

Piotra Łobacz, Kinga Dobrzańska

Instytut Lingwistyki Uniwersytetu Adama Mickiewicza
Poznań

Opis akustyczny głosek sybilantnych w wymowie dzieci przedszkolnych*

An Acoustic Description of Preschool Children Sibilants

Słowa kluczowe: polskie głoski sybilantne, opis akustyczny.

Key words: Polish sibilant consonants, acoustic description.

Streszczenie

Przedstawiono wyniki analiz akustycznych polskich spółgłosek sybilantnych (trących i zwarto-trących) w wymowie dzieci przedszkolnych. Stwierdzono zależność doskonalenia artykulacji trących sybilantów od wieku dzieci. Spółgłoski zwartotrące są doskonalone wolniej i w sposób bardziej przypadkowy niż odpowiednie głoski trące.

Summary

This study presents the results of acoustic analyses of Polish sibilant consonants (fricative and affricate) as pronounced by preschool children. The gradual approximation of fricative sibilants to the adult norm has been found to depend on the age. This process is both more random and slower in the affricates than the fricatives.

I. UWAGI WSTĘPNE

Przedmiotem niniejszej pracy jest interpretacja, z punktu widzenia fonetyki akustycznej, procesu doskonalenia wymowy spółgłosek trących i afrykat charak-

* Praca została wykonana w ramach projektu badawczego nr 996/H01/95/09, finansowanego przez KBN.

teryzujących się silnym szumem, wypowiedzianych przez dzieci w wieku od czterech do siedmiu lat. System fonologiczny języka polskiego zawiera 12 takich spółgłosek [s, z, ʃ, ʒ, ɕ, ʑ, ʦ, ʣ, ʧ, ʤ, ʦ̥, ʣ̥]. Na tle innych języków jest to dość znaczna liczba [por. np. Maddieson 1984]. Zbadanie mechanizmów ich nabywania przez dzieci wydaje się szczególnie interesujące, ponieważ w powszechnej opinii niektóre głoski sybilantne uchodzą za jedne z najtrudniejszych dźwięków [por. np. Gallagher, Shrinner 1975; Ingram (i in.) 1980; Stephens (i in.) 1986; Daniloff (i in.) 1991]. Doskonalenie ich artykulacji rozciąga się nawet na kilka lat.

Zdaniem Smoczyńskiego [1955] opanowanie przez dziecko niezbędnych opozycji artykulacyjnych umożliwiających m.in. wytworzenie określonych głosek o silnym szumie następuje dopiero w trzecim roku życia [np. opozycji: zębowość – dźwiękowość, miękkość – twardość]. Wczesne pojawienie się u dzieci głosek spalatalizowanych [ʃ, ʒ] [por. Zarębina 1965] uznać należy za zjawisko mniej typowe, podobnie jak całkowicie poprawną artykulację ząbówego silnego szumu u dziecka niespełna trzyletniego, które dysponowało poza tym ograniczonym systemem spółgłoskowym [por. Łobacz 1996]. Ingram (i in.) [1980] stwierdzają, że wymowa głosek [ʃ, ʒ, ɕ] przyswojona zostaje na przełomie trzeciego i czwartego roku życia, podobnie zdaniem Brajerskiego [1967] polskie dzieci dopiero powyżej trzeciego roku życia osiągają możliwość utworzenia szczeliny dźwiękowej i palatalnej, co pozwala na rozróżnienie artykulacyjne [ʃ, ʒ] od [ɕ, ʑ]. Żaden z tych autorów nie wypowiada się, w jakim stopniu artykulacje dzieci w tym wieku są zgodne z tzw. dorosłą normą. Wojnarowska [1987] obserwuje u dzieci przedszkolnych zjawisko tzw. seplenienia fizjologicznego zamiast prawidłowych artykulacji [s, ʒ, ʦ, ʣ], co potwierdzone zostało również wcześniejszymi badaniami przekrojowymi [Demel 1959; Bartkowska 1968; Mystkowska 1959; 1970; Zarębina 1980].

Można wyróżnić cztery zasadnicze typy dziecięcej artykulacji głosek o silnym szumie: (1) zgodną z dorosłą normą, (2) dentalizowaną, (3) boczną oraz (4) cofniętą¹. Typ (2) i (4) są szczególnie częste w wymowie dzieci i dotyczą nie tylko języka polskiego [por. Daniloff (i in.) 1991; Stephens (i in.) 1986]. Z polskich badaczy wymowę cofniętą zaobserwowała Zarębina [1980] u dziecka w wieku 4,8 roku.

Ponieważ wielu autorów wskazuje, że okres od 4 do 7 lat jest szczególnie istotny dla procesu doskonalenia artykulacji głosek sybilantnych, badania eksperymentalne referowane w obecnej pracy koncentrować się będą na tym właśnie przedziale wiekowym [por. m.in.: Fry 1966; Kaczmarek 1966; Menyuk 1971;

¹ Termin „dentalizacja” został tutaj użyty zgodnie z regułami fonetyki ogólnej i nie odnosi się do polskiej nazwy charakterystycznej cechy artykulacyjnej sybilantności polegającej na zbliżeniu zuchwy do górnych siekaczy, lecz do artykulacji międzyzębowej lub ściśle zębowej (czubek języka przylega do granicy między górnymi a dolnymi siekaczami maksymalnie zbliżonymi do siebie).

Blanchard 1972; Kent 1976; Fourcin 1978; Haller 1979; Wojnarowska 1987; Ingram 1989].

II. MATERIAŁ BADAWCZY I ZASTOSOWANA METODOLOGIA POMIAROWA

W niniejszej pracy wykorzystano materiał zarchiwizowany do badań wymowy dzieci przedszkolnych szczegółowo opisany w *Polskiej fonologii dziecięcej*, rozdz. 4 [Łobacz 1996]. Z przeprowadzonego testu nazywania, obejmującego 75 wyrazów wybrano dla obecnych potrzeb 28 następujących jednostek: *sale, kosa, kasa, zęby, koza, wazy, szale, kasza, wiesza, żaba, waży, wieża, Basie, Kasia, w lesie, bazie, tacka, pieniądze, taczka, ciało, bućka, w lecie, budzik, dozorca, łyczeczka, dziewczynka, drzewo, część*.

Wyraz „drzewo”, który ma dwie uznane realizacje normatywne [dʒevo, dʒʒevo], w obecnym materiale uwzględniono ze względu na jego (niepoprawną) wymowę przez poznańskie dzieci ([dʒʒevo]), co pozwala na lepsze zrównoważenie udziału poszczególnych grup spółgłosek w badanych przykładach. Ponieważ widma dźwięcznych i bezdźwięcznych spółgłosek sybilantnych o tym samym miejscu artykulacji nie różnią się charakterystycznym kształtem ich obwiedni – powyżej ok. 1000 Hz – lecz całkowitą energią akustyczną oraz ponieważ dzieci często ubezdźwięczniają (choćby częściowo) głoski dźwięczne, zdecydowano się obie klasy głosek potraktować łącznie i otrzymano następujące ich liczebności w badanym materiale językowym:

s, z (S)	7	ʦ, ʣ (C)	3
ʃ, ʒ (SZ)	8	ʧ, ʤ (CZ)	5
ɕ, ʑ (SI)	5	ʦ̥, ʣ̥ (CI)	7

W nawiasach okrągłych znajdują się symbole stosowane dla każdej z par spółgłosek we wszystkich zestawieniach liczbowych i rycinach. Niska liczba ząbówych afrykat wynika z ograniczonej częstości występowania głoski [dʒ] w języku polskim. Zrezygnowano z wyrazu „scyzoryk”, występującego w pierwotnym badaniu, mimo że zawierał spółgłoskę [ʦ], ponieważ w nagłosowej zbitce dochodziło do różnych procesów asymilacyjnych utrudniających jednoznaczność segmentację tego wyrazu lub do całkowitej elizji drugiej spółgłoski.

Wykorzystano nagrania czterech czterolatek, sześciorga pięcioletków, czwórki sześciolatek oraz pięciorga dzieci w wieku 7 lat. Szczegółowo metoda przeprowadzania sesji nagranych została opisana przez Łobacz [1996].

