

**Olga Matuszkin**

**PERCEPCYJNE  
GRANICE MIĘDZYSEGMENTALNE  
W POŁĄCZENIACH CVCVC  
ZE SPÓŁGŁOSKAMI [m, n, ŋ, r, l, w, j]**

**13 / 1982**

P. 269

**WARSZAWA 1982**

ISSN 0208-5658

Praca wpłynęła do Redakcji dnia 23 listopada 1981 r.

Zarejestrowana pod nr 13/1982



57061



Na prawach rękopisu

---

Instytut Podstawowych Problemów Techniki PAN

Nakład 130 egz. Ark.wyd.3,7. Ark.druk. 4,25.

Oddano do druku w kwietniu 1982 r.

Nr zamówienia 138/0/82 Z - 87

---

Warszawska Drukarnia Naukowa, Warszawa,  
ul.Śniadeckich 8

Olga Matuszkińska  
Pracownia Fonetyki Akustycznej  
IPPT PAN

PERCEPCYJNE GRANICE MIĘDZYSEGMENTALNE  
W POŁĄCZENIACH CVCVC ZE SPÓŁGŁOSKAMI  
[m, n, ŋ, r, l, w, j]<sup>1</sup>

Streszczenie

Ustalono granice percepcyjne dla fonemów [m, n, ŋ, r, l, w, j]. Miejsce tych granic i czas trwania minimalnych segmentów percepcyjnych zależy od rodzaju dźwięku i jego pozycji w logatonie. W nagłosie decyduje informacja zawarta we fragmentach środkowo-końcowych, w śródgłosie decyduje informacja zawarta w fragmentach początkowych badanych dźwięków. Stwierdzono dużą wartość informacji w końcowych ich fragmentach. Informacja dotycząca [ŋ], [j] i [w] jest zawarta w dużym stopniu w początkowych fragmentach następujących samogłosek. Stwierdzono informacyjną wartość elementu wokalicznego dla większości tych dźwięków w pozycji wygłosowej. Percepcja [r] jest ściśle związana ze strukturą polisegmentalną tego fonemu. Doświadczenie wykonano na materiale 1062 bodźców z udziałem 20 słuchaczy.

1. Wstęp.

Obecna praca stanowi kolejny etap badań nad percepcją spółgłosek polskich, a ściślej nad ustaleniem granic percepcyjnych pomiędzy spółgłoskami a samogłoską [e]. Wyniki próby ustalenia granic percepcyjnych dla spółgłosek trących zawiera praca [17], dla spółgłosek zwartych praca [13]. W roku bieżącym wykonano dodatkowy eksperyment, który miał uzupełnić i uściślić wyniki otrzymane dla spółgłosek trących oraz zwartych. Wyniki tego

<sup>1</sup> Praca wykonana w ramach problemu węzłowego 06.9.

eksperymentu zostaną opisane w osobnej publikacji.

Niniejsza praca jest poświęcona badaniu granic percepcyjnych dla następujących polskich fonemów :

- spółgłoska nosowa dwuwargowa [m],
- " nosowa przedniojęzykowo-zębowa [n],
- " nosowa dźwięczna-śródjęzykowa [ɲ],
- " boczna dźwięczna [ɹ],
- " drżąca przedniojęzykowo-dźwięczna [r],

oraz tzw. samogłoski niesylabiczne [j] i [w], które "mają odpowiednio te same cechy widmowe co samogłoski [i, u] ...

Istotna różnica między [j] i [w] z jednej strony a [i] oraz [u] z drugiej, polega jedynie na tym, że w pozycji, w której mogą wystąpić tak jedno jak i drugie [j, w] są krótsze. Jednakże bezwzględna długość [j, w] waha się w szerokich granicach" ([11], str. 116). Dokładny opis wymienionych głosek od strony parametrów artykulacyjno-akustycznych, a także cech dystyngtywnych umieszczono np. w pracach [3, 5, 8, 12, 14, 15], a także w odpowiednich rozdziałach prac [9, 10, 11, 21, 23]. Cele obecnego etapu pracy pokrywają się z zadaniami poprzednich etapów i są następujące :

1. ustalenie czy istnieją granice percepcyjne pomiędzy spółgłoską płynną a sąsiadującą samogłoską i jaki jest stopień zgodności z tzw. granicami fizycznymi pomiędzy dźwiękami ustalonymi na podstawie zapisów oscylograficznych i spektrograficznych,
2. sformułowanie reguł określających stopień ważności poszczególnych fragmentów dźwięku dla jego prawidłowej identyfikacji,
3. ustalenie czasu trwania minimalnego segmentu przebiegu akustycznego, który można zidentyfikować jako głoskę należącą do określonego fonemu, w danym przypadku jako jedną ze spółgłosek płynnych.
2. Materiał fonetyczny wykorzystany w eksperymencie składa się z logatomów o budowie CVCVC, w których jako C występują spółgłoski [m, n, ɲ, l, r, w, j], jako V - [e], pod względem parametrów fonetyczno-akustycznych najbardziej neutralna samogłoska polska. Badane były zatem następujące logatomy :

memem  
nenen  
nepep  
lelel  
rerer  
wewew  
jejej

Materiał nagrano na taśmie magnetofonowej w studiu bezechowym Pracowni Fonetyki Akustycznej w Poznaniu przy użyciu magnetofonu ZK-240 (19 cm/s) w wykonaniu jednej osoby (głos męski, niski, fonetycznie wyszkolony). Logatomy były odczytywane w jednakowym tempie, monotonnie przy  $F_0$  ok. 100 Hz.

### 3. Metoda przygotowania testów percepcyjnych.

3.1. Zapisane na taśmie logatomy zostały skopiowane na magnetofonach ZK-240 i Ferrograph, następnie taśmę rozcięto, utworzono pętlę z każdego logatomu oraz wgrano przed każdym logatodem impulsy startowe. Następnym etapem było przetwarzanie analogowo-cyfrowe przy użyciu Cyfrowego Rejestratora typ 7502<sup>1</sup>. Przy użyciu interface'u otrzymane taśmy perforowane mogły być wczytane do pamięci Cyfrowego Rejestratora, a następnie po 1000-krotnym obniżeniu prędkości przetwarzania cyfrowo-analogowego zostały wykonane oscylogramy za pomocą Rejestratora analogowego typ PSG-101.

3.2. Granice pomiędzy dźwiękami na oscylogramach przeprowadzono tradycyjnie (por. reguły podane w [1], [7]). W miejscach wątpliwych korzystano dodatkowo z pomocy spektrogramów wykonanych na aparaturze Kay Electric Sona-Graph. Wyznaczenie granic w logatomach zawierających [w] i [j] narażało pewne trudności, dlatego na otrzymanych w wyniku pracy wykresach w szczególnie wątpliwych przypadkach (także i dla innych dźwięków) zaznaczono również miejsca, gdzie ewentualnie mogłaby przebiegać granica przy innej interpretacji zapisów oscylograficznych oraz spektrogramów. Jest to na ogół jeden okres po prawej lub lewej stronie od zaznaczonych granic.

<sup>1</sup> Dokładny opis metody przygotowania testów percepcyjnych podany jest w pracach [17], [13].

Podziału przebiegów oscylograficznych na krótsze fragmenty dokonano według tych samych kryteriów, jakie były stosowane poprzednio : 1/ w segmentach dźwięcznych jako jednostkę czasu kwantyzacji przyjęto okres tonu krtaniowego. Dotyczy to w tej części pracy wszystkich przebiegów oprócz [r], 2/ przebieg akustyczny [r] był segmentowany na a/ fragmenty szumowe (przeważnie w całości) b/ fragmenty zwarcia (w dwóch przypadkach fragment zwarcia został podzielony na dwie części), c/ fragmenty wokaliczne podzielone z kolei na poszczególne okresy.

W celu uniknięcia niepożądanych efektów przy segmentacji jako zasadę stosowano regułę, że początek segmentu (miejsce, gdzie przebiega cięcie) stanowi najmniejsza ujemna wartość po przejściu przez zero na skwantowanej osi czasu.

3.3. Następnie na podstawie segmentacji sporządzonej na oscylogramach dokonano segmentacji zapisu cyfrowego stosując jako kryterium pomocnicze skalę czasową odpowiadającą na taśmie 50  $\mu$ s na jeden bajt.

3.4. Przetwarzanie sygnału cyfrowo-analogowe (ale już w postaci poszczególnych pożądaných fragmentów sygnału o czasie trwania od 10-20 ms do maksymalnie około 340 ms) odbywało się za pomocą minikomputera MERA 303 (co opisano w pracy [13], str. 6-7), głównie według ustalonego poprzednio schematu (praca [17], rozdz. 3, etap 6a), umożliwiającego zwiększenie sygnału w kierunku zgodnym z przebiegiem czasowym lub też w kierunku odwrotnym (od końca sygnału ku początkowi) przy uwzględnieniu wszystkich pozycji spółgłosek w badanych logotamach.

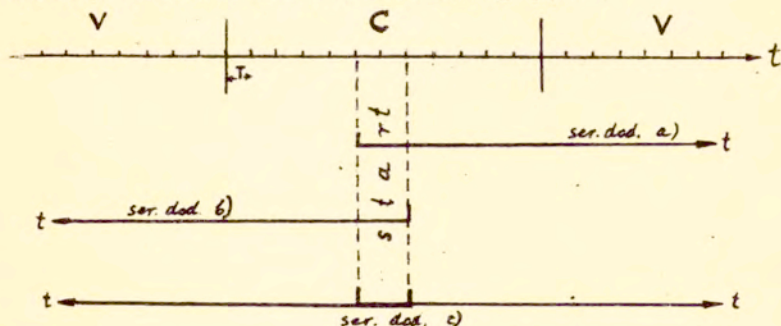
Tak jak poprzednio, oznaczono poszczególne serie otrzymanych bodźców jako  $\underline{C}^e$ ,  $\overset{e}{C}$ ,  $\overset{e}{C}$ ,  $\overset{e}{C}$ ,  $\overset{e}{C}$ ,  $\overset{e}{C}$ , gdzie "C" oznacza spółgłoskę, [e] u góry - sąsiednią samogłoskę, strzałka  $\rightarrow$  oznacza: czas trwania sygnału wydłuża się zgodnie z osią czasową, wspólny początek, obcięte fragmenty po prawej stronie, strzałka  $\leftarrow$  - - czas trwania sygnału wydłuża się odwrotnie do przebiegu czasowego, wspólny koniec, obcięte fragmenty po lewej stronie sygnału.

Krok wydłużania czasu trwania sygnału związany z wielkością  $F_0$  wynosi w tym materiale około 10 ms dla wszystkich dźwięków

oprócz [r], gdzie momenty cięcia były uzależnione od struktury tego dźwięku.

Badany materiał w odniesieniu do spółgłosek nosowych i płynnych został rozszerzony o kilka nowych serii w porównaniu z trącymi i zwartymi. Jedną z nowych serii jest związana ze specyfiką artykulacji spółgłosek tej grupy w pozycji wygłosowej, z tym mianowicie, że każdy logatom kończy się elementem wokalicznym oznaczanym w pracy jako ([a]<sup>1</sup>) powstającym w wyniku zwolnienia toru ustnego w pozycji wygłosowej. Element ten jest widoczny na spektrogramach i oscylogramach, a także słyszalny jako samogłoskopodobny, kiedy jest wycięty. Również spektrogramy [wewew] i [jejeje] wskazują na obecność krótkiego elementu wokalicznego w wygłosie, który jest znacznie słabszy niż w przypadku np. spółgłosek nosowych, a który był odbierany przez słuchaczy odpowiednio jako [u, i] bądź [ə] w obu przypadkach. W badanym materiale nie stwierdzono takiego elementu jedynie dla [r] wygłosowego (patrz rys. 1)<sup>2</sup>.

W związku z tym dla [r] takiej serii brakuje. Oprócz tego dążąc do bardziej dokładnej odpowiedzi na sformułowane we wstępie pytania wykonano kilka serii dodatkowych, mianowicie: dla [l] nagłosowego, [n] śródgłosowego oraz [m] wygłosowego, co można przedstawić schematycznie następująco:



<sup>1</sup> Por. [6], str. 117 i następne, gdzie tożsawisko jest opisane jako typowe dla polskich głosek "sonornych".

<sup>2</sup> Być może został on zmazany przy wgrywaniu na pętlę impulsu startowego potrzebnego do procesu wyzwiania przetwarzania analogowo-cyfrowego.

gdzie C - oznacza badaną spółgłoskę, V - sąsiednie samogłoski z których jedna jest nieobecna odpowiednio dla C nagłosowego lub wygłosowego, T - okres  $F_0$  według podanych wyżej reguł stanowiący krok wydłużania czasu trwania bodźca, linią przerywaną zaznaczony jest środek spółgłoski stanowiący czasowy punkt odniesienia i wspólny dla wszystkich serii, strzałką - kierunek narastania czasu trwania bodźca (dla serii "c" jak widać ma miejsce jednoczesne narastanie czasu trwania sygnału w obie strony).

3.5. W wyniku segmentacji powstało 1062 sygnałów. Doświadczenie poprzednich etapów pracy wykazało ujemny wpływ na jakość ocen długich ciągów złożonych z podobnie brzmiących bodźców. Można było przewidywać, że fragmenty wycięte ze spółgłosek nosowych i płynnych w większości będą miały podobne brzmienie (co mogło zasugerować słuchaczom odpowiedzi na zasadzie domyslenia się), dlatego jednocześnie przygotowano dodatkowe serie (identyczne z opisanymi wcześniej) dla spółgłosek [f, x, v], z których każda występowała w pozycji nagłosowej, śródgłosowej i wygłosowej oraz powtórzono (dokonano nowego przetwarzania cyfrowo-analogowego i wszystkich dalszych manipulacji) kilka serii dla spółgłosek zwartych:  $\underline{e}_k^e$ ,  $\underline{e}_k$ ,  $\underline{e}_p^e$ ,  $\underline{e}_p$ ,  $\underline{e}_t^e$ ,  $\underline{e}_t^1$  przy wykorzystaniu już istniejących z poprzednich etapów pracy taśm perforowanych.

Dodatkowe bodźce w liczbie 414 zostały w sposób losowy wymieszane z materiałem odsłuchowym dla nosowych i płynnych (po uprzednim skopiowaniu wszystkich otrzymanych nagrań), następnie całość materiału zrandomizowano i ułożono w nowe ciągi po 50 bodźców w serii. Jak poprzednio, każdy bodziec został po konwersji cyfrowo-analogowej zarejestrowany na taśmie magnetofonowej 3-krotnie z krótkimi przerwami (około 1 s), a podczas rozcinania oraz montażu taśm starano się zachowywać w miarę równe (5-6-sekundowe) odstępy między sąsiednimi "trójkami", aby umożliwić słuchaczom zapisanie odpowiedzi. Każda z serii (których było 30) trwała około 15 minut (oprócz ostatniej o połowę krótszej).

<sup>1</sup> Uwzględnienie tego materiału było potrzebne także w celu uzupełnienia i sprawdzenia niektórych wyników otrzymanych poprzednio.



#### 4. Przeprowadzenie eksperymentu percepcyjnego.

Odsłuchy przeprowadzono w wytłumionym studio Pracowni Fonetyki Akustycznej w Poznaniu. Cała powierzchnia studia była równomiernie nagłośniona przez cztery odpowiednio umieszczone głośniki, dzięki czemu można było założyć, że każdy ze słuchaczy odbierał sygnały akustyczne w identycznych warunkach. Jednorazowo w odsłuchach brało udział 1-5 osób, odsłuchując w trakcie jednej "sesji" 2-4 serii z kilku lub kilkunastominutową przerwą w połowie na odpoczynek. Jeżeli słuchacze byli zmęczeni to przerwa miała miejsce po każdej serii. Na czas przerwy podawano na głośniki muzykę rozrywkową.

W doświadczeniu wzięło udział 20 osób (pracowników PFA, studentów oraz uczniów ostatnich klas liceum), z których tylko 3 osoby miały fonetyczne przygotowanie. Jak poprzednio, słuchacze zostali uprzedzeni, że będą słyszeli wycięte z ciągłej mowy dźwięki, które należy zaklasyfikować do któregoś fonemu języka polskiego. Słuchacze byli uprzedzeni o możliwości usłyszenia krótkich fragmentów głosek albo połączeń dwu lub trzech dźwięków. Zaproponowano wybór ze wszystkich spółgłosek polskich oraz zapowiedziano, że we wszystkich fragmentach pierwotnie jest ta sama samogłoska, ale ma ona różne brzmienie w zależności od pewnych warunków i słuchacze są proszeni o jak najdokładniejsze notowanie swojego wrażenia sudytorycznego. Ponieważ większość biorących udział w doświadczeniu osób nie znała transkrypcji fonetycznej ustalono zapis ortograficzny, przyjmując osobny znak diakrytyczny (apostrof przy spółgłosce) dla spółgłosek miękkich. Podano też kilka możliwości oznaczania samogłoski, między innymi znak (·) pod literą na oznaczenie samogłoski przymkniętej, Instrukcjom towarzyszyły demonstracja nagrań i krótkie ćwiczenia. Tak jak poprzednio proszono o konsekwentne stosowanie jednego znaku dla oznaczenia jednego dźwięku (aby odczytanie odpowiedzi nie mogło być dwuznaczne). W badanym materiale dotyczy to przede wszystkim [ɲ] - w ortografii "ń" albo "ni".

