

Analiza akustyczna polskich spółgłosek trących bezdźwięcznych realizowanych przez dzieci niesłyszące

Acoustic analysis of Polish voiceless fricatives produced by deaf children

Janusz Kleśta

Uniwersytet im. Adama Mickiewicza, Poznań

Streszczenie

Opracowanie skomputeryzowanych narzędzi diagnostycznych oraz dydaktycznych służących rewalidacji mowy w przypadku dzieci z uszkodzonym słuchem jak również konstrukcja systemów automatycznego rozpoznawania mowy, przeznaczonych dla osób głuchych wymaga wyczerpującej wiedzy z zakresu parametrów akustycznych charakteryzujących mowę osób niesłyszących. W obecnych badaniach analizie akustycznej poddano polskie spółgłoski trące bezdźwięczne, realizowane przez dzieci o głębokim ubytku słuchu, a uzyskane wartości parametrów poddano następnie analizie statystycznej. Wyniki analizy dyskryminacyjnej wskazują na możliwość wykorzystania badanych tutaj parametrów akustycznych w procesie konstrukcji narzędzi, wyspecjalizowanych w rozpoznawaniu oraz rewalidacji mowy osób niesłyszących.

Słowa kluczowe: fonetyka akustyczna, surdologopedia, rewalidacja mowy.

Summary

The development of computer-based diagnostic and teaching aids used in the revalidation of hearing-impaired children's speech and the construction of automatic speech recognition systems intended for deaf people require exhaustive knowledge in the area of deaf speech acoustic parameters. In the current research, Polish voiceless fricatives produced by profoundly-deaf children were analyzed acoustically and the parameter values obtained were then analyzed statistically. The results of the discriminant analysis show that the acoustic parameters under investigation can be utilized in the construction of tools specialized in the recognition and revalidation of deaf speech.

Key words: acoustic phonetics, deaf speech, speech revalidation.

Przedstawione w niniejszym artykule analizy stanowią fragment dysertacji doktorskiej [Kleśta 2002], objętej grantem promotorskim KBN nr 1 H01D 020 18

Wprowadzenie

Malejące rozmiary produktów współczesnej elektroniki w połączeniu ze stale rosnącą złożonością ich budowy powodują, iż coraz częściej niewystarczające okazują się w ich przypadku dotychczasowe, konwencjonalne sposoby sterowania, oparte na wszelkiego rodzaju regulatorach natury mechanicznej, stanowiących istotny element hamujący w procesie dalszej miniaturyzacji sprzętu. W związku z powyższym coraz liczniej podejmowane są próby stworzenia nowego narzędzia komunikacji w układzie człowiek–maszyna, którego realizacja sprzętowa wykazywałaby się minimalnymi gabarytami, a prostota obsługi eliminowała konieczność specjalnego przeszkolenia. Wydaje się, że narzędziem spełniającym te wymagania mógłby stać się interfejs, oparty na mowie naturalnej jako kodzie oraz systemie automatycznego jej rozpoznawania jako dekodera sygnału przesyłanego pomiędzy człowiekiem a maszyną [por. Shannon (i in.) 1949], jednakże

konstrukcja tego typu interfejsu wymaga wyczerpującego opisu akustycznego dla sygnału mowy produkowanego przez te wszystkie grupy potencjalnych mówców, których przedstawiciele mieliby stać się jego użytkownikami.

Tymczasem, mimo gwałtownego ostatnimi czasy rozwoju fonetyki akustycznej, nadal niewiele powstało w naszym kraju prac, których temat stanowiłaby akustyczna charakterystyka mowy realizowanej przez całkiem licznie reprezentowaną grupę społeczną, jaką stanowią ludzie głusi¹. Nieliczne badania związane z analizą mowy osób niesłyszących, prowadzone w naszym kraju, koncentrują się przy tym w głównej mierze na cechach prozodycznych sygnału [np. Mikiel (i in.) 1979; Demenko 1987; Karczewska (i in.) 2000; Sieńkowska (i in.) 2000; Stankiewicz (i in.) 2000], natomiast brakuje prawie całkowicie prac poruszających problemy jego segmentacji, czy analizy widmowej [por. Łobacz, Francuzik, Szalkowska 2002; Trochymiuk 2003]. Tymczasem niewystar-

¹ Jak podaje Szczepankowski [1994], niesłyszący w Polsce to około 45 tysięcy osób, w tym kilka tysięcy dzieci w wieku przedszkolnym i szkolnym, przy czym liczba ta uwzględnia wyłącznie osoby o ubytku słuchu wyższym niż 80 dB.

czająca wiedza na temat segmentalnych parametrów mowy realizowanej przez dzieci niesłyszące uniemożliwia również konstrukcję narzędzi diagnostycznych oraz rewalidacyjnych dla tej grupy mówców, wykorzystujących szerokie możliwości współczesnych metod automatycznej analizy sygnału mowy.

W chwili obecnej surdologopeda dokonujący diagnozy jakości mowy osoby niesłyszącej zdany jest zatem w głównej mierze na informacje dostępne w wyniku zindywidualizowanej oceny percepcyjnej analizowanego sygnału, podczas gdy, jak pisze Kapusta [2000], subiektywna możliwość ucha ludzkiego i całego systemu percepcji słuchowej człowieka na rozmaite nieprawidłowości w strukturze i przebiegu analizowanego sygnału jest bardzo zróżnicowana. Niekiedy bowiem duże różnice w formie sygnału mogą być subiektywnie niedostrzegane lub akceptowane jako granice normy, podczas gdy w innym przypadku nawet drobne różnice pomiędzy sygnałami stanowią kryterium do ich różnicowania i są podstawą dyskwalifikacji jednego z nich jako formy patologicznej. Poza tym pamiętać należy, jak stwierdza dalej ten sam autor (*ibid.*), iż każdy przypadek mowy patologicznej jest w gruncie rzeczy unikatowy i niepodobny do innych przypadków, nawet przy tym samym rodzaju patologii. Tymczasem człowiek (np. lekarz) potrafi głównie różnicować oceny pacjentów w zakresie wykrycia podstawowej różnicy między mową poprawną i patologiczną, natomiast w przypadku, kiedy zachodzi potrzeba pogłębionej analizy porównawczej, słuch człowieka okazuje się bezradny.

Jeśli chodzi o kolei o konstrukcję automatycznych narzędzi do rewalidacji mowy, zwrócić należy uwagę, iż podjęto już w naszym kraju próby opracowania programów komputerowych, wykorzystujących automatyczną analizę akustyczną sygnału mowy do celów ćwiczeń nad poprawnością wymowy na poziomie segmentalnym u kilkuletnich dzieci o nieuszkodzonym słuchu. Przykładem takich narzędzi może być pakiet gier logopedycznych Grocholewskiego, stosowany do indywidualnej pracy dziecka słyszącego nad jakością własnej wymowy, przy wykorzystaniu odpowiednio dostosowanych do wieku mówcy technik wizualizacji mowy. Wydaje się, że nic nie stoi na przeszkodzie, aby wykorzystać wspomniane techniki w procesie konstrukcji podobnych narzędzi dla dzieci niesłyszących (czego pierwsze próby zostały już zresztą podjęte), tym bardziej, iż, jak pisze Łobacz [1987, str. 105], „projektowanie urządzeń technicznych oraz przygotowanie programu pedagogicznego dla osób głuchych stanowi jeden z ważniejszych problemów związanych z wizualizacją mowy”. Zasadniczym jednak wymogiem, jaki stawia budowa wspomnianych narzędzi, jest dostęp do odpowiedniego opisu akustycznego mowy dzieci głuchych, który dostarczałby wiedzy na temat głównych źródeł zniekształcenia sygnału w jej przypadku, charakteryzowanych w kategoriach parametrów akustycznych.

Celem badań, których fragment stanowi podstawę niniejszego artykułu, było zatem obok opracowania wstępnej charakterystyki akustycznej mowy dzieci głuchych odnalezienie takich parametrów akustycznych sygnału z nią związanego, na których oprzeć można skuteczną identyfikację poszczególnych segmentów fonetycznych i które w związku z powyższym mogą także odgrywać kluczową rolę w procesie percepcji sygnału produkowanego przez mówców niesłyszących. Identyfikacja odpowiednich wielkości stanowić

mogłaby bowiem punkt wyjściowy dla konstrukcji pomocy diagnostyczno-dydaktycznych, które przy zastosowaniu współczesnych metod automatycznej analizy oraz wizualizacji mowy wykorzystywane byłyby do samodzielnej, bądź kontrolowanej przez surdologopedę, pracy dzieci głuchych nad zrozumiałością własnych wypowiedzi, poprzez doskonalenie poprawności produkcji na poziomie segmentalnym. Poszukiwanie takich parametrów odbywało się w oparciu o wielkości akustyczne oraz metody statystyczne, stosowane wspólnie do celów automatycznego rozpoznawania mowy niezaburzonej, a także przy wykorzystaniu wybranych partii informacji dostępnych w wyniku analizy spektrograficznej, stanowiącej w dalszym ciągu jedno z podstawowych narzędzi w zakresie badań nad mową zaburzoną [por. Ball (i in.) 1997].

Przedmiot analiz

Niniejszy artykuł przedstawia wyniki badań obejmujących realizację spółgłosek trących bezdźwięcznych. Tę grupę głosek postanowiono poddać analizie głównie z tego względu, że należą do niej sybilanty, których artykulacja, jak podaje Emiluta-Roza [1994], sprawia dzieciom niedosłyszącym największe trudności, przede wszystkim ze względu na ograniczone możliwości ich odbioru słuchowego w związku z kumulacją sygnału w wysokim pasmie częstotliwości oraz utrudniony odbiór wzrokowy, jak również konieczność wykonania precyzyjnych ruchów narządów artykulacyjnych w trakcie ich produkcji [por. także Gallagher (i in.) 1975; Daniloff (i in.) 1980; Ingram (i in.) 1980; Stephens (i in.) 1986; Łobacz (i in.) 1999].

Dodatkowym czynnikiem, który zdecydował o wyborze tej właśnie grupy głosek, był fakt, iż poważne problemy związane z ich poprawną identyfikacją pojawiają się także w przypadku podejmowanych prób automatycznego rozpoznawania mowy niezaburzonej [por. Jassem 1995a; Jassem 1995b], co spowodowane jest wciąż niedokładną znajomością parametrów akustycznych, odpowiedzialnych w największym stopniu za zróżnicowanie barwy głosek z opisywanej grupy.

Materiał badawczy

Wszelkie analizy przeprowadzone w ramach obecnych badań oparte zostały na materiale zgromadzonym dzięki uprzejmości i współpracy wychowanków oraz pracowników Ośrodka Szkolno-Wychowawczego dla Dzieci Niesłyszących przy ulicy Bydgoskiej w Poznaniu. W czasie, kiedy gromadzony był materiał badawczy (1999 rok), szkoła podstawowa opisywanej placówki posiadała strukturę identyczną jak ówczesna szkoła masowa na tym samym poziomie, to znaczy obejmowała klasy od 1 do 8, a także dodatkowo klasę zerową.

