

Analiza akustyczna polskich spółgłosek nosowych realizowanych przez dzieci niesłyszące¹

Acoustic analysis of Polish nasals produced by deaf children

Janusz Kleśta

Instytut Językoznawstwa, Uniwersytet im. Adama Mickiewicza w Poznaniu
ul. Międzychodzka 5, 60-371 Poznań

janklest@amu.edu.pl

Abstract

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 nasals 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 nasals produced by deaf people involve similar identification difficulty as their realizations by normally hearing speakers.

1 Wstęp

Malejące rozmiary produktów współczesnej elektroniki w połączeniu ze stale rosnącą złożonością ich budowy powodują, iż coraz częściej niewystarczającymi 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 charakteryzowałaby się minimalnymi gabarytami, a prostota obsługi eliminowała konieczność specjalnego przeszkolenia. Wydaje się, że narzędziem spełniającym powyższe 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 et al. 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

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

realizowanej przez całkiem licznie reprezentowaną grupę społeczną, jaką stanowią ludzie głusi². 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 et al. 1979, Demenko 1987, Karczewska et al. 2000, Sieńkowska et al. 2000, Stankiewicz i Włoch 2000), natomiast brakuje prawie całkowicie prac poruszających problemy jego segmentacji czy analizy widmowej. Tymczasem niewystarczająca wiedza na temat segmentalnych parametrów mowy realizowanej przez dzieci niesłyszące uniemożliwia również całkowicie konstrukcję narzędzi diagnostycznych oraz rewalidacyjnych dla powyższej 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 rozbieżności pomiędzy sygnałami stanowią kryterium do ich różnicowania i są podstawą do 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 z 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 dla 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 powyżej techniki w procesie konstrukcji podobnych narzędzi dla dzieci niesłyszących (czego pierwsze próby zostały już zresztą podjęte), tym bardziej, iż, jak pisze Łobacz (1987: 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 też kontrolowanej przez surdologopedę, pracy dzieci głuchych nad zrozumiałością własnych wypowiedzi, poprzez doskonalenie poprawności produkcji na poziomie segmentalnym. Poszukiwanie wspomnianych parametrów odbywało się w oparciu o wielkości akustyczne oraz metody statystyczne stosowane współcześnie dla celów automatycznego rozpoznawania mowy niezaburzonej, a także przy wykorzystaniu wybranych partii informacji, dostępnych w wyniku

² 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.

analizy spektrograficznej, stanowiącej w dalszym ciągu jedno z podstawowych narzędzi w zakresie badań nad mową zaburzoną (por. Ball et al. 1997).

2 Przedmiot analiz

Niniejszy artykuł przedstawia wyniki badań obejmujących realizację spółgłosek nosowych. Powyższą grupę głosek postanowiono poddać analizie ze względu na ograniczoną w dalszym ciągu wiedzę z zakresu ich struktury akustycznej, w związku ze znaczną chwiejnością w wartościach parametrów akustycznych oraz słabym zróżnicowaniem widmowym w obrębie tej grupy dźwięków mowy (por. Jassem 1973).

Jak stwierdza Dukiewicz (1967), znaczne zróżnicowanie osobnicze w zakresie parametrów akustycznych charakteryzujących spółgłoski nosowe związane jest przede wszystkim z udziałem komory nosowej w procesie ich artykulacji. Główną rolę odgrywają tutaj takie czynniki jak anatomiczne różnice w kształcie i rozmiarze tej komory u różnych osób, istnienie całego szeregu możliwych pozycji pośrednich, jeśli chodzi o położenie podniebienia miękkiego, decydującego o stopniu sprzężenia komory nosowej z komorą ustną, jak również obecność w przejściach nosowych elementów blokujących, zwiększających tłumienie rezonansu nosowego (obrzmienie ścianek wewnętrznych w wyniku przeziębienia, nagromadzenie śluzu i zanieczyszczeń, jak również różna u poszczególnych osób ilość włosów rosnących w nozdrzach).

Choć jednak „różnice między widmami stadium ustalonego poszczególnych spółgłosek nosowych nie stanowią dostatecznej podstawy ich rozróżnialności słuchowej” (Dukiewicz 1967: 66), „widma stadium ustalonego [m], [n], [ŋ], [ɲ] posiadają [...] pewne wspólne cechy charakterystyczne, różne dla każdej z tych spółgłosek” (op. cit.: 66). W toku obecnych analiz podjęta zostanie próba stwierdzenia, czy różnice w wartościach odpowiednich parametrów widmowych dla poszczególnych głosek nosowych, w przypadku realizacji powyższych dźwięków mowy przez dzieci niesłyszące, są istotne statystycznie.

3 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 miało miejsce w takich przypadkach, kiedy dane dziecko było uprzednio wychowankiem szkoły masowej lub innego ośrodka szkolno-wychowawczego. W związku z powyższym, 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 Panasiuk 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 i Harris (1999:

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 dla celów obecnych badań wyboru 37 mówców, których jedynym udokumentowanym upośledzeniem był głęboki ubytek słuchu oraz u których nie zaobserwowano zaburzeń natury emocjonalnej czy też znacznych odstępstw od intelektualnej normy rozwojowej.