Ustalono, że bodźce krótsze (albo część bodźca) można zapisywać, stosując dwupoziomowy zapis (krótszą część u góry). Dopuszczono też możliwość odpowiedzi alternatywnych, braku odpo-

wiedzi w przypadkach sygnałów niemożliwych do rozpoznania oraz krótkich komentarzy.

Mimo tych wszystkich przygotowań i instrukcji dobór słuchaczy i postać odpowiedzi są najbardziej problematycznymi zagadnieniami, choć bardzo istotnymi w tego rodzaju doświadczeniu. Ustalając reguły notowania odpowiedzi nie można przewidzieć wszystkich ewentualności. Osobom, które nigdy nie miały do czynienia z fonetyką i nie brały udziału w doświadczeniach odsłuchowych z tak specyficznymi sygnałami jakimi są krótkie bodźce wycięte z mowy, brakuje umiejętności i możliwości przekazywania na papierze swoich wrażeń słuchowych. W ten sposób powstają zapisy w postaci połączeń fonemów, np. "dwyt" w odpowiedzi na bodziec o czasie trwania około 30 ms stanowiący zaledwie część jednej ze spółgłosek (takie odpowiedzi były interpretowane jako "brak" odpowiedzi).

Analiza odpowiedzi słuchaczy mogłaby sugerować, że ich notowanie poprzedzone jest (szczególnie w "trudnych" przypadkach) kilkakrotnym powtórzeniem (albo raczej próbą powtórzenia) bodźca przez słuchacza tak zwaną "mową wewnętrzną" i właśnie w tym momencie odgrywa ważną rolę fakt, czy osoba odsłuchująca jest wykształcona fonetycznie, czy nie, czy jej aparat artykulacyjny jest o tyle wyćwiczony, że potrafi w miarę dokładnie odtworzyć motorycznie usłyszany sygnał<sup>1</sup>. Wydaje się bowiem, że w "trudnych" przypadkach jest notowany przez słuchacza raczej produkt własnej działalności artykulacyjnej, nie zawsze dokładnie naśladowujący usłyszany bodziec. Często dotyczy to "nienormalnie" krótkich bodźców, kiedy z powodu braku w bodźcu istotnej informacji nie kojarzy się on słuchaczowi z żadnym ze znanych mu fonemów, przy czym każdy ma w takich sytuacjach swoje faworyzowane odpowiedzi i często - rzecz niemal paradoksalna - tym dłuższe im krótszy był bodziec.

##### 5. Opracowanie wyników eksperymentu.

Odpowiedzi otrzymane w wyniku eksperymentu percepcyjnego zostały przepisane w jednolitej (międzynarodowej) transkrypcji

<sup>1</sup> To przypuszczenie kojarzyłoby się z "motoryczną teorią" percepcji mowy, która, jak pokazały badania ostatnich lat nie znalazła potwierdzenia (np. [4], str. 33).

fonetycznej z zachowaniem większości indywidualnych szczegółów zapisów (dwupoziomowości, komentarza), następnie zderandomizowane i ułożone wzdłuż osi czasu przebiegów akustycznych sygnałów. Odpowiedzi dotyczące spółgłosek trących i zwartych za pomocą których był "urozmaïcony" materiał odsłuchowy zostały na tym etapie wyeliminowane z dalszej analizy.

Na podstawie otrzymanych zestawień w układzie czasowym wyznaczono częstości poprawnych rozpoznań poszczególnych spółgłosek jako funkcję czasu obcięcia kolejnych sygnałów percepcyjnych. Dane przedstawiono na wykresach w układzie współrzędnych czas - procent poprawnych odpowiedzi (rys.3-35). Przebieg powstałych empirycznych krzywych percepcji posłużył za podstawę ustalenia granic percepcyjnych między samogłoską a badanymi spółgłoskami. Grubszą linią pionową zaznaczono granicę fizyczną pomiędzy dźwiękami, krótkie linie pionowe w sąsiedztwie z granicą fizyczną na niektórych wykresach (z napisem "trans") oznaczają ewentualne inne miejsca granicy fizycznej, linią przerywaną zaznaczono granice percepcyjne. Dla [r] są także zaznaczone poszczególne segmenty, z których ten dźwięk się składa. Strzałka → albo ← oznacza moment pierwszego cięcia i kierunek narastania czasu trwania sygnału. Skala na osi poziomej (oznaczona przez m.c.) oznacza przebieg czasowy sygnału, gdzie jedna działka równa się w przybliżeniu 10 ms - oprócz wykresów dotyczących [r], gdzie ta skala jest aktualna tylko dla fragmentów samogłoskowych. Pod skalą czasową oznaczono (ręcznie) numery poszczególnych momentów cięcia w logatomach (co ułatwia porównywanie wykresów ze sobą). U dołu podano procent poprawnych odpowiedzi dla każdego z momentów obcinania. Wykresy zostały wykonane dla wszystkich serii we wszystkich badanych logatomach. Zawierają one główne wyniki przeprowadzonej pracy. Tak jak w pracach poprzednich etapów ([17], rozdz. 5), szukając kryterium określenia granic percepcyjnych przyjęto, że krzywa percepcyjna jest funkcją ilości zawartej informacji w każdym z momentów cięcia sygnału. Opierając się na istniejących badaniach sygnałów akustycznych przyjmujemy, że w sygnale mowy istnieją takie odcinki, w których następuje pewne zdarzenie decydujące o percepcji określonego dźwięku mowy. Jeżeli bodziec jest

rozpoznawany z poprawnością  $p\%$ , a następny z poprawnością  $p + x\%$ , to dzieje się to dzięki zwiększeniu się ważnej informacji (zdarzeń akustycznych) zawartej w tym dodatkowym fragmencie (por. [20]). Jeżeli zdarzenie to powoduje gwałtowny przyrost liczby poprawnych identyfikacji, to krzywa percepcji na danym odcinku wzrasta najszybciej (tj. zakładając ciągłość funkcji percepcji jej pochodna osiąga maksimum). Przyjęto zatem, że granica percepcyjna określająca minimalny segment poprawnego odbioru danego dźwięku przebiega przez koniec odcinka, na którym empiryczna krzywa percepcji osiągnęła największy przyrost, o ile w punkcie tym udział poprawnych identyfikacji jest nie mniejszy niż 50 %. W tych przypadkach, gdzie maksymalna prędkość wzrostu krzywej percepcji była prawie jednakowa na kilku kolejnych odcinkach, granice percepcyjne wytyczono przez końce odcinka, na którym po raz pierwszy otrzymano co najmniej 50 % poprawnych identyfikacji.

Takie reguły były przyjęte na pierwszym etapie pracy i zostały utrzymane również na obecnym celem nadania pracy jednolitości, jakkolwiek (dopóki są wykorzystywane jako kryteria porównawcze czas trwania fragmentów oraz procent poprawnych identyfikacji) nadają się one bardziej do porównywania wyników uzyskanych dla tej samej spółgłoski (np. w różnych pozycjach), czy też spółgłosek jednej klasy (np. tylko dźwięcznych, czy też tylko trących) itd. Celem większej porównywalności danych istnieje możliwość wprowadzenia za Cistovic i autorami cytowanych przez nią prac ([4], str. 84) "zakresu nieokreślonego". Jest to interwał wartości parametrów bodźców zawarty pomiędzy 0,25 a 0,75 odpowiedzi dotyczących określonego fonemu na funkcji identyfikacji. Pojęcie to jest stosowane w badaniach dotyczących poszukiwań granic między fonemami w przestrzeni cech (np. wśród samogłosek), kiedy "granica pomiędzy fonemami odpowiada wartości parametrów bodźca, przy których dwa sąsiadujące fonemy są wybierane z jednakowym prawdopodobieństwem ([4], str. 84).

Problem ten na obecnym etapie pracy nie został rozstrzygnięty.

W interpretacji wyników odsłuchów za prawidłowe uznano te odpowiedzi, w których występuje poszukiwana spółgłoska, nawet

jeżeli pojawia się ona w połączeniu z innymi fonemami. Pojawienie się innych zapisów obok poszukiwanej spółgłoski jest przeważnie wynikiem albo obcięcia sygnału (kiedy ma miejsce raptowna zmiana poziomu sygnału) albo też obecności resztek sąsiadującej samogłoski. W przypadku uzyskania zapisu kilkogłoskowego zwracano uwagę, aby kolejność dźwięków w zapisie była zgodna z ich kolejnością w logatomie. W ten sposób np. nie zostały uznane za prawidłowe odpowiedzi w rodzaju [je] dla fonemu [j], o ile bodziec był wycięty ze stacjonarnej (środkowej) części jednego z [e] w logatomie [jeje] (w serii  $e_j e$ ). Tego rodzaju zapis uznano raczej za charakterystykę jakości fragmentu samogłoski [e] (jako przymkniętej).

#### 6. Wyniki eksperymentu.

Wszystkie otrzymane dane przedstawiono w formie skondensowanej w Tabeli I umieszczonej na końcu pracy.

##### 6.1. Spółgłoski nosowe.

W licznych badaniach dotyczących spółgłosek nosowych - w tym również dla języka polskiego - stwierdzono, że dla ich rozpoznawania automatycznego oraz ich percepcji słuchowej znaczenie posiada kontekst fonetyczny [18, 22]. Okazało się również, że ważne informacje zawarte są w nieustalonych fragmentach częstotliwości formantowych w obrębie tych spółgłosek, a także w ugięciach formantowych samogłosek [2, 8, 16, 19, 20], chociaż jeszcze Malecot ([16], str. 281) wskazywał na trudności przy dokładnym określeniu miejsca, gdzie kończy się część ustalona a rozpoczyna tzw. "przejście"<sup>1</sup>.

Wyniki obecnej pracy potwierdzają niektóre wnioski zawarte w wymienionych pracach.

##### a/ logatom [memem]

Analiza odpowiedzi słuchaczy dla serii  $\underline{e}_m e$  oraz  $\underline{e}_m^{(a)}$  (rys. 3) wskazuje na wysoką zawartość informacji niezbędnej dla percepcji [m] we fragmentach przejściowych pomiędzy poprzedzającym [e] a następującym [m] i w początkowych fragmentach [m]. Długość

<sup>1</sup> Zgodnie z przyjętą terminologią w pracy używane jest określenie "fragmenty nieustalone dźwięków" albo też "fragmenty przejściowe" jako odpowiednik ang. "transition" albo ros. "переход", "переходный сегмент".

fragmentu przed granicą percepcyjną jest mniejsza dla [m] wygłosowego - 10 ms początkowej części [m] w tej pozycji daje 75 % identyfikacji, podczas gdy początkowe 30-40 ms decyduje o prawidłowej (na poziomie 95-100 %) identyfikacji [m] interwokalicznego.

W nagłosie - seria  $m^e$  - (rys. 4) cały fragment tej spółgłoski łącznie z częścią końcową tuż przy granicy jest nierozpoznawalny, ale dodanie już 10 ms z fragmentu nieustalonego [e] pozwala na przeprowadzenie granicy percepcyjnej (75 %). Aby osiągnąć 90-100 % identyfikacji trzeba było dodać następne 20-40 ms samogłoski. Błędy na całym fragmencie spółgłoskowym ([n, n̄, bv, bv̄, di, dn, dz̄]) wskazują na nieokreśloność brzmienia części ustalonej, po dodaniu fragmentów samogłoskowych nadal [m] jest mylone z [n]. Można by wyciągnąć z tego wniosek, że na błędnej identyfikacji [m] zaważył brak poprzedzającej samogłoski lub fakt obcięcia ugięcia formantowego następującej samogłoski, jednakże taka interpretacja nie byłaby zgodna z wynikami uzyskanymi dla [m] w innych seriach obecnej pracy. Wydaje się, że w naszym materiale przyczyną takizłej identyfikacji całego fragmentu nagłosowego [m] były raczej cechy akustyczne tej spółgłoski - jest ona krótsza (ok. 2 razy) od śródgłosowego i tym bardziej wygłosowego [m] i ma mniej wyraźną strukturę formantową (rys. 2).

W seriach z wsteczną filtracją czasową (rys. 5,6) właśnie dla [m] wygłosowego otrzymano najkorzystniejsze wyniki, bowiem granica percepcyjna (75 % poprawnych odpowiedzi) powstaje już po dodaniu 10 ms części nieustalonej [m] pomiędzy spółgłoską a elementem wygłosowym [a]. Natomiast dla identyfikacji [m] nagłosowego i śródgłosowego w tych seriach na poziomie odpowiednio 70 i 60 % jest potrzebne dodanie do fragmentu następującej samogłoski ok. 40 ms identyfikowanej spółgłoski, który to odcinek zawiera tak fragmenty przejściowe jak i części fragmentów ustalonych (dla  $m^e$  jest to połowa czasu trwania tej spółgłoski). Nagłosowe [m] jest i tu identyfikowane gorzej od śródgłosowego i nagłosowego.

Dowodów na dużą zawartość informacji w części końcowej [m] wygłosowego dostarcza porównanie powyższych wyników z wynikami

uzyskanymi dla serii  $\underline{e}_m$  (przy obciętym [ə]) (rys. 7).

Granica percepcyjna (75 %) powstaje już na fragmencie końcowym o czasie trwania 29 ms, a na fragmencie o długości ok. 60 ms (po pierwszym obniżeniu się krzywej) procent poprawnych identyfikacji ustala się na poziomie 95 - 100 %.

Krótkie serie dodatkowe (rys. 8a,b) zawierają [m] wygłosowe. Uzyskane dla nich wyniki stanowią potwierdzenie wysokiej zawartości informacyjnej fragmentów ustalonych tej spółgłoski<sup>1</sup>, nawet dosyć krótkich : dla fragmentów o długości 30-40 ms wyciętych ze środka spółgłoski we wszystkich trzech seriach otrzymano 90-100 % poprawnych identyfikacji. Wynik ten jest zgodny z obserwacjami zawartymi w pracy [16], z których także wynika, że stadium ustalone [m] wycięte z mowy naturalnej (w odróżnieniu od innych nosowych) jest dobrze identyfikowane przez odbiorców. Autor tłumaczy to powszechną używalnością tej spółgłoski w języku. W pracy [8] opisano eksperyment, w którym nieco wyższy w porównaniu z innymi nosowymi procent poprawnych identyfikacji (jakkolwiek w sumie niski : 30 %) otrzymano również dla stadium ustalonego [m] (także wyciętego z mowy naturalnej), ale autor pracy porównując wyniki otrzymane dla różnych osób uważa, "że na oceny odbiorców większy wpływ wywierały obiektywne cechy widma, niż takie czynniki zewnętrzne jak np. większe lub mniejsze "osłuchanie" się z danym dźwiękiem, wynikające z częstości jego występowania" ([8], str. 28). Porównanie wyników otrzymanych dla [m] nagłosowego i wygłosowego (patrz także spektrogram - rys. 2) wydaje się potwierdzać ten wniosek.

Na wykresach  $\underline{m}^e$ ,  $\underline{m}^e$ ,  $\underline{m}$  (rys. 5,7) obserwuje się wyraźne opadanie krzywej percepcyjnej do 50-65 % w odległości 50-70 ms od prawej granicy fizycznej (w seriach  $\underline{m}^e$  i  $\underline{m}^e$  na korzyść [n]), w serii  $\underline{m}$  natomiast otrzymano w tym miejscu [ɹ] 2 razy, [j] - 3, [bi] - 1 oraz 4 razy - brak odpowiedzi) co na obecnym etapie pracy jest trudne do wyjaśnienia.

<sup>1</sup> Serie dodatkowe nie umożliwiają niestety porównania zawartości informacyjnej stadium ustalonego tej spółgłoski w różnych pozycjach.

b/ logatom [nenen]

Obserwując krzywe percepcyjne dla serii o normalnym przebiegu czasowym (rys. 9, 10) można stwierdzić, że tak jak w przypadku [m] fragment przejściowy pomiędzy [e] a [n] jest bardziej informacyjny dla [n] wygłosowego (które także ma najbardziej wyraźną strukturę formantową w porównaniu z dwoma innymi pozycjami) 20 ms początkowej części [n] daje 90 % identyfikacji dla [n] wygłosowego, 50 % dla śródgłosowego. W przypadku braku poprzedzającego [e] (rys. 9) wydłuża się minimalny segment percepcyjny, a ogólny procent poprawnych identyfikacji jest także niższy dla [n] nagłosowego.

Krzywa percepcyjna na całej długości [n] we wszystkich tych seriach nie jest ciągła, tylko w kilku punktach osiąga 100 % w przypadku [n] wygłosowego. Najczęstsze błędy to [m] i [b] w nagłosie oraz [m, n] w pozostałych pozycjach, przy czym odpowiedzi [n] dawała zawsze ta sama osoba (która w wielu przypadkach zaznaczyła również miękkość dla [m] w logatomie [memem]). Dla serii z wsteczną filtracją czasową (rys. 11, 12) 85-90 % poprawnych identyfikacji otrzymano na fragmencie zawierającym 20 ms końcowej części [n] oraz część następującej samogłoski. Większość błędnych odpowiedzi przed granicą percepcyjną - to zwarte [b, d] z przewagą [d], zaś po granicy percepcyjnej - sporadyczne pomyłki z [m].