Wiek dzieci rozpoczynających naukę w szkole podstawowej przy ulicy Bydgoskiej wynosił zazwyczaj sześć lub siedem lat, choć niektórzy wychowankowie uczęszczali do powyższej szkoły dopiero od ósmego, dziewiątego, dziesiątego, czy nawet jedenastego roku życia, co najczęściej występowało w takich przypadkach, kiedy dane dziecko było uprzednio wychowankiem szkoły masowej lub innego ośrodka szkolno-wychowawczego. W związku z tym, poszczególne klasy szkoły podstawowej przy ulicy Bydgoskiej nie były

homogeniczne ze względu na wiek uczniów, jak ma to zazwyczaj miejsce w szkołach powszechnych.

Wychowankowie ośrodka byli mieszkańcami Poznania lub innych miast i miejscowości województwa wielkopolskiego, przy czym dzieci należące do drugiej z wymienionych grup zamieszkiwały w przyszkolnym internacie, w którym proces wychowawczy przebiegał w tzw. „postaci złagodzonej” [por. Hoffman 1987]. Większość kadry dydaktycznej ośrodka stanowili nauczyciele słyszący, którzy w kontaktach z dziećmi stosowali w miarę możliwości metodę totalnej komunikacji [por. Krakowiak (i in.) 1992; Szczepankowski 1994], nie posługując się jednak jako narzędziem podstawowym językiem migowym, którego znajomość u większości z nich była ograniczona. Język migowy stanowił natomiast podstawowy środek komunikacji pomiędzy dziećmi zamieszkującymi w internacie.

Ponieważ Hoffmann [1987] podaje, że uszkodzenie słuchu nie zawsze jest ograniczone jedynie do jego ubytku, bowiem częstokroć towarzyszą mu pewne dodatkowe upośledzenia, takie jak zaburzenia sprawności psychicznej, afazja, zaburzenia charakterologiczne, zaburzenia w funkcjonowaniu innych zmysłów, bądź uszkodzenia narządu ruchu, a Okalidou et al. [1999, str. 396] stwierdzają, że przy doborze materiału do badań nad mową dzieci głuchych niezwykle istotne jest, aby „*the deaf speakers had no other documented handicaps*”, przy współpracy wychowawców Ośrodka dla Dzieci Niesłyszących dokonano do celów obecnych badań wyboru 37 mówców, których jedynym udokumentowanym upośledzeniem był głęboki ubytek słuchu oraz u których nie zaobserwowano zaburzeń natury emocjonalnej czy znacznych odstępstw od intelektualnej normy rozwojowej.

Charakterystyka mówców

Badana grupa dzieci obejmowała uczniów klas od trzeciej do ósmej i charakteryzowała się w miarę równomiernym rozkładem, jeżeli chodzi o płeć oraz klasę, przy czym również na poziomie poszczególnych klas liczba mówców obu płci była podobna.

W przypadku 32 spośród 37 badanych dzieci ubytek słuchu dla lepszego ucha, określony na podstawie ostatniego badania audiometrycznego, był wyższy niż 90 dB, czyli kwalifikował badanych do grupy osób z „głęboką utratą słuchu” zarówno na podstawie klasyfikacji biap (Międzynarodowe Biuro Audiofonologii) [por. Hoffmann 1987], jak i w oparciu o normy ANSI (*American National Standards*) oraz ISO (*International Standardization Organization*) [por. Pruszewicz 2000a; Pruszewicz 2000b]. Średni ubytek w lepszym uchu, obliczony dla powyższej grupy dzieci, wyniósł 97 dB, przy czym najwyższą wartość, jaką osiągnął, to 113 dB. Jedynie piątka spośród wybranych do badań dzieci wykazywała ubytek słuchu w lepszym uchu niższy niż 90 dB, przy czym w przypadku trojga mówców wyniósł on 80 dB, natomiast u pozostałych dwojga 70 dB. Powyższe wyniki badania audiometrycznego kwalifikują wymienionych pięciu mówców do grupy osób o „wysokim stopniu utraty słuchu” według aktualnej klasyfikacji biap oraz do grupy osób o „nasilonym ubytku słuchu” według norm ANSI oraz ISO. Pamiętać jednak należy, że jeszcze nie tak dawno jako dolną granicę tak zwanej głuchoty całkowitej przyjmowano ubytek na poziomie 70 dB [por. Góralówna 1994].

Wszystkie badane dzieci posiadały uszkodzenie słuchu typu odbiorczego [por. Krakowiak 1995], spowodowane rozmaitymi czynnikami natury genetycznej, prenatalnej, perinatalnej lub też związane z chorobą w wieku wczesnodziecięcym albo też z toksycznym środkiem farmakologicznym podanym w tym okresie [por. etiologiczna klasyfikacja uszkodzeń słuchu w: Hoffmann 1987]. Należy jednak zwrócić uwagę, że w przeszło połowie przypadków (20 mówców) niemożliwe było ustalenie faktycznej przyczyny uszkodzenia słuchu, co jest zjawiskiem często wspomnianym w literaturze.

Jak powszechnie wiadomo, niezwykle istotną rolę ze względu na możliwości rozwoju mowy u osób niesłyszących odgrywa wiek, w którym nastąpiła utrata słuchu [por. Góralówna 1994]. Jak wynika ze zgromadzonych danych, 35 spośród 37 opisywanych dzieci utraciło słuch w tak zwanym okresie prelingwalnym, czyli przed ukończeniem drugiego roku życia [por. Góralówna 1994], a jedynie u dwójki z nich nastąpiło to nieznacznie później, w tak zwanym okresie wczesnolingwalnym. Należy jednak zwrócić uwagę, że napotkane trudności z jednoznacznym ustaleniem przyczyny utraty słuchu uniemożliwiały także w wielu przypadkach precyzyjne określenie momentu, w którym utrata ta nastąpiła.

Zaledwie trójka spośród badanych uczniów zaopatrzona została w aparat słuchowy przed ukończeniem trzeciego roku życia, co z pewnością stanowi bardzo niekorzystny czynnik, jeśli chodzi o rozwój funkcji słuchowych w ramach analizowanej grupy mówców. Większość spośród badanych dzieci otrzymała aparaty słuchowe pomiędzy 3 a 7 rokiem życia, przy czym w przypadku przeszło połowy z nich nastąpiło to w wieku 3-4 oraz 6-7 lat, a więc w momencie pójścia do przedszkola lub szkoły podstawowej. Ośmioro dzieci aparatowanych było jednostronnie, natomiast pozostałe obustronnie.

Jak wiadomo, niezwykle istotną rolę, nie tylko dla rozwoju mowy, lecz również dla ogólnego rozwoju intelektualno-emocjonalnego dziecka głuchego, odgrywa odpowiednio wczesne podjęcie czynności rehabilitacyjnych [por. Eckert 1994; Gałkowski 1994; Góralówna 1994]. Spośród analizowanej grupy dzieci jedynie połowa poddana została jakimkolwiek formom rehabilitacji przed rozpoczęciem nauki w szkole podstawowej dla niesłyszących, przy czym rehabilitacja ta polegała w większości przypadków na korzystaniu z pomocy logopedy (11 osób) lub poradni dla dzieci głuchych (6 osób).

Uważa się także [por. Krakowiak (i in.) 1992; Pietrzak 1994; Prillwitz 1996], że odpowiednio wczesne zaoferowanie dziecku głuchemu możliwości komunikacji za pomocą języka migowego umożliwia mu rozwój na poziomie porównywalnym z rozwojem dzieci słyszących, zarówno w wymiarze intelektualnym, jak i komunikacyjno-językowym. W związku z tym dość istotną rolę w możliwości rozwoju mowy u danego dziecka wydaje się również odgrywać poziom znajomości języka migowego, szczególnie we wczesnym wieku, na etapie rozpoczęcia nauki w szkole specjalnej. Jak już wcześniej stwierdzono, język migowy nie stanowi co prawda podstawowego narzędzia komunikacji w procesie kształcenia niesłyszących w poznańskiej szkole przy ulicy Bydgoskiej, jednakże jest on bardzo chętnie stosowany przez uczące się tam dzieci w ramach ich wzajemnych kontaktów, szczególnie w przypadku uczniów przebywających w internacie, co stanowi podstawowe źródło rozwoju wspomnianego języka wśród badanej grupy. Oprócz możliwości nauki języka migo-

wego w środowisku niesłyszących rówieśników, przeszło jedna czwarta badanych uczniów (10 osób) miała także szanse na jego wczesną akwizycję w środowisku domowym (rodzice bądź rodzeństwo danego dziecka to także osoby niesłyszące).

Dla trzydziściorga pięciorga spośród trzydziściorga siedmiorga badanych dzieci przeprowadzone zostały również na którymś z etapów edukacji szkolnej testy psychologiczno-rozwojowe. Jak wynika z uzyskanych rezultatów, sprawność intelektualna zdecydowanej większości uczniów, których mowę poddawano analizie w ramach obecnych badań, mieściła się w granicach odpowiedniej normy, natomiast nieznaczne opóźnienia rozwojowe odnotowano jedynie w przypadku trojga dzieci.

Materiał nagraniowy

Ponieważ Rakowska [1994, str. 128] stwierdza, iż „rozwój świadomości językowej jedenastoletniego dziecka głuchego, ucznia szkoły specjalnej, nie przekracza poziomu umiejętności językowych trzy- czteroletniego dziecka słyszącego”, po konsultacjach z wychowawcami poznańskiego Ośrodka, w ramach prowadzonych badań postanowiono wykorzystać 40-wyrazowy obrazkowy test artykulacyjny dla słyszących dzieci żłobkowych w wieku od 2 do 3 lat, skonstruowany przez M. Golanowską z Wyższej Szkoły Pedagogiki Specjalnej w Warszawie [Golanowska 1999].

Realizacja nagrań

Sesje nagraniowe z udziałem 37 dzieci zrealizowano w Ośrodku Szkolno-Wychowawczym dla Dzieci Niesłyszących przy ulicy Bydgoskiej w Poznaniu. Każda z sesji składała się z dwóch części: w pierwszej z nich poszczególne elementy leksykalne wydobywane były, zgodnie z założeniami zastosowanego testu artykulacyjnego, przy użyciu serii obrazków (test nazywania), natomiast w części drugiej nagrywane osoby odczytywały listę wyrazową testu (test sprawności czytania). Z zasady unikano w przebiegu nagrań metody imitacyjnej, choć ten sposób uzyskiwania próbek mowy nie jest sprzeczny z założeniami teoretycznymi, dotyczącymi realizacji testów artykulacyjnych [por. Grunwell 1987; Krajna 1998].

O ile przeczytanie listy wyrazowej nie sprawiało trudności żadnemu z nagrywanych dzieci, w związku z czym zgromadzony w ten sposób kompletny korpus danych obejmował 1480 jednowyrazowych próbek mowy (37 mówców x 40 jednostek leksykalnych), nie wszystkie wyrazy udało się uzyskać stosowaną w pierwszej części nagrania metodą „dobrowolnego wymuszenia”. Ogółem 37 nagrywanych dzieci nie zrealizowało w ramach testu nazywania 319 elementów leksykalnych, tj. ok. 20% wszystkich wyrazów, przy czym na uwagę zasługuje fakt, że żadne z 37 dzieci nie zdołało zrealizować w ramach obrazkowego testu nazywania całego 40-elementowego zestawu wyrazów².