3.1 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 biap³ (por. Hoffmann 1987) jak i w oparciu o normy ANSI⁴ oraz ISO⁵ (por. Pruszewicz 2000a, Pruszewicz 2000b). Średni ubytek w lepszym uchu, obliczony dla powyższej grupy dzieci, wyniósł 97 dB, przy czym najwyższa 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 3 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 różnymi czynnikami natury genetycznej, prenatalnej, perinatalnej, lub też związane z chorobą, która miała miejsce w wieku wczesnodziecięcym albo też z toksycznym środkiem farmakologicznym podanym w powyższym 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 wspominanym 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 miała miejsce.

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 posługiwało się aparatem jednousznym, natomiast reszta stosowała aparat dwuuszny.

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

³ biap – Międzynarodowe Biuro Audiofonologii.

⁴ ANSI – American National Standards.

⁵ ISO – International Standardization Organization.

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 Panasiuk 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 powyższym, dość istotną rolę dla 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 migowego w środowisku niesłyszących rówieśników, przeszło jedna czwarta badanych uczniów (10 osób) miała także szansę 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.

3.2 Materiał nagraniowy

Ponieważ Rakowska (1994: 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).

3.3 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 powyższego 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 wyrazów⁶.

Na wspomniane powyżej 319 nie zrealizowanych elementów leksykalnych składały się naturalnie również niektóre spośród wykorzystywanych w ramach niniejszych badań wyrazów, zawierających w przypadku prawidłowej artykulacji spółgłoskę nosową, a mianowicie:

⁶ Omówienie przyczyn związanych z częstym brakiem realizacji niektórych elementów leksykalnych znaleźć można w Klešta (2002).

- *kieszęć* (głoska [ɲ]) — brak realizacji w przypadku 30 mówców,
- *nie* (głoska [ɲ]) — brak realizacji w przypadku 30 mówców,
- *jedzenie* (głoska [ɲ]) — brak realizacji w przypadku 23 mówców,
- *ziemia* (głoska [mʲ]) — brak realizacji w przypadku 22 mówców,
- *miau miau* (głoska [mʲ]) — brak realizacji w przypadku 7 mówców,
- *dzwonek* (głoska [n]) — brak realizacji w przypadku 6 mówców,
- *myje się* (głoska [m]) — brak realizacji w przypadku 5 mówców,
- *nos* (głoska [n]) — brak realizacji w przypadku 2 mówców,
- *pociągi* (głoska [ɲʲ]) — brak realizacji w przypadku 2 mówców oraz
- *zielony* (głoska [n]) — brak realizacji w przypadku 1 mówcy.

3.4 Wybór materiału do analizy akustycznej

W celu dokonania wyboru materiału nagraniem poddawane ostatecznie analizie akustycznej w ramach obecnych badań, postanowiono posłużyć się kryterium wykorzystywanym już we wcześniejszych pracach (np. Levitt 1972, Fornby et al. 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: 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 nagrany obejmujący wypowiedzi tych osób, dla których w ramach testów percepcyjnych odnotowano poziom zrozumiałości przekraczający pewien minimalny pułap (por. np. Levitt 1972, Fornby et al. 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 uzyskała najslabsze wyniki, tzn. nie więcej niż 10% bodźców wyrazowych rozpoznanych poprawnie (por. Kleśta 2003a). Tak dobrane ósmiosobowe grupy mówców były w pełni zrównoważone pod względem płci, co oznacza, że w każdej z nich znalazły się cztery dziewczynki oraz czterech chłopców.

3.5 Transkrypcja analizowanego materiału

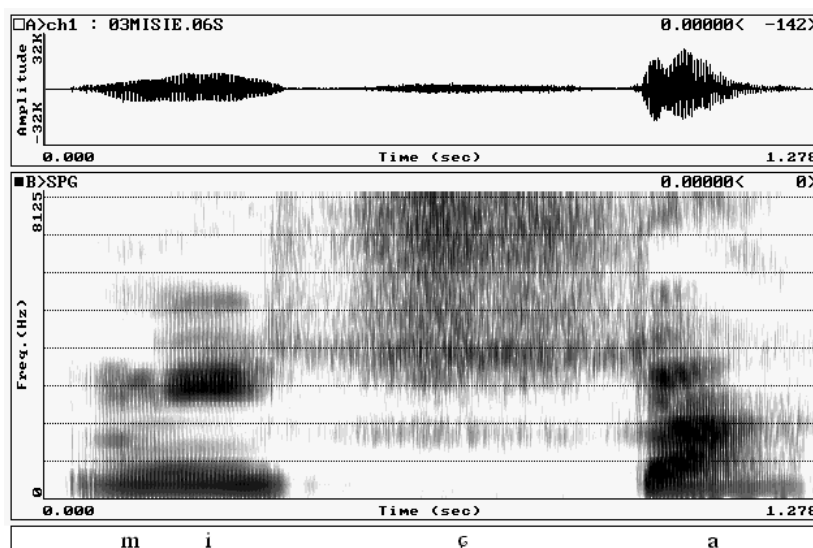
Kolejny etap badań stanowiło opracowanie transkrypcji zgromadzonego materiału⁷. Choć analizie poddawana była mowa zaburzona, będąca następstwem głębokiego niedosłuchu, faktycznie interesowały nas jedynie takie jej parametry, które mają decydujący wpływ na identyfikację poszczególnych głosek jako realizacji odpowiednich fonemów z zakresu inwentarza odbiorców o słuchu niezaburzonym i w ten 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 (por. § 3.4), 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,

⁷ 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 et al. (2000).

a oczekiwaną reakcją - ich zapis przy pomocy symboli alfabetu IPA. Przeprowadzoną transkrypcję można zatem określić mianem typowego testu wyrazistości (por. Jassem 1973).