Porównanie serii  $e_n^{(a)}$  i  $e_n$  wskazuje na ważność końcówki [a] dla identyfikacji [n] (niższy procent poprawnych identyfikacji na granicy percepcyjnej, nieciągłość krzywej) ale stwierdzić można, że identyfikacja w przypadku obcięcia [a] jest możliwa. Na żadnym z fragmentów począwszy od 30 ms nie otrzymano identyfikacji poniżej 55 %, a fragment nieco dłuższy niż połowa spółgłoski daje 90 % identyfikacji.

Analiza wyników dla serii dodatkowych zawierających śródgłosowe [n] pokazuje, że fragment środkowy o czasie trwania ok. 60 ms (połowa długości spółgłoski) jest odbierany poprawnie w 65 %. Jednoczesne wydłużenie tego fragmentu w obie strony podnosi procent poprawnych identyfikacji a fragment zawarty pomiędzy granicami fizycznymi (przy obciętych sąsiednich samogłoskach) daje 100 % identyfikacji.



Dwie pozostałe serie dodatkowe wskazują także na większą zawartość informacyjną fragmentów narastania i wygasania, niż samej części ustalonej [n] (co również zgodne jest z wynikami opisanymi w pracy [8]), chociaż również w tych seriach [n] jest identyfikowane w 60-65 % na podstawie środkowych fragmentów o czasie trwania 30-60 ms.

W porównaniu z [m] wygłosowym stadium ustalone [n] jest odbierane gorzej.

c/ logatom [ɲeɲɲ]

W doświadczeniu opisanym w pracy [8], (str. 31-34 a także 47-55) wykazano znaczenie ugięć formantowych a także obecności stadiów początkowego i końcowego dla percepcji [ɲ] w porównaniu z innymi nosowymi. W pierwszym z tych doświadczeń stadium ustalone spółgłoski [ɲ] pozbawionej fragmentu początkowego i końcowego w ogóle nie było identyfikowane jako [ɲ]. Potwierdzają to również wyniki uzyskane w niniejszej pracy dla serii  $r^e$  (rys. 13), gdzie wszystkie odpowiedzi na całym fragmencie [ɲ] wskazują na nieokreśloność brzmienia bodźców ([m̄, bv, mvu, η, u, im, dzu, dv, b̄, bwu, bwm, bu]). Krzywa percepcji podnosi się raptownie do 100 % na ostatnim okresie [ɲ] przed granicą fizyczną (na spektrogramie jedynie w tym ostatnim okresie pojawiła się energia w górnym zakresie częstotliwości).

W przypadku [ɲ] interwokalicznego (rys. 14) dużą zawartość informacyjną posiada fragment początkowy. Informacje odnoszące się do spółgłoski zawarte w końcowych fragmentach poprzedzających samogłosek wydają się dotyczyć wyłącznie jej miękkości: fragment samogłoski wraz z 20-milisekundowym fragmentem początkowej części [ɲ] daje wrażenie słuchowe [e, é, ê, eî, ej, ei, i]. Bardzo dużą rolę w percepcji [ɲ] odgrywają fragmenty nieustalone następującej samogłoski (w tym elemencie wygłosowego [ə]). Wskazują na to wyniki serii  $r^e$ ,  $e_r^e$  i  $e_r^{(a)}$  (rys. 15, 16), gdzie 80-90 % identyfikacji [ɲ] następuje przed granicą fizyczną, tuż po niej, albo jednocześnie - czyli w dużym stopniu na podstawie informacji zawartej w początkowych fragmentach samogłosek.

Część samogłoskowa, na której rozpoczyna się rozpoznawanie oraz wzrasta ilość poprawnych odpowiedzi wynosi ok. 30 ms.

Wiadomo z wielu prac (patrz np. [2], str. 54, a także [7]), że czas trwania fragmentu "i-podobnego" w "miękkich" spółgłoskach wynosi od 30-40 ms do 100 ms i więcej. Z wykresów 15, 16 widać, że pewne informacje dotyczące miejsca i sposobu artykulacji miękkiej spółgłoski zawarte są w początkowych 10-20 ms tego fragmentu. Wniosek ten w znacznie mniejszym stopniu odnosi się do pozostałych spółgłosek nosowych. Porównanie wykresów 16 i 17 pozwala stwierdzić, że decydująca i wystarczająca informacja dla identyfikacji [ɲ] wygłosowego zawarta jest pomiędzy fragmentem końcowym [ɲ] a elementem [ə]. Końcowe [ɲ] pozbawione elementu [ə] jest często mylone z innymi nosowymi, mimo to jest rozpoznawalne w odróżnieniu od [ɲ] w nagłosie (w tym przypadku również spółgłoska wygłosowa ma bardziej wyraźną strukturę formantową).

#### Podsumowanie dla spółgłosek nosowych.

1. Stadium ustalone spółgłosek nagłosowych jest identyfikowane gorzej niż stadium ustalone spółgłosek wygłosowych (świadczy o tym porównanie z seriami, w których obcięto [ə]). Może to sugerować, że fragmenty końcowe spółgłosek nosowych zawierają więcej informacji od fragmentów początkowych. Wszystkie wygłosowe spółgłoski nosowe mają wyraźniejszą strukturę formantową od spółgłosek nagłosowych.

2. Stadium ustalone (krótkie fragmenty wycięte ze środka spółgłoski) jest bardzo istotne dla [m]. Procent identyfikacji fragmentów środkowych [n] rośnie w miarę dodawania do tych fragmentów części początkowej i końcowej. Cały odcinek [n] (przy obciętych sąsiednich samogłoskach) daje ok. 100 % identyfikacji.

3. Obcięcie elementu wokalicznego [ə] po spółgłoskach wygłosowych nie uniemożliwia ich identyfikacji, jakkolwiek utrudnia odbiór [ɲ], w mniejszym stopniu [n], a w najmniejszym [m].

4. Poprzedzająca samogłoska na ogół nie zawiera informacji o miejscu i sposobie artykulacji następującej spółgłoski nosowej.

5. Ugięcia formantowe następującej samogłoski (także elementu wygłosowego [ə]) zawierają znaczną ilość informacji o poprzedzającej spółgłosce. W przypadku [ɲ] zawierają one całość

(albo prawie całość) informacji niezbędnej dla percepcji tej spółgłoski.

6. W odniesieniu do [m] i [n] dużą rolę odgrywają fragmenty początkowe i końcowe tych spółgłosek, a w przypadku [m] duże znaczenie posiada także stadium ustalone.

7. Miejsce granicy percepcyjnej oraz czas trwania minimalnych segmentów percepcyjnych (patrz Tabl. I) są uzależnione, jak to wynika z poprzednich punktów od pozycji spółgłosek nosowych w wypowiedzi, a także od rodzaju spółgłoski, szczególnie jeśli została ona pozbawiona kontekstu samogłoskowego.

#### 6.2. Spółgłoski [l,r] i samogłoski niesylabiczne [w,j].

Fonemy [l,r] tworząc razem z fonemami nosowymi grupę spółgłosek sonornych różnią się od nich brakiem nosowości oraz rodzajem przegrody (zwężenia) w torze głosowym. Charakterystyka akustyczna polskich fonemów [l,r] umieszczona jest w wymienionych we wstępie pracach, między innymi w pracy [12]. Analizie akustycznej polskich tzw. "samogłosek niesylabicznych" [w,j] w opozycji do samogłosek [u,i] poświęcona jest praca [15].

Od strony percepcyjnej wymienione fonemy języka polskiego nie były do tej pory badane. Percepcja spółgłosek była przedmiotem badań na gruncie innych języków np. percepcja rosyjskich sonantów [l,r,j] przy zastosowaniu różnych metod i sposobów ich skracania [7], czy też angielskich [w,l,r,j] np. w pracach [3], [14], w których ustalono cechy akustyczne istotne dla różniczenia tych spółgłosek.

a/ logatom [lelel]

Spółgłoska [l] w pozycji postwokalicznej (rys. 18) jest identyfikowana na podstawie jej początkowego fragmentu nieustalonego (ok. 20 ms), który zawiera całą niezbędną informację. Fragmenty poprzedzającej samogłoski nie zawierają istotnej dla percepcji [l] informacji, natomiast dodanie początkowych 10 ms spółgłoski pozwala na określenie miejsca artykulacji (odpowiedzi [e,ed,et,eth]).

W przypadku [l] nagłosowego dopiero fragment o długości 70 ms (całe [l] wynosi 91 ms) pozwalał je zidentyfikować w 75 %. Stadium ustalone [l] pozwala na określenie miejsca artykulacji oraz dźwięczności, natomiast reszta istotnych informacji zawarta

jest w końcowym fragmencie [l].

Identyczne wyniki otrzymano podając do oceny fragmenty wycięte z [l] nagłosowego w seriach dodatkowych. Potwierdzają one znaczenie końcowych fragmentów [l] nagłosowego dla jego identyfikacji. Natomiast obcięcie ok. 40 ms od początku [l] nie wpływa na jego percepcję. Przy obcięciu stadium wygasania [l] jest identyfikowane na poziomie poniżej 50 %. Wydłużanie środkowego fragmentu [l] jednocześnie w obu kierunkach także wskazuje na znaczenie informacyjne ostatnich okresów [l] - fragment od początku granicy fizycznej przy obciętej samogłosce daje 100 % identyfikacji [l].

Analiza serii  $\underline{l^e}$ ,  $\underline{e_1^e}$  i  $\underline{e_1^{(a)}}$  również potwierdza wniosek o zasadniczej roli, jaką odgrywa w percepcji końcowy fragment [l] o długości 10-30 ms (rys. 19,20). Samogłoska następująca po [l] śródgłosowym i element samogłoskowy [ə] po [l] wygłosowym wydają się nie zawierać informacji istotnych dla percepcji [l]. W przypadku [l] nagłosowego 55 % identyfikacji otrzymano na granicy fizycznej, z tym, że decydującą rolę w percepcji odgrywa jednak końcowe 30 ms spółgłoski.

Z danych otrzymanych w serii  $\underline{e_1}$  (rys. 21) przy obciętych elementach wokalicznym wynikałoby, że w przypadku [l] wygłosowego nieobecność [ə] odgrywa mniejszą rolę, niż w przypadku spółgłosek nosowych. Procent identyfikacji na granicy percepcyjnej jest wprawdzie niższy, niż w serii  $\underline{e_1}$ , ale też fragment podany do oceny jest nieco krótszy (w serii  $\underline{e_1^e}$  - 45 ms, tu - 27,5 ms). Nasuwa się przypuszczenie, że obcięcie elementu wokalicznego przyczynia się do braku ciągłości krzywej percepcyjnej (błędne odpowiedzi po granicy percepcyjnej: [zɨ, ɨ, dʒ, b, r<sub>w</sub>]). Mimo to można stwierdzić dosyć wysoki procent identyfikacji na podstawie tylko końcowego fragmentu [l], co jest potwierdzeniem poprzednich obserwacji.

#### Podsumowanie dla spółgłoski [l]

1. [l] postwokaliczne jest odbierane na podstawie początkowego fragmentu spółgłoski o długości 20 ms podanego do odsłuchów wraz z fragmentem poprzedzającej samogłoski. Sam fragment samogłoskowy, jak stwierdzono, nie zawiera informacji pozwalającej na identyfikację [l].

2. Stwierdzono dużą zawartość informacyjną stadium wygasania tej spółgłoski, np. : [l] nagłosowe jest odbierane na podstawie informacji zawartej w stadium wygasania.

3. Stadium ustalone przy obciętych stadium wygasania jest źle identyfikowane.

4. Element [ə] pełni mniejszą rolę w percepcji [l] niż spółgłosek nosowych.

5. Ugięcia formantowe następującej samogłoski w przypadku [l] nagłosowego pozwalają w pewnej mierze na identyfikację [l]. Nie stwierdzono w naszym materiale takiej informacji dla [l] interwokalicznego i wygłosowego.

b/ logatom [rerer]

Nawet pobieżne spojrzenie na wykresy krzywych percepcyjnych dla [r] nasuwa wniosek o dużej zawartości informacji dotyczącej [r] mieszczącej się w poprzedzającej samogłosce (rys. 21,22). [r] postwokaliczne jest odbierane przez połowę słuchaczy wyłącznie na podstawie tej informacji. W przypadku  $\underline{e_r}$  na granicy fizycznej pomiędzy końcem poprzedzającej samogłoski a początkiem następującego [r] otrzymano 85 % identyfikacji, dodanie do fragmentu [e] segmentów zwarcia i szumowego pozwala na 100 % percepcji [r]<sup>1</sup>.

Minimalny segment percepcyjny dla [r] nagłosowego (w serii  $\underline{r^e}$ ) wynosi 63 ms, co obejmuje następujące segmenty (licząc od początku spółgłoski) - segment szumowy, segment wokaliczny składający się z dwóch okresów, segment zwarcia i następnym szumowy.

Wykresy krzywych percepcyjnych w seriach  $\underline{r^e}$ ,  $\underline{e_r^e}$  i  $\underline{e_r}$  (rys. 24,25) sugerowałyby, że [r] może być dobrze (w 100 %) zidentyfikowane na podstawie fragmentu spółgłoski (w przybliżeniu jej drugiej połowy) zawierającego przedostatni segment wokaliczny (albo jego część) zwarcia, segment szumowy i ostatni segment wokaliczny. Ponieważ w naszym nagraniu brakuje segmentu

<sup>1</sup> Nagłosowe [r] jest w naszych nagraniach trójударowe, śródgłosowe i wygłosowe - dwuударowe.

wygłosowego [ə] w przypadku [r] wygłosowego (por. Wstęp str. 5) nie pozwoliło to dokładnie zbadać znaczenia następującej samogłoski (także [ə]) dla percepcji poprzedzającego [r]. W każdym razie w przypadku interwokalicznego [r] testy percepcyjne nie wykazują zawartości informacji o [r] w obrębie następującej samogłoski, natomiast w przypadku nagłosowego [r] - pewną jej ilość. Granica percepcyjna znajduje się tu bliżej granicy fizycznej pomiędzy [r] a następującą samogłoską, niż dla [r] interwokalicznego, jakkolwiek do stuprocentowej identyfikacji są również niezbędne segmenty conajmniej jednego zwarcia i następujący szumowy a także chociaż część poprzedzającego to zwarcie segmentu wokalicznego.

#### Podsumowanie dla spółgłoski [r]

1. Ugięcia formantowe poprzedzającej samogłoski zawierają znaczną ilość informacji dotyczącej [r].
2. Percepcja [r] jest ściśle związana ze strukturą polisegmentalną tej spółgłoski : do stuprocentowej identyfikacji [r] nagłosowego potrzebne są segmenty : szumowy, wokaliczny, segment zwarcia i następny szumowy, do stuprocentowej identyfikacji [r] postwokalicznego potrzebne są segmenty : zwarcia i szumowy.
3. W przypadku obcięcia fragmentów od początku [r] w dowolnej pozycji do stuprocentowej identyfikacji potrzebne są końcowe fragmenty zawierające jedno zwarcie, jeden segment szumowy plus conajmniej część poprzedzającego fragmentu wokalicznego.
4. Niewielka ilość informacji dotyczącej [r] nagłosowego zawarta jest w ugięciach następującej samogłoski. Nie stwierdzono w naszym materiale takiej informacji odnoszącej się do [r] interwokalicznego i wygłosowego (por. punkt 5 str. 15).

#### c/ logatom [wewew]

W przypadku [w] bardzo trudno jest analizować wyniki doświadczenia, ponieważ zaznaczone granice fizyczne (por. Wstęp str.3) można tu uważać raczej za orientacyjne. W związku z tym za orientacyjny trzeba też uznać czas trwania minimalnych segmentów percepcyjnych podany w tabeli na końcu pracy, gdzie punktem odniesienia jest granica między dźwiękami. Istotną rolę w percepcji odgrywa obecność poprzedzającej samogłoski (rys. 26,27).

W przypadku [w] nagłosowego potrzebna jest połowa jego długości aby zostało ono odebrane w 75 %. [w] postwokaliczne jest odbierane na podstawie krótszego niż w poprzednim przypadku fragmentu od początku spółgłoski, a 90-100 % identyfikacji otrzymano już w pierwszej połowie trwania tego dźwięku. Ugięcia formantowe poprzedzającej samogłoski zawierały nieznaczną część informacji. Błędne odpowiedzi przed granicą percepcyjną i po niej - to głównie [u] zamiast [w] nagłosowego, zaś w przypadku interwokalicznego i wygłosowego [w] są to [e, e<sup>u</sup>, er, el, e<sup>o</sup>] przed granicą percepcyjną, które po granicy zamieniają się w [e<sup>u</sup>].

Z analizy serii  $\underline{w^e}$ ,  $\underline{w^e}$ ,  $\underline{w^{(a)}}$  (rys. 28,29,30)<sup>1</sup> wynika, że cała albo prawie cała informacja potrzebna do identyfikacji poprzedzającego samogłoskę [w] jest zawarta w stadium nieustalonym następujących samogłosek, a granice percepcyjne powstają tu "wcześniej" od fizycznych. Błędne odpowiedzi w przypadku [w], po którym następuje [e] to głównie [e], natomiast w przypadku [w] wygłosowego - głównie [u], czasem w połączeniu ze zwartymi.