Na wspomniane powyżej 319 niezrealizowanych elementów leksykalnych składały się naturalnie również niektóre spośród wykorzystywanych w ramach niniejszych badań wyrazów, zawierających w przypadku prawidłowej artykulacji głoskę tracą bezdźwięczną, a mianowicie:

- *kieszeń* (głoska [ʃ]) – brak realizacji przez 30 mówców,
- *fartuszek* (głoski [f] oraz [ʃ]) – brak realizacji przez 22 mówców,
- *wąsy* (głoska [s]) – brak realizacji przez 9 mówców,
- *ciasto* (głoska [s]) – brak realizacji przez 8 mówców,
- *hau hau* (głoska [x]) – brak realizacji przez 6 mówców,
- *siedzi* (głoska [c]) – brak realizacji przez 6 mówców,
- *myje się* (głoska [c]) – brak realizacji przez 5 mówców,
- *fotel* (głoska [f]) – brak realizacji przez 4 mówców,
- *lew* (głoska [f]) – brak realizacji przez 4 mówców,
- *łóżko* (głoska [ʃ]) – brak realizacji przez 3 mówców,
- *nos* (głoska [s]) – brak realizacji przez 2 mówców,
- *pies* (głoska [s]) – brak realizacji przez 1 mówcę.

Wybór materiału do analizy akustycznej

Celem dokonania wyboru materiału nagraniowego, podawanego ostatecznie analizie akustycznej w ramach obecnych badań, postanowiono posłużyć się kryterium wykorzystywanym już we wcześniejszych pracach [np. Levitt 1972; Fornby (i in.) 1981], opartym na stopniu zrozumiałości analizowanego sygnału, „*measured by playing randomized recordings of the children's speech to normal-hearing listeners not previously exposed to the speech of the deaf*” [Levitt 1972, str. 39]. W tym celu przeprowadzono odpowiednie testy percepcyjne, których szczegółowy opis oraz wyniki znaleźć można w Kleśta [2003a].

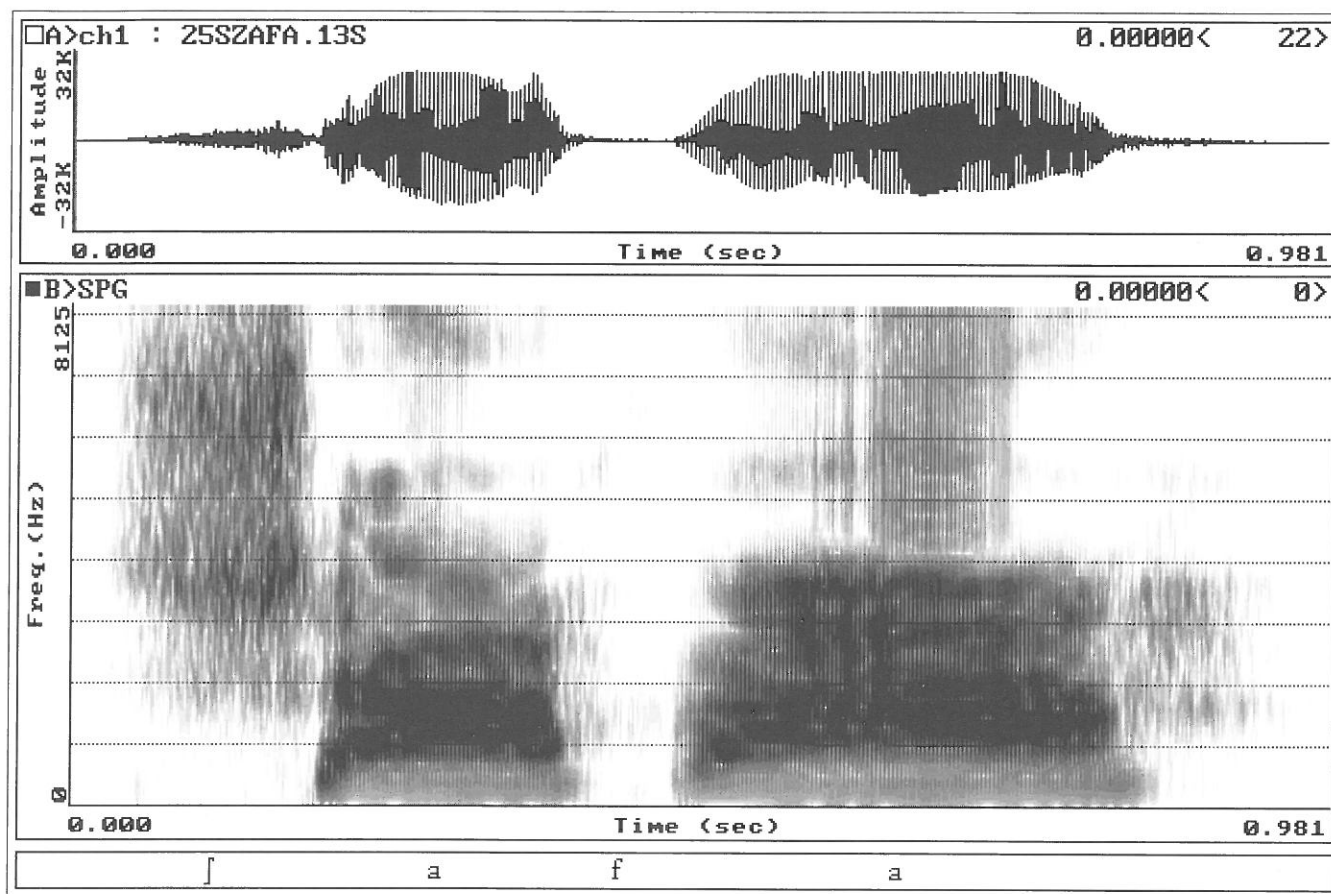
Choć w badaniach nad strukturą akustyczną mowy dzieci głuchych, opisywanych w literaturze, do etapu analiz akustycznych zakwalifikowany zostaje najczęściej wyłącznie materiał nagraniowy obejmujący wypowiedzi tych osób, dla których w ramach testów percepcyjnych odnotowano poziom zrozumiałości przekraczający pewien minimalny pułap [por. np. Levitt 1972; Fornby (i in.) 1981], w ramach obecnych badań analizom akustycznym postanowiono poddać zarówno materiał zrealizowany przez mówców wykazujących się najwyższą zrozumiałością mowy, jak i tych, w przypadku których obliczony poziom zrozumiałości był najniższy. Do pierwszej z wymienionych grup zakwalifikowano 8 dzieci, w przypadku których w ramach przeprowadzonych testów percepcyjnych prawidłowo rozpoznana została przez słuchaczy przynajmniej 1/3 nadanych bodźców wyrazowych, natomiast do grupy drugiej zakwalifikowano ósemkę dzieci, która w ramach testów percepcyjnych uzyskiwała najslabsze wyniki, tzn. nie więcej niż 10% bodźców wyrazowych rozpoznanych poprawnie [por. Kleśta 2003a]. Tak dobrane ośmioosobowe grupy mówców były w pełni zrównoważone pod względem płci, co oznacza, że w każdej z nich znalazły się cztery dziewczynki oraz czterech chłopcy.

Transkrypcja analizowanego materiału

Kolejny etap badań stanowiło opracowanie transkrypcji zgromadzonego materiału³. Choć analizie poddawana była mowa zaburzona, będąca następstwem głębokiego niedosłuchu, faktycznie interesowały nas jedynie takie jej parametry, które mają decydujący wpływ na identyfikację poszczególnych głosek jako realizacji odpowiednich fonemów z zakresu inwentarza odbiorców o słuchu niezaburzonym i w ten

² Omówienie przyczyn związanych z częstym brakiem realizacji niektórych elementów leksykalnych znaleźć można w Kleśta [2002].

³ Szczegółowy opis procedur zastosowanych w celu uzyskania transkrypcji zgromadzonego materiału, jak również omówienie problemów związanych z transkrypcją mowy zaburzonej, znaleźć można w Kleśta [2002], a także Andruszka i wsp. [2000].



Ryc. 1. Oscylogram oraz spektrogram realizacji wyrazu „szafa”, transkrybowanej jako /ʃafa/, mówca o wysokim poziomie zrozumiałości, test obrazkowy

sposób oddziałują na poziom zrozumiałości mówców głuchych przez mówców słyszających.

Choć więc transkrybenci proszeni byli w rzeczywistości o wykonanie wąskiej transkrypcji fonetycznej, uwzględniającej szereg zjawisk artykulacyjnych charakterystycznych dla mowy dzieci głuchych, najbardziej interesujące z punktu widzenia obecnych badań było przede wszystkim zakwalifikowanie przez nich poszczególnych usłyszanych głosek jako realizacji odpowiednich fonemów z ich własnego inwentarza. Analizowana pod tym kątem transkrypcja fonetyczna nie była zatem niczym innym, jak kolejnym testem percepcyjnym, którego przedmiot stanowiły wyrazy zrealizowane przez dzieci niesłyszące, natomiast zadaniem odpowiednio przeszkolonych w tym przypadku słuchaczy [por. Kleśta 2002] był odbiór oraz rozpoznanie kolejnych fonemów, a oczekiwaną reakcją – ich zapis przy pomocy symboli alfabety IPA. Przeprowadzoną transkrypcję można zatem określić mianem typowego testu wyrazistości [por. Jassem 1973, str. 281].

Na podstawie wykonanej transkrypcji do etapu analiz akustycznych zakwalifikowano wyłącznie te realizacje, które percypowane były przez transkrybentów jako zgodne z zamierzonymi, czyli jako realizacje odpowiednich fonemów trących bezdźwięcznych, przy czym identyfikacja segmentów akustycznych związanych z ich artykulacją w strukturze ana-

lizowanego sygnału na podstawie spektrogramu oraz oscylogramu nie sprawiała większych trudności (por. ryc. 1).

Analizowane parametry akustyczne

Do tradycyjnie badanych parametrów spektralnych w przypadku analizy spółgłosek trących bezdźwięcznych należą częstotliwości formantów F1-F4⁴. Poza tym, jak podaje Łobacz [1998, str. 144], „poszukiwania kryteriów akustycznych pozwalających na poprawne rozpoznawanie wszystkich dziewięciu głosek szumowych języka polskiego, różniących się pięcioma miejscami artykulacji doprowadziło Jassem [1995] do wykorzystania w klasyfikacji tych głosek trzech następujących momentów centralnych: średniej ważonej częstotliwości widma, skośności widma oraz kurtozy (stro-kości)”. Wymienione momenty centralne, obliczane na podstawie uśrednionego widma FFT analizowanych segmentów, opisywane są następującymi wzorami [por. Forrest (i in.) 1988; Jassem 1995b]:

M1 — Pierwszy moment widma (średnia ważona częstotliwości)

$$M1 = \frac{\sum_j \hat{O}_j f_j}{\sum_j \hat{O}_j}$$

⁴ Jak podaje Jassem: „te skupienia energii w widmie, które odpowiadają czterem pierwszym formantom, wykazują regularności pozaindywidualne”, podczas gdy „częstotliwości [...] maksimów widma przy wyższych częstotliwościach są silnie zróżnicowane osobniczo” [1973, str. 223].