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 nosowych, przy czym identyfikacja segmentów akustycznych związanych z ich artykulacją w strukturze analizowanego sygnału na podstawie spektrogramu oraz oscylogramu nie sprawiała większych trudności (por. rycina 1).



Rycina 1: Oscylogram oraz spektrogram realizacji wyrazu „misie”, transkrybowanej jako /m i ɕ a/, mówca o wysokim poziomie zrozumiałości, test obrazkowy

4 Analizowane parametry akustyczne

Do tradycyjnie badanych parametrów spektralnych w przypadku analizy spółgłosek nosowych należą wartości pierwszych pięciu bądź sześciu formantów, szerokości wstęg tych formantów oraz względne poziomy ich amplitud, jak również, co jest charakterystyczne dla tej grupy głosek, częstotliwości antyformantów (por. Malecot 1956, Dickson 1962, Kacprowski 1963, Martony 1964, Fujimura 1965, Dukiewicz 1967, Fant 1970, Jassem 1973, Pickett 1980, Stevens 1998 i in.).

Choć jednak w literaturze (por. Clark et al. 1990) zaleca się czasami ostrożność, jeśli chodzi o stosowanie analizy LPC do badania spółgłosek nosowych, właśnie ze względu na fakt, iż „most LP analyses today are models of resonances only, not antiresonances” (Kent et al. 1992: 73), podczas gdy “the vocal tract does introduce antiresonances, especially in the production of nasal and lateral speech sounds” (op. cit.: 73), Oppenheim (1982: 138) stwierdza, iż „jest uzasadnione przypuszczenie, że wiele ważnych właściwości transmitancji kanału głosowego można wyrazić za pomocą modelu zawierającego tylko bieguny”. Jak podaje ten sam autor (1982: 146), „jeśli obwiednia widma sygnału mowy zawiera bieguny jak i zera, niezbędna jest wówczas przy analizie aproksymacja kanału głosowego za pomocą transmitancji odpowiednio dużego rzędu, tak by można wyrazić zera obwiedni wyłącznie za pomocą biegunów”. W metodzie tej „stosuje się analizę z predykcją liniową, w której korzysta się z układu transmitancji zawierającej wyłącznie bieguny” (op. cit.: 146).

W oparciu o powyższe stwierdzenia jak również wyniki wcześniejszych badań Domagały (1988), w ramach obecnych badań analizom statystycznym postanowiono zatem poddać ostatecznie wyłącznie częstotliwości szczytów w obwiedni widma, wyznaczanych metodą predykcji liniowej 20 rzędu, przy czym uzyskane przy jej pomocy dane na temat szerokości wstęg

formantowych posłużyły eliminacji ewentualnych błędów analizy statystycznej, stanowiącej integralny element metody LPC⁸.

W celu wykonania odpowiednich analiz akustycznych, jak również uzyskania spektrogramów oraz oscylogramów, których analiza wzrokowa umożliwiła identyfikację segmentów akustycznych związanych z realizacją odpowiednich fonemów w analizowanym sygnale mowy, posłużono się cyfrowym analizatorem akustycznym CSL50 model 4300B⁹ firmy KAY Elemetrics, zaimplementowanym w postaci karty dźwiękowej, współpracującej z odpowiednim oprogramowaniem za pośrednictwem komputera klasy PC.

5 Wyniki analizy akustycznej

Ponieważ przeprowadzone testy percepcyjne (por. § 3.4) 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 artykulacyjnych odnotowanych w transkrypcji fonetycznej materiału zgromadzonego powyższymi dwiema metodami (por. § 3.5) 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 (por. § 3.3 oraz § 3.4).

Analiza akustyczna spółgłosek nosowych w ramach obecnych badań obejmowała realizacje trzech fonemów występujących w języku polskim, a mianowicie /m/, /n/ oraz /ɲ/. Zrezygnowano natomiast z badania parametrów akustycznych realizacji fonemu /ŋ/ (por. system fonologiczny Jassema (1997), stworzony specjalnie na potrzeby technologii mowy), ponieważ w zakresie zastosowanego testu odnotowano zaledwie jedno jego wystąpienie (w wyrazie /p o ɸ̃ o ŋ ʃ i/), i to wyłącznie pod warunkiem przyjęcia odpowiedniej interpretacji fonologicznej, dotyczącej realizacji grafemów „ą” oraz „ę”, których wymowa w dalszym ciągu budzi szereg kontrowersji (por. Jassem 1997, Krajna 1998, Kleśta 2002). Dodatkowym argumentem przemawiającym za rezygnacją z badania wspomnianej spółgłoski był także fakt, że w swym jedynym wystąpieniu pojawiała się ona w kontekście palatalnym.

Materiał badawczy dla analizowanych spółgłosek nosowych obejmował zatem cztery realizacje fonemu /m/ (wyrazy /m a m a/ (dwukrotnie), /d o m/, /m i j e c e/) oraz po trzy realizacje fonemów /n/ (wyrazy /n o s/, /z e l o n i/, /d̥z v o n e k/) i /ɲ/ (wyrazy /j e d̥z e ɲ e/, /ɲ e/, /c j e ʃ e ɲ/). Ze względu na kontekst palatalny postanowiono natomiast wyłączyć z analizy parametrów akustycznych realizacje fonemu /m/ w wyrazach /m i c e/, /m j a w m j a w/ oraz /z e m j a/.