Z listy podanych powyżej wyrazów wynika, że analizowane spółgłoski znajdowały się w bardzo różnorodnym kontekście: w nagłosie przed samogłoską (np. *zęby*), w nagłosie przed spółgłoską (np. *świeca*), w śródgłosie między samogłoskami (np. *bucik*), w śródgłosie jako pierwszy element grupy

spółgłoskowej (np. *taczka*), w śródgłosie jako drugi element grupy spółgłoskowej (np. *dziewczynka*) oraz w wygłosie jako pierwszy i drugi element grupy spółgłoskowej.

Od lat sześćdziesiątych przyjmowano, że najważniejszą cechą identyfikacyjną spółgłosek trących są segmenty szumowe, a nie tzw. ugięcia formantowe sąsiednich samogłosek [por. Marteny 1962]. Względna niezależność koartykulacyjna trących była przeciwstawiana silnemu wpływowi kontekstowemu właściwemu identyfikacji spółgłosek zwartych. Soli [1981] wykazał, że szczyty w widmie trących w rejonie drugiego formantu samogłosek (w zakresie od 1,5 kHz do ok. 2 kHz) nie są właściwymi cechami charakterystycznymi tych spółgłosek, a jedynie wynikiem samogłoskowej koartykulacji. Yeni-Komshian i Soli [1981] wykazali, że wszystkie szczyty widmowe [s, z, ʃ, ʒ] znajdują się o ok. 100-300 Hz wyżej przed [i] niż przed [a, u]. Podobne rezultaty uzyskali na podstawie odrębnych metod pomiarowych Sereno (i in.) [1987]. Jassem [1979] wykazał, że w przypadku klasyfikacji spółgłosek trących na podstawie środków ciężkości ich widm lepsze rezultaty klasyfikacyjne osiąga się dla poszczególnych kontekstowych wariantów tych spółgłosek niż dla fonemów. Richter i Domagała [1995] na podstawie statystycznej analizy dyskryminacyjnej ustalili, że najsilniejszy wokaliczny wpływ kontekstowy na spółgłoski [s, ʃ, ʒ] wywiera obustronne sąsiedztwo [u]. W przypadku percepcji spółgłosek sybilantnych przez dzieci przedszkolne stwierdzono, że stadia nie ustalone sąsiednich samogłosek są istotnymi wskazówkami identyfikującymi barwę sybilantów dla dzieci 3-5-letnich, natomiast dla dzieci 7-letnich najważniejszy jest kształt samego widma [por. np. Nittrouer, Studdert-Kennedy 1986]. Istotna jest również rola kontekstu w przypadku artykulacji spółgłosek sybilantnych przez dzieci. Trzylatki bardziej poprawnie artykułowały [s] w bezpośrednim kontekście spółgłoskowym niż w otoczeniu samogłosek, podobnie dzieci starsze (7-letnie) częściej poprawnie realizują [s] w pojedynczych zbitkach niż w izolacji [por. Stephens (i in.) 1987].

Przyjęte kryteria metodologiczne w niniejszej pracy traktują różnorodność kontekstu za okoliczność pozytywną. Na obecnym etapie badań poszukuje się przeciętnego wzorca artykulacyjnego głosek sybilantnych w wymowie dzieci przedszkolnych i możliwie najprostszych reguł identyfikujących poszczególne fonemy.

W pierwszych badaniach polskiej mowy dziecięcej wykorzystujących metodologię fonetyki akustycznej wykazano, że trzyletnie dzieci nie mają jeszcze opanowanej poprawnej realizacji spółgłosek sybilantnych [Łobacz 1996]. Najgorzej przyswojonym dźwiękiem była spółgłoska [s], a najlepiej spółgłoski palatalne [ʃ, ʒ]. Ponieważ przedmiotem wymienionej pracy była ogólna charakterystyka fonetyczno-akustyczna mowy dziecięcej, uwzględniająca m.in. samogłoski i głoski nosowe, zastosowana wówczas metodologia analizy akustycznej sygnału mowy pozwalała na pomiary podstawowych parametrów wid-

mowych w zakresie od 0 do 6 kHz. Z punktu widzenia charakterystyki spółgłosek sybilantnych w wymowie dzieci za optymalny uznano zakres częstotliwości od 0 do 12 kHz. Wymagało to ponownej dygitalizacji wypowiedzi dziecięcych zarchiwizowanych wcześniej, którą przeprowadzono jeszcze raz dla potrzeb obecnych badań. Analiza pomiarowa została wykonana za pomocą systemu edycji i analizy sygnału akustycznego SONOLAB polskiej produkcji [por. Łobacz 1996]. Mimo że dla analizy mowy dziecka wykorzystuje się raczej niższe wartości rzędu predykcji liniowej [por. np. Lieberman 1980; Bergem 1988], to w przypadku badań głosek sybilantnych, charakteryzujących się szumem w wysokich częstotliwościach, wybrano rząd wynoszący 24 [por. Repp 1985].

Na podstawie tzw. spektrogramów szerokostęgowych całego wyrazu przeprowadzono jego segmentację i dokonano optymalnego wyboru miejsca dla analizy widmowej wybranej głoski trącej. Poszukiwania optimum w przypadku nie ustabilizowanej artykulacji dziecięcej nie oznaczało najbardziej środkowego fragmentu głoski. Dla każdej głoski wykonywano zawsze trzykrotną analizę widmową, ale tylko jedną z nich, wnoszącą najwięcej informacji, wykorzystywano w dalszych procedurach obliczeniowych. Główną podstawą pomiarów było widmo gładzone cepstralnie, pomocniczo wykorzystywano widmo LPC oraz widmo uśrednione długiego transientu. Łącznie wykonano ok. 2000 analiz widmowych, z których tylko 532 wykorzystano w interpretacji wymowy dzieci. Gdyby realizacje głosek trących przez wszystkie występujące w badaniu dzieci umożliwiały analizę akustyczną, chociażby ze względu na wystarczającą dynamikę sygnału, to należałoby uwzględnić 665 widm (37 głosek x 19 dzieci). Liczba ta w rzeczywistości jest znacznie niższa i dla poszczególnych grup wiekowych wynosi odpowiednio:

	spółgłoski trące	afrykaty
4-latki	54	37
5-latki	101	66
6-latki	75	57
7-latki	85	57

III. SŁUCHOWA OCENA GŁOSEK SYBILANTNYCH A KSZTAŁT OBWIEDNI WIDMA

Analiza słuchowa była dodatkowym czynnikiem wykorzystanym dla porównania barwy spółgłosek z ich charakterystycznymi cechami widmowymi i wykonana została jedynie przez eksperymentatora. Zaobserwowano takie procesy, jak: substytucje, seplenienie międzyzębowe, palatalizacja, np. [si,ʃi] (szczególnie w grupie czterolatków), udziąłowienie (cofnięcie artykulacji) oraz ubezdźwięcznienie. W celu porównania analizy słuchowej z ogólnym kształtem widma przyjęto trzy następujące typy wymowy dziecięcej:

- (1) artykulacja poprawna, jak u mówcy dorosłego,
- (2) dentalizacja – czyli obecność energii w niskich partiach widma z płaską obwiednią,
- (3) artykulacja cofnięta – względnie silna energia w niskich częstotliwościach i jeszcze silniejsza w wyższych rejonach częstotliwości w przypadku [s] oraz silna energia 2 i 3 maksimum występujących blisko siebie i znacznie słabsza w innych rejonach częstotliwości w przypadku [ʃ] (por. § 1). U badanych dzieci nie stwierdzono na podstawie odsłuchu artykulacji bocznej, manifestującej się na widmie obniżeniem dolnego zakresu szumu do 1 kHz i obecnością antyformantu pomiędzy drugim a trzecim maksimum w widmie.

Czwartą cechą było ubezdźwięcznienie, którego ocenę słuchową konfrontowano z zapisem oscylograficznym.

Klasyfikację ze względu na kształt obwiedni widma przeprowadzono osobno dla trzech miejsc artykulacji: ząbówego, zadziąsłowego i w ograniczonym zakresie (ze względu na brak artykulacji cofniętej) dla palatalnego. W poniższych zestawieniach wszystkie dane liczbowe przedstawiono w procentach.

