

LILIANNA KONOPSKA¹, JERZY SAWICKI²

¹Uniwersytet Szczeciński, Katedra Pedagogiki Specjalnej

²Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

Badania parametru VOT w realizacjach polskich bezdzwięcznych i dźwięcznych fonemów zwarto-wybuchowych*

Studies of the VOT parameter in realizations of Polish voiceless
and voiced plosive phonemes

STRESZCZENIE

W pracy przedstawiono wyniki badań fonetyczno-akustycznych dotyczące pomiaru parametru VOT (ang. *Voice Onset Time*) w nagłosowych realizacjach bezdzwięcznych i dźwięcznych fonemów zwarto-wybuchowych /p/, /t/, /k/, /b/, /d/, /g/ u 18 osób z normatywną wymową w wieku od 5 do 18 lat.

Słowa kluczowe: VOT, spółgłoski zwarto-wybuchowe, dzieci.

SUMMARY

The paper presents the results of phonetic and acoustic studies concerning the VOT (*Voice Onset Time*) parameter in the word-initial realizations of voiceless and voiced plosive phonemes /p/, /t/, /k/, /b/, /d/, /g/ in 18 speakers with standard pronunciation, aged from 5 to 18 years.

Key words: VOT, plosive consonants, children.

WSTĘP

Spółgłoski zwarte (ang. *stop consonants, plosives*), zwane również zwarto-wybuchowymi, są jedynym rodzajem spółgłosek, które występują w każdym

* Praca finansowana ze środków budżetowych na naukę w latach 2010–2013 jako projekt badawczy o numerze N N104 084639.

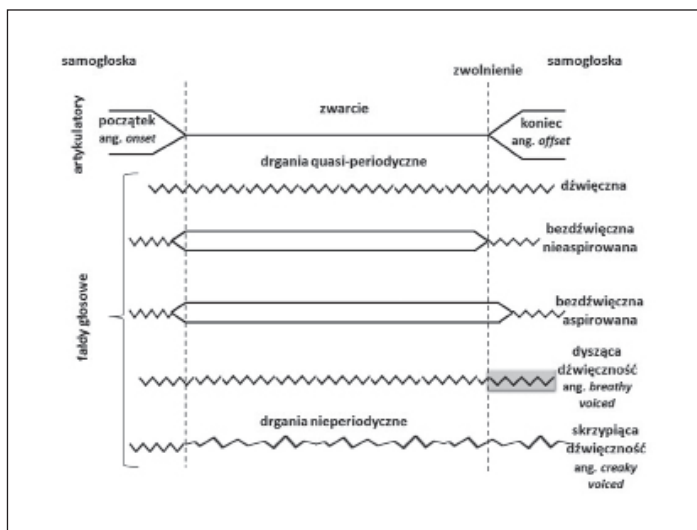
znany języku świata. Mogą być realizowane we wszystkich częściach toru głosowego, począwszy od warg a skończywszy na głośni (Ladefoged, Maddieson, 2008, 13–15). Wśród wielu możliwych miejsc artykulacji jedne są bardziej preferowane niż pozostałe. Według danych UCLA (UPSID)¹ z 1990 r., występowanie głosek zwartych ze względu na miejsce artykulacji w 393 badanych językach kształtowało się malejąco według następującego porządku: dentalno-alweolarnie², dwuwargowe, welarne, głośniowe, podniebienne, języczkowe, retrofleksyjne, postalweolarnie, wargowo-zębowe, językowo-wargowe. Uniwersalne granice obszaru artykulacji spółgłosek zwartych wyznaczone są przez wargi i języczek³ (Henton i in., 1992). Wśród wszystkich spółgłosek zwartych najczęściej występują bezdźwięczne nieaspirowane, następnie dźwięczne, potem zaś bezdźwięczne aspirowane. Jeżeli język posiada tylko jedną serię spółgłosek zwartych, są to zawsze płucne nieaspirowane bezdźwięczne spółgłoski zwarte. Z reguły najbardziej powszechne miejsca artykulacji spółgłosek zwartych łączą się z najbardziej powszechnymi typami fonacji i mechanizmami wywoływania przepływu strumienia powietrza artykulacyjnego, zatem za najczęściej występującą spółgłoskę zwartą w badanych językach uznano nieaspirowaną bezdźwięczną spółgłoskę [t] (dentalną/alweolarną) (Henton i in., 1992).