5.1 Częstotliwości formantów

W celu wyznaczenia częstotliwości formantowych spółgłosek nosowych w obecnych badaniach posłużono się analizą LPC 20 rzędu, stosowaną wraz z metodą autokorelacji, preemfazą o wartości 0,9 oraz oknem analizującym Hamminga o szerokości 20 ms. Ponieważ, jak podaje Dukiewicz (1967: 22), „praktycznie pełna informacja o cechach istotnych stadium ustalonego spółgłosek nosowych mieści się w paśmie do około 5000Hz”, analizie poddano wyłącznie wartości pierwszych pięciu szczytów, które występują w podanym powyżej zakresie częstotliwości. 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. Kleśta 2002).

⁸ Szczegółowy opis procedur pomiarowych znaleźć można w Kleśta (2002).

⁹ CSL - Computer Speech Laboratory.

5.1.1 Fonem / m /

Średnie wartości formantów w realizacjach fonemu /m/ wyniosły odpowiednio 353Hz dla F1 (odchylenie standardowe = 141Hz), 1339Hz dla F2 (SD = 307Hz), 2414Hz dla F3 (SD = 327Hz), 3382Hz dla F4 (SD = 474Hz) oraz 4403Hz dla F5 (SD = 496Hz).

Jak wykazała analiza wariancji, częstotliwości pierwszego formantu były w sposób statystycznie istotny ($p < 0,001$) wyższe w wymowie dziewcząt, natomiast częstotliwości drugiego formantu ($p < 0,001$) — w wymowie chłopców (por. tabela 1). W przypadku pozostałych formantów nie odnotowano statystycznie istotnych różnic ze względu na płeć oraz poziom zrozumiałości mówców.

Tabela 1: Średnie wartości formantów w realizacjach fonemu /m/ w Hz, w zależności od płci mówców

FORMANT	PŁEĆ MÓWCÓW	
	DZIEWCZĘTA	CHŁOPCY
F1***	393	320
F2***	1269	1398
F3	2398	2427
F4	3333	3422
F5	4380	4422

5.1.2 Fonem / n /

Średnie wartości formantów w realizacjach fonemu /n/ wyniosły odpowiednio 383Hz dla F1 (odchylenie standardowe = 167Hz), 1440Hz dla F2 (SD = 361Hz), 2574Hz dla F3 (SD = 301Hz), 3435Hz dla F4 (SD = 389Hz) oraz 4777Hz dla F5 (SD = 424Hz).

Jak wykazała analiza wariancji, częstotliwości pierwszego formantu były w sposób statystycznie istotny ($p < 0,0003$) wyższe w wymowie dzieci o niskim poziomie zrozumiałości, natomiast częstotliwości drugiego formantu ($p < 0,0047$) — w wymowie dzieci o wysokim poziomie zrozumiałości (por. tabela 2). Jeśli chodzi o płeć, zarówno formant czwarty jak i piąty osiągał wyższe wartości w wymowie chłopców, przy czym odnotowane różnice były istotne statystycznie na poziomie odpowiednio $p < 0,0006$ oraz $p < 0,0027$ (por. tabela 3).

Tabela 2: Średnie wartości formantów w realizacjach fonemu /n/ w Hz, w zależności od poziomu zrozumiałości mówców

FORMANT	POZIOM ZROZUMIAŁOŚCI	
	WYSOKI	NISKI
F1***	354	477
F2**	1489	1278
F3	2587	2532
F4	3438	3428
F5	4823	4617

Tabela 3: Średnie wartości formantów w realizacjach fonemu /n/ w Hz, w zależności od płci mówców

FORMANT	PŁEĆ MÓWCÓW	
	DZIEWCZĘTA	CHŁOPCY
F1	416	359
F2	1398	1470
F3	2539	2599
F4***	3299	3533
F5**	4646	4870

5.1.3 Fonem /ɲ/

Średnie wartości formantów w realizacjach fonemu /ɲ/ wyniosły odpowiednio 328Hz dla F1 (odchylenie standardowe = 53Hz), 1186Hz dla F2 (SD = 411Hz), 2563Hz dla F3 (SD = 229Hz), 3469Hz dla F4 (SD = 506Hz) oraz 4603Hz dla F5 (SD = 468Hz).

Zgodnie z wynikami analizy wariancji, wartości formantów F1, F2 oraz F5 były w sposób statystycznie istotny (odpowiednio na poziomie $p < 0,0002$, $p < 0,0083$ oraz $p < 0,0052$) wyższe w wymowie dziewcząt (por. tabela 4), natomiast niemożliwa była statystyczna analiza częstotliwości formantów ze względu na poziom zrozumiałości mówców, ponieważ realizacje fonemu /ɲ/ percypowane jako zgodne z zamierzonymi odnotowano wyłącznie w artykulacjach dzieci o wysokiej zrozumiałości mowy.

Tabela 4: Średnie wartości formantów w realizacjach fonemu /ɲ/ w Hz, w zależności od płci mówców

FORMANT	PŁEĆ MÓWCÓW	
	DZIEWCZĘTA	CHŁOPCY
F1***	358	298
F2**	1362	1011
F3	2534	2593
F4	3538	3401
F5**	4810	4384

5.2 Analiza dyskryminacyjna

Poddawany analizie dyskryminacyjnej wektor zmiennych, charakteryzujący realizacje poszczególnych fonemów nosowych, złożony był z częstotliwości pięciu pierwszych szczytów formantowych, których wartości średnie w badanym materiale mieściły się w zakresie od 300Hz do 4800Hz, natomiast korpus analizowanych danych obejmował trzy klasy obiektów (por. tabela 5).