Z analizy wyników testu percepcyjnego a także z porównania wykresów  $\underline{w^{(a)}}$  i  $\underline{w}$  (rys. 30,31) wynikałoby, że występuje tu krótki  $\alpha$ -podobny fragment wokaliczny, który jak już wspomniano we Wstępie często był odbierany przez słuchaczy jako [a] albo [u]. Wskazywałby na to także fakt, że minimalny segment percepcyjny [w] przy obciętych końcowych okresach jest dłuższy, a także fakt pokrywania się granic percepcyjnej i fizycznej w serii  $\underline{w^{(b)}}$  tzn. kiedy fragment wokaliczny był obcięty, podobnie jak to miało miejsce w przypadku [w] nagłosowego i śródgłosowego. Na wykresach (30,31) zaskakuje sam przebieg krzywej percepcyjnej - [w] jest bardzo często odbierane jak zbliżony do niego fonem [u]. Błędy takie nie występują w seriach  $\underline{w^e}$ ,  $\underline{w^e}$ , a więc tam gdzie po [w] następuje samogłoska [e].

<sup>1</sup> Jako fragment samogłoskopodobny [a] zostały wydzielone 3 końcowe okresy - na podstawie obserwacji zapisu oscylograficznego, chociaż element ten jest mało wyraźny na oscylogramie i prawie niewidoczny na spektrogramie.

### Podsumowanie dla [w]

1. [w] nagłosowe jest identyfikowane na podstawie fragmentów obejmujących środek i koniec tego dźwięku.

2. [w] postwokaliczne jest identyfikowane na podstawie fragmentów początkowych.

3. Bardzo duża zawartość informacji o poprzedzającym [w] mieści się w ugięciach formantowych i początkowych fragmentach następujących samogłosek, co umożliwia identyfikację [w] wyłącznie na podstawie tej informacji.

4. Duże znaczenie informacyjne posiada element wygłosowy [ə].

5. [w] jest bardzo często mylone ze zbliżonym do niego fonemem [u] - przede wszystkim w seriach  $w^e$ ,  $e_w^{(a)}$  i  $e_w$ .

d/ logatom [jejej]

Podczas wyznaczania granic percepcyjnych pomiędzy [j] a [e] wyniknęły podobne trudności jak w przypadku [w], a więc i tu oznaczone na wykresach granice trzeba uważać za orientacyjne.

Interpretacja wyników odsłuchowych w przypadku [j] jest bardzo podobna do wyników otrzymanych dla [w], co potwierdza przynależność tych dwóch dźwięków do jednej klasy.

Tak jak w przypadku [w] dużą rolę odgrywa obecność poprzedzającej samogłoski, wówczas gdy [j] jest identyfikowane na podstawie początkowych fragmentów (rys. 32) z tym, że 10-milisekundowy fragment [j] następujący bezpośrednio po samogłosce ma większą zawartość informacyjną, niż odpowiedni fragment [w]. Natomiast końcowe fragmenty poprzedzającego [e] wydają się nie zawierać informacji, która umożliwiałaby identyfikację [j]. Błędne odpowiedzi [l], które obniżyły krzywą w seriach  $e_j^e$  i  $e_j^{(a)}$  w obrębie całego [j] należą do tej samej osoby i stanowią właściwość indywidualną tej osoby.

W nagłosie krzywa percepcyjna rośnie w miarę dodawania kolejnych okresów [j], a o percepcji decydują fragmenty przejściowe (końcowe fragmenty [j] i początkowe fragmenty [e]) i to głównie prawdopodobnie należące do [e].

Cały fragment [j] przed granicą percepcyjną dostarcza 100-procentowej informacji o miękkości: otrzymano tu błędne odpowiedzi [di, bi, pi, ji, vi, dzi, d', i] itd.

Następująca samogłoska (jej fragmenty nieustalone) zawiera



bardzo dużo informacji potrzebnej do identyfikacji [j]. W seriach  $j^e$  i  $e_j^e$  (rys. 33) granice percepcyjne powstają przed granicą fizyczną pomiędzy [j] a następującą samogłoską. Dodanie krótkich końcowych fragmentów [j] umożliwia jego 100-procentowy odbiór. Odpowiedzi przed granicą percepcyjną (w obrębie samogłoski) wskazują, że odbierano bodziec jako samogłoskę przymkniętą: oprócz [e] występują [é, ɛ, <sup>i</sup>e, i, e<sup>j</sup>, e<sup>ɛ</sup>] i oczywiście odpowiedzi [j<sup>e</sup>], na podstawie których wykreślono krzywą percepcyjną. [j] identyfikowano (w małym procencie) na fragmentach samogłoskowych oddalonych od granicy j/e nawet o 100 ms.

Wygłosowe [j] w serii  $e_j^{(a)}$  (rys. 34) odbierane jest gorzej od dwóch poprzednich (obraz spektrograficzny tego dźwięku jest także mniej wyraźny). Jest ono wprawdzie identyfikowane na dosyć wysokim poziomie, ale zanotowano dużą ilość pomyłek z fonemem [i], co świadczy o bliskości [j] i [i]. Spotyka się także sporadyczne odpowiedzi [je], [je]. Obcięcie końcowego fragmentu odbieranego przez słuchacza jako [ə] albo [i] czy też [ɛ] (rys. 35) znacznie zmniejsza procent identyfikacji [j] wygłosowego. Decyduje tu jak się wydaje, nie czas trwania obciętego segmentu, o który został skrócony fragment [j], lecz zawarta w nim informacja. [j] pozbawione elementu wygłosowego jest mylone z [i] aż do momentu dodania końcowych fragmentów poprzedzającej samogłoski. Błędne odpowiedzi [l'] obniżają krzywą percepcyjną nawet po dodaniu fragmentów samogłoskowych. Są to odpowiedzi otrzymane od jednej (zawsze tej samej) osoby (por. str. 20).

#### Podsumowanie dla [j].

1. [j] nagłosowe jest identyfikowane na podstawie fragmentów nieustalonych pomiędzy [j] a następującym [e].
2. [j] postwokaliczne jest identyfikowane na podstawie krótkich fragmentów (od 10 do 30-40 ms) początkowych [j].
3. Następująca samogłoska zawiera bardzo dużo informacji o poprzedzającym [j], szczególnie jej fragmenty nieustalone, przy czym również dla fragmentów samogłoskowych oddalonych od granicy o 80-100 ms można spotkać odpowiedź [j].
4. Dla [j] wygłosowego dużą wartość informacyjną posiada

następujący po nim element [ə].

5. Znaczna ilość identyfikacji [i] zamiast [j] przede wszystkim w seriach  $j^e$ ,  $e_j(ə)$  i  $e_j$  (por. wyniki dla [w]) świadczy o podobieństwie tych dwóch fonemów. Częste są połączenia innych spółgłosek z [i] (szczególnie w nagłosie).

#### Podsumowanie ogólne. Wnioski.

W licznych dotychczas prowadzonych badaniach nad cechami akustycznymi i dystynktywnymi fonemów którym była poświęcona obecna praca, podkreśla się szczególne znaczenie ugięć formantowych z punktu widzenia zawartej w nich informacji dla prawidłowego brzmienia, rozróżnialności i wyrazistości tych dźwięków. Np. angielskie głoski płynne [w, j, r, l] jak wykazano w pracy [3] rozróżniane są w obrębie tej klasy dzięki ugięciom  $F_2$  i  $F_3$ . Znaczenie percepcyjne fragmentów narastania i przede wszystkim fragmentów wygasania dla polskich fonemów [m, n] udowodniono w badaniach opisanych w monografii [8].

Obecne badanie polegające na ocenie słuchowej fragmentów o różnym czasie trwania wyciętych z sygnału mowy, potwierdzając wyniki wcześniejszych badań umożliwiło ustalenie granic percepcyjnych badanych dźwięków [m, n, r, l, w, j] i czasu trwania ich minimalnych segmentów percepcyjnych. Zastosowana metoda - wydłużania czasu trwania sygnału zgodnie z przebiegiem czasowym (przy obciętych fragmentach środkowych i końcowych) a także odwrotnie do przebiegu czasowego (przy obciętych fragmentach początkowych i środkowych) - pozwala ocenić i porównać poszczególne fragmenty badanych dźwięków z punktu widzenia informacji istotnej dla ich percepcji.

Szczegółowe wnioski zamieszczono w odpowiednich rozdziałach zawierających omówienie wyników oraz w Tabeli na końcu pracy. Uogólniając te wnioski można stwierdzić, że :

1. Miejsce wyznaczonych na podstawie ocen słuchowych granice percepcyjnych, stopień ich zgodności lub niezgodności z tzw. granicami fizycznymi oraz czas trwania minimalnych segmentów percepcyjnych zależą

a/ od rodzaju dźwięku : np. [m] wygłosowe w naszym materiale zawiera dużo istotnej informacji w krótkich fragmentach środkowych, [n] śródgłosowe mogło być percypowane w 100 % przy

obciętych sąsiednich samogłoskach, natomiast warunkiem percepcji [ɣ] jest obecność sąsiadujących samogłosek, z których następująca zawiera dużo informacji pozwalającej zidentyfikować [ɣ] już w obrębie samogłoski (w seriach z wsteczną filtracją czasową). Z kolei percepcja [r] w każdej pozycji jest związana z jego strukturą polisegmentalną.

b/ od pozycji badanego dźwięku w sygnale mowy : np. dźwięki w pozycji nagłosowej są odbierane na podstawie fragmentów albo końcowych albo końcowych w połączeniu ze środkowymi, natomiast w pozycji postwokalicznej dla prawidłowej identyfikacji wystarcza na ogół początkowy fragment dźwięku obejmujący od 10 do 50 ms. Jedynie w przypadku [r] i w mniejszym stopniu również [w] odnotowano prawidłową identyfikację na podstawie końcowych fragmentów poprzedzającej samogłoski.

c/ od kierunku, w którym był wydłużany czas trwania bodźców: Stosując dwa sposoby wydłużania czasu trwania bodźców stwierdzono zawartość informacji umożliwiającej percepcję danej głoski we fragmentach narastania lub wygasania spółgłosek np. interwokaliczne [m,n,l] (wówczas granica percepcyjna powstaje po granicy fizycznej, czyli w obrębie spółgłoski) albo we fragmentach sąsiadujących samogłosek (kiedy to informacja w nich zawarta okazała się wystarczająca dla odebrania danego dźwięku w warunkach kiedy w rzeczywistości był on obcięty np. serie :  $\overleftarrow{e}^{(a)}$ ,  $\overleftarrow{w}^e$ ,  $\overleftarrow{w}^e e$ ,  $\overleftarrow{j}^e$ ,  $\overleftarrow{r}^e e$ ,  $\overleftarrow{r}^e$ ).

2. Percepcja głosek pozbawionych kontekstu nie tylko semantycznego, ale także niekiedy fonetycznego odbywa się na podstawie poszczególnych fragmentów tych głosek, co pozwala określić stopień ważności każdego z fragmentów dźwięku dla jego prawidłowej identyfikacji. W przypadku gdy nie jest obcięta poprzedzająca samogłoska, spółgłoska płynna jest na ogół odbierana na podstawie jej krótkich fragmentów początkowych, nawet jeżeli poprzedzająca ją samogłoska nie zawiera informacji istotnych dla percepcji spółgłoski.

Spółgłoska w nagłosie jest często odbierana na podstawie jej fragmentów obejmujących środek i koniec o zróżnicowanym czasie trwania, a także na podstawie fragmentów nieustalonych następującej samogłoski. Wyjątek stanowią nagłosowe [r] i [w],

w mniejszym stopniu - [n], które były identyfikowane na podstawie ich pierwszej połowy (odpowiednio 100 %, 75 % i 65 %).

W przypadku obcięcia początkowych części głosek (obejmujących odcinki o różnej długości) w tej czy innej pozycji percepcja jest możliwa na podstawie albo tylko początkowych fragmentów następujących samogłosek (jak w przypadku logatomów z [ɲ, w, j]), albo końcowych fragmentów spółgłosek podawanych wraz z fragmentem następującej samogłoski (czy też elementu wokalicznego w wygłosie).

Obcięcie elementu wokalicznego w wygłosie (przy jednoczesnym obcięciu poprzedzającej samogłoski i początkowych fragmentów spółgłoski wygłosowej) pogarsza percepcję [ɲ, j], w mniejszym stopniu pozostałych głosek, dla percepcji których informacja zawarta w następującej samogłosce odgrywa mniejszą rolę.

Za pomoc w wykonaniu różnych etapów niniejszej pracy dziękuję mgr. W. Mikielowi, mgr. J. Imiończycowi, mgr. R. Ciarkowskiemu, p. E. Szubińskiej, p. B. Stróżyk oraz p. A. Mossakowskiemu.

Tablica I

Logatom	spółgłoski		Czas trwania ms				Długość odłanka przed granicą percepcyjną ms	Procent poprawnych identyfikacji na granicy percepcyjnej	Najwyższy procent poprawnych identyfikacji	Czas trwania minimalnego segmentu, na którym osiągnięto najwyższy procent poprawnych identyfikacji
	1	2	3	4	fragmentu uwzględnionego w serii					
					poprzedzającej samogłoski	spółgłoski następującej samogłoski				
1	2	3	4	5	6	7	8	9	10	
memem	$\begin{matrix} e \\ \leftarrow \\ m \end{matrix}$	148	-	80	68	80+21	75	100	80+48	
	$\begin{matrix} e \\ \leftarrow \\ m \end{matrix}$	278	89	120	69	89+30	85	100	89+40	
	$\begin{matrix} e \\ \leftarrow \\ m \end{matrix}$	273	79	135	59	79+11	75	100	79+29	
	$\begin{matrix} e \\ \leftarrow \\ m \end{matrix}$	148	-	80	68	38+68	70	90	80+68	
	$\begin{matrix} e \\ \leftarrow \\ m \end{matrix}$	258	69	120	69	40+69	60	95 100	89+69 19+120+69	
	$\begin{matrix} e \\ \leftarrow \\ m \end{matrix}$	273	79	135	59	10+59	75	100	29+59	
		214	78	135	-	88	75	100	77	

1 Podkreślona liczba oznacza długość odłanka samogłoskowego

1	2	3	4	5	6	7	8	9	10	
menen	$\overline{n^e}$	154	-	93	61	52	65	100	$93+10$	
	$e_n^e$	296	87	128	81	$87+20$	50	90	$87+49$	
	$e_n^{(a)}$	264	88	125	51	$88+20$	90	100	$87+128+21$	
	$\overline{n^e}$	173	-	93	80	$21+80$	85	100	$41+80$	
	$e_n^e$	296	88	128	80	$20+80$	90	100	$31+80$	
	$e_n^{(a)}$	264	89	125	50	$19+50$	90	100	$29+50$	
	$e_n$	213	88	125	-	-	28	60	90	76
	$\overline{n^e}$	182	-	122	60	60	122	100	100	$78+125$
	$e_n^e$	308	77	144	87	87	$77+29$	100		
	$e_n^{(a)}$	286	87	131	68	68	$87+28$	75	100	$87+47$
reper	$\overline{n^e}$	182	-	122	60	50	90	100	$10+60$	
	$e_n^e$	299	77	144	78	$9+78$	80	100	$38+78$	
	$e_n^{(a)}$	286	87	131	68	68	80	100	$10+68$	
	$\overline{n^e}$	218	87	131	-	-	56	75	95	102
	$e_n$								100	131

1	2	3	4	5	6	7	8	9	10
lelel	$l^e$	151	-	91	60	70	75	100	102
	$e_l^e$	254	57	128	69	<u>57+20</u>	100		
	$e_l^{(s)}$	236	77	123	36	<u>77+10</u>	75	100	<u>77+20</u>
	$l^e$	151	-	91	60	<u>21+60</u>	65	100	<u>32+60</u>
	$e_l^e$	264	67	128	69	<u>10+69</u>	65	100	<u>30+69</u>
	$e_l^{(s)}$	236	77	123	36	<u>9+36</u>	80	100	<u>21+36</u>
	$e_l$	200	77	123	-	28	60	100	37
	$r^e$	215	-	115	100	63	100		
	$e_r^e$	322	136	69	117	<u>136</u>	50	100	<u>136+21</u>
	$e_r$	186	128	58	-	<u>118</u>	50	100	<u>128+24</u>
rerer	$r^e$	215	-	114	101	<u>8+101</u>	60	100	<u>19+101</u>
	$e_r^e$	283	116	69	98	<u>30+98</u>	100		
	$e_r$	207	126	79	-	50	85	100	55