(1) Spółgłoski ząbówego:

(a) artykulacja prawidłowa	49%
(b) substytucja	10%
(c) dentalizacja (seplenienie)	16%
(d) artykulacja cofnięta	25%

Udział prawidłowych artykulacji zwiększa się z wiekiem, ale istotny próg procesu doskonalenia artykulacji przypada na 6 rok życia dziecka. Wartości poprawnych artykulacji dla poszczególnych grup wiekowych wynoszą odpowiednio: czterolatki – 33%, pięcioletki – 30%, sześciolatki – 60% oraz siedmioletki 69% wszystkich, w danym przedziale wiekowym, realizacji spółgłosek ząbówych.

Do najczęstszych substytucji należała zamiana [z] na [c], [ʃ] na [tʃ], oraz [s] na [c]. Wystąpiła również sporadycznie zamiana [ʃ] na spółgłoskę zwartą [t]. Artykulacja cofnięta świadczy o skłonności dzieci do palatalizowania i mimo że częściej występuje u dzieci młodszych, to dziewięciokrotnie pojawiła się jeszcze u dzieci najstarszych.

Ubezdźwięcznienia obliczono niezależnie od poprawności artykulacji ponadkrtaniowej. Stanowią one 24% wszystkich realizacji sybilantów ząbówych i występują częściej u dzieci młodszych. Kilka z nich jednak pojawiło się w wymowie 7-latków.

(2) Spółgłoski zadziąsłowe:

(a) artykulacja prawidłowa	56%
----------------------------	-----

(b) substytucja	13%
(c) dentalizacja	26%
(d) artykulacja cofnięta	5%

Do najczęściej występujących substytucji zaliczyć należy: [ʃ] – [s], [ʒ] – [z] oraz [tʃ] – [ʃ]. Dwukrotnie nastąpiło zastąpienie afrykaty [tʃ] spółgłoską zwartą [t]. Niski procent artykulacji cofniętych nie tyle dowodzi niskiej częstości występowania tego zjawiska, co raczej trudności jednoznacznej interpretacji obwiedni widma. Słuchowo znacznie częściej stwierdzano zjawisko słabszej lub silniejszej palatalizacji. W wielu pracach dotyczących dyskryminacji spółgłosek trących na podstawie charakterystycznych parametrów widmowych rozróżnialność [ʃ] i [c] była najgorsza [por. np. Jassem 1995].

Ubezdźwięcznienie dotyczyło ok. 20% wszystkich spółgłosek zadziąsłowych dźwięcznych i najczęściej występowało dla spółgłoski [ʒ] w wyrazie „żaba”.

(3) Spółgłoski prepalatalne:

(a) artykulacja prawidłowa	77%
(b) substytucja	1%
(c) dentalizacja	22%

Określenie uprzednionej międzyzębowej artykulacji spółgłosek palatalnych wyłącznie na podstawie kształtu obwiedni widma było trudne ze względu na ich występowanie m.in. w przynajmniej jednostronnym sąsiedztwie samogłoski [i], co powodowało obecność w widmie silnego maksimum w zakresie 2-3 kHz. Nie można było jednoznacznie ocenić, czy jest to wynik uprzednionej artykulacji, czy wpływ kontekstu samogłoskowego. Jeśli maksimum w omawianym zakresie było najwyższym szczytem w widmie, charakteryzującym się w wysokich częstotliwościach niską energią, to takie zjawisko zaliczano do dentalizacji. Artykulacja dentalizowana występowała także, gdy widmo charakteryzowało się względnie płaskim kształtem i niską całkowitą energią.

Zaobserwowano tylko dwie substytucje: [d] oraz [ʒ] zamiast [dʒ]. Ubezdźwięcznienia były częstym zjawiskiem i wystąpiły w 44% wszystkich realizacji głosek normatywnie dźwięcznych. Szczególnie miało to miejsce w przypadku wyrazu „budzik”. Z kolei w wyrazie „bucik” kilkakrotnie wystąpiło udźwięcznienie afrykaty zrealizowane przez dzieci sześciolatnie, co oznacza, że mimo znanych dzieciom i bardzo różnych desygnatów, wymowa obu wyrazów była przez dzieci utożsamiana.

Kształt widm spółgłosek sybilantnych w wymowie polskich dzieci przedszkolnych charakteryzuje się wyższą amplitudą niskich częstotliwości w porównaniu z analogicznymi widmami dla mówców dorosłych.

IV. CHARAKTERYSTYCZNE PARAMETRY WIDMOWE GŁOSEK SYBILANTNYCH

Badania fonetyczno-akustyczne polskich spółgłosek trących prowadzone są od ponad 40 lat. W każdym z nich próbuje się wskazać najbardziej relewantne cechy akustyczne dla poszczególnych spółgłosek i stworzyć model ich rozpoznawania na podstawie wybranych parametrów akustycznych. Najczęściej wykorzystywano w tym celu: poziomy formantów szumowych, częstotliwości formantów oraz dolny zakres szumu. Poszukiwano relacji pomiędzy cechami widmowymi a cechami dystynktywnymi, wyznaczano parametry statystyczne (np. pierwszy, trzeci i czwarty centralny moment widma, czyli średnią ważoną częstotliwości, oraz jego skośność i kurtozę [por. np. Jassem 1979; Jassem 1995]).

Liczba cech widmowych wykorzystywanych w automatycznej klasyfikacji polskich spółgłosek szumowych może wynosić nawet kilkanaście parametrów widmowych, przy czym dla każdego z nich wyznacza się jeszcze kilka zmiennych statystycznych. Stopień rozbudowania aparatu pomiarowo-statystycznego zależy od liczby uwzględnionych mówców i różnorodności materiału językowego – liczby możliwych kontekstów [por. np. Domagała, Richter 1994].

Z drugiej strony badania porównawcze przeprowadzone przez Narteya [1982] dla ok. 10 języków wykazały, że w danym języku można dokonać poprawnego rozróżnienia spółgłosek trących tylko na podstawie dwóch parametrów – częstotliwości dwóch szczytów widmowych. Ten sam autor stwierdził jednak, że dla głosek szumowych występujących w polskim języku należy uwzględnić cztery parametry akustyczne [por. szerzej także Łobacz 1998].

Porównanie dla amerykańskiej odmiany języka angielskiego analizy widmowej spółgłosek [s, ʃ] wymówionych przez mówców dorosłych i przez dzieci wykazało następujące różnice: (1) znacznie silniejszą energię w niskich rejonach częstotliwości u mówców dziecięcych, (2) przesunięcie w górę w skali częstotliwości występowania silnej energii u dzieci, co jest uwarunkowane fizjologicznie rozmiarem ponadkrtaniowego toru artykulacyjnego, oraz (3) znacznie większy wpływ u dzieci kontekstu samogłoskowego [por. McGowan, Nittrouer 1988]. Cechę tę stwierdzono również dla polskich przedszkolaków – por. wyżej punkt III. Stosunki amplitudowe wykorzystywano także jako parametr w pracy dotyczącej rozwoju artykulacji /s/ [por. Stephens (i in.) 1986].

W obecnej pracy poszukuje się jak najprostszego modelu rozróżnienia spółgłosek sybilantnych w wymowie dzieci. Nieustabilizowanie artykulacji nawet u tego samego dziecka, liczne odchylenia od wymowy normatywnej, zróżnicowany wiek dzieci i różnorodny kontekst powodowały, że poszukuje się najbardziej istotnych wskazówek pozwalających na scharakteryzowanie odrębności wymowy

dziecięcej w stosunku do wymowy normatywnej mówców dorosłych. Za najważniejszy w obecnych badaniach uznano czynnik rozwojowy.

Z akustycznego punktu widzenia o naturze sybilantności świadczy względnie silna energia w wysokich rejonach widma. Z kolei wpływ kontekstu zaznacza się przeważnie w zakresie częstotliwości typowym dla formantu drugiego sąsiadujących samogłosek, czyli w niższych partiach widma [por. Yeni-Komshian, Soli 1981]. Wobec powyższego w obecnej pracy dokonano pomiarów czterech najwyraźniejszych szczytów w widmie występujących powyżej 2 kHz (por. punkt II). Uzyskane wyniki przedstawiono osobno dla spółgłosek trących i dla afrykat.

Klasyfikacja prezentowanych poniżej danych liczbowych autentycznej wymowy dzieci jest zgodna z systemem fonologicznym dorosłego użytkownika języka.