Tabela 5: Średnie częstotliwości formantów (Hz) w realizacjach fonemów nosowych

FONEM	FORMANT				
	F1	F2	F3	F4	F5
/m/	353	1339	2414	3382	4403
/n/	383	1440	2574	3435	4777
/ɲ/	328	1186	2563	3469	4603

Jak wynika z przedstawionej powyżej tabeli, w realizacjach fonemu /n/ wszystkie pięć formantów osiągnęło wyższe wartości średnie niż w realizacjach fonemu /m/. Jeśli zaś chodzi o realizacje fonemu /ɲ/, to wyróżniały się one niższą niż w przypadku pozostałych dwóch fonemów nosowych częstotliwością formantów pierwszego oraz drugiego, przy czym zanotowana różnica była szczególnie wyraźna w odniesieniu do F2. Pozostałe formanty w realizacjach powyższego fonemu osiągnęły natomiast wartości średnie zbliżone do częstotliwości charakterystycznych dla fonemu /n/.

W tabelach 6 oraz 7 przedstawiono współczynniki korelacji połączonych w obrębie grup oraz korelacji całkowitych dla badanych parametrów akustycznych. Jak widać, wartości F1 w spółgłoskach nosowych wydają się być w zdecydowanie mniejszym stopniu powiązane z częstotliwościami innych szczytów widmowych niż ma to miejsce w odniesieniu do pozostałych par formantów, przy czym w żadnym z przypadków współczynniki korelacji nie wykazują wartości tak wysokich, by na etapie analizy dyskryminacyjnej uzasadnione było pominięcie któregośkolwiek z formantów jako zmiennej redundantnej.

Tabela 6: Matryca korelacji połączonych w obrębie grup

FORMANT	F2	F3	F4	F5
F1	0,098	0,167	0,095	0,071
F2		0,466	0,400	0,510
F3			0,595	0,558
F4				0,617

Tabela 7: Matryca korelacji całkowitych

FORMANT	F2	F3	F4	F5
F1	0,119	0,177	0,096	0,095
F2		0,461	0,390	0,509
F3			0,591	0,589
F4				0,598

Ponieważ zbiór danych poddawanych analizie dyskryminacyjnej¹⁰ obejmuje trzy klasy obiektów (dla trzech fonemów nosowych) a maksymalna liczba zmiennych dyskryminacyjnych (zwanych inaczej zmiennymi kanonicznymi lub pierwiastkami) jest mniejsza o jeden od liczby klas (por. Klecka 1980), zmienne początkowe (F1 – F5) reprezentowane będą w ramach poniższej analizy statystycznej przy pomocy dwóch nieskorelowanych zmiennych kanonicznych - P1 oraz P2. Tabela 8 przedstawia wartości współczynników dyskryminacyjnych, obrazujących stopień, w jakim wartości obu zmiennych kanonicznych związane są z wartościami poszczególnych zmiennych początkowych, oraz skumulowany zakres, w jakim powyższe pierwiastki kanoniczne odpowiedzialne są za całkowitą zmienność w analizowanym zbiorze danych.

¹⁰ Opis podstaw analizy dyskryminacyjnej wraz z przykładem zastosowania w badaniach z dziedziny fonetyki akustycznej znaleźć można w Jassem (1998).

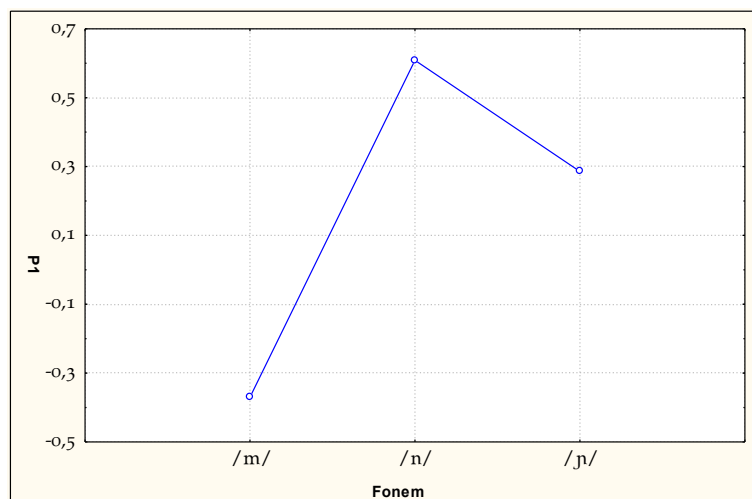
Tabela 8: Wartości współczynników standaryzowanych dla zmiennych dyskryminacyjnych P1 i P2 oraz skumulowany zakres, w jakim zmienne te odpowiedzialne są za całkowitą zmienność w analizowanym zbiorze danych

ZMIENNA POCZĄTKOWA	ZMIENNA DYSKRYMINACYJNA	
	P1	P2
F1	0,113	-0,362
F2	-0,283	-1,037
F3	0,459	0,542
F4	-0,729	0,280
F5	1,125	-0,043
SKUMUL.	0,764	1,000