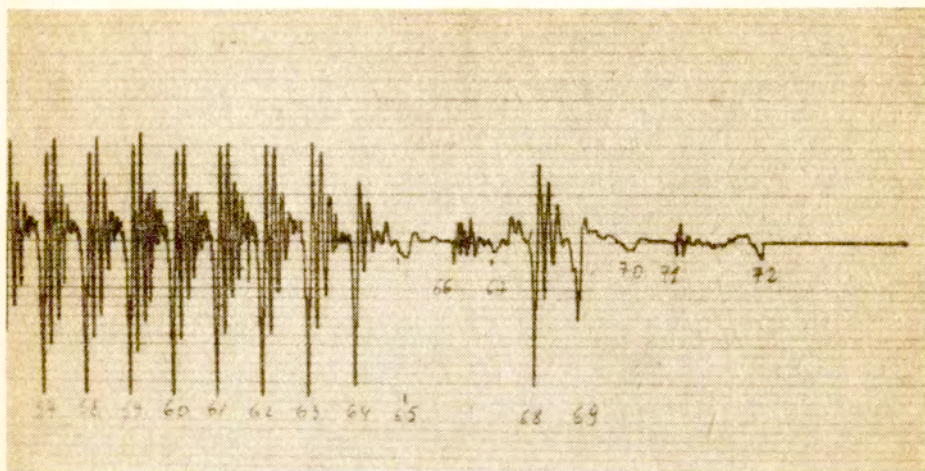
1	2	3	4	5	6	7	8	9	10
wewew	$\overrightarrow{e_w}$	169	-	78	91	38	75	100	65
	$\overrightarrow{e_w}$	339	109	100	130	<u>109+19</u>	95	100	<u>109+39</u>
	$\overrightarrow{e_w(a)}$	260	101	129	30	30	85	100	<u>101+60</u>
	$\overrightarrow{e_w}$	169	-	78	91	<u>60</u>	80	100	<u>11+91</u>
	$\overrightarrow{e_w}$	279	90	100	89	<u>52</u>	85	100	<u>69</u>
	$\overrightarrow{e_w(a)}$	240	81	129	30	<u>30</u>	55	100	<u>30+30</u>
	$\overrightarrow{e_w}$	201	72	129	-	-	59	90 100	<u>99</u> <u>21+129</u>
	$\overrightarrow{e_j}$	204	-	94	110	73	50	100	<u>94+30</u>
	$\overrightarrow{e_j}$	317	87	101	129	<u>87+10</u>	55	95 100	<u>87+30</u> <u>87+69</u>
	$\overrightarrow{e_j(a)}$	228	68	99	51	<u>68+10</u>	50	100	<u>68+40</u>
jejeje	$\overrightarrow{e_j}$	204	-	94	110	<u>92</u>	60	100	<u>110</u>
	$\overrightarrow{e_j}$	337	87	101	149	<u>132</u>	80	100	<u>11+149</u>
	$\overrightarrow{e_j(a)}$	247	88	109	50	<u>20+50</u>	75	95 100	<u>70+50</u> <u>2+109+50</u>
	$\overrightarrow{e_j}$	177	68	109	-	109	75	95	<u>2+109</u>



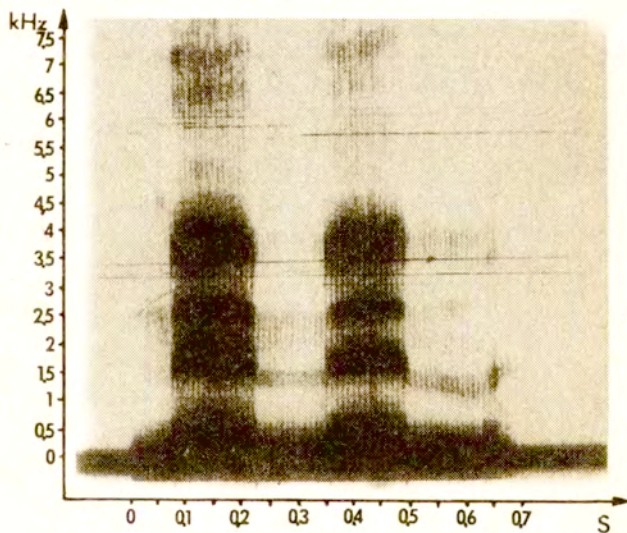
## BIBLIOGRAFIA

1. BONDARKO, L.W., Oscyllografičeskij analiz reči, Leningrad 1965.
2. CEMIEL, G.I., Opoznavanije rečewych signalow, Moskwa 1971.
3. O'CONNOR, J.D., GERSTMAN, L.J., LIBERMAN, A.M., DELATTRE, P.C., COOPER, F.S., Acoustic Cues for the Perception of Initial [w,j,r,l] in English, Word, vol. 13, N 1, str. 24-43, 1957.
4. ČISTOVIČ, L.A., /red./, Fizjologia reči, Wosprijatije reči čelowiekom, Leningrad 1976.
5. DELATTRE, P., From Acoustic Cues to Distinctive Features, *Phonetica* 18, str. 198-230, 1968.
6. DEJUSKA, M., Fonetyka polska. Artykulacje głosek polskich, Warszawa-Kraków 1981.
7. DUKELSKI, N.J., Principy segmentacji rečewego potoka, Moskwa-Leningrad 1962.
8. DUKIEWICZ, L., Polskie głoski nosowe, Warszawa 1967.
9. FANT, G., Akustičeskaja teorija rečeobrazowanija, str. 139-164, Moskwa 1964.
10. JASSEM, W., Podstawy fonetyki akustycznej, Warszawa 1973.
11. JASSEM, W., Mowa a nauka o łączności, Warszawa 1974.
12. KOSIEL, U., Acoustic-phonetic parameters of the Polish frictionless consonants, [l] and [r], *Speech Analysis and Synthesis*, vol. II, str. 135-161, Warszawa 1970.
13. KRAWIEC, D., MATUSZKINA, O., MYTKOWSKI, K., Percepcyjne granice międzysgmentalne w logatomach o budowie CVCV/C/ ze spółgłoską zwartą, *Prace IPPT* 63/1978, Warszawa 1978.
14. LISKER, L., Minimal Cues for Separating [w,r,l,j] in Intervocalic Position, *Word*, vol. 13, N 2, str. 256-267, 1957.
15. ŁCBACZ, P., An Acoustic Analysis of the oppositions /j:i/ and /w:u/ in present-day Polish, *Speech Analysis and Synthesis*, vol. II, str. 119-134, Warszawa 1970.
16. MALECOT, H., Acoustic cues for nasal consonants, *Language*, vol. 32, N 2, str. 274-284, 1956.

17. MATUSZKINA, O., MIKIEL, W., Percepcyjne granice między-segmentalne w połączeniach CVCV/C/ ze spółgłoską trącą, Prace IPPT 25/1979, Warszawa 1979.
18. De MORI, R., GUBRYNOWICZ, R., LAFACE, P., Inference of a Knowledge Source for the Recognition of Nasals in Continuous Speech, IEEE, vol. ASSR-27, N 5, str. 538-550, 1979.
19. NAKATA, K., Synthesis and Perception of Nasal Consonants, JASA, vol. 31, N 6, 1959.
20. ÖHMAN, S., On the Perception of Swedish Consonants in Intervocalic Position, S.T.L. Report, N 25, Stockholm, 1962.
21. SAPOŹKOW, M.A., Sygnał mowy w telekomunikacji i cybernetyce, Warszawa 1966.
22. WANG, W.S-Y., FILLMORE, CH.J., Intrinsic Cues and Consonant Perception, Journal of Speech and Hearing Research, vol. 4, N 2, 1961.

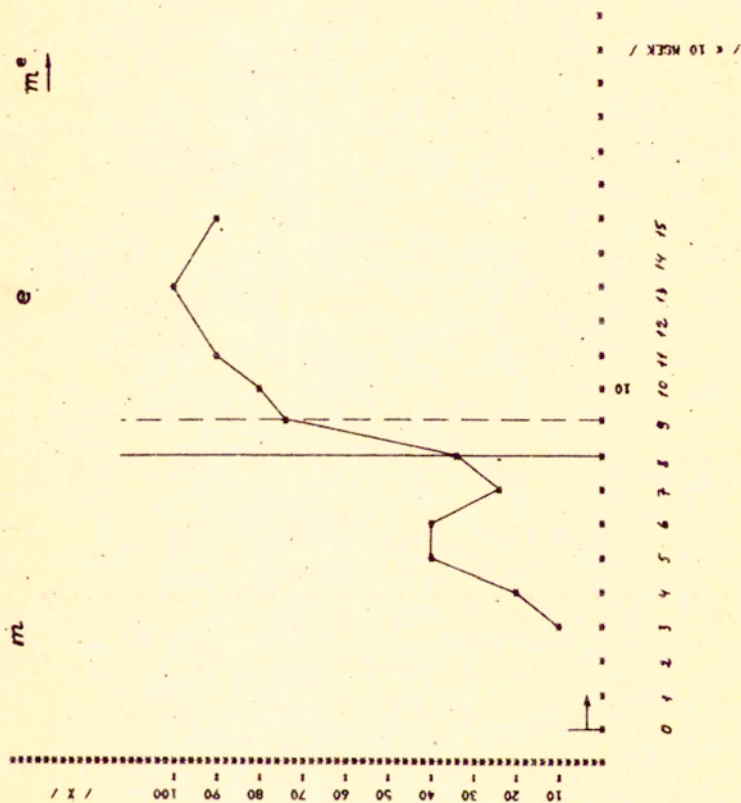


Rys. 1. Fragment oscylogramu, [r] wygłosowe

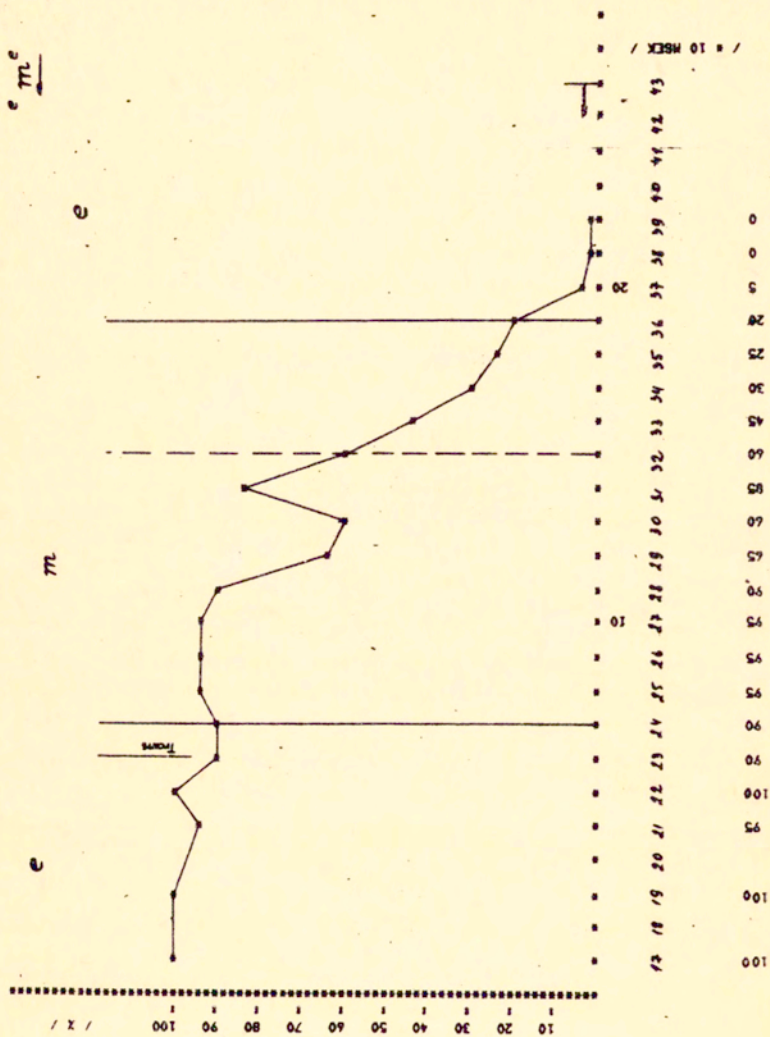


Rys. 2. Spektrogram logatomu [memem]

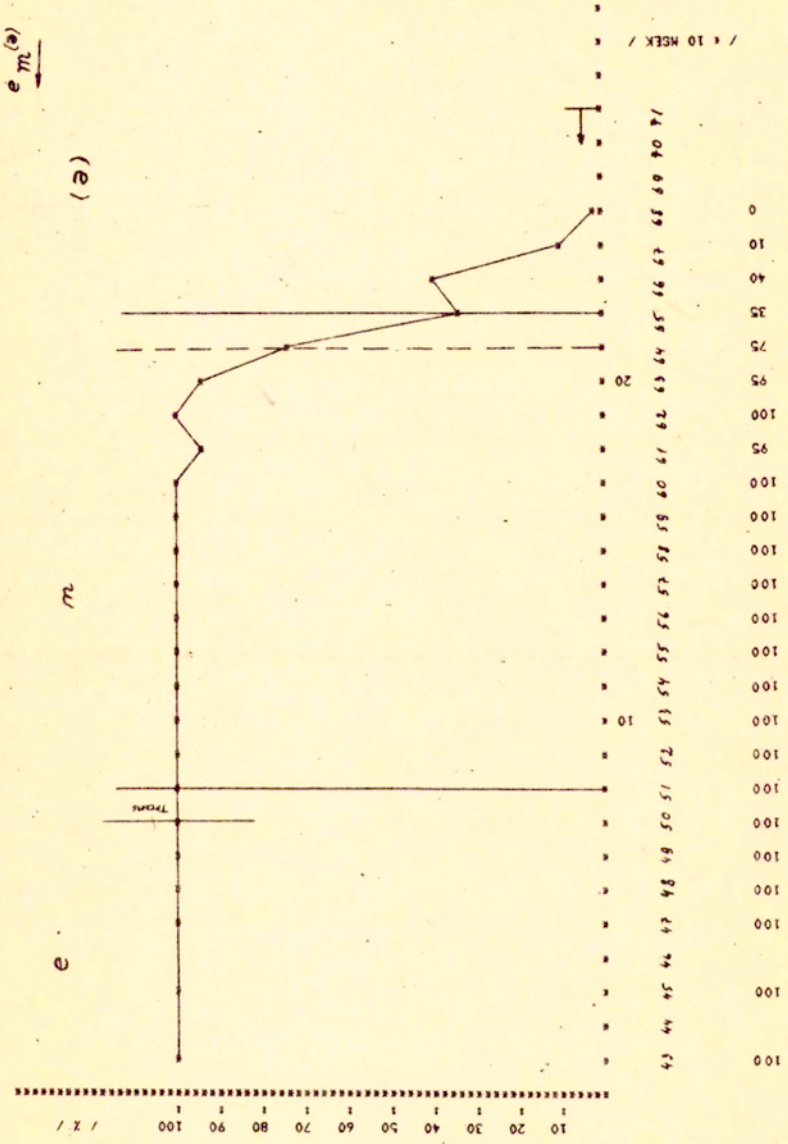




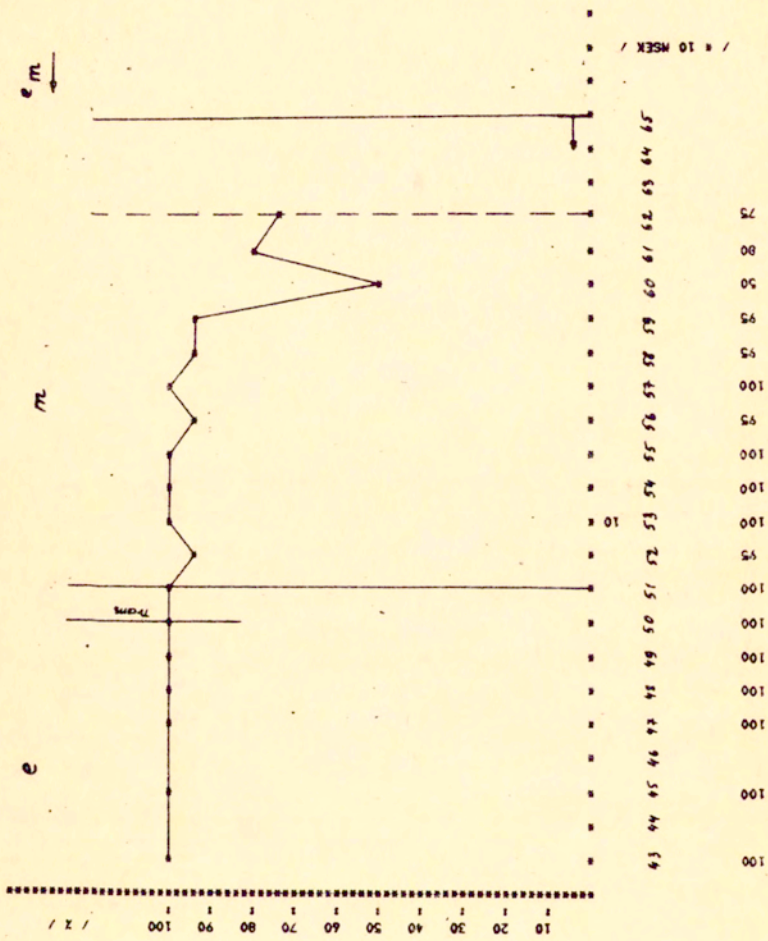
Rys. 4. [m<sub>max</sub>], [m] nagiosowe, seria m<sub>g</sub>