1. Spółgłoski sybilantne trące

Dla 315 realizacji spółgłosek [s, z, ʃ, ʒ, ʃ, ʒ, ʃ, ʒ] pogrupowanych w trzy zbiory (por. punkt I) wyznaczono wartości częstotliwości czterech najwyższych maksimum w obwiedni widma (wszystkie powyżej 2000 Hz). Jeśli dwa kolejne szczyty o takich samych amplitudach znajdowały się blisko siebie, wyznaczano ich średnią częstotliwość.

W tab. 1 przedstawiono średnie arytmetyczne wartości częstotliwości dla poszczególnych par spółgłosek i grup wiekowych dzieci.

Przeciętne wartości częstotliwości czterech maksimum dla wszystkich grup wiekowych pozwalają na następującą charakterystykę spółgłosek:

- (1) częstotliwość wszystkich maksimum dla spółgłosek [ʃ, ʒ] przypada w najniższych zakresach;
- (2) wysoka wartość częstotliwości maksimum 1 odróżnia [ʃ, ʒ] od pozostałych spółgłosek;
- (3) najwyższe wartości częstotliwości maksimum 2, 3 i 4 mają spółgłoski [s, z].

Przedziały modalne poszczególnych maksimum dla porównywanych spółgłosek przedstawiono w tab. 2. Wspólne zakresy występują dla: 1 maksimum spółgłosek [s, ʃ] oraz 3 maksimum spółgłosek [ʃ, ʒ]. Wartości maksimum 2 i 4 układają się regularnie według kolejności: spółgłoski szumiące, ciszące, syczące. Wszystkie cztery maksima są istotne dla rozróżnienia tych trzech grup trących.

Przedstawione powyżej wartości średnie częstotliwości czterech maksimum oraz przedziały modalne wynoszące 500 Hz wskazują na znaczny rozrzut danych pomiarowych.

Tab. 1. Wartości częstotliwości maksimów widma (LPC) dla spółgłosek S SZ SI i grup wiekowych dzieci

Spółgłoska	Wiek (w latach)	Maksimum 1	Maksimum 2	Maksimum 3	Maksimum 4
S	4	2589	4823	6601	8035
	5	2479	4899	6538	8081
	6	2243	4997	6850	7798
	7	2092	4781	6372	7794
	$\bar{x}$	2341	4875	6574	7918
SZ	4	2553	3980	5474	6829
	5	2421	4061	5384	6743
	6	2325	3999	5360	6898
	7	2150	3682	4959	6483
	$\bar{x}$	2341	3923	5255	6714
SI	4	2938	4290	5614	6801
	5	2918	4327	5827	7425
	6	2921	4215	5781	7360
	7	2789	4151	5433	6939
	$\bar{x}$	2885	4246	5680	7241

Tab. 2. Przedziały modalne dla czterech maksimów widm trących spółgłosek sybilantnych

Maksimum	Spółgłoski	Przedział Modalny (w Hz)	Liczebność pomiarów (w %)
1	S	2000-2500	47
	SZ	2000-2500	48
	SI	2500-3000	53
2	S	5000-5500	36
	SZ	3500-4000	34
	SI	4000-4500	38
3	S	6000-7500*	70
	SZ	5000-5500	51
	SI	5000-5500	30
4	S	7500-8000	34
	SZ	6500-7000	34
	SI	7000-7500	30

* Dla maksimum 3 spółgłosek grupy S otrzymano rozkład dwumodalny z przedziałami 6000-6500 oraz 7000-7500. Ponieważ liczebność przedziału pomiędzy dwoma modalnymi była niższa tylko o 1%, uznano go za równoważny, stąd przedział modalny wynosi w tym przypadku 1500 Hz.

Tab. 3. Współczynnik zmienności spółgłosek trących

Spółgłoska	Wiek (w latach)	Maksimum 1	Maksimum 2	Maksimum 3	Maksimum 4
S	4	19,50	10,56	10,76	7,52
	5	17,36	13,87	11,92	8,40
	6	13,36	11,37	8,43	6,56
	7	14,61	10,96	10,20	7,50
SZ	4	17,70	9,93	10,36	6,95
	5	22,24	13,47	7,76	8,65
	6	13,13	10,50	8,77	7,51
	7	15,98	10,37	9,47	9,65
SI	4	8,55	9,81	7,68	4,08
	5	10,76	9,09	8,79	7,81
	6	17,57	14,80	13,38	8,83
	7	14,05	12,17	8,58	8,67

W celu oszacowania tego rozrzutu wyznaczono współczynnik zmienności, którego wartości przedstawiono w tab. 3. Wysokość współczynnika zmienności maleje wraz ze wzrostem częstotliwości, co oznacza m.in. brak silniejszego wpływu otaczającego kontekstu. Na znaczne wartości tego współczynnika ma również wpływ nie zrównoważona liczebność pomiarów dla poszczególnych maksimów, związana z rozchwianiem artykulacji i nierówną reprezentacją grup wiekowych dzieci, oraz inne czynniki przypadkowe.

Tab. 4. Wartości testu F dla zmiennych niezależnych (spółgłoski – zmienna 1 oraz wieku – zmienna 2) oraz ich interakcje dla 4 maksimów spółgłosek [s, ʃ, ɕ] (wartości istotne oznaczono gwiazdką)