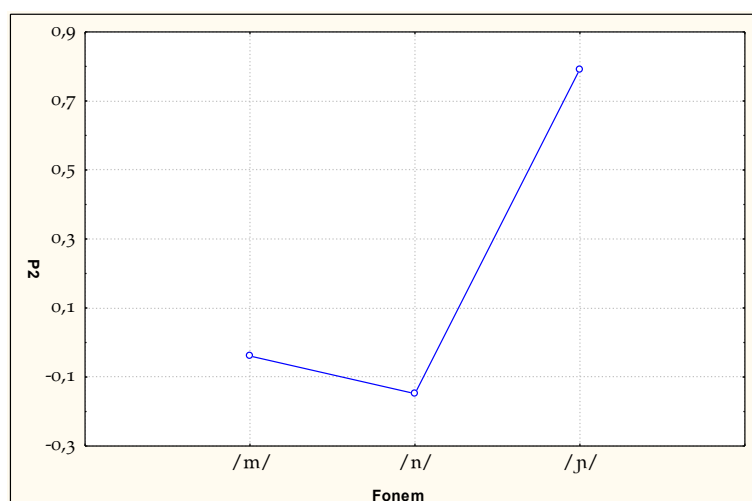
Jak wynika z powyższej tabeli oraz z przedstawionej poniżej matrycy korelacji pomiędzy zmiennymi dyskryminacyjnymi i początkowymi (por. tabela 9), wartość pierwiastka P1, odpowiedzialnego za 76% zmienności w zbiorze, w największym stopniu związana jest z częstotliwością formantu F5, natomiast wartość pierwiastka P2 — z częstotliwością formantu F2. Jeżeli weźmiemy pod uwagę średnie częstotliwości formantów dla realizacji poszczególnych fonemów nosowych, przedstawione w tabeli 5, okaże się, że wartości zmiennej kanonicznej P1 w największym stopniu różnicują realizacje fonemu /m/ od realizacji fonemu /n/, natomiast wartości zmiennej kanonicznej P2 odpowiedzialne są w głównej mierze za zróżnicowanie pomiędzy realizacjami powyższych fonemów niepalatalnych a realizacjami fonemu /ɲ/, co też obrazują odpowiednio ryciny 2 oraz 3.

Tabela 9: Matryca korelacji pomiędzy zmiennymi dyskryminacyjnymi oraz początkowymi

ZMIENNA POCZĄTKOWA	ZMIENNA DYSKRYMINACYJNA	
	P1	P2
F1	0,172	-0,349
F2	0,226	-0,730
F3	0,540	0,140
F4	0,135	0,127
F5	0,795	-0,123



Rycina 2: Średnie wartości zmiennej kanonicznej P1 dla realizacji poszczególnych fonemów nosowych



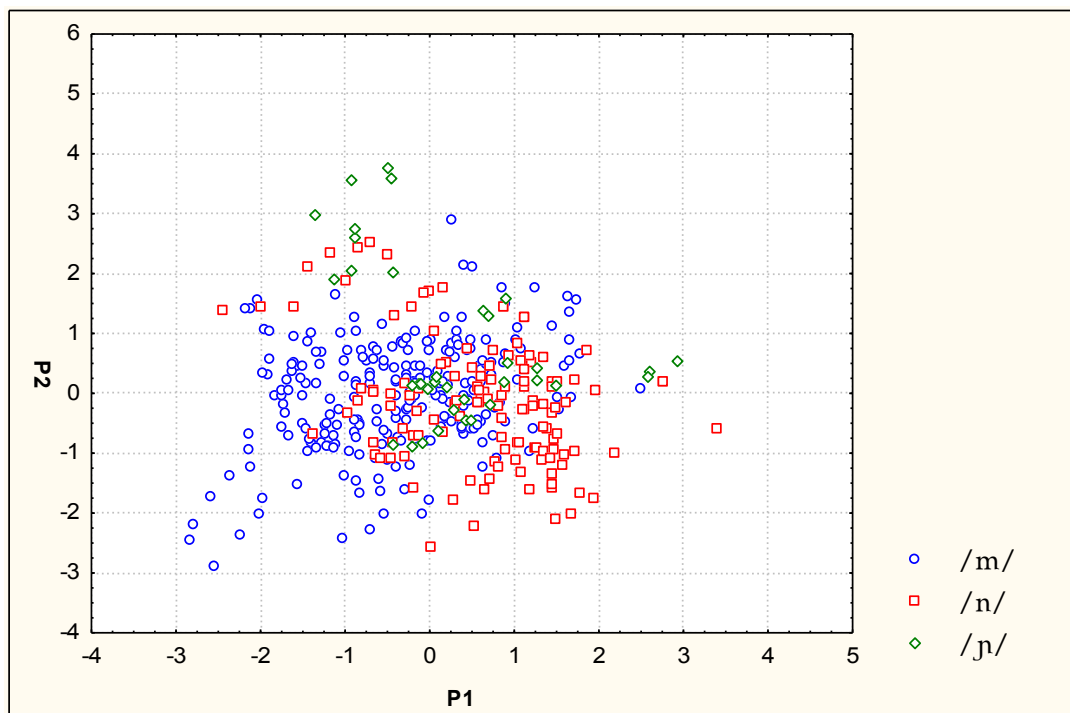
Rycina 3: Średnie wartości zmiennej kanonicznej P2 dla realizacji poszczególnych fonemów nosowych

Jak wskazują wysokie wartości statystyki Lambda Wilksa¹¹, przedstawione w tabeli 10, nawet przy uwzględnieniu obu zmiennych kanonicznych poszczególne klasy w badanym zbiorze są stosunkowo słabo rozróżnialne, co obrazuje także wykres rozrzutu na rycinie 4, przedstawiający wszystkie przypadki z badanych grup na płaszczyźnie, której współrzędne stanowią wartości zmiennych dyskryminacyjnych.