Rys. 5. [memm], [m] Aróggiosowe, seria 00

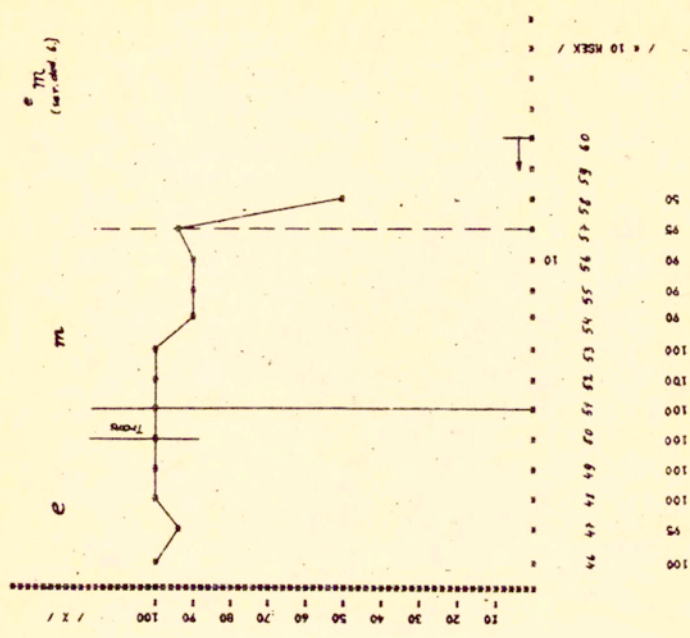
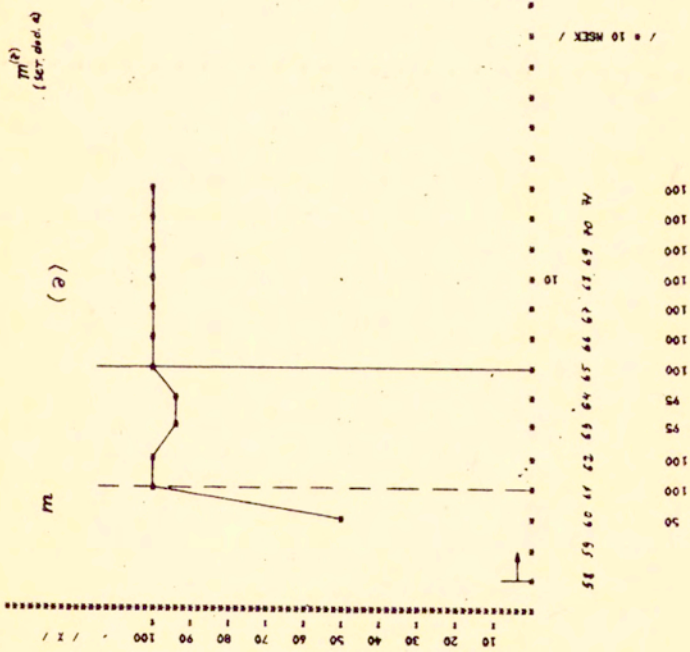


Rys. 6. [aenem], [a] wygłosowe, seria  $e m(a)$



Rys. 7. [memem], [m] wygłosewie, seria e<sub>m</sub>





Rys. 8 a, b, [eeme], [e] wygłosowe, serie dodatkowe a, b ( $\frac{a^{(2)}}{m}$ ,  $\frac{e}{m}$ )



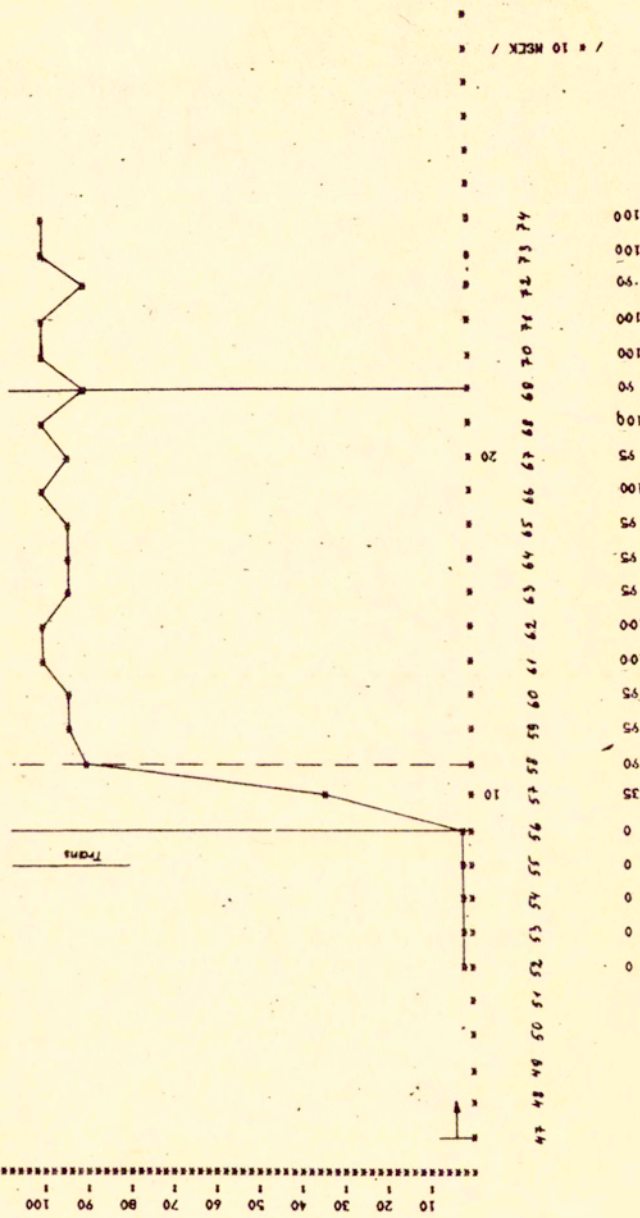
$\epsilon_{\text{H}}^{(a)}$

(a)

$\mu$

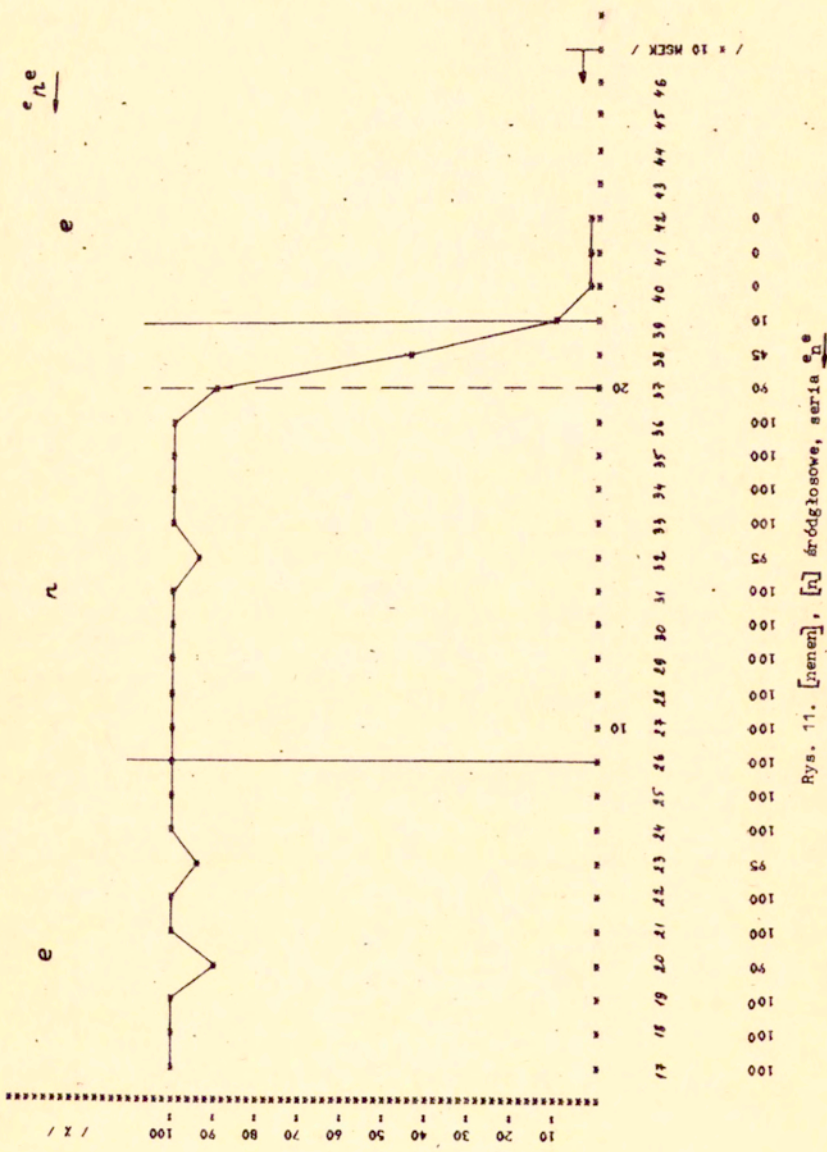
0

/ 2 /



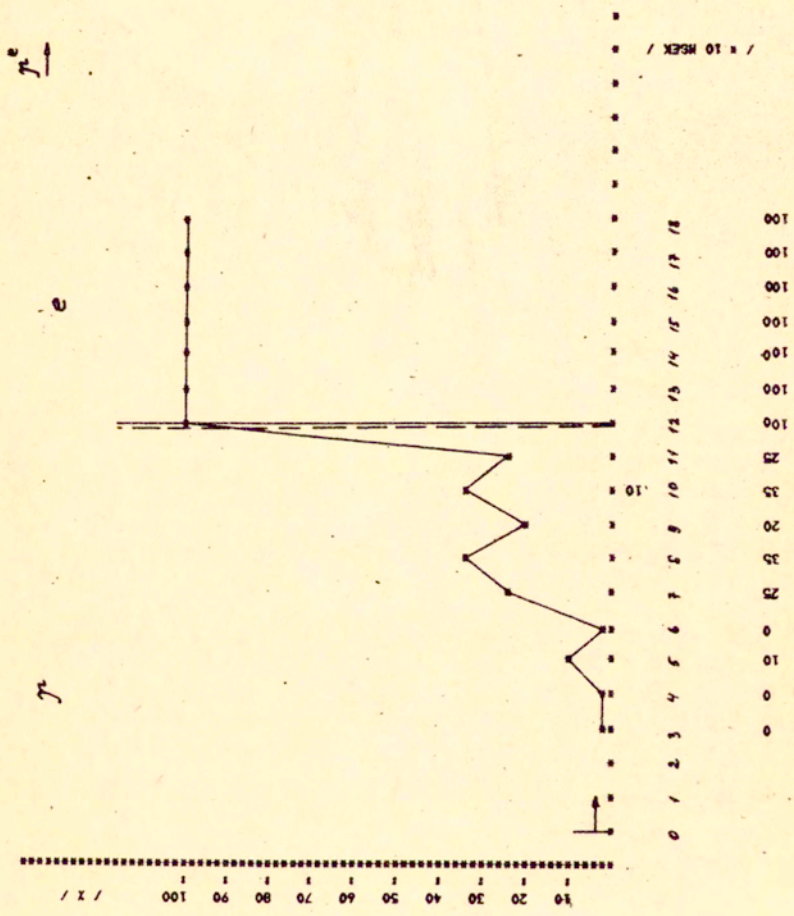
/ 10 MCKK /

Rys. 10. [nener], [n] wygłosowe, seria  $\epsilon_{\text{H}}^{(a)}$

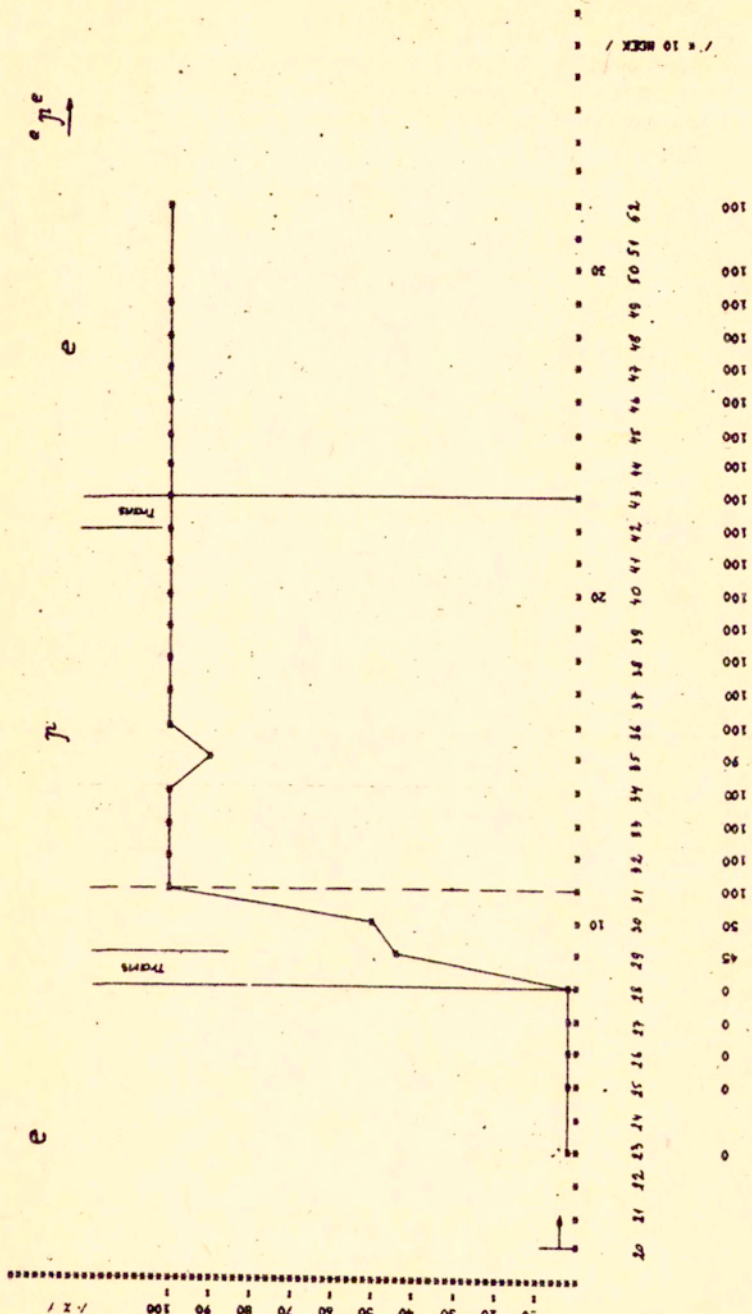


Rys. 11.  $[n_{\text{men}}]$ ,  $[n]$  dróglisowe, seria  $e_n^e$

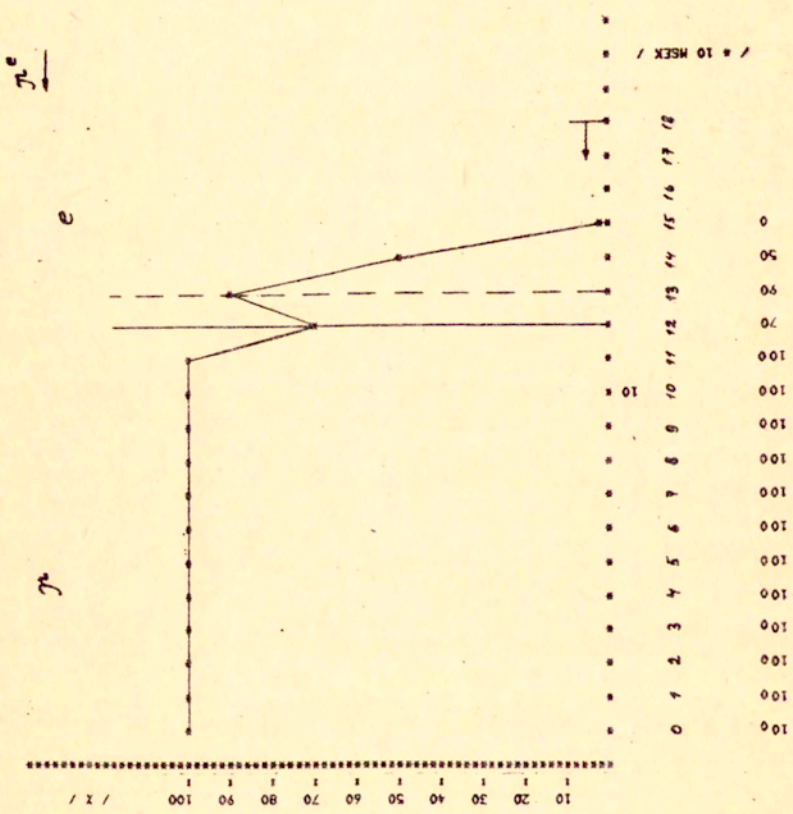




Rys. 13. [r], [x] negatywne, seria r<sup>e</sup>



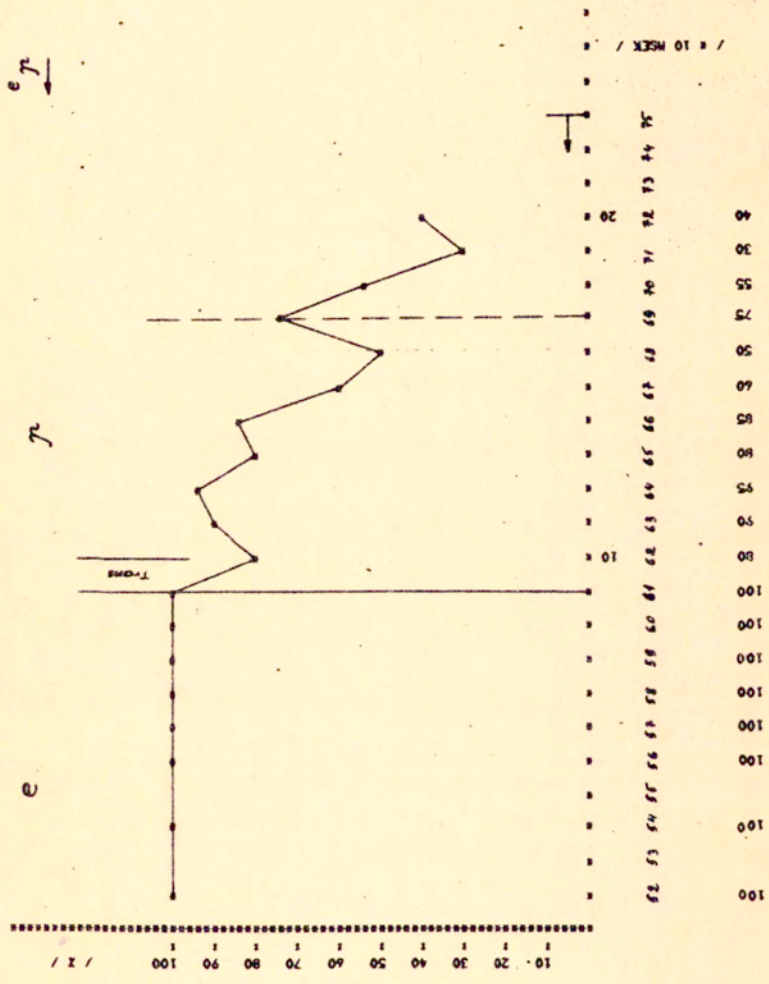
Rys. 14. [p], [e] drógiowe, seria p<sup>e</sup>