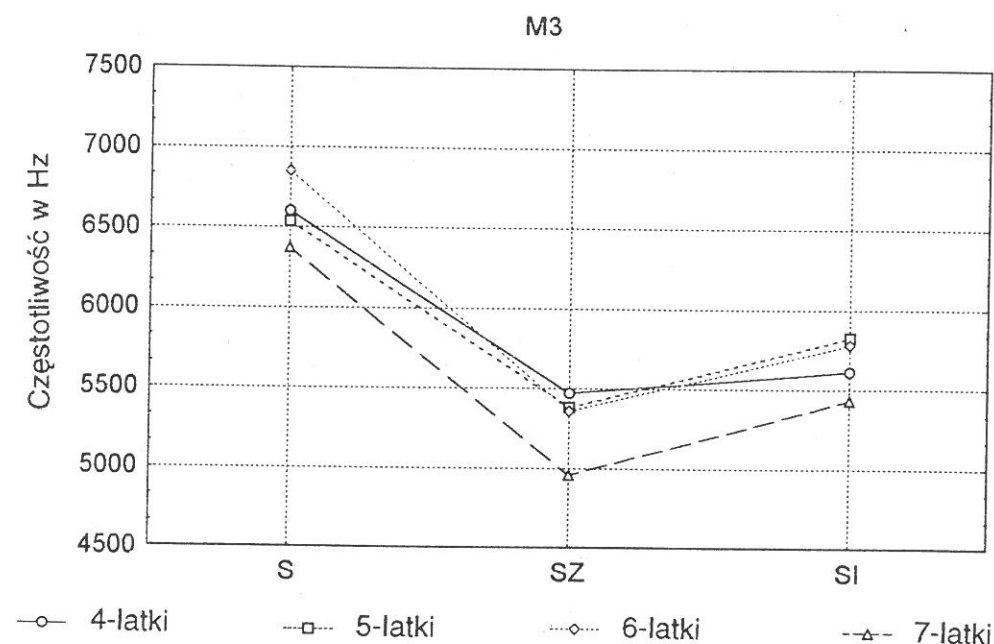
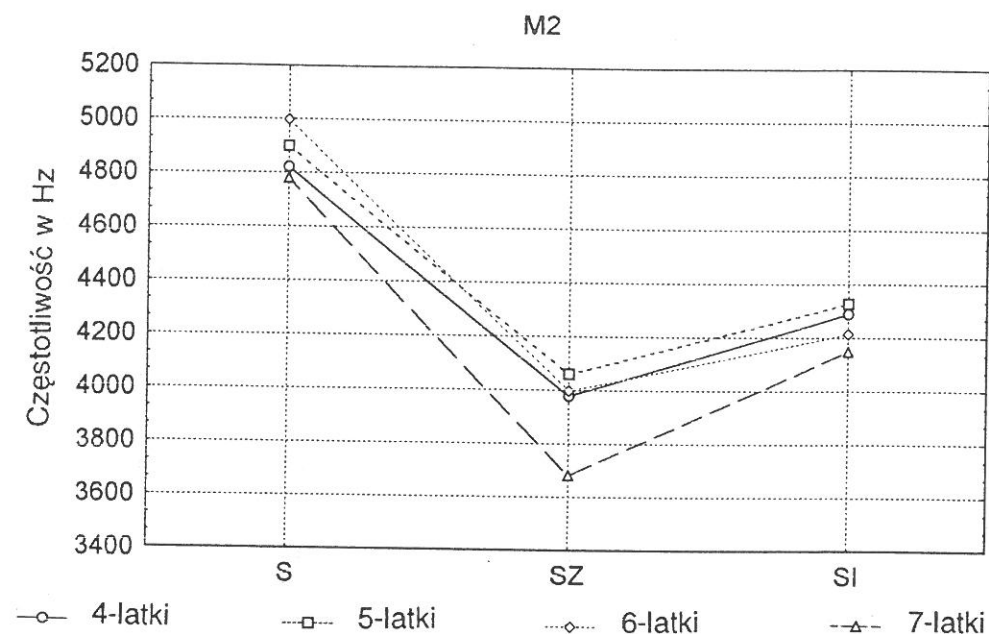
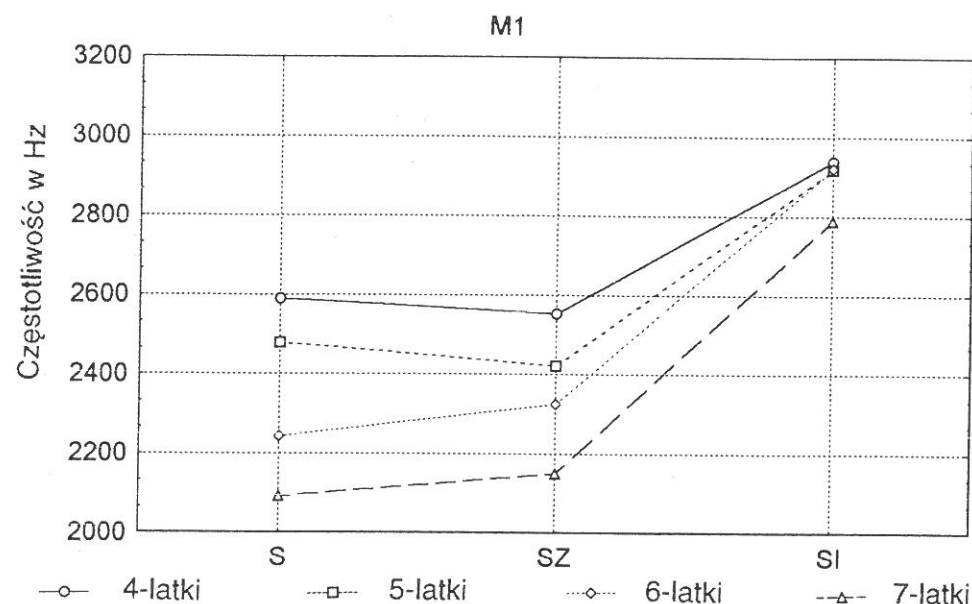
Maksimum	Zmienne	F	P
M1	1*	43,79	0,000
	2*	8,60	0,000
	1×2	0,87	0,514
M2	1*	80,15	0,000
	2*	3,19	0,024
	1×2	0,62	0,712
M3	1*	112,26	0,000
	2*	6,94	0,000
	1×2	0,82	0,554
M4	1*	98,87	0,000
	2*	4,79	0,003
	1×2	1,59	0,150

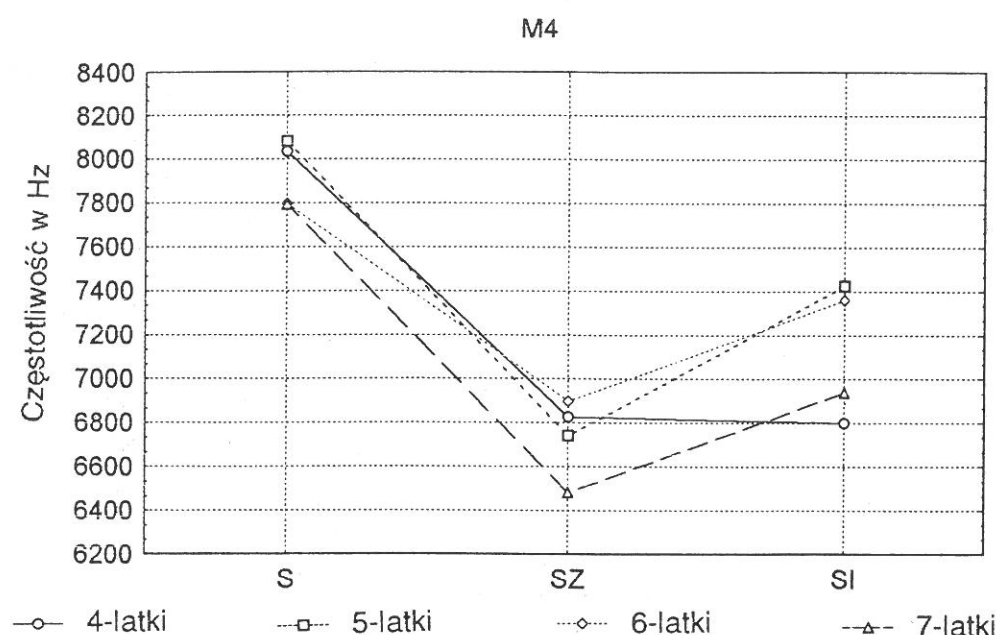
Pomimo dość dużego rozrzutu analizowanych danych zdecydowano się na przeprowadzenie dwuczynnikowej analizy wariancji. H_0 sformułowano następująco: różnice częstotliwości pomiędzy maksimami dla poszczególnych spółgłosek są przypadkowe. Zmiennymi niezależnymi były spółgłoski oraz wiek dzieci. W tab. 4 przedstawiono wartości testu F oraz poziom jego istotności dla poszczególnych zmiennych niezależnych oraz ich interakcji, osobno dla każdego maksimum.

Obie zmienne niezależne okazały się istotne dla wszystkich czterech maksimów widmowych. Nieistotne okazały się natomiast interakcje między nimi. Istotność czynnika wieku jest ważna dla ustalenia normy rozwojowej.

Na ryc. 1 a-d przedstawiono wartości poszczególnych szczytów widmowych dla określonych grup wiekowych dzieci. Największy rozrzut wartości w zależności od wieku dzieci zaobserwowano dla maksimum 1 spółgłosek [s, ʃ] (ryc. 1 a). Wysokość tego maksimum w skali częstotliwości obniża się wraz ze wzrostem wieku dzieci. Wartości maksimum 2 (ryc. 1 b) są wprawdzie najniższe dla dzieci najstarszych, ale najwyższe występują dla pięciolatków (dotyczy to wszystkich spółgłosek). Podobne zależności występują dla maksimum 3 (ryc. 1 c). Nie stwierdzono różnic dla maksimum czwartego spółgłoski [ʃ] oraz [ç] u najmłodszych dzieci.

Na podstawie przedstawionych wyników należy stwierdzić, że wymowa dzieci siedmioletnich jest najbardziej ustalona i wartości wszystkich analizowanych parametrów są najniższe. Wyniki dla pozostałych grup wiekowych świadczą o znacznym braku stabilizacji artykulacyjnej, wobec czego próg normy rozwojowej należy dla polskich trących spółgłosek sybilantnych ustalić na wiek powyżej 6 lat.





Ryc. 1. Porównanie wartości częstotliwości czterech szczytów obwiedni widm spółgłosek trących dla czterech grup wiekowych dzieci (a – maksimum 1, b – maksimum 2, c – maksimum 3 oraz d – maksimum 4)

W tab. 5 przedstawiono średnie wartości częstotliwości czterech szczytów widmowych wyznaczonych osobno według płci dzieci. Dla niektórych maksimumów wartości dla chłopców są niższe niż dla dziewczynek, ale w dwóch przypadkach (maksimum 1 [s] oraz maksimum 3 [ʃ]) wartości dla dziewczynek były niższe. Wartości testu F dla przeprowadzonej analizy wariancji okazały się dla zmiennej niezależnej płci dzieci nieistotne.

