHATTUCK-HUFNAGEL, S., STEVENS, K.N., WESTBURY, J.R., ZUE, V.W., Studies of speech production and perception, M.I.T. Progress Report No 122, 147-156, 1980.
- 5 BROWMAN, C.P., Tip of the Tongue and Slip of the Ear. Implications for Language Processing, Working Papers in Phonetics No 42, U.C.L.A., 1978.
- 6 COHEN, A., Speech, percepts, and Linguistic forms, Models for the Perception of Speech and Visual Forms (Wathen-Dunn W. red.), Cambridge, Massachusetts, London, 320-325, 1967.
- 7 COHEN, A., The word as a processing unit in speech perception, Progress Report, Institute of Phonetics University of Utrecht, Vol. 5, No 1, 33-47, 1980.
- 8 COLE, R.A., JAKIMIK, J., COOPER, W.E., Segmenting speech into words, JASA, Vol. 67, No 4, 1323-1332, 1980.
- 9 FORSTER, K.J., Accessing the mental lexicon, New Approaches to Language Mechanisms (Wales R.J., Walker E.C.T. red.), Amsterdam, North-Holland, 257-287, 1976.
- 10 FORSTER, K.J., Levels of processing and the structure of language processor, Sentence Processing (Cooper, W.E., Walker, E.C.T. red.), Lawrence Erlbaum Associates, Hillsdale, New Jersey, 27-85.
- 11 FOSS, D.J., SWINNEY, D.A., On the psychological reality of the phoneme : perception, identification and consciousness, Journal of Verbal Learning and Verbal Behaviour, Vol. 12, 246-257, 1973.
- 12 FROMKIN, V., Tips of the slung - or - to err is human, Working Papers in Phonetics, Vol. 14, 40-79, 1970.

- 13 FROMKIN, V.A. red., *Speech Errors as Linguistic Evidence*, Mouton, The Hague, 1973.
- 14 GROSJEAN, F., *Spoken word recognition processes and the gating paradigm*, *Perception and Psychophysics*, Vol. 28, No 4, 267-283, 1980.
- 15 JASSEM, W., KRZYŃSKO, M., DYCZKOWSKI, A., *Klasyfikacja samogłosek polskich na podstawie częstotliwości formantów*, *Prace IPPT* 64/72, 1972.
- 16 KAROŃSKI, M., CALIŃSKI, T., *Grupowanie obiektów wielocechowych na podstawie odległości euklidesowych*, *Roczniki AR w Poznaniu*, LXIV, *Algorytmy Biometryczne i Statystyczne*, zesz. 2, 117-129, 1974.
- 17 KLATT, D.H., *Speech perception : a model of acoustic-phonetic analysis and lexical access*, *Journal of Phonetics*, Vol. 7, No 3, 279-312, 1979.
- 18 KLATT, D., *Lexical representations for speech production and perception*, *The Cognitive Representation of Speech* (Myers T., Laver, J., Anderson, J. red.), North-Holland, 11-36, 1981.
- 19 LANCE, G.M., WILLIAMS, W.T., *A general theory of classificatory sorting strategies*, *Hierarchical systems*, *Computer Journal* Vol. 9, 373-380, 1967.
- 20 LIEBERMAN, P., *Some effects of semantic and grammatical context on the production and perception of speech*, *Language and Speech*, Vol. 6, 172-183, 1963.
- 21 ŁOBACZ, P., *Percepcyjne klasyfikacja spółgłosek polskich*, *Prace IPPT* 57/79, 1979.
- 22 ŁOBACZ, P., *Processing and Decoding the Signal in Speech Perception*, *Prace IPPT* 5/81, 1981.
- 23 ŁOBACZ, P., JASSEM, W., *Fonotaktyczna analiza mówionego tekstu polskiego*, *Biuletyn PTJ*, XXXII, 179-197, 1974.
- 24 MARCUS, S.M., *Associative coding and word boundary detecting*, *IPO Annual Progress Report* No. 16, 49-56, 1981 a.
- 25 MARCUS, S.M., *ERIS : context-sensitive coding in speech perception*, *Journal of Phonetics*, Vol. 9, 197-220, 1981 b.
- 26 MARLSEN-WILSON, W.D., *Sequential Decision During Spoken Word Recognition*, *Psychonomic Society Meetings in San Antonio*, 1978, manuskrypt .

- 27 MARLSEN-WILSON, W.D., WELSH, A., Processing interactions and lexical access during word recognition in continuous speech, *Cognitive Psychology*, Vol. 10, 29-63, 1978.
- 28 MARUSZEWSKI, M., NOWAKOWSKA, M., Próba eksperymentalnego badania "przezroczystości słowa dla znaczenia", *Studia Psychologiczne*, Vol. 10, 1970.
- 29 MORTON, J., Interaction of information in word recognition, *Psychological Review*, Vol. 76, 165-178, 1969.
- 30 MORTON, J., A functional model for memory, *Models of Human Memory*, (Norman, D.A. red.), Academic Press, New York, 203-254, 1971.
- 31 NOOTEBOOM, S.G., More attention for words in speech communication research., *Frontiers of Speech Communication Research* (Lindblom B., Ohman, S. red.), Academic Press, London, New York, San Francisco, 203-211, 1979.
- 32 NOOTEBOOM, S.G., Speech rate and segmental perception on the role of words in phoneme identification, *The Cognitive Representation of Speech* (Myers, T., Laver, J., Anderson, J. red.), North-Holland, 143-150, 1981 a.
- 33 NOOTEBOOM, S.G., Lexical retrieval from spoken words : beginnings vs endings, *Journal of Phonetics*, Vol. 9, 407-424, 1981 b.
- 34 OTTEVANGER, J.B., Detection of mispronunciations in relation to the point where a word can be identified, *Progress Report, Institute of Phonetics University of Utrecht*, Vol. 6, No 2, 44-69, 1980.
- 35 OTTEVANGER, J.B., The function of the recognition point in the perception of isolated mispronounced words, *Progress Report, Institute of Phonetics University of Utrecht*, Vol. 6, No 2, 48-69, 1981.
- 36 OTTEVANGER, J.B., Word recognition in non-constraining context in comparison with recognition of isolated words, *Progress Report, Institute of Phonetics University of Utrecht*, Vol. 7, No 2, 41-56, 1982.
- 37 PERKAL, J., *Matematyka dla przyrodników i rolników*, Cz. II, PWN, Warszawa, 1963.
- 38 POLS, L.C.W., *Spectral Analysis and Identification of Dutch Vowels in Monosyllabic Words*, Institute for Perception TNO, Soesterberg, 1977.
- 39 REPP, B.H., On levels of description in speech research, *Status Report on Speech Research SR-65*, Haskins

Laboratories, 217-222, 1981.

- 40 VAISSIERE, J., Speech recognition programs as models of speech perception, *The Cognitive Representation of Speech* (Myers, T., Laver, J., Anderson, J. red.), North-Holland, 443-457, 1981.
- 41 WICKELGREN, W.A., Context-sensitive coding in speech recognition, articulation and development, *Information Processing in the Nervous System* (Leibovic, K.N. red.), Springer Verlag, 85-95, 1969.
- 42 WISHART, D., An algorithm for hierarchical classifications, *Biometrics*, Vol. 25, No 1, 165-170.
- 43 ZAREBINA, M., Rola wyrazów w słowniku i tekście, *Język polski*, L, 33-46.
- 44 ZAREBINA, M., Najczęstsze wyrazy polszczyzny mówionej, *Język polski*, I.I, 336-347.