Tabela 10: Wartości Lambda Wilksa przed oraz po usunięciu pierwszego pierwiastka

USUNIĘTE PIERWIĄSTKI	LAMBDA WILKSA	STATYSTYKA CHI-KWADRAT	STOPNIE SWOBODY	POZIOM P
0	0,78	100,44	10	0,000000
1	0,94	24,89	4	0,000053

¹¹ 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).



Rycina 4: Wszystkie badane przypadki na płaszczyźnie P1 x P2

Jak wykazały ostatecznie wyniki analizy dyskryminacyjnej (por. tabela 11), średnia poprawność klasyfikacji w analizowanym zbiorze danych, w oparciu o wykorzystywane zmienne, wyniosła zaledwie 56%, przy czym zdecydowanie najgorzej identyfikowane były realizacje fonemu palatalnego /ɲ/. Jak wskazują odległości kwadratowe Mahalanobisa, przedstawione w tabeli 12, w pięciowymiarowej przestrzeni zmiennych początkowych (F1 – F5), względne odległości statystyczne pomiędzy badanymi klasami obiektów miały we wszystkich trzech parach podobne wartości średnie.

Tabela 11: Macierz klasyfikacji spółgłosek nosowych

	PROCENT POPRAWNOŚCI			
		m	n	ɲ
m	57,50	138	51	51
n	58,91	26	76	27
ɲ	36,11	9	14	13
RAZEM	56,05	173	141	91

Tabela 12: Kwadraty odległości Mahalanobisa pomiędzy poszczególnymi klasami

	n	ɲ
m	0,977	1,130
n		0,996

6 Podsumowanie

Jak wykazała przeprowadzona analiza dyskryminacyjna, prawidłowa identyfikacja polskich spółgłosek nosowych, realizowanych przez osoby niesłyszące, przy wykorzystaniu częstotliwości pierwszych pięciu szczytów formantowych, nastęrcza podobnych trudności, jak ma to miejsce w przypadku mowy niezaburzonej.

Warto jednakże zwrócić uwagę na odnotowaną w niniejszych badaniach możliwość wykorzystania częstotliwości drugiego formantu do względnie skutecznego odróżniania realizacji fonemu nosowego /ɲ/ od realizacji niepalatalnych spółgłoskowych fonemów nosowych /m/ oraz /n/, a także na fakt, iż największą rolę w różnicowaniu realizacji ostatnich dwóch spośród wymienionych powyżej fonemów nosowych wydaje się odgrywać wartość formantu F5. Spostrzeżenia powyższe znaleźć mogą zastosowanie nie tylko w ramach konstrukcji skomputeryzowanych narzędzi rewalidacyjnych, przeznaczonych dla osób niesłyszących, ale również w systemach automatycznego rozpoznawania mowy niezaburzonej.

Fakt, iż w ramach prowadzonych badań żadne z dzieci o niskim poziomie zrozumiałości ani razu nie zrealizowało prawidłowo fonemu /ɲ/ wydaje się również nasuwać wniosek natury dydaktycznej, iż w toku kształcenia dzieci niesłyszących konieczne jest zwrócenie szczególnej uwagi na doskonalenie wymowy spółgłoski nosowej palatalnej.