M3 — Trzeci moment widma (skośność)

$$M3 = \frac{\sum_j \hat{O}_{ij} (f_j - \bar{f})^3}{\sum_j \hat{O}_{ij}}$$

M4 — Czwarty moment widma (stromość)

$$M4 = \frac{\sum_j \hat{O}_{ij} (f_j - \bar{f})^4}{\sum_j \hat{O}_{ij}} - 3$$

gdzie: f = częstotliwość
 i = liniowe natężenie widmowe
 j = kolejne punkty na skali częstotliwości

W ramach obecnych badań analizom statystycznym poddano wartości osiągane przez sześć parametrów akustycznych, a mianowicie częstotliwości pierwszych pięciu szczytów widmowych w zakresie do ok. 8000 Hz oraz pierwszy moment centralny widma. Z analizy statystycznej momentów widmowych M2 oraz M3 zrezygnowano, ponieważ z badań Jassem [1995b] wynika, iż pomimo pewnych własności dyskryminacyjnych powyższych dwóch parametrów akustycznych, „for the fricatives, the first moment is virtually the only reliably discriminant variable” (str. 93).

Zastosowane metody analizy akustycznej⁵

Całość analiz akustycznych, objętych opisywanymi badaniami, zrealizowano przy użyciu cyfrowego analizatora akustycznego CSL50 model 4300B⁶ firmy KAY Elemetrics, zaimplementowanego w postaci karty dźwiękowej, współpracującej z odpowiednim oprogramowaniem za pośrednictwem komputera klasy PC.

Jak już wcześniej wspomniano, wzrokowa analiza zawartości spektrogramów oraz oscylogramów uzyskanych przy pomocy powyższego oprzyrządowania posłużyła wyłącznie identyfikacji segmentów akustycznych związanych z realizacją odpowiednich fonemów w analizowanym sygnale mowy, natomiast do celów wyznaczenia wartości badanych parametrów akustycznych zastosowano odpowiednie techniki analizy cyfrowej.

Celem uzyskania danych służących obliczeniu wartości pierwszego momentu centralnego widma zastosowana została 128-punktowa analiza FFT, natomiast w pomiarach wartości pierwszych pięciu szczytów formantowych badanych segmentów mowy wykorzystywana była analiza LPC 14 rzędu, przy czym uzyskane przy jej pomocy dane na temat szerokości wstęg formantów posłużyły eliminacji ewentualnych błędów analizy statystycznej, stanowiącej integralny element metody LPC.

Wyniki analizy akustycznej

Ponieważ przeprowadzone testy percepcyjne dały podobne rezultaty, jeśli chodzi o poziom zrozumiałości sygnału mowy produkowanego przez dzieci niesłyszące w ramach obrazkowego testu nazywania oraz czytanej listy wyrazowej,

a także ze względu na duże podobieństwo zjawisk artykulatoryjnych odnotowanych w transkrypcji fonetycznej materiału zgromadzonego powyższymi dwiema metodami oraz na fakt, iż na podstawie wcześniejszych badań nad strukturą akustyczną samogłosek realizowanych przez tę samą grupę mówców [por. Francuzik i Szalkowska 2001] stwierdzono, że sposób gromadzenia materiału nagraniowego pozostaje bez większego wpływu na parametry widmowe poszczególnych głosek, postanowiono, że w ramach obecnych badań parametry akustyczne głosek realizowanych w teście obrazkowym oraz czytanej liście wyrazowej rozpatrywane będą wspólnie. W toku wstępnej analizy statystycznej dane rozdzielone zostały natomiast ze względu na płeć mówców oraz ich poziom zrozumiałości.

Analizie akustycznej poddane zostały zarówno realizacje fonemów obejmujących polskie głoski sybilantne /s/, /ʃ/ oraz /ç/, jak i pozostałe dwie bezdźwięczne głoski trące występujące w języku polskim, czyli /x/ oraz /f/. Stosunkowo najszerszy materiał badawczy dostępny był w odniesieniu do fonemów /s/ (pięć wystąpień), a także /ʃ/ oraz /f/ (po cztery wystąpienia), które w przypadku prawidłowej artykulacji realizowane były w wyrazach: /tʃasto/, /ser/, /pjes/, /nos/, /võsi/ (fonem /s/), /ʃafa/, /wuʃko/, /cjesep/, /fartuʃek/ (fonem /ʃ/) oraz /lef/, /ʃafa/, /fotel/ i /fartuʃek/ (fonem /f/). Spółgłoska /ç/ pojawiała się w teście trzykrotnie (wyrazy /mice/, /miʃece/ oraz /cedzi/), natomiast fonem /x/ zaledwie dwukrotnie, przy czym oba wystąpienia miały miejsce w onomatopei /xawxaw/.

Częstotliwości formantów

W celu wyznaczenia pierwszych pięciu szczytów w obwiedni widma posłużono się analizą LPC 14 rzędu, stosowaną wraz z metodą autokorelacji, preemfazą o wartości 0,9 oraz oknem analizującym Hamminga o szerokości 20 ms. W przypadku każdej głoski dokonywano trzech pomiarów, w takich punktach czasowych, aby parametry widmowe wyznaczane były na początku, na końcu oraz w środkowej części badanego segmentu, przy jednoczesnym uniknięciu pomiarów w obrębie ugięć formantowych [por. Jassem 1998].

Fonem /s/

Zgodnie z tym, co podaje Jassem [1973], w spółgłoskach trących sybilantnych poziom pierwszego formantu jest zazwyczaj tak niski, że w większości wypadków formant ten nie jest w ogóle widoczny w widmie. Fakt ten znalazł potwierdzenie również w obecnych badaniach, w których wyznaczenie częstotliwości F1 w realizacjach fonemu /s/ możliwe było w przypadku zaledwie 15% analiz. Średnia wartość powyższego formantu wyniosła w zgromadzonym materiale 478 Hz, przy czym zaznaczyć trzeba, że wszystkie realizacje, w których możliwa była jego ekstrakcja, pochodziły z nagrań głosów żeńskich.

Żadnych trudności nie sprawiało natomiast wyznaczenie częstotliwości pozostałych formantów spółgłoski trącej [s], a ich wartości średnie wyniosły odpowiednio 2117 Hz dla F2 (odchylenie standardowe = 485 Hz), 3616 Hz dla F3 (SD = 428 Hz), 4892 Hz dla F4 (SD = 502 Hz) oraz 6148 Hz dla F5 (SD = 518 Hz). Choć wartość średnia każdego z tych czterech

⁵ Szczegółowy opis zastosowanych metod analizy akustycznej oraz samych procedur pomiarowych znaleźć można w Kleśta [2002].

⁶ CSL – Computer Speech Laboratory.

rech formantów wyższa była w przypadku dziewcząt niż chłopców (por. tab. 1), przeprowadzona analiza wariancji wykazała, iż na poziomie $p < 0.01$ różnice te były statystycznie istotne jedynie w przypadku F2 ($p < 0.0039$). Dla wartości żadnego z badanych formantów nie stwierdzono natomiast statystycznie istotnych różnic ze względu na poziom zrozumiałości mówców.

Tab. 1. Średnie wartości formantów w realizacjach fonemu /s/ w Hz, w zależności od płci mówców

Formant	Płeć mówców	
	Dziewczęta	Chłopcy
F2**	2205	1932
F3	3651	3537
F4	4923	4823
F5	6174	6088

Fonem /j/

W przypadku fonemu /j/, w żadnej z realizacji nie było możliwe automatyczne wyznaczenie wartości pierwszego formantu, który zgodnie z danymi teoretycznymi [por. Kudela 1968; Jassem 1973] powinien przypadać w obrębie częstotliwości odnotowanych dla F1 głoski [s]. Wyznaczenie częstotliwości pozostałych formantów metodą LPC odbywało się natomiast w sposób całkowicie automatyczny, a ich wartości średnie wyniosły odpowiednio 2263 Hz dla F2 (odchylenie standardowe = 417 Hz), 3600 Hz dla F3 (SD = 410 Hz), 4601 Hz dla F4 (SD = 447 Hz) oraz 6084 Hz dla F5 (SD = 436 Hz).

Choć częstotliwości średnie powyższych czterech formantów wyższe były w artykulacjach dzieci o niskiej zrozumiałości mowy (por. tab. 2), w przypadku żadnego z nich analiza wariancji nie wykazała, na poziomie $p < 0.01$, statystycznie istotnych różnic w osiąganych wartościach ze względu na poziom zrozumiałości czy też płeć mówców (por. tab. 3).

Tab. 2. Średnie wartości formantów w realizacjach fonemu /j/ w Hz w zależności od poziomu zrozumiałości mówców

Formant	Poziom zrozumiałości	
	Wysoki	Niski
F2	2217	2452
F3	3553	3790
F4	4549	4806
F5	6033	6285

Tab. 3. Średnie wartości formantów w realizacjach fonemu /j/ w Hz w zależności od płci mówców

Formant	Płeć mówców	
	Dziewczęta	Chłopcy
F2	2308	2158
F3	3642	3503
F4	4639	4511
F5	6016	6243

Fonem /c/

W przypadku fonemu /c/, podobnie jak dla /j/, niemożliwe było w żadnej z realizacji automatyczne wyznaczenie wartości pierwszego formantu, który zgodnie z danymi teoretycznymi [por. Kudela 1968; Jassem 1973] przypadać powinien w okolicach od 400 do 600 Hz. Wyznaczone w spo-

sób automatyczny częstotliwości pozostałych formantów wyniosły natomiast średnio 2390 Hz dla F2 (odchylenie standardowe = 428 Hz), 3810 Hz dla F3 (SD = 386 Hz), 4773 Hz dla F4 (SD = 386 Hz) oraz 6134 Hz dla F5 (SD = 452 Hz), przy czym jedynie w przypadku F3 analiza wariancji wykazała statystycznie istotne różnice ($p < 0.045$) pomiędzy wartościami odnotowanymi w wymowie chłopców oraz dziewcząt (por. tab. 4).

Tab. 4. Średnie wartości formantów w realizacjach fonemu /c/ w Hz w zależności od płci mówców

Formant	Płeć mówców	
	Dziewczęta	Chłopcy
F2	2413	2363
F3**	3697	3941
F4	4768	4779
F5	6051	6230

Fonem /x/

Choć, jak podaje Jassem [1973], poziom pierwszego formantu w takich głoskach trących, jak [h], [χ] czy [x], powinien być stosunkowo wysoki, w ramach obecnych badań również w przypadku realizacji fonemu /x/ wielokrotnie niemożliwe było automatyczne wyznaczenie częstotliwości F1 przy pomocy zastosowanej metody LPC. Odpowiednie wartości uzyskano zaledwie w 38% analiz, a ich średnia wyniosła 946 Hz, co sugeruje (podobnie jak średnia częstotliwość F3 – por. poniżej), że w większości przypadków realizowane były warianty języczkowe lub gardłowe [por. Jassem 1973]. Wartości średnie pozostałych formantów w realizacjach badanego fonemu wyniosły odpowiednio 1836 Hz dla F2 (odchylenie standardowe = 324 Hz), 3264 Hz dla F3 (SD = 510 Hz), 4465 Hz dla F4 (SD = 437 Hz) oraz 5947 Hz dla F5 (SD = 902 Hz).

Statystycznie istotne różnice (na poziomie $p < 0.000002$) odnotowano wyłącznie w przypadku formantu F5, który przyjmował znacznie wyższe wartości w realizacjach dzieci o niskim poziomie zrozumiałości mowy niż w artykulacjach mówców o wysokiej zrozumiałości (odpowiednio 6622 Hz oraz 5557 Hz). Tab. 5 zawiera średnie wartości formantów dla realizacji badanego fonemu w zależności od płci mówców.