2. Spółgłoski sybilantne zwarto-trące

W tab. 6 przedstawiono przeciętne wartości częstotliwości czterech maksimumów trzech par spółgłosek zwarto-trących dla każdej grupy wiekowej badanych dzieci.

Zestawione w tab. 6 wartości średnie dla wszystkich grup wiekowych pozwalają na rozróżnienie afrykat ze względu na miejsce artykulacji na podstawie pierwszych trzech maksimumów widmowych. Wartości maksimum 4 wykazują te same relacje pomiędzy spółgłoskami co częstotliwości maksimum 3. Częstotliwość maksimum 1 jest najwyższa dla [tʃ, dʒ] – przeciętna różnica w stosunku do M1 [tʃ, dʒ].

Tab. 5. Średnie wartości częstotliwości maksimumów spółgłosek trących w zależności od płci dzieci

Spółgłoska	Płeć	Maksimum 1	Maksimum 2	Maksimum 3	Maksimum 4
S	chłopcy	2370	4824	6503	7785
	dziewczeta	2265	4953	6657	8068
SZ	chłopcy	2319	3931	5308	6705
	dziewczeta	2330	3941	5244	6771
SI	chłopcy	2872	4206	5634	7101
	dziewczeta	2904	4276	5717	7293

oraz M1 [tʃ, dʒ] wynosi ponad 1000 Hz. Podobna tendencja występuje dla maksimum 2, choć odległość na skali częstotliwości pomiędzy M2 [tʃ, dʒ] a M2 [ts, dz] uległa znacznemu zmniejszeniu. M3 oraz M4 mają wyższe wartości o ok. 600 Hz dla spółgłosek ząbówowych i prepalatalnych niż dla zadziastowych.

Ze względu na znaczne zróżnicowania osobnicze w zakresie poprawności artykulacyjnej poszczególnych dzieci, zróżnicowania kontekstowe i liczebnościowe dla poszczególnych grup wiekowych wyznaczono nie tylko średnią arytmetyczną częstotliwości czterech maksimumów widma, lecz także, podobnie jak w przypadku spółgłosek trących, przedziały modalne. Zakresy modalne przedstawiono w tab. 7.

Tab. 6. Średnie wartości częstotliwości maksimumów spółgłosek zwarto-trących dla grup wiekowych

Spółgłoska	Wiek (w latach)	Maksimum 1	Maksimum 2	Maksimum 3	Maksimum 4
C	4	3028	4625	6505	8330
	5	2464	4689	6617	7536
	6	2251	4822	6157	7507
	7	2219	4754	6490	7851
	$\bar{x}$	2474	4731	6435	7790
CZ	4	2578	4150	5770	7190
	5	2337	3906	5790	7217
	6	2293	4281	6007	7354
	7	2243	4181	5779	7299
	$\bar{x}$	2333	4116	5843	7274
CI	4	3523	5156	6682	8118
	5	3502	5081	6578	7739
	6	3747	5051	6311	7728
	7	3456	5009	6479	7823
	$\bar{x}$	3554	5066	6507	7834

Na podstawie porównania przedziałów modalnych o szerokości 500 Hz można określić wysoką rolę dyskryminacyjną M1 dla spółgłosek prepalatalnych,

najniższą wartość M2 dla spółgłosek zadziąsłowych, odrębne dla trzech porównywanych par zakresy M3 oraz wspólny zakres M4 dla spółgłosek ząbówowych i prepalatalnych, co może świadczyć o dwóch tendencjach typowych dla wymowy dziecięcej: palatalizacji spółgłosek ząbówowych i uprzednionej artykulacji silnie szumowej spółgłosek prepalatalnych.

Tab. 7. Przedziały modalne dla czterech maksimów widm sybilantnych afrykat

Maksimum	Spółgłoski	Przedział modalny w Hz	Liczebność pomiarów w %
1	C	2000-2500	33
	CZ	2000-2500	55
	CI	3500-4000	37
2	C	5000-5500	30
	CZ	4000-4500	29
	CI	5000-5500	43
3	C	6500-7000	40
	CZ	5500-6000	36
	CI	6000-6500	38
4	C	7500-8000	41
	CZ	7000-7500	46
	CI	7500-8000	31

Tab. 8. Wartości współczynnika zmienności dla spółgłosek zwarto-trących

Spółgłoska	Wiek (w latach)	Maksimum 1	Maksimum 2	Maksimum 3	Maksimum 4
C	4	11,35	9,89	8,69	7,16
	5	18,44	12,19	5,83	6,29
	6	16,21	12,18	7,94	7,45
	7	18,35	15,72	11,98	10,87
CZ	4	12,32	16,15	2,78	7,86
	5	9,53	11,50	3,21	4,53
	6	15,74	12,38	8,60	6,83
	7	13,69	12,45	10,09	7,05
CI	4	9,11	5,82	8,99	5,47
	5	16,90	15,65	9,98	7,12
	6	11,91	9,99	11,29	5,19
	7	14,20	11,82	9,90	6,29

W tab. 8 przedstawiono wartości współczynnika zmienności dla porównywanych par spółgłosek i grup wiekowych dzieci. Podobnie jak w przypadku

spółgłosek trących, wystąpiła tendencja mniejszych wartości tego współczynnika dla wyższych maksimów.

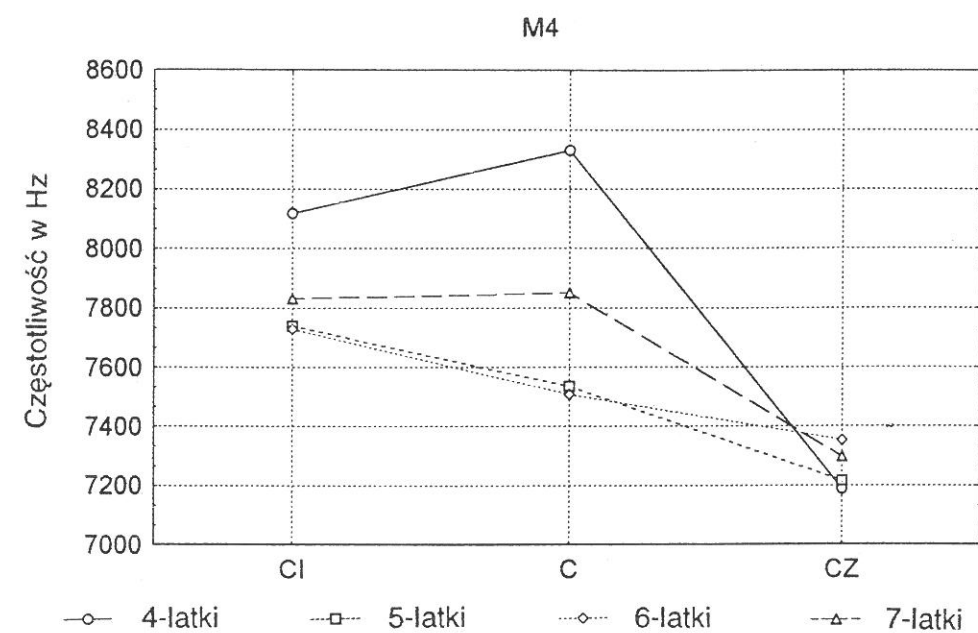
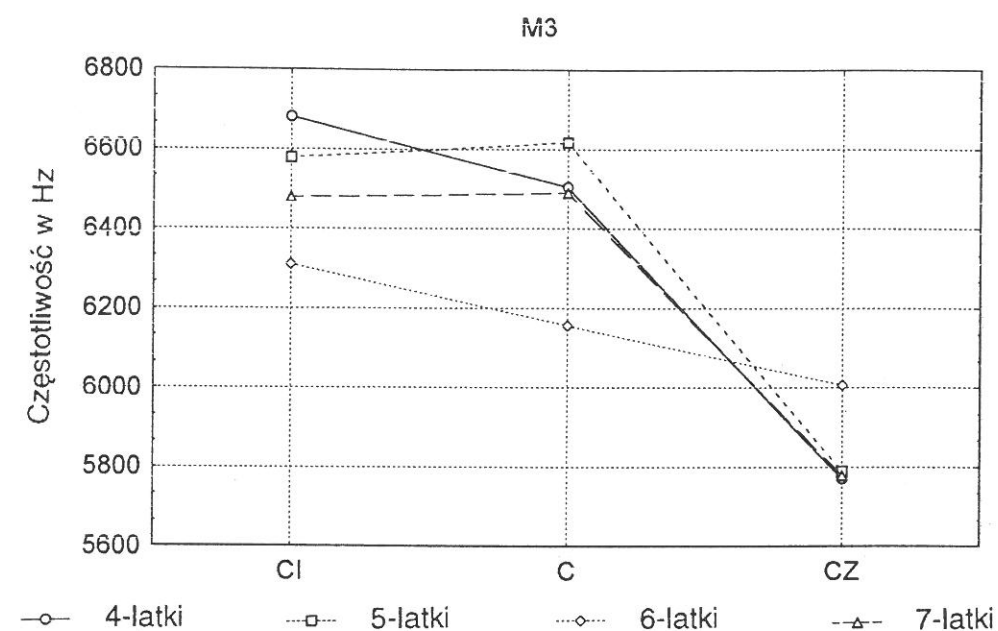
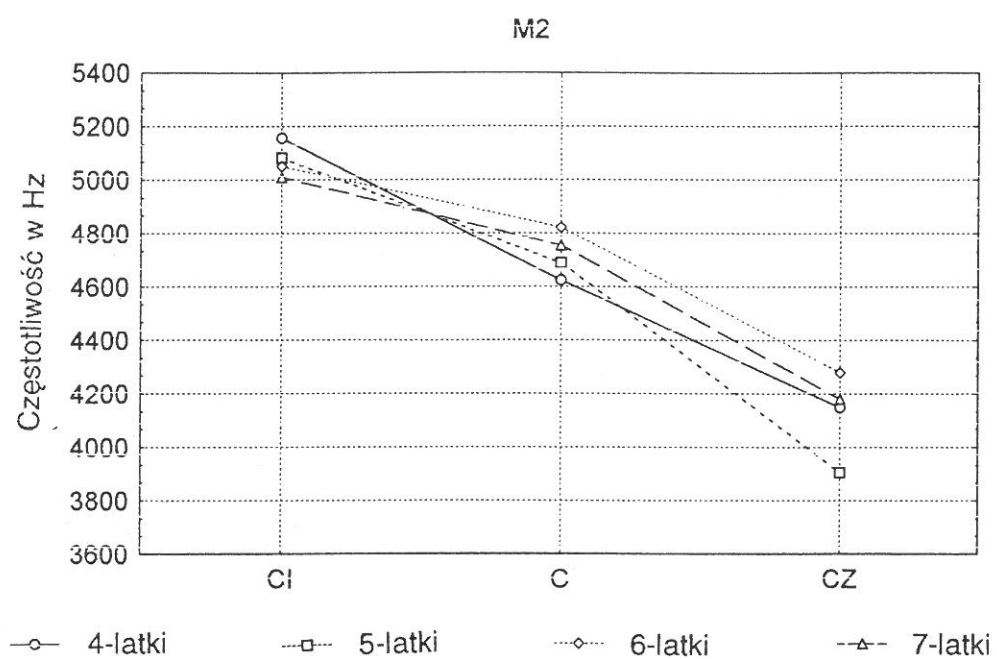
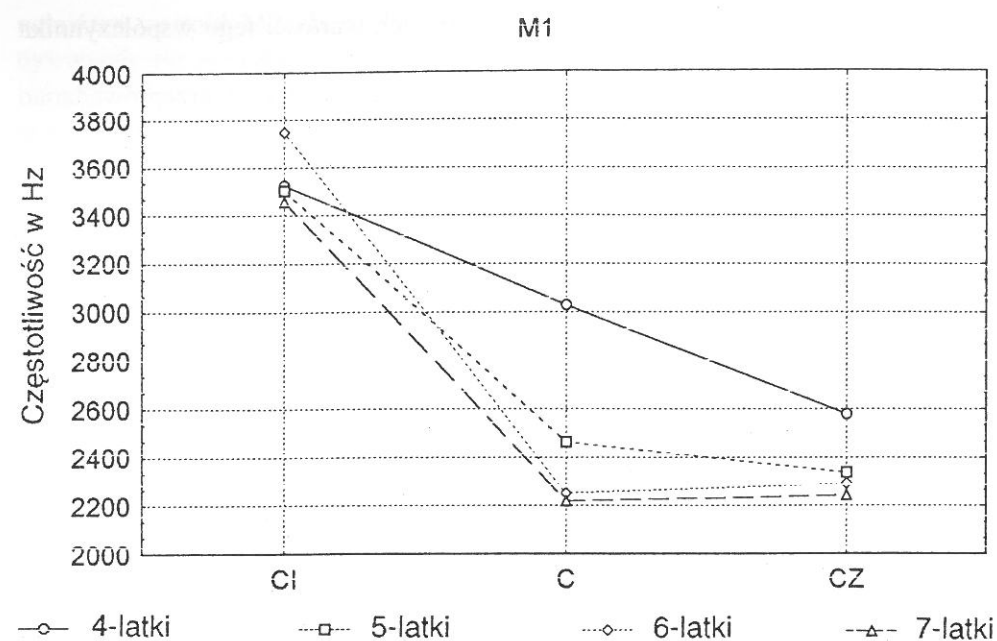
Pomimo wysokich wartości współczynnika zmienności przeprowadzono dwuczynnikową analizę wariancji dla zmiennych niezależnych: spółgłosek oraz wieku dzieci. Wartości testu F dla obu zmiennych oraz ich interakcji wraz z poziomem istotności dla wszystkich czterech maksimów przedstawiono w tab. 9.

Tab. 9. Wartości testu F dla zmiennych niezależnych (spółgłoski – zmienna 1 i wieku – zmienna 2) oraz ich interakcji dla czterech maksimów ts , tʃ , tc (wartości istotne oznaczono gwiazdką)

Maksimum	Zmienne	F	P
M1	1*	148,12	0,000
	2*	4,80	0,003
	1 × 2	2,47	0,025
M2	1*	48,90	0,000
	2*	0,61	0,610
	1 × 2	0,67	0,673
M3	1*	30,09	0,000
	2*	0,82	0,483
	1 × 2	1,49	0,189
M4	1*	22,51	0,000
	2*	3,22	0,024
	1 × 2	1,54	0,166

Wartość testu F okazała się istotna dla wszystkich czterech maksimów spółgłoskowych (dla pierwszej zmiennej niezależnej). Wiek dzieci okazał się istotny tylko w przypadku M1 oraz M4. Dla maksimum pierwszego istotna jest również interakcja pomiędzy spółgłoskami a wiekiem badanych dzieci.

Na ryc. 2 przedstawiono zależność wartości czterech maksimów od wieku badanych dzieci. Mimo że analiza wariancji wykazała brak zależności pomiędzy wiekiem dzieci a wartościami M2 oraz M3, na ryc. 2 uwzględniono również dane dla tych szczytów obwiedni widma. Na podstawie ryc. 2b można stwierdzić, że wartości M2 są prawie takie same dla wszystkich czterech grup wiekowych, natomiast w przypadku M3 (ryc. 2 c) różnica dotyczy tylko 6-letnich dzieci. Dla M1 (ryc. 2 a) występuje obniżanie się wartości częstotliwości w zależności od wieku, ale zasadniczo dotyczy ono różnicy pomiędzy najmłodszymi dziećmi a pozostałymi. Na ryc. 2 c uwidacznia się brak różnicy pomiędzy dziećmi 5- i 6-letnimi.