W artykulacji spółgłosek zwartych z płucnym mechanizmem przepływu powietrza wyróżnia się kilka zasadniczych faz: początek zwarcia, kiedy jeden artykulator zbliża się do drugiego; zwarcie, kiedy złączone artykulatory blokują przepływ powietrza, oraz zwolnienie zwarcia, kiedy artykulatory odsuwają się od siebie i nagromadzone powietrze zostaje gwałtownie uwolnione (plozja). Zmiany aktywności krtaniowej zachodzące w fazie formowania, trwania oraz zwolnienia zwarcia mogą być – jako czynności, które rozdzielają – porównywalnie ważne w procesie audytywnego postrzegania spółgłosek zwartych. Segment zwarcia z następującą plozją i szumem aspiracyjnym są najważniejszymi nośnikami informacji o spółgłosce zwartej (Henton i in., 1992). Użytkownicy poszczególnych języków stosują różne stany aktywności krtaniowej (głośni), w konsekwencji także różny przepływ powietrza przez głośnię i tor głosowy, a także kombinacje tych połączeń, by osiągnąć fonologiczne dla danego języka zróżnicowanie spółgłosek zwartych (ryc. 1.).

¹ Uniwersytet Kalifornijski, Los Angeles (ang. *University of California, Los Angeles*, UCLA). Baza danych: *UCLA Phonological Segment Inventory Databas*, UPSID). Baza ta zawiera dane na temat dystrybucji 919 różnych segmentów w 451 językach świata.

² W przytoczonej pracy wyszczególniono jeden wspólny obszar dla zwartych spółgłosek dentalnych i alweolarnych, zgodnie z zasadami obliczeń stosowanymi w UPSID (Henton i in., 1992).

³ Z wyłączeniem artykulacji wargowo-zębowych i językowo-wargowych (Henton i in., 1992).



Ryc. 1. Schematyczna struktura czasowa poszczególnych faz artykulacji spółgłosek zwartych w pozycji interwokalicznej z wyszczególnieniem zróżnicowanych stanów aktywności krtaniowej i przepływu strumienia powietrza artykulacyjnego⁴

Bezsporne zatem jest, że w procesie tworzenia spółgłosek zwartych w poszczególnych językach, także zależnie od miejsca spółgłoski w wyrazie i fazy jej artykulacji, mogą pojawić się zróżnicowane typy fonacji, np.: preaspiracja⁵ (ang. *preaspired*) – język islandzki, gaelicki; dźwięczność (ang. *voiced*) – większość języków; bezdźwięczność (ang. *voiceless*) – większość języków; skrzypiąca dźwięczność (ang. *creaky voiced*) – języki hausa i mazatecki; sztywna dźwięczność (ang. *stiff voiced*) – język jingpho i koreański; aspiracja (ang. *aspirated*) – język duński i tajski; dysząca dźwięczność (ang. *breathy voiced*) – język hindi i marathi; luźna dźwięczność (ang. *slack voiced*) – język jawajski i język wu. Zróżnicowany może być również mechanizm przepływu strumienia powietrza i jego kierunek, np. płucny egresywny (zdecydowana większość języków), ale także krtaniowy ingresywny (np. język sandhi, a w grupie języków afrykańskich język hausa i khoisan). Spółgłoski zwarte mogą też być odmiennie inicjowane w fazie tworzenia zwarcia (np. inicjalna glottalizacja, inicjalna nazalizacja⁶), choć generalnie czynności te są mniej urozmaicone i uważane za mniej ważne niż

⁴ Opracowano i zmodyfikowano na podstawie: Henton, Ladefoged, Maddieson (1992, 49, 66).

⁵ W pozycji interwokalicznej lub końcowej (Henton i in., 1992).

⁶ Prenazalizacja pojawia się najczęściej z dźwięcznymi głoskami zwartymi w wielu językach, w których żadne dźwięczne spółgłoski zwarte nie kontrastują ze spółgłoskami inicjalnie nosowymi, może to być sposób, by ułatwić utrzymanie dźwięczności spółgłoski zwartej (Henton i in., 1992).

zwarcie i/lub czynności kończące zwarcie, w których to fazach występują istotne fonemiczne zjawiska (Henton i in., 1992; Ladefoged, Maddieson, 2008)⁷.

PARAMETR VOT

Wśród możliwych fonetycznych wskazówek, które sygnalizują w obrębie spółgłosek zwartych fonologiczną korelację dźwięczności, dominującym sposobem kontrastowania są różne stany głośniowej aktywności. Jednym z najbardziej mierzalnych parametrów fonetyczno-akustycznych spółgłosek zwartych jest wyrażony w domenie czasu moment rozpoczęcia dźwięczności (wibracji fałdów głosowych) w stosunku do momentu uwolnienia zwarcia spółgłoski (Henton i in., 1992; Ladefoged, Maddieson, 2008). Przyjmuje się, że idea *Voice Onset Time*, czyli czasowej relacji istniejącej pomiędzy momentem zwolnienia zwarcia spółgłoski zwarto-wybuchowej a momentem rozpoczęcia periodycznych drgań fałdów głosowych sięga końca XIX w. i została wyrażona przez H. Adjarian w opisie realizacji ormiańskich spółgłosek zwartych⁸. Natomiast historia VOT, jako parametru użytego w syntezie mowy, sięga lat 50. XX w. i prac badawczych prowadzonych w