Tab. 5. Średnie wartości formantów w realizacjach fonemu /x/ w Hz w zależności od płci mówców

Formant	Płeć mówców	
	Dziewczęta	Chłopcy
F2	1806	1880
F3	3184	3379
F4	4445	4494
F5	5967	5918

Fonem /f/

W przypadku realizacji fonemu /f/ automatyczne wyznaczenie częstotliwości pierwszego formantu metodą LPC okazało się równie trudne, w realizacjach pozostałych fonemów trących bezdźwięcznych. Odpowiednie wartości uzyskano zaledwie w 22% przeprowadzonych analiz, a ich średnia wyniosła 746 Hz. Średnie wartości pozostałych formantów w realizacjach fonemu /f/ wyniosły w badanym materiale odpowiednio 1776 Hz dla F2 (odchylenie standar-

dowe = 424 Hz), 3385 Hz dla F3 (SD = 455Hz), 4800 Hz dla F4 (SD = 564 Hz) oraz 6244 Hz dla F5 (SD = 574 Hz).

Jak wykazała analiza wariancji, statystycznie istotne różnice ($p < 0,0014$) dotyczyły jedynie formantu drugiego, który w wymowie dzieci o niskim poziomie zrozumiałości osiągał wartości średnio o 240 Hz wyższe niż w przypadku mówców o wysokiej zrozumiałości (odpowiednio 1946 Hz oraz 1706 Hz). Tab. 6 zawiera średnie wartości formantów dla realizacji badanego fonemu w zależności od płci mówców.

Tab. 6. Średnie wartości formantów w realizacjach fonemu /f/ w Hz w zależności od płci mówców

Formant	Płeć mówców	
	Dziewczęta	Chłopcy
F2	1753	1789
F3	3419	3365
F4	4852	4769
F5	6320	6201

Analiza dyskryminacyjna

Ponieważ wyznaczenie wartości pierwszego formantu sprawiało duże trudności w przypadku wszystkich głosek trących bezdźwięcznych, dalszym analizom statystycznym postanowiono poddać wyłącznie częstotliwości szczytów widmowych F2-F5, których wartości średnie w badanym materiale mieściły się w zakresie od 1750 Hz do 6500 Hz (por. tab. 7).

Jak wynika z tab. 7, w realizacjach fonemu /x/ wszystkich pięć formantów osiągało niższe wartości średnie niż w przypadku sybilantów, natomiast w realizacjach fonemu /f/

Tab. 7. Średnie częstotliwości formantów (Hz) w realizacjach fonemów trących bezdźwięcznych

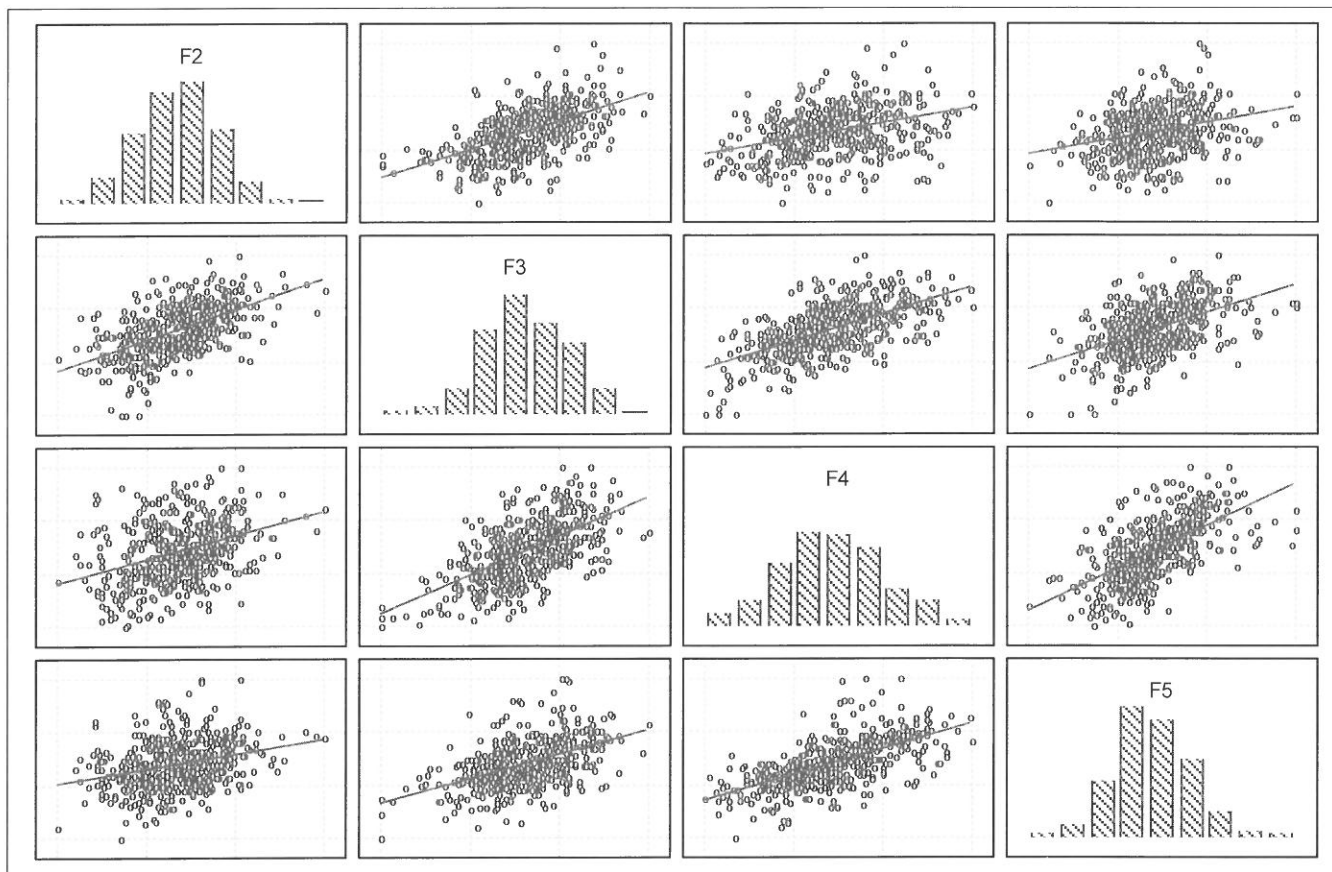
Fonem	Format			
	F2	F3	F4	F5
/s/	2117	3616	4892	6148
/ʃ/	2263	3600	4601	6084
/ç/	2390	3810	4773	6134
/x/	1836	3264	4465	5947
/f/	1776	3385	4800	6244

stan ten dotyczył jedynie formantów F2 oraz F3. Z kolei w ramach spółgłosek trących sybilantnych częstotliwości średnie poszczególnych formantów układały się w rozmaitej kolejności, jednakże z wyjątkiem F2 najniższe wartości obserwowano zawsze w przypadku głoski [ʃ].

W tab. 8 oraz 9 przedstawiono współczynniki korelacji połączonych w obrębie grup oraz korelacji całkowitych dla badanych parametrów akustycznych. Jak widać, mimo iż częstotliwości niektórych formantów w głoskach trących bezdźwięcznych są ze sobą w znacznym stopniu powiązane (por. ryc. 2), współczynniki korelacji w żadnym z przypadków nie osiągnęły wartości tak wysokich, by na etapie analizy dyskryminacyjnej uzasadnione było pominięcie któregośkolwiek z formantów jako zmiennej redundantnej [por. Jassem 1998].

Tab. 8. Matryca korelacji połączonych w obrębie grup

Formant	F3	F4	F5
F2	0,485647	0,430231	0,359281
F3		0,639003	0,530914
F4			0,618981



Ryc. 2. Wykresy korelacji całkowitych dla badanych parametrów akustycznych

Tab. 9. Matryca korelacji całkowitych

Formant	F3	F4	F5
F2	0,556932	0,363503	0,287761
F3		0,613055	0,494940
F4			0,624022

Zbiór danych poddawanych analizie dyskryminacyjnej⁷ obejmuje zatem pięć klas (dla pięciu głosek trących bezdźwięcznych), w ramach których wszystkie przypadki opisywane są przy pomocy czterech zmiennych (częstotliwości formantów F2-F5). Ponieważ maksymalna liczba zmiennych dyskryminacyjnych (zwanych inaczej zmiennymi kanonicznymi lub pierwiastkami) jest mniejsza o jeden od liczby klas [por. Klecka 1980], w obecnym eksperymencie zmienne początkowe (F1-F5) reprezentowane będą przy pomocy czterech nieskorelowanych zmiennych kanonicznych (P1-P4), uporządkowanych malejąco zgodnie ze stopniem, w jakim odpowiedzialne są one za zmienność w badanym zbiorze (por. tab. 11). Jako że zmienne kanoniczne (P1-P4) stanowią kombinacje liniowe zmiennych początkowych (F2-F5), charakteryzują się one odpowiednimi współczynnikami, które w różnym stopniu wiążą ich wartości z wartościami poszczególnych zmiennych początkowych (por. tab. 10).

Tab. 10. Wartości współczynników standaryzowanych dla poszczególnych zmiennych dyskryminacyjnych oraz skumulowany zakres, w jakim zmienne te odpowiedzialne są za całkowitą zmienność w analizowanym zbiorze danych

Zmienna początkowa	Zmienna dyskryminacyjna			
	P1	P2	P3	P4
F2	-0,847	-0,081	0,122	-0,787
F3	-0,617	0,244	-0,364	1,167
F4	0,619	1,056	0,682	-0,382
F5	0,359	-0,365	-1,139	-0,394
Skumul.	0,808	0,973	0,999	1,000

Jak widać w tab. 10, wartość pierwiastka P1, odpowiedzialnego za 81% zmienności w zbiorze, w największym stopniu związana jest z częstotliwością formantu F2, wartość pierwiastka P2 – z częstotliwością F4, natomiast wartość pierwiastka P3 – z częstotliwością formantu F5, przy czym łącznie te trzy zmienne kanoniczne odpowiedzialne są za praktycznie całość zmienności w badanym zbiorze. Stopień powiązania wartości poszczególnych zmiennych dyskryminacyjnych z wartościami odpowiednich zmiennych początkowych potwierdza również przedstawiona w tab. 11 matryca korelacji.