Ryc. 2. Porównanie wartości częstotliwości czterech szczytów obwiedni widm spółgłosek zwarto-trących dla czterech grup wiekowych dzieci (a – maksimum 1, b – maksimum 2, c – maksimum 3 oraz d – maksimum 4)

V. UWAGI KOŃCOWE

We wszystkich pracach dotyczących charakterystyki akustycznej segmentu szumowego afrykat zwraca się uwagę na ścisłą odpowiedniość tego segmentu do frakcji spółgłosek trących o tym samym miejscu artykulacji. Niniejsze badania tej odpowiedniości nie potwierdzają. Jej brak jest dowodem znacznie wolniejszego doskonalenia przez dzieci artykulacji głosek zwarto-trących w porównaniu z sybilantami trącymi. Uzyskane wyniki pomiarowe potwierdzają bowiem więcej podobieństw dziecięcych artykulacji do normatywnej wymowy mówców dorosłych w obrębie spółgłosek trących niż afrykat. Szczególnym odstępstwem od wymowy normatywnej jest brak rozróżnienia pomiędzy [tʂ, dʂ] a [tʂ, dʂ] ze względu na obecność silnego szumu w wyższych częstotliwościach (M2-M4). Ponieważ otrzymany wynik odbiega znacznie od powszechnego przekonania, że spółgłoski prepalatalne (w tym [tʂ]) są najszybciej z omawianego zbioru głosek przyswojone przez dzieci, należałoby dla spółgłosek zwartotrących przeprowadzić dalsze badania. Dla obecnych danych mieszanie się szeregu syczącego i ciszącego w obrębie afrykat świadczy o braku ustabilizowania się normatywnej wymowy dla wszystkich przedziałów wiekowych dzieci przedszkolnych. Wpływ na obecne wyniki mogła mieć również niska liczebność w materiale językowym ząbów afrykat. Ponieważ przeprowadzono nagrania ponad 200 dzieci w wieku przedszkolnym podczas 100-wyrazowego testu artykulacyjnego [por. Krajna, Bryndal – w druku], które także dowodzą znacznego nieustabilizowania wymowy w obrębie polskich afrykat, szczegółowe badania tej kwestii są szczególnie pożądane.

Bibliografia

- Bartkowska T. (1968): Rozwój wymowy dziecka przedszkolnego jako wynik oddziaływań wychowawczych rodziny i przedszkola. Warszawa: PZWS.
- Bergem D. R. van, Pols L. C. W., Kopmans van Beinum F. J. (1988): Perceptual normalization of the vowels of a man and a child in various context. „Speech Communication” 1, 1-20.
- Blanchard I. (1972): Phonemic analysis of speech development; the crucial fifth year. Proceedings of the Seventh International Congress of Phonetic Sciences (Rigault A., Charbonneau R., eds.). The Hague-Paris-Mouton, 1091-1094.
- Brajerski T. (1967): Rozwój sprawności artykulacyjnej kilkorga dzieci ze środowiska inteligentnego. „Logopedia” 1, 36-40.
- Daniloff R. G., Wilcox K., Stephens M. I. (1991): An acoustic-articulatory description of children's defective /s/ productions. Readings in clinical spectrography of speech (Baken R. J., Daniloff R. G., eds.). San Diego: Singular Publishing Group, Inc. and Kay Elemetrics Corp., 476-492.
- Demel M. (1959): Kilka liczb i uwag dotyczących mowy i głosu dzieci szkolnych. „Życie Szkoły” 2, 9-19.
- Domagała P., Richter L. (1994): Automatyczna klasyfikacja spółgłosek trących języka polskiego na bazie optymalizacji przestrzeni parametrów. Prace IPPT 9. Warszawa: IPPT PAN.

- Fourcin A. J. (1978): Acoustic patterns and speech acquisition. The development of communication (Watson N., Snow K., eds.). Chichester-New York-Brisbane-Toronto: John Wiley & Sons, 47-74.
- Fry D. B. (1966): The development of the phonological system in the normal and the deaf child. The genesis of language (Smith F., Miller G. A., eds.). Cambridge, Mass.: MIT Press, 187-206.
- Gallagher T. M., Shriner T. H. (1975): Articulatory inconsistencies in the speech of normal children. „Journal of Speech and Hearing Research” 18, 168-175.
- Haller R. A. (1979): Some competencies influencing phoneme acquisition in children. Current issues in the phonetic sciences. Proceeding of the IPS 77 Congress (Hollien H., Hollien P., eds.). Part II. Amsterdam: John Benjamin B.V., 995-960.
- Ingram D. (1989): First language acquisition, method, description and explanation. Cambridge: Cambridge University Press.
- Ingram D., Christiensen L., Veach S., Webster B. (1980): The acquisition of word-initial fricatives and affricates in English by children between 2 and 6 years. W: Child phonology. Vol. 1 (Yeni-Komshian G. H., Kavanagh J. F., Ferguson C. A., eds.). London-Toronto-Sydney-San Francisco-New York: Academic Press, 169-192.
- Jassem W. (1979): Classification of fricative spectra using statistical discriminant functions. frontiers of speech communication research (Lindblom B., Öhman S., eds.). London-New York-San Francisco: Academic Press, 77-91.
- Jassem W. The Acoustic parameters of Polish voiceless fricatives: An analysis of variance. „Phonetica” 52, 251-258.
- Kaczmarek L. (1966): Nasze dziecko uczy się mowy. Lublin: Wyd. Lubelskie.
- Kent R. D. (1976): Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies. „Journal of Speech and Hearing Research” 19, 421-447.
- Krajna E., Bryndal M. (1999): 100-wyrazowy test artykulacyjny. Analiza słuchowa nagrań i próba normalizacji testu. „Audiofonologia” 14, 137-174.
- Lieberman P. (1980): On the development of vowel production in young children. W: Child phonology Vol. 1 (Yeni-Komshian G. M., Kavanagh J. F., Ferguson C. A., eds.). New York-London-Toronto-Sydney, San Francisco: Academic Press, 113-142.
- Łobacz P. (1996): Polska fonologia dziecięca. Warszawa: Energeia.
- Łobacz P. (1998): Sybilantność. W: Scripta Manent (Puppel S., ed.). Poznań: Motivex, 135-154.
- Maddieson I. (1984): Patterns of sounds. Cambridge-London-New York-New Rochelle-Melbourne-Sydney: Cambridge University Press.
- Martony J. (1962): On the synthesis and perception of voiceless fricatives. Speech transmission Laboratory Quarterly Progress and Status Report 1/1962. Stockholm, 25-28.
- McGowan R., Nittrouer S. (1988): Differences in fricative production between children and adults. Evidence from acoustic analysis of [ʃ] and [s]. „Journal of the Acoustic Society of America” 83, 229-236.
- Menyuk P. (1971): The acquisition and development of language. Cambridge, Mass.: University Press.
- Mystkowska H. (1959): Kształcenie wymowy dzieci w przedszkolu. Ćwiczenia ortofoniczne. Warszawa.
- Mystkowska H. (1970): Właściwości mowy dziecka 6-, 7-letniego. Warszawa: PZWS.
- 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.
- Nittrouer S., Studdert-Kennedy M. (1986): The role of coarticulatory effects in the perception of fricatives by children and adults. Haskins Laboratories Status Report on Speech Research Sr-88 73-94.

- Repp B. H. (1985): Some observations on the development of the anticipatory articulation. Haskins Laboratories Status Report on Speech Research Sr-84 19-25.
- Richter L., Domagała P. (1995): Wariantywność kontekstowa spółgłosek trących bezdźwięcznych języka polskiego. *Prace IPPT* 7. Warszawa: IPPT PAN.
- Sereno J., Baum S., Mearan G., Lieberman P. (1987): Acoustic analyses and perceptual data on anticipatory labial coarticulation in adults and children. „*Journal of the Acoustic Society of America*” 81, 512-519.
- Smoczyński P. (1955): Przyswajanie przez dziecko podstaw systemu językowego. Łódź.
- Soli S. (1981): Second formants in fricatives: Acoustic consequences of fricative-vowels coarticulation. „*Journal of the Acoustic Society of America*” 70, 976-984.
- Stephens M. J., Hoffmann P. R., Daniloff R. G. (1986): Phonetic characteristics of delayed /s/ development. „*Journal of Phonetics*” 14, 247-256.
- Wojnarowska W. (1987): O pojęciu normy w rozwoju mowy dziecka. Materiały dydaktyczne dla logopedów (Wójtowicz J., ed.). Warszawa: Wydawnictwa UW, 101-117.
- Yeni-Komshian G., Soli S. (1981): Recognition of vowels from information in fricatives: Perceptual evidence of fricative-vowel coarticulation. „*Journal of the Acoustic Society of America*” 70, 966-975.
- Zarębina M. (1965): Kształtowanie się systemu dźwiękowego dziecka. Wrocław–Warszawa–Kraków: „Ossolineum”.
- Zarębina M. (1980): Język polski w rozwoju jednostki. Analiza tekstów dzieci do wieku szkolnego. Rozwój semantyczny języka dziecka. Kraków: Wydawnictwo Naukowe Wyższej Szkoły Pedagogicznej.