Poza tym 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 trących (por. Kleśta 2003b), zwartych, zwartotrących oraz samogłosek niesylabicznych. Uzupełnione w ten sposób dane, w połączeniu z rezultatami wspomnianych w niniejszym artykule badań Francuzik i Szalkowskiej (2001), 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.
- Ball, M.J., Code, Ch. 1997. *Instrumental Clinical Phonetics*. London: Whurr Publishers Ltd.
- Clark, J., Yallop, C. 1990. *An Introduction to Phonetics and Phonology*. Cambridge, Massachusetts: Blackwell Publishers Inc.
- Demenko, G. 1987. Wizualizacja częstotliwości podstawowej mowy. w: W. Jassem (Red.) *Wizualizacja mowy i jej zastosowania*. Warszawa: IPPT PAN. str. 65–88.
- Dickson, D.R. 1962. An acoustic study of nasality. *Journal of Speech and Hearing Research*, vol. 5, nr 2, str. 103–111.
- Domagała, P. 1988. Wielowymiarowa statystyczna analiza parametrów LPC segmentów fonetycznych w typowych połączeniach. *Prace IPPT*, 21/98.
- Dukiewicz, L. 1967. *Polskie głoski nosowe — analiza akustyczna*. Warszawa: PWN.
- Eckert, U. 1994. Znaczenie i warunki rozwoju języka słownego w kształceniu osób z wadami słuchu. w: S. Grabias (Red.) *Głuchota a język*. Lublin: UMCS. str. 63–70.
- Fant, G. 1970. *Acoustic Theory of Speech Production*. The Hague: Mouton.
- Fornby, C., Monsen, R.B. 1981. Long-term average speech spectra for normal and hearing-impaired adolescents. *Journal of the Acoustical Society of America*, vol. 71, No 1, Jan 1982, str. 196–202.
- Francuzik, K., Szalkowska, E. 2001. *Częstotliwości formantowe samogłosek w mowie dzieci niesłyszących*. Praca magisterska, materiały Instytutu Językoznawstwa UAM w Poznaniu.
- Fujimura, O. 1965. Formant-Antiformant Structure of Nasal Murmurs. *Speech Communication Seminar*. Stockholm.
- Galkowski, T. 1994. Stan rehabilitacji oraz potrzeby osób z uszkodzonym słuchem. w: S. Grabias (Red.) *Głuchota a Język*. Lublin: UMCS. str. 11–31.
- 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*. Lublin: UMCS. str. 59–62.
- Grunwell, P. 1987. *Clinical Phonology*. London: Croom Helm.
- Hoffmann, B. 1987. *Surdopedagogika*. Warszawa: PWN.
- Jassem, W. 1973. *Podstawy Fonetyki Akustycznej*. Warszawa: PWN.
- Jassem, W. 1987. *Wizualizacja mowy i jej zastosowania*. Warszawa: IPPT PAN.
- Jassem, W. 1997. Zrównoważone częstościowo i fonetycznie polskie listy wyrazowe. w: W. Jassem, Cz. Basztura (Red.) *Technologia Mowy i Języka*, tom 1, str. 71–100.
- 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*, tom 2, str. 9–45.
- Kacprowski, J. 1963. An Approach to the Synthesis of Polish Nasal Consonants by Means of the 'Terminal-Analog' Speech Synthesizer. *Proceedings of Vibration Problems*, tom 3, str. 235–254.
- Kapusta, M. 2000. *Wykorzystanie sieci Kohonena do wizualizacji mowy patologicznej*. Praca doktorska, materiały Akademii Górniczo-Hutniczej w Krakowie.
- Karczewska, L.B., Gubrynowicz, R., Karczewski, R. 2000. Analiza porównawcza konturów melodycznych w mowie dzieci głuchych. *Audiofonologia*, tom XVI, str. 5–20.
- Kent, R.D., Read, Ch. 1992. *The Acoustic Analysis of Speech*. San Diego, California: Singular Publishing Group Inc.
- 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. *Investigationes Linguisticae*, tom X, <http://www.staff.amu.edu.pl/~inveling>.
- Kleśta, J. 2003b. Analiza akustyczna polskich spółgłosek trących bezdźwięcznych realizowanych przez dzieci niesłyszące. *Audiofonologia*, tom XXVI, str. 105–118.
- Krajna, E. 1998. Lingwistyczne podstawy opracowania logopedycznego testu artykulacyjnego. *Investigationes Linguisticae*, tom IV, str. 5–22.
- Krakowiak, K. 1995. Fonogesty jako narzędzie formowania języka u dzieci z uszkodzonym słuchem. *Komunikacja językowa i jej zaburzenia*, tom 9, Lublin: UMCS.
- Krakowiak, K., Panasiuk, M. 1992. Umiejętności komunikacyjne dziecka z uszkodzonym słuchem. *Komunikacja językowa i jej zaburzenia*, tom 3, Lublin: UMCS.
- Levitt, H. 1972. Acoustic analysis of deaf speech using digital processing techniques. *IEEE*, vol. AU-20, No 1, str. 35–41.
- Łobacz, P. 1987. Wizualna Interpretacja Obrazów Spektrograficznych Mowy. w: W. Jassem (Red.) *Wizualizacja Mowy i Jej Zastosowania*. Warszawa: IPPT PAN. str. 89–134.
- Łobacz, P. 2002. Wstępna charakterystyka fonetyczna mowy wybranej grupy dzieci niesłyszących. *w druku*.
- Malecot, A. 1956. Acoustic cues for nasal consonants. *Language*, tom 32, nr 2, str. 274–284.
- Martony, J. 1964. The role of formant amplitudes in synthesis of nasal consonants. *Speech Transmission Laboratory, Quarterly Progress and Status Report*, 3/1964, Stockholm, str. 28–31.
- Mikiel, W., Żarnecki, P., Thuchowski, W., Komorowska, A. 1979. Parametry akustyczne mowy dzieci głuchych. *Otwarte Seminarium z Akustyki*, tom 26.

- 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), July 1999, str. 394–410.
- Oppenheim, A.V. 1982. *Sygnaly Cyfrowe*. Warszawa: Wydawnictwa Naukowo-Techniczne.
- Pickett, J.M. 1980. *The Sounds of Speech Communication*. Baltimore: University Park.
- Pietrzak, W. 1994. Różne drogi poznawania języka przez dzieci z wadą słuchu. w: S. Grabias (Red.) *Głuchota a Język*. Lublin: UMCS. str. 95–104.
- 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*. Poznań: Wydawnictwo AM. str. 324–338.
- Pruszevicz, A. 2000b. Diagnostyka zaburzeń słuchu u dzieci. w: *Zarys Audiologii Klinicznej*. Poznań: Wydawnictwo AM. str. 338–346.
- Rakowska, A. 1994. Natywizm Chomsky'ego a uczenie się języka przez głuchych. w: S. Grabias (Red.) *Głuchota a Język*. Lublin: UMCS. str. 115–130.
- Shannon, C.E., Weaver, W. 1949. *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.
- Sieńkowska, H., Gubrynowicz, R., Gałkowski, T. 2000. Nauka intonacji w mowie dzieci głuchych. *Audiofonologia*, tom XVII, str. 41–66.
- Stankiewicz, K., Włoch, K. 2000. *Cechy prozodyczne w mowie osób niesłyszących*. Praca magisterska, materiały Instytutu Językoznawstwa UAM w Poznaniu.
- Stevens, K.N. 1998. *Acoustic Phonetics*. Cambridge: MIT Press.
- Szczepankowski, B. 1994. Stosunek środowiska niesłyszących do edukacji oralnej — badania, które trzeba podjąć. w: S. Grabias (Red.) *Głuchota a Język*. Lublin: UMCS. str. 105–114.