Jak wskazują wysokie wartości statystyki Lambda Wilksa⁸, przedstawione w tab. 12, nawet przy uwzględnieniu wszystkich czterech zmiennych kanonicznych poszczególne klasy w badanym zbiorze są stosunkowo słabo rozróżnialne, natomiast po usunięciu pierwszych dwóch pierwiastków odpowiednie różnice nie są już istotne statystycznie nawet na poziomie $p=0,05$. Słabą rozróżnialność analizowanych klas obrazuje także wykres rozrzutu na ryc. 3, przedstawia-

Tab. 11. Matryca korelacji pomiędzy zmiennymi dyskryminacyjnymi oraz początkowymi dla badanego zbioru danych

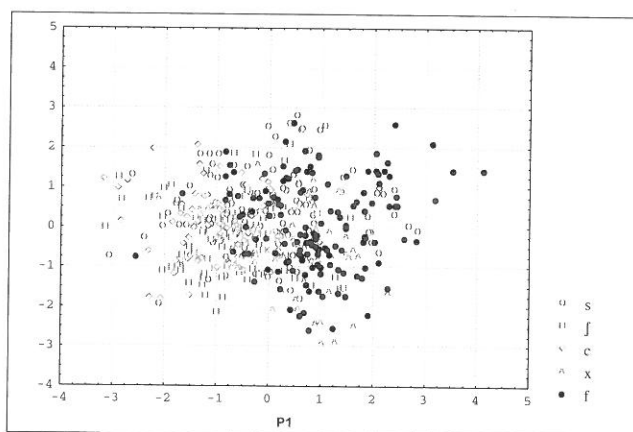
Zmienna początkowa	Zmienna dyskryminacyjna			
	P1	P2	P3	P4
F2	-0,751	0,361	-0,171	-0,526
F3	-0,442	0,685	-0,474	0,332
F4	0,083	0,951	-0,203	-0,218
F5	0,110	0,389	-0,866	-0,293

Tab. 12. Wartości Lambda Wilksa po usunięciu kolejnych pierwiastków

Usunięte pierwiastki	Lambda Wilksa	Statystyka chi-kwadrat	Stopnie swobody	Poziom p
0	0,59	253,77	16	< 0,000000
1	0,89	55,71	9	< 0,000000
2	0,98	8,01	4	< 0,091316
3	1,00	0,31	1	< 0,574730

jący wszystkie przypadki z badanych grup na płaszczyźnie, której współrzędne stanowią wartości dwóch najistotniejszych zmiennych dyskryminacyjnych.

Przewidywana słaba rozróżnialność badanych klas znalazła potwierdzenie w wynikach analizy dyskryminacyjnej. Jak widać w tab. 13, średnia poprawność klasyfikacji w analizowanym zbiorze danych, w oparciu o wykorzystywane zmienne, wyniosła zaledwie 41%, przy czym osiągała ona podobne wartości w przypadku wszystkich grup. Jak wynika z odległości kwadratowych Mahalanobisa (por. tab. 14), będących wykładnikiem względnych odległości statystycznych pomiędzy poszczególnymi klasami w czterowymiarowej przestrzeni zmiennych początkowych (F2-F5), zdecydowanie najmniejszy dystans dzielił realizacje fonemu /j/ od realizacji fonemu /c/, co stwierdzone zostało wcześniej także w badaniach dla mowy niezaburzonej [por. Jassem 1995b]. Niewielkie odległości odnotowano także pomiędzy realizacjami fonemów /x/ oraz /f/, natomiast stosunkowo duże odległości wydają się dzielić realizacje /f/ od /j/ oraz /c/, a także /x/ od /c/.



Ryc. 3. Wszystkie badane przypadki na płaszczyźnie P1 x P2

⁷ Opis podstaw analizy dyskryminacyjnej wraz z przykładem zastosowania w badaniach z dziedziny fonetyki akustycznej znaleźć można w Jassem [1998].

⁸ Statystyka Lambda Wilksa całkowitej dyskryminacji jest obliczana jako stosunek wyznacznika macierzy wariancji/kowariancji wewnątrzgrupowej do wyznacznika macierzy wariancji/kowariancji całkowitej. Lambda Wilksa może przyjmować wartości w zakresie od 0 (doskonała dyskryminacja) do 1 (brak dyskryminacji).

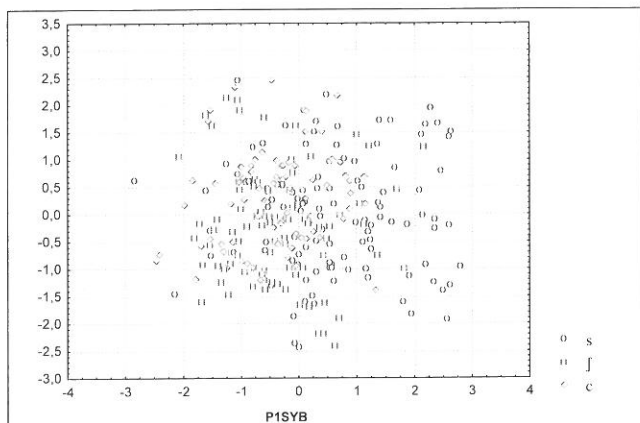
Tab. 13. Macierz klasyfikacji dla spółgłosek trących bezdźwięcznych

	Procent poprawności	s	ʃ	ç	x	f
s	31,62	37	14	29	13	24
ʃ	42,53	11	37	23	13	3
ç	45,95	15	19	34	5	1
x	44,62	10	5	4	29	17
f	42,96	24	5	8	44	61
Razem	40,82	97	80	98	104	106

Tab. 14. Kwadraty odległości Mahalanobisa pomiędzy poszczególnymi klasami

	ʃ	ç	x	f
s	0,996791	1,132657	1,132543	1,061655
ʃ		0,271915	1,295386	2,786452
ç			2,47897	3,749028
x				0,849909

Ze względu na uzyskane rezultaty postanowiono sprawdzić, czy poprawność klasyfikacji wzrośnie, jeżeli założymy, że możliwe jest wstępne rozdzielenie badanego zbioru danych na głoski trące sybilantne oraz niesybilantne w oparciu o inne parametry akustyczne, jak na przykład całkowita energia szumu [por. Harris 1958; Nartey 1982; Pongweni 1984; Balise (i in.) 1994; Jassem 1998; Łobacz 1998].



Ryc. 4. Rozrzut głosek sybilantnych na płaszczyźnie P1SYB x P2SYB

Jak widać na ryc. 4, ogólna rozróżnialność głosek sybilantnych w oparciu o nowe zmienne kanoniczne (P1SYB, P2SYB) nieznacznie wzrosła, jednakże, mimo iż poprawność ich klasyfikacji zwiększyła się średnio o przeszło jedną trzecią (por. tab. 15), w dalszym ciągu była stosunkowo niska i wyniosła 54%. Odległości kwadratowe Mahalanobisa (por. tab. 16) ponownie potwierdziły, iż względny dystans dzielący realizacje fonemu /ʃ/ od realizacji fonemu /ç/ jest znacznie mniejszy niż w przypadku pozostałych dwóch par spółgłosek trących sybilantnych.

Tab. 15. Macierz klasyfikacji dla głosek sybilantnych

	Procent poprawności	s	ʃ	ç
s	55,56	65	23	29
ʃ	56,32	16	49	22
ç	47,30	15	24	35
Razem	53,60	96	96	86

Tab. 16. Kwadraty odległości Mahalanobisa pomiędzy klasami głosek sybilantnych

	ʃ	ç
s	0,973561	1,052068
ʃ		0,31413

W przypadku głosek trących niskoszumowych poprawność klasyfikacji wzrosła natomiast prawie o połowę i wyniosła 65% (por. tab. 17), przy czym należy zwrócić uwagę, że jedyna zmienna kanoniczna w analizie dyskryminacyjnej tych głosek (P1NISK) najsilniej skorelowana była z wartościami formantu czwartego (por. tab. 18).

Tab. 17. Macierz klasyfikacji dla głosek trących niskoszumowych

	Procent poprawności	f	x
f	61,97	88	54
x	70,77	19	46
Razem	64,73	107	100

Tab. 18. Korelacje pomiędzy zmienną dyskryminacyjną P1NISK a poszczególnymi zmiennymi początkowymi

	P1NISK
F2	-0,159
F3	0,346
F4	0,769
F5	0,549

Pierwszy moment centralny

Analizie poddano wartości pierwszego momentu centralnego (M1), obliczane przy użyciu wzoru zamieszczonego w paragrafie 4 na podstawie uśrednionego widma FFT dla środkowych 50% analizowanego segmentu akustycznego.

Wartości pierwszego momentu centralnego widma w ramach wszystkich analizowanych realizacji fonemów trących bezdźwięcznych rozciągały się w zakresie od 3800 Hz do 6700 Hz, przy czym najniższą wartość minimalną odnotowano w przypadku fonemów /f/ oraz /x/ (odpowiednio 3859 Hz oraz 4097 Hz), natomiast najwyższą wartość maksymalną – w przypadku fonemu /s/ (6651 Hz).

Aby sprawdzić, czy wartości pierwszego momentu widmowego zależne są od płci oraz poziomu zrozumiałości mówców, przeprowadzono analizę wariancji, w której wyniku stwierdzono, iż na poziomie $p < 0,1$ statystycznie istotne różnice zaobserwować można było jedynie w ramach realizacji fonemu /x/, kiedy M1 przyjmował wyższe wartości w wymowie dzieci o niskim poziomie zrozumiałości (por. tab. 19 i 20).

Tab. 19. Wartości średnie M1 (Hz) w realizacjach poszczególnych fonemów oraz poziom istotności różnic ze względu na płeć mówców

Fonem	Płeć mówców		Poziom istotności
	Dziewczęta	Chłopcy	
/s/	5679	5807	$p < 0,4199$
/ʃ/	5677	5538	$p < 0,1371$
/ç/	5767	5755	$p < 0,9084$
/x/	4713	4632	$p < 0,6146$
/f/	5221	5473	$p < 0,1510$

Tab. 20. Wartości średnie M1 (Hz) w realizacjach poszczególnych fonemów oraz poziom istotności różnic ze względu na poziom zrozumiałości mówców

Fonem	Poziom zrozumiałości		Poziom istotności
	Wysoki	Niski	
/s/	5593	5900	$p < 0,0343$
/ʃ/	5607	5770	$p < 0,1123$
/ç/	5768	5724	$p < 0,7698$
/x/**	4491	4983	$p < 0,0004$
/f/	5350	5465	$p < 0,5387$

Tab. 21 zawiera średnie wartości pierwszego momentu widmowego w realizacjach poszczególnych fonemów, obliczone dla ogółu badanych mówców, wraz z odpowiednimi odchyleniami standardowymi. Powyższy parametr akustyczny stanowił będzie jedyną zmienną w ramach przeprowadzanej w obecnym paragrafie analizy statystycznej głosek trących bezdźwięcznych, obejmującej pięć klas obiektów.

Tab. 21. Wartości średnie M1 w realizacjach poszczególnych fonemów (Hz) wraz z odchyleniami standardowymi

Fonem	M1	Odchylenie standardowe
/s/	5719	451
/ʃ/	5640	224
/ç/	5762	265
/x/	4679	351
/f/	5384	596

Jak wynika z tab. 21, wartości średnie M1 dla poszczególnych fonemów sybilantnych są w badanym materiale prawie identyczne. Jeżeli weźmiemy dodatkowo pod uwagę przedstawione wartości odchylenia standardowego, wydaje się, że jedynym fonemem, dla którego analizowany parametr mógłby stanowić skuteczną cechę identyfikacyjną, jest /x/.

Przypuszczenie to znalazło potwierdzenie w wynikach analizy dyskryminacyjnej (por. tab. 22), zgodnie z którymi poprawność klasyfikacji fonemu /x/ w oparciu o wartość pierwszego momentu widmowego wyniosła 91%, jednakże w przypadku pozostałych fonemów była ona tak niska, że dla ogółu grup osiągnęła zaledwie 34%. Uwagę zwraca wyjątkowo niska poprawność klasyfikacji dla realizacji fonemu /s/, wynikająca z faktu, że pierwszy moment centralny w ich przypadku przyjmował w badanym materiale wartości z bardzo szerokiego zakresu (od 4610 Hz do 6651 Hz). Związane jest to najprawdopodobniej z faktem, iż spółgłoska [s], jak stwierdza Łobacz [2001], jest bardzo podatna na zniekształcenia w wielu zaburzeniach mowy, również w mowie osób niesłyszących od urodzenia, czy w przypadku głuchoty nabytej. Mimo jednak wielu różnych sposobów jej artykulacji, spotykanych szczególnie wtedy, gdy mamy do czynienia z mową zaburzoną, czy nie do końca ukształtowaną [por. Krajna (i in.) 1999], z percepcyjnego punktu widzenia jest głoską dobrze odbieraną, ponieważ słuchacz jest przygotowany w jej przypadku na szeroki zakres wariantów allofonicznych [Łobacz 2001].