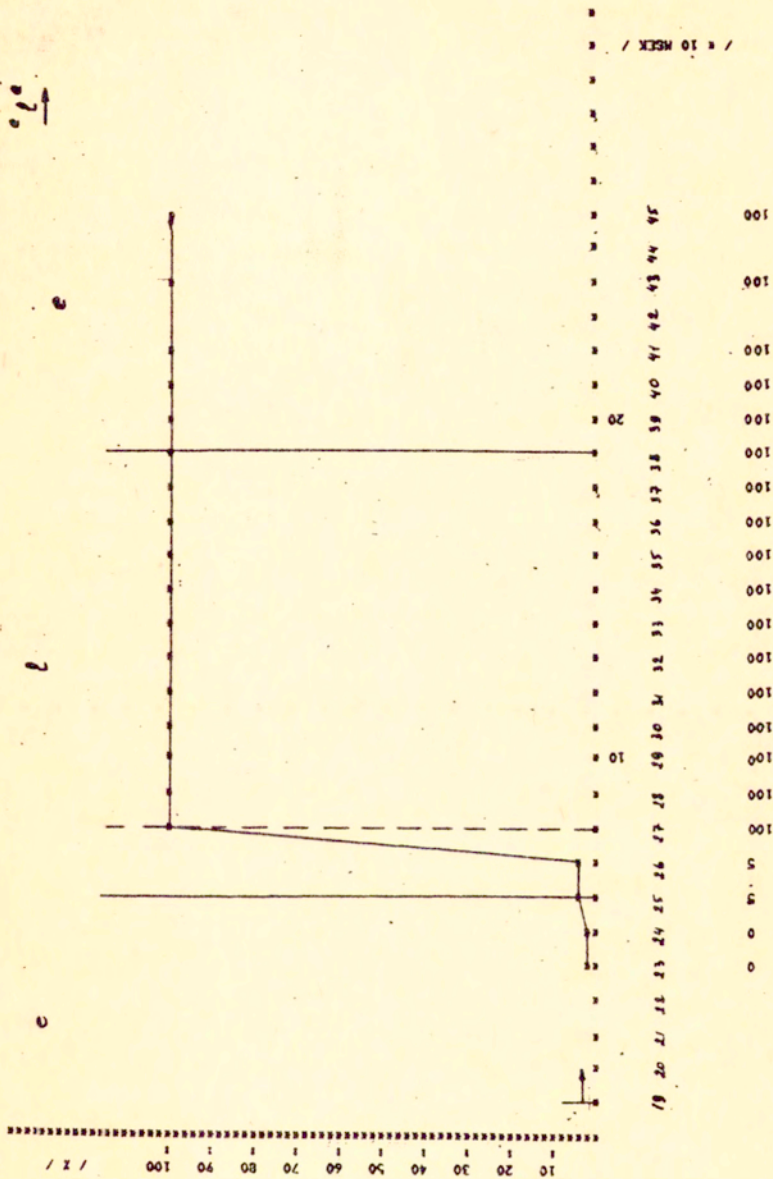


Rys. 15. [J^e]naglosowe, seria J^e

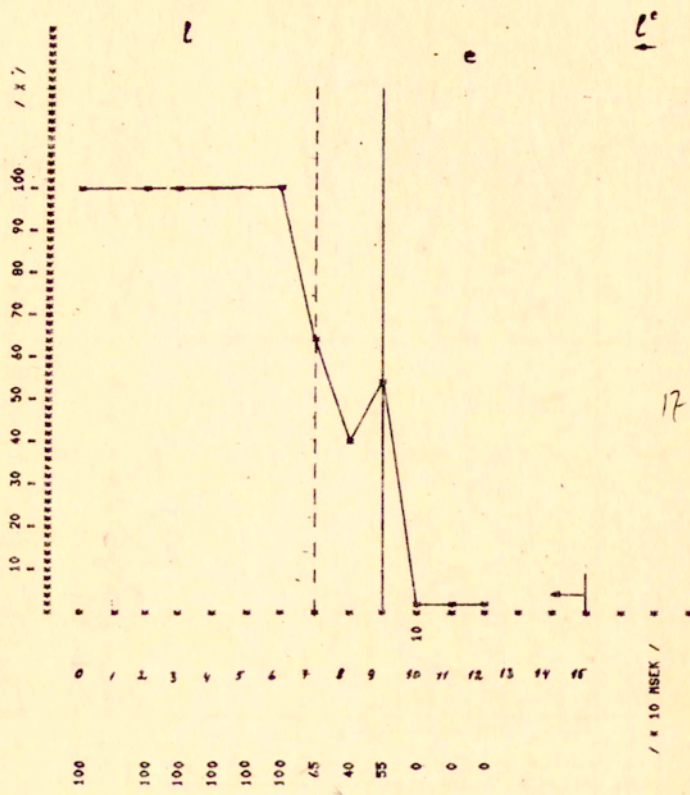






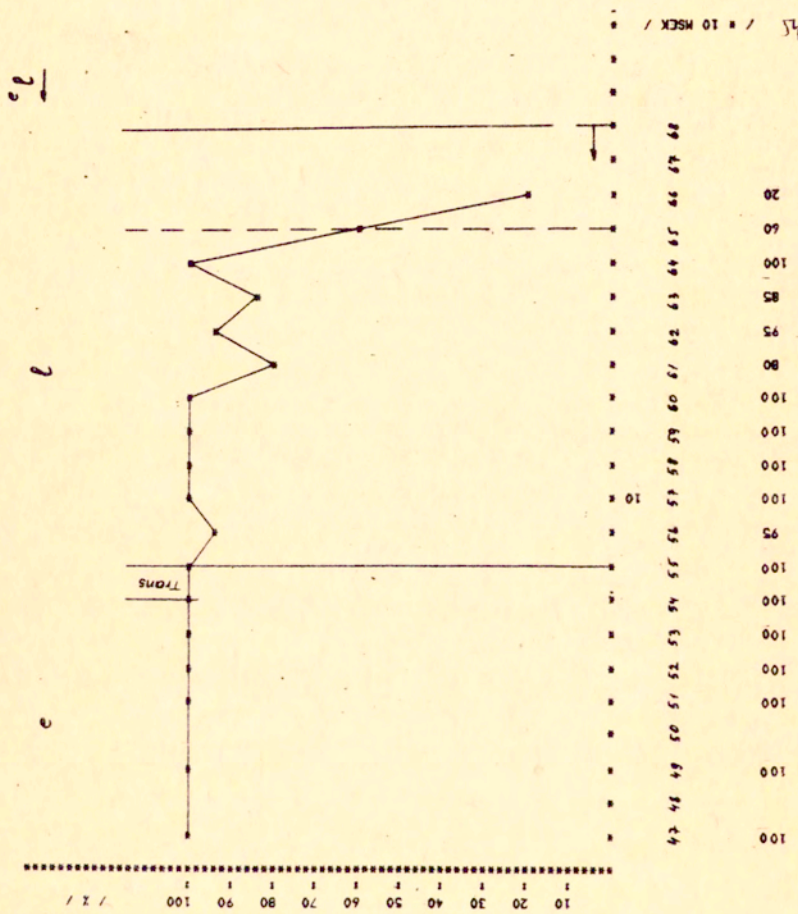


Rys. 18. [lele], [1] drógiosowe, seria  $\theta_{10}$

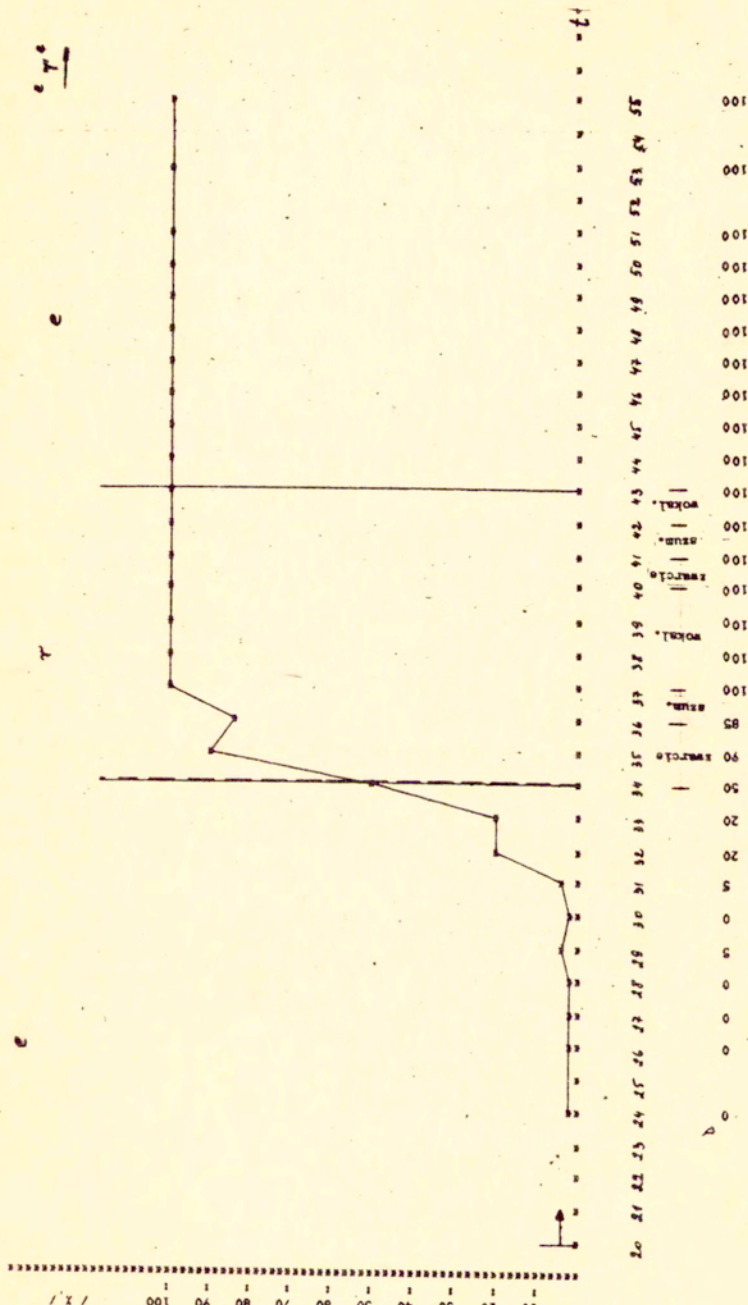


19. [lelel], [l] nagłosowe, seria 1<sup>o</sup>

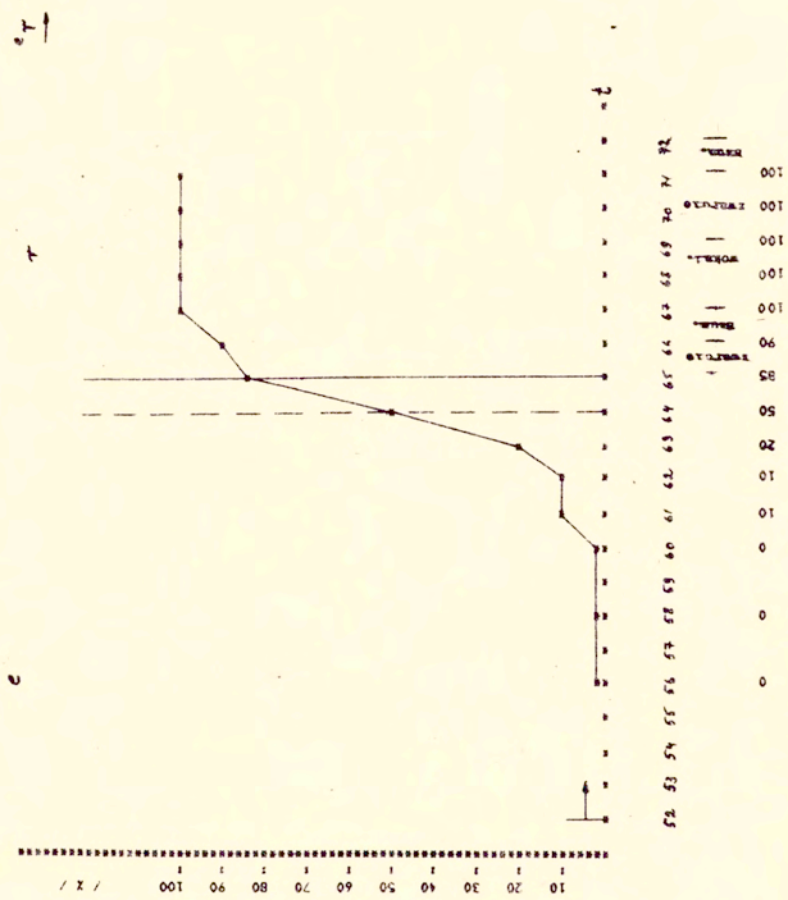




Rys. 21. [1e1e], [1] wygłosowe, seria 0<sub>1</sub>

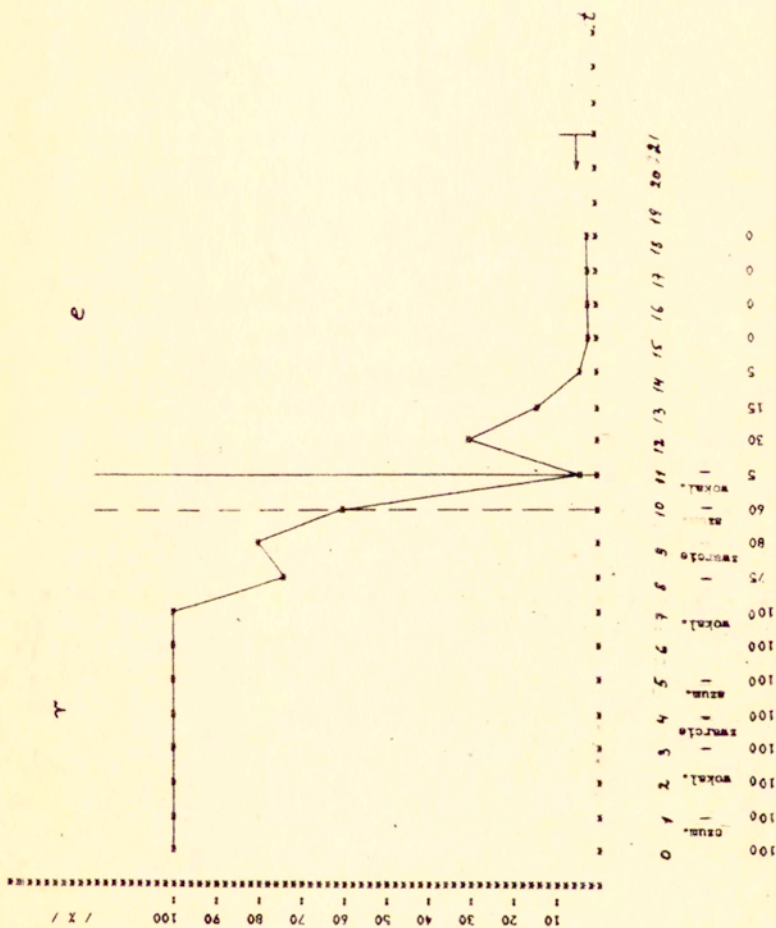


Rys. 22. [rere], [r] óródgiosowe, seria r<sup>o</sup>



rys. 23. [rere], [r] wygłosowe, seria 1.

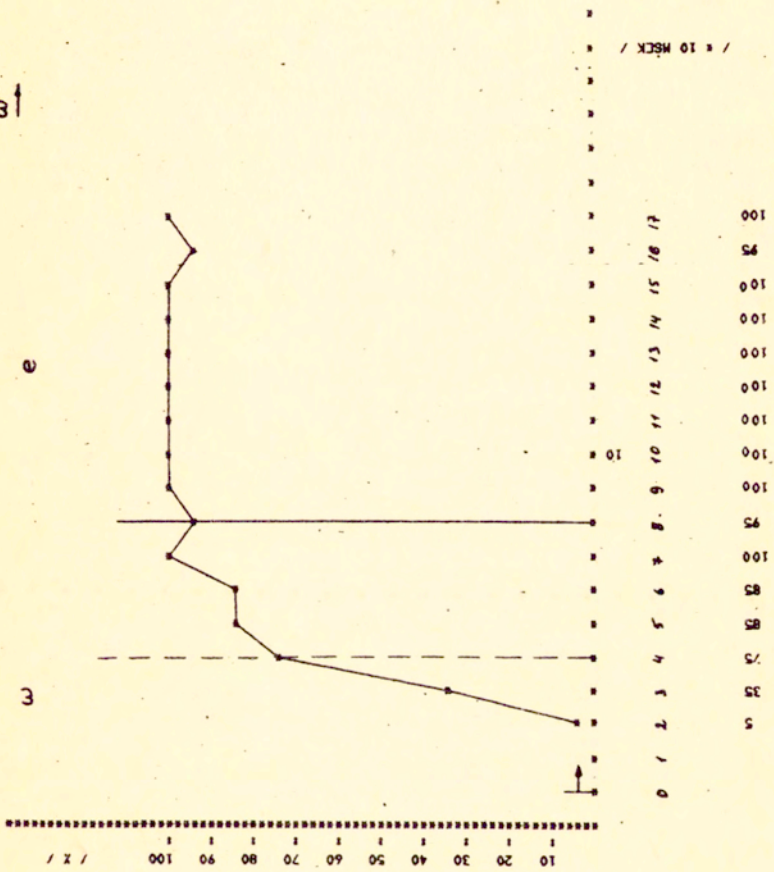




Rys. 24. [r̥] negacje, [r] negacje, seria r<sup>e</sup>

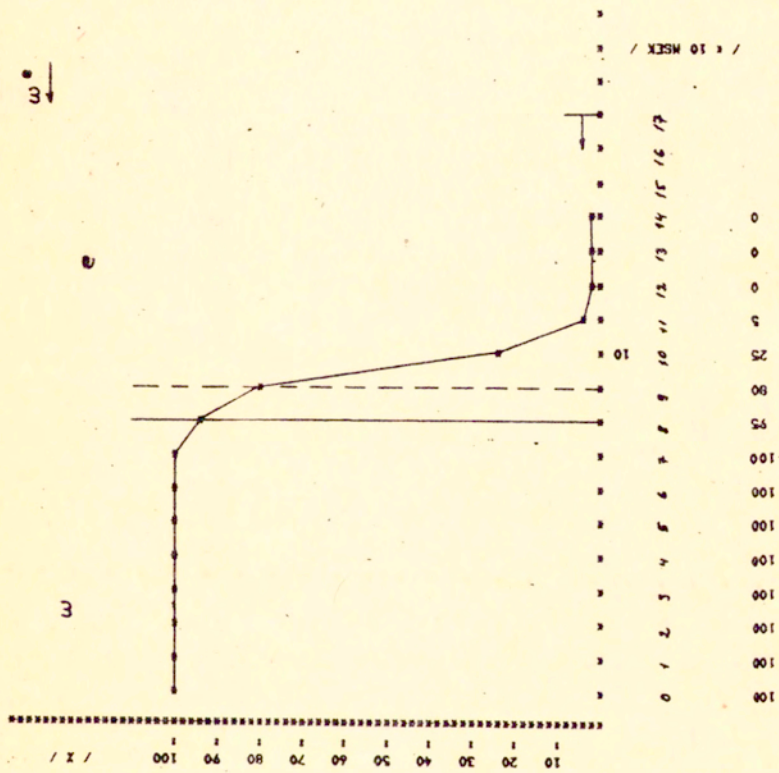


31

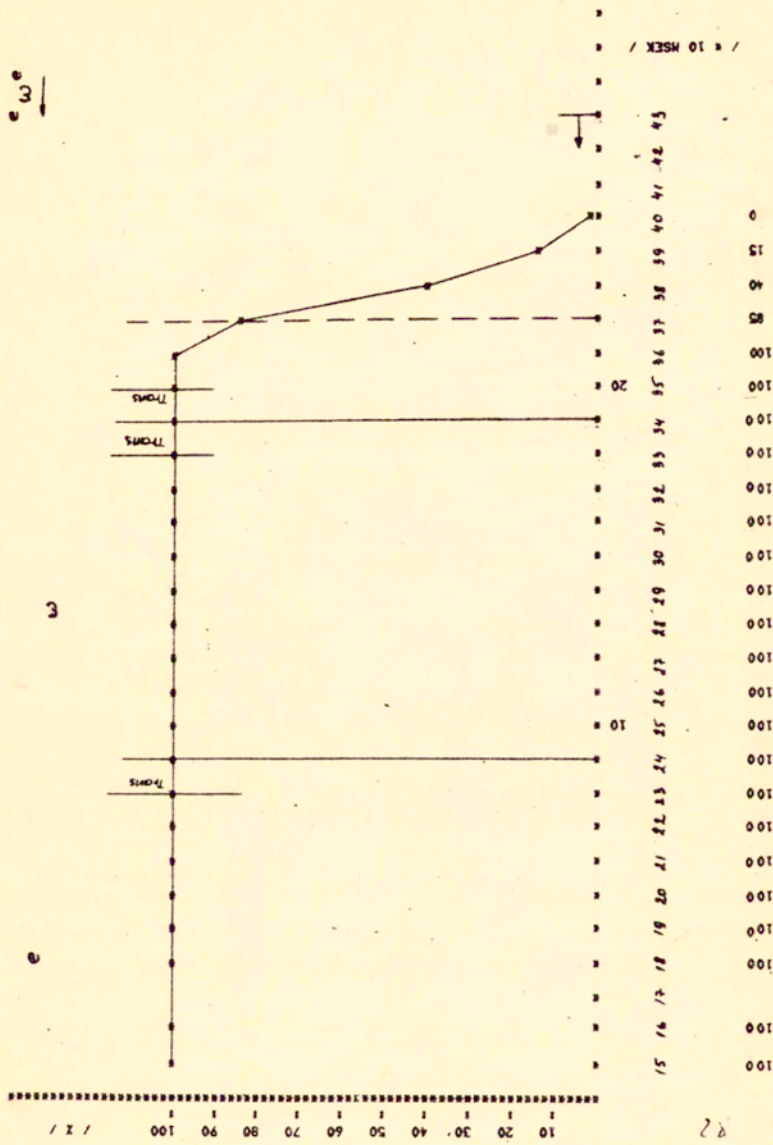


Rys. 26. [view], [w] neglosowe, seria V





Rys. 28. [view], [w] nagłosowe, seria  $v^e$



Rys. 29. [wewew], [w] średniowe, serie  $\omega_w^0$



3 ↓

10 MSEC /

3

THRESH

e

100 90 80 70 60 50 40 30 20 10 / 1 /

↓

42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63

10

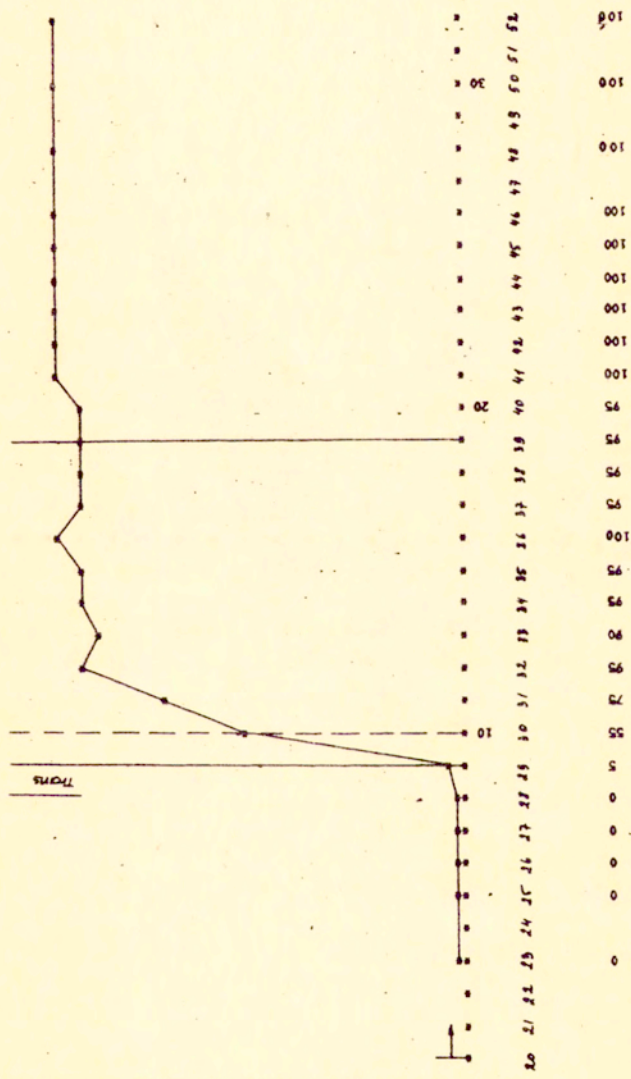
100 100 98 90 93 85 90 96 90 60 50 55 15 30 28

Rys. 31. [weview], [v] wygłosowe, seria 6<sub>w</sub>



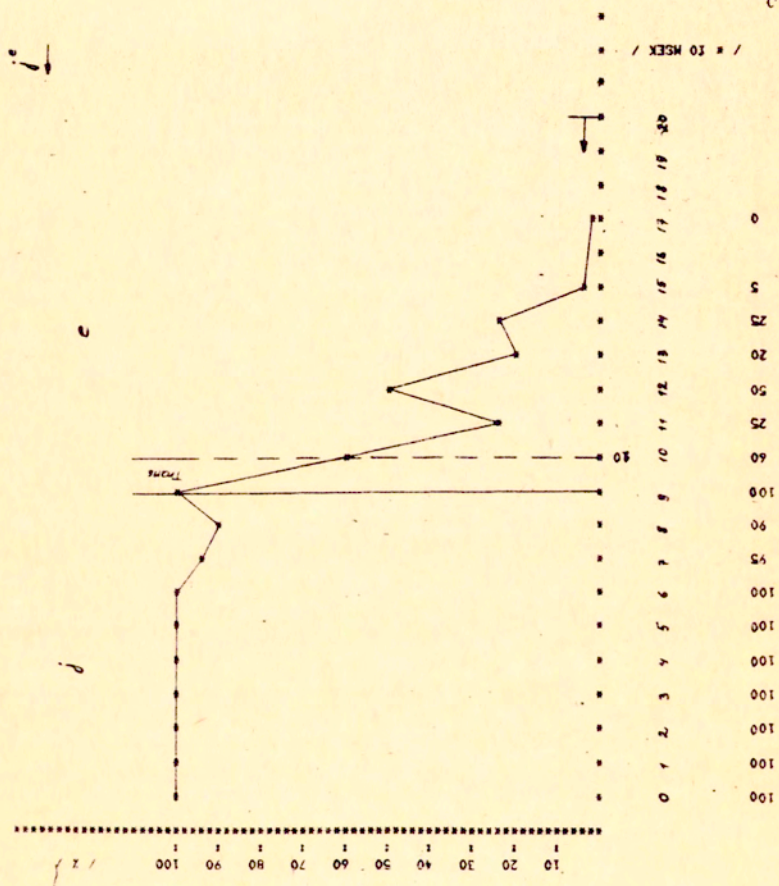
↑  
j

10 20 30 40 50 60 70 80 90 100 / X /

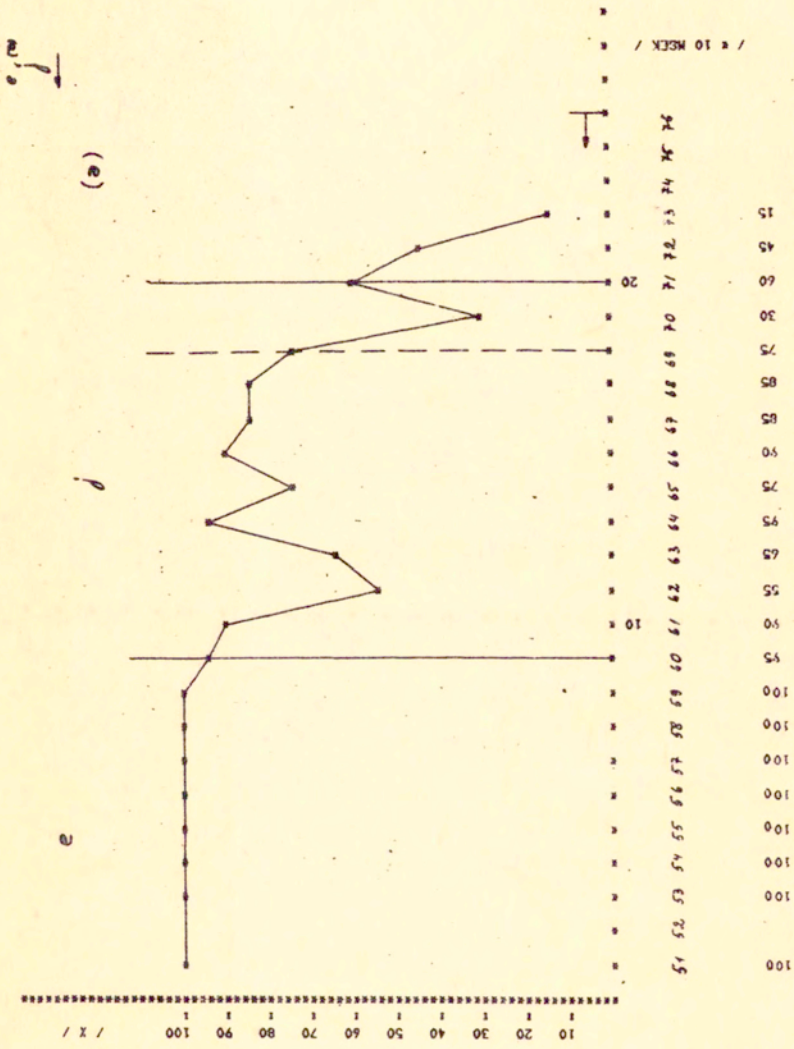


Rys. 32. [jeje], [j] średniosowe, seria 3

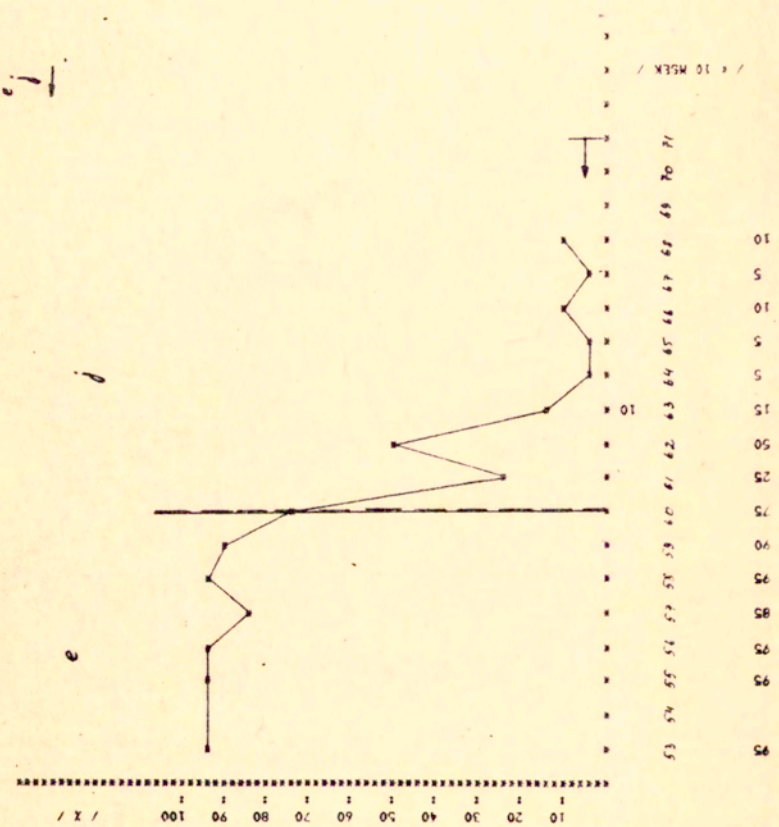
/ \* 10 MSEK /



Rys. 53. [jeje], [j] nagłosowe, seria j<sup>o</sup>



Rys. 34. [deted], [j] wygłosowe, seria  $e_j^{10}$



Rys. 35. [e(e)], [j] wygłosowe, seria e