Postanowiono w takim wypadku sprawdzić, czy pierwszy moment widmowy stanowić może przynajmniej wiarygodny parametr, rozróżniający głoski sybilantne od niskoszumowych. Jak wykazała przeprowadzona analiza (por. tab. 23), taki sposób wykorzystania wartości M1 prowadził do 72% poprawności klasyfikacji. Stwierdzono także, że w przy-

padku uprzedniego zakwalifikowania danej głoski jako niskoszumowej na podstawie innego parametru akustycznego wartość pierwszego momentu widmowego umożliwia stosunkowo skuteczną jej identyfikację jako realizację odpowiedniego fonemu (por. tab. 24).

Tab. 22. Macierz klasyfikacji dla spółgłosek trących bezdźwięcznych

	Procent poprawności	s	ʃ	ç	x	f
s	7,69	3	8	18	3	7
ʃ	23,33	5	7	9	0	9
ç	48,00	2	7	12	0	4
x	90,48	0	1	0	19	1
f	29,41	4	8	14	10	15
Razem	33,73	14	31	53	32	36

Tab. 23. Macierz klasyfikacji dla dwóch klas obiektów: głoski sybilantne oraz niskoszumowe

	Procent poprawności	Sybil.	Niskosz.
Sybilanty	80,85	76	18
Niskoszumowe	61,11	28	44
Razem	72,29	104	62

Tab. 24. Macierz klasyfikacji dla dwóch klas obiektów: realizacje fonemu /f/ oraz realizacje fonemu /x/

	Procent poprawności	f	x
f	80,39	41	10
x	90,48	2	19
Razem	83,33	43	29

Jednoczesne wykorzystanie wartości formantów oraz pierwszego momentu centralnego

W związku z rezultatami uzyskanymi w poprzednich dwóch paragrafach postanowiono sprawdzić, jak skuteczna będzie identyfikacja realizacji poszczególnych fonemów trących bezdźwięcznych w przypadku jednoczesnego uwzględnienia częstotliwości formantów F2 – F5 oraz wartości pierwszego momentu widmowego. W tym celu przeprowadzono ponownie analizę dyskryminacyjną zgromadzonego zbioru danych, złożonego z pięciu klas obiektów (dla pięciu fonemów trących bezdźwięcznych), w oparciu o wektor zbudowany ze wspomnianych pięciu zmiennych początkowych.

Przed wszystkim zauważono (por. tab. 25 i 26), że wartości pierwszego momentu widmowego są prawie całkowicie nieskorelowane z częstotliwością któregośkolwiek z badanych formantów, co oznacza, że wykorzystanie powyższego parametru akustycznego powinno w znacznym stopniu poszerzyć charakterystykę poszczególnych przypadków, dostępną na podstawie samej struktury formantowej. Stwierdzono także, że przy tak zbudowanym wektorze zmiennych początkowych pierwsza z czterech zmiennych dyskryminacyjnych, odpowiedzialna za 64% zmienności w analizowanym zbiorze danych (por. tab. 27), najbardziej skorelowana jest właśnie z wartościami M1. Ponieważ średnie wartości powyższej zmiennej kanonicznej dla realizacji poszczególnych fonemów wskazują (por. ryc. 5), iż zmienna ta odpowiada przede wszystkim za rozróżnienie pomiędzy głoskami sybilantnymi oraz niskoszumowymi, oznacza to, że w wykorzystywanym wektorze danych taką właśnie funkcję dyskryminacyjną odgrywa w głównej mierze pierwszy moment widmowy.

Tab. 25. Korelacje połączone w obrębie grup

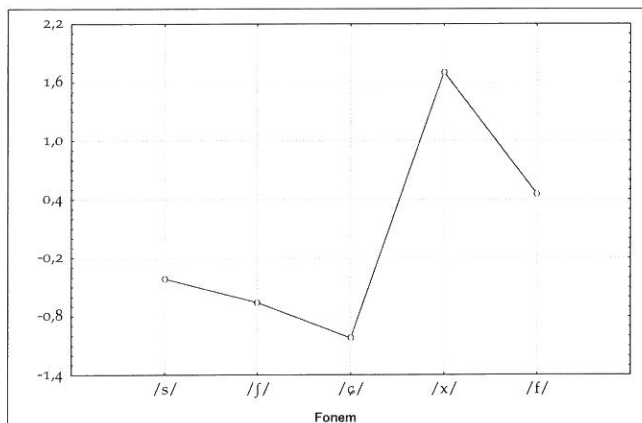
Formant	M1
F2	0,037209
F3	0,158065
F4	0,179580
F5	0,027504

Tab. 26. Korelacje całkowite

Formant	M1
F2	0,175046
F3	0,288905
F4	0,219953
F5	0,041394

Tab. 27. Matryca korelacji pomiędzy zmiennymi dyskryminacyjnymi i początkowymi oraz skumulowany zakres, w jakim poszczególne pierwiastki kanoniczne odpowiedzialne są za całkowitą zmienność w zbiorze danych

Zmienna początkowa	Zmienna dyskryminacyjna			
	P1	P2	P3	P4
F2	-0,293	0,138	0,687	0,492
F3	-0,307	0,311	0,431	0,686
F4	-0,050	0,809	0,217	0,540
F5	0,058	0,611	0,751	0,010
M1	-0,874	0,376	-0,211	-0,221
Skumul.	0,644	0,936	0,992	1,000



Ryc. 5. Średnie wartości zmiennej kanonicznej P1 dla realizacji poszczególnych fonemów trących bezdźwięcznych

Jak wykazały ostatecznie wyniki analizy dyskryminacyjnej (por. tab. 28), łączne wykorzystanie częstotliwości formantów wraz z wartościami pierwszego momentu centralnego widma podniosło średnią poprawność klasyfikacji głosek trących bezdźwięcznych do około 50%, a więc o prawie jedną czwartą w stosunku do analizy opartej na samych częstotliwościach formantów oraz o prawie połowę w stosunku do analizy opartej wyłącznie na wartościach M1. Nie zmienia to jednak faktu, że osiągnięta poprawność klasyfikacji pozostaje w dalszym ciągu o około 30% gorsza od średnich wyników uzyskiwanych przy pomocy różnych modeli dyskryminacyjnych dla mowy niezaburzonej [por. Jassem 1995a; Jassem 1995b; Jassem 1998].

Na podstawie oddzielnych analiz stwierdzono również, iż dodatkowe wykorzystanie parametru M1 w żadnym stopniu nie podwyższa poprawności klasyfikacji w obrębie samych

głosek sybilantnych, w stosunku do wyników uzyskiwanych wyłącznie na podstawie o wartości formantów, choć powoduje oczywiście zwiększenie skuteczności dyskryminacyjnej w stosunku do modelu opartego wyłącznie na wartościach pierwszego momentu widmowego, w którego ramach głoski sybilantne były w ogóle nierozróżnialne (poprawność równa 33%).

Tab. 28. Macierz klasyfikacji dla spółgłosek trących bezdźwięcznych

	Procent poprawności	s	ʃ	ç	x	f
s	37,25	38	12	26	7	19
ʃ	53,57	11	45	20	1	7
ç	45,07	16	21	32	0	2
x	79,03	0	3	0	49	10
f	43,36	34	14	9	24	62
Razem	48,92	99	95	87	81	100

Jeżeli chodzi o rozróżnienie głosek niskosumowych, uwzględnienie wartości M1 podwyższa (por. tab. 29) średnią poprawność identyfikacji w stosunku do wyników uzyskiwanych w ramach modelu opartego na samych częstotliwościach formantów i powoduje, iż osiąga ona poziom podobny do tego, jaki odnotowano w przypadku prób identyfikacji przy wykorzystaniu wyłącznie wartości pierwszego momentu widmowego.

Tab. 29. Macierz klasyfikacji dla głosek trących niskosumowych

	Procent poprawności	f	x
f	81,82	117	26
x	85,48	9	53
Razem	82,93	126	79

Podsumowanie

Jak wykazały przeprowadzone analizy dyskryminacyjne, wykorzystanie takich parametrów akustycznych, jak częstotliwości formantów czy wartość pierwszego momentu widmowego, umożliwia identyfikację polskich głosek trących bezdźwięcznych realizowanych przez osoby niesłyszące z poprawnością średnio o około 30% niższą od wyników odnotowywanych dla mowy niezaburzonej. Wy tłumaczyć można to znacznie większym rozchwianiem wartości tych parametrów w przypadku mowy, której nadawca ma silnie ograniczone możliwości audytywnej weryfikacji produkowanego przez siebie sygnału.

Zauważyć też trzeba, że bez względu na dobór parametrów widmowych, wykorzystywanych w analizie dyskryminacyjnej badanych spółgłosek, najlepsze rezultaty identyfikacyjne uzyskiwano zawsze w przypadku realizacji fonemu /x/, natomiast najgorsze w przypadku realizacji fonemu /s/. Ponieważ zaś w badaniach nad automatycznym rozpoznawaniem mowy niezaburzonej wykorzystywane tutaj parametry stosowane są z porównywalnym powodzeniem w przypadku identyfikacji pierwszej z powyższych głosek, natomiast z nieporównywalnie większą skutecznością (średnio około 90%) w odniesieniu do drugiej z nich, wnioskować można także, iż właśnie fonem /x/ realizowany jest przez badane dzieci głuche w sposób najbliższy realizacjom mówców słyszących, podczas gdy realizacje fonemu /s/ w wykonaniu tych dzieci w największym stopniu odbiegają od realizacji charaktery-

stycznych dla mówców o nieuszkodzonym słuchu lub przy najmniej wykazują największe rozchwanie artykulacyjne.

Warto też zwrócić uwagę na odnotowany w niniejszych badaniach prawie całkowity brak korelacji pomiędzy pierwszym momentem widmowym a wartościami formantów w przypadku głosek trących bezdźwięcznych, jak również na wykazaną tutaj możliwość wykorzystania pierwszego z powyższych parametrów akustycznych do względnie skutecznego rozróżniania bezdźwięcznych sybilantów od spółgłosek [f] oraz [x], czy też dla celów rozróżniania głosek niskosumowych pomiędzy sobą, co znaleźć może zastosowanie w ramach konstrukcji skomputeryzowanych narzędzi rewalidacyjnych, przeznaczonych dla osób niesłyszących.

Ponadto z przeprowadzonych analiz wynika, iż w prawidłowych realizacjach badanych fonemów, to znaczy takich, które percypowane były przez transkrybentów w sposób zgodny z zamierzeniami mówców, brak jest wśród badanych dzieci regularnych różnic w wartościach poszczególnych parametrów widmowych ze względu na płeć oraz poziom zrozumiałości mówców.

Wydaje się, że naturalną kontynuację obecnych badań stanowić powinna analiza akustyczno-statystyczna realizacji pozostałych grup fonemów w wymowie dzieci niesłyszących, a mianowicie spółgłosek nosowych [por. Kleśta 2003b], zwartych, zwartotrących oraz samogłosek niesylabicznych. Uzupełnione w ten sposób dane, w połączeniu z rezultatami wspominanych w niniejszym artykule badań [Francuzik, Szalkowska 2001; Łobacz, Francuzik, Szalkowska 2002; Trochymiuk 2003], umożliwiłyby bowiem opracowanie pełniejszej charakterystyki widmowej mowy polskich dzieci głuchych oraz wykorzystanie jej w przyszłości jako podstawy do konstruowania systemów automatycznego rozpoznawania powyższej mowy, nawet gdyby systemy te ograniczone miały być do niewielkiego tylko zbioru wyrazów.

Bibliografia

- Andruszka A., Dembińska S., Goralewska M., Pawełek P., Piasecka J. [2000]. Fonetyczna charakterystyka wyrazów w wypowiedziach dzieci niesłyszących. Praca magisterska, materiały Instytutu Językoznawstwa UAM w Poznaniu.
- Balise R. T., Diehl R. L. [1994]. Some distributional facts about fricatives and perceptual explanation. „Phonetica” 51, 99-110.
- Ball M. J., Code Ch. [1997]. Instrumental Clinical Phonetics, London: Whurr Publishers Ltd.
- Daniłoff R. G., Wilcox K., Stephens M. O. [1980]. An acoustic-articulatory description of children's defective /s/ productions. „Journal of Communication Disorders” 13, 347-363.
- Demenko G. [1987]. Wizualizacja częstotliwości podstawowej mowy, W: W. Jassem (red.). Wizualizacja mowy i jej zastosowania. 65-88. Warszawa: IPPT PAN.
- Eckert U. [1994]. Znaczenie i warunki rozwoju języka słownego w kształceniu osób z wadami słuchu. W: S. Grabias (red.). Głuchota a Język. 63-70. Lublin: UMCS.
- Emiluta-Roży D. [1994]. Wymowa dzieci niedosłyszących. W: S. Grabias (red.). Głuchota a Język. 177-184. Lublin: UMCS.
- Fornby C., Monsen R. B. [1981]. Long-term average speech spectra for normal and hearing-impaired adolescents. „Journal of the Acoustical Society of America” 1(71), 196-202.
- Forrest K., Weismer G., Milenkovic P., Dougal R. [1988]. Statistical analysis of word initial voiceless obstruents, preliminary data. „Journal of the Acoustical Society of America” 84, 115-123.
- Francuzik K., Szalkowska E. [2001]. Częstotliwości formantowe samogłosek w mowie dzieci niesłyszących. Praca magisterska, materiały Instytutu Językoznawstwa UAM w Poznaniu.
- Gallagher T. M., Shriner T. H. [1975]. Articulatory inconsistencies in the speech of normal children. „Journal of Speech and Hearing Research” 18, 168-175.
- Gałkowski T. [1994]. Stan rehabilitacji oraz potrzeby osób z uszkodzonym słuchem. W: S. Grabias (red.). Głuchota a język. 11-31. Lublin: UMCS.
- Golanowska M. [1999]. Test artykulacyjny dla dzieci w wieku 2;0 do 3;0 lat. Materiały Instytutu Językoznawstwa UAM w Poznaniu.
- Góralówna M. [1994]. Czynniki decydujące o rozwoju językowym dziecka z uszkodzonym słuchem. W: S. Grabias (red.). Głuchota a język. 59-62. Lublin: UMCS.
- Grunwell P. [1987]. Clinical Phonology. London: Croom Helm.
- Harris K. S. [1958]. Cues for the Discrimination of American English Fricatives in Spoken Syllables. „Language and Speech” 1, 1-7.
- Hoffmann B. [1987]. Surdopedagogika. Warszawa: PWN.
- Ingram D., Christensen L., Veach S., Webster B. [1980]. The acquisition of word-initial fricatives and affricates in English by children between 2 and 6 years. W: G.H. Yeni-Komshian, J. F. Kavanagh, C. A. Ferguson (red.). Child Phonology, vol. 1. 169-192. London-Toronto-Sydney-San Francisco-New York: Academic Press.
- Jassem W. [1973]. Podstawy Fonetyki Akustycznej. Warszawa: PWN.
- Jassem W. [1987]. Wizualizacja mowy i jej zastosowania. Warszawa: IPPT PAN.
- Jassem W. [1995a]. The Acoustic Parameters of Polish Voiceless Fricatives: An Analysis of Variance. „Phonetica” 52, 251-258.
- Jassem W. [1995b]. An acoustic and statistical analysis of the continuous spectra of voiceless Polish plosives and fricatives. „Zeitschrift für empirische Textforschung” 2, 87-96.
- Jassem W. [1998]. An Acoustical Linear-Predictive and Statistical Discriminant Analysis of Polish Fricatives and Affricates. W: W. Jassem, Cz. Basztura, K. Jassem (red.). Technologia Mowy i Języka, t. 2. 9-45.
- Kapusta M. [2000]. Wykorzystanie sieci Kohonena do wizualizacji mowy patologicznej, praca doktorska, materiały Akademii Górniczo-Hutniczej w Krakowie.
- Karczeńska L. B., Gubrynowicz R., Karczewski R. [2000]. Analiza porównawcza konturów melodycznych w mowie dzieci głuchych. „Audiofonologia” 16, 5-20.
- Klecka W. R. [1980]. Discriminant Analysis. Beverly Hills, N.J.
- Kleśta J. [2002]. Charakterystyka akustyczna cech segmentalnych mowy dzieci niesłyszących, dysertacja doktorska. Materiały UAM w Poznaniu.
- Kleśta J. [2003a]. Percepcyjna ocena zrozumiałości mowy realizowanej przez dzieci niesłyszące poddawane kształceniu w szkole specjalnej. W druku.
- Kleśta J. [2003b]. Analiza akustyczna polskich spółgłosek nosowych realizowanych przez dzieci niesłyszące. W druku.
- Krajna E. [1998]. Lingwistyczne podstawy opracowania logopedycznego testu artykulacyjnego. W: J. Pogonowski (red.). Investigationes Linguisticae, t. 4. 5-22.
- Krajna E., Bryndal M. [1999]. 100-wyrazowy test artykulacyjny. Analiza słuchowa nagrań i próba normalizacji testu. „Audiofonologia” 14, 137-174.
- Krakowiak K. [1995]. Fonogesty jako narzędzie formowania języka u dzieci z uszkodzonym słuchem, Komunikacja językowa i jej zaburzenia, cz. 9. Lublin: UMCS.
- Krakowiak K., Panasiuk M. [1992]. Umiejętności komunikacyjne dziecka z uszkodzonym słuchem, Komunikacja językowa i jej zaburzenia, t. 3. Lublin: UMCS.
- Kudela K. [1968]. Spectral Analysis of Polish Fricative Consonants, W: W. Jassem (red.). Speech Analysis and Synthesis, tom 1. 93-188. Warszawa: PWN.
- Levitt H. [1972]. Acoustic analysis of deaf speech using digital processing techniques. „IEEE” 1(AU-20), 35-41.
- Łobacz P. [1987]. Wizualna Interpretacja Obrazów Spektrograficznych Mowy. W: W. Jassem (red.). Wizualizacja Mowy i Jej Zastosowania. Warszawa: IPPT PAN, 89-134.

- Łobacz P. [1998]. Sybilantność. W: S. Puppel (red.). *Scripta Manent*. 135-154. Poznań: UAM.
- Łobacz P. [2001]. Wymowa patologiczna a norma fonetyczna w świetle analizy akustycznej. W: S. Grabias (red.). *Zaburzenia Mowy*. 189-215. Lublin: Wydawnictwo UMCS.
- Łobacz P. [2002]. Wstępna charakterystyka fonetyczna mowy wybranej grupy dzieci niesłyszących. „*Scripta Neophilologica Posnaniensia*” 4, 29-50.
- Łobacz P., Dobrzańska K. [1999]. Opis akustyczny głosek sybilantnych w wymowie dzieci przedszkolnych. „*Audiofonologia*” 14, 5-26.
- Łobacz P., Francuzik K., Szalkowska E. [2002]. Acoustic-phonetic description of hearing-impaired children's vowels in polish. „*Psychology of Language and Communication*” 6(2), 3-31.
- Mikiel W., Żarnecki P., Tłuchowski W., Komorowska A. [1979]. Parametry akustyczne mowy dzieci głuchych. „*Otwarte Seminarium z Akustyki*” 26.
- Nartey J. N. A. [1982]. On Fricative Phones and Phonemes: Measuring the Phonetic Differences Within and Between Languages. *UCLA Working Papers in Phonetics* 55, Los Angeles.
- Okalidou A., Harris K. S. [1999]. A comparison of intergestural patterns in deaf and hearing adult speakers: implications from an acoustic analysis of disyllables. „*Journal of the Acoustical Society of America*” 106(1), 394-410.
- Pietrzak W. [1994]. Różne drogi poznawania języka przez dzieci z wadą słuchu. W: S. Grabias (red.). *Głuchota a język*. 95-104. Lublin: UMCS.
- Pongweni A. J. C. [1984]. An Acoustic Study of Karagana Fricatives. „*Zeitschrift für Phonetik Sprachwissenschaft und Kommunikationsforschung*” 37, 328-347.
- Prillwitz S. [1996]. Język, komunikacja i zdolności poznawcze niesłyszących. Warszawa: Wydawnictwa Szkolne i Pedagogiczne.
- Pruszevicz A. [2000a]. Diagnostyka ogólna zaburzeń słuchu. W: *Zarys Audiologii Klinicznej*. 324-338. Poznań: Wydawnictwo AM.
- Pruszevicz A. [2000b]. Diagnostyka zaburzeń słuchu u dzieci. W: *Zarys Audiologii Klinicznej*. 338-346. Poznań: Wydawnictwo AM.
- Rakowska A. [1994]. Natywizm Chomsky'ego a uczenie się języka przez głuchych. W: S. Grabias (red.). *Głuchota a język*. 115-130. Lublin: UMCS.
- Shannon C. E., Weaver W. [1949]. *The mathematical theory of communication*. Urbana: University of Illinois Press.
- Sierkowska H., Gubrynowicz R., Gałkowski T. [2000]. Nauka intonacji w mowie dzieci głuchych. „*Audiofonologia*” 17, 41-66.
- Stankiewicz K., Włoch K. [2000]. Cechy prozodyczne w mowie osób niesłyszących. Praca magisterska, materiały Instytutu Językoznawstwa UAM w Poznaniu.
- Stephens M. J., Hoffmann P. R., Daniloff R. G. [1986]. Phonetic characteristics of delayed /s/ development. „*Journal of Phonetics*” 14, 247-256.
- Szczepankowski B. [1994]. Stosunek środowiska niesłyszących do edukacji oralnej – badania, które trzeba podjąć. W: S. Grabias, (red.). *Głuchota a Język*. 105-114. Lublin: UMCS.
- Trochymiuk A. [2003]. Realization of the voiced-voiceless contrast by hearing-impaired children. „*Studia Phonetica Posnaniensia*”. W druku.

Adresy do korespondencji

Janusz Kleśta
 Uniwersytet im. Adama Mickiewicza
 Instytut Językoznawstwa
 ul. Międzychodzka 5, 60-371 Poznań
 e-mail: janklest@amu.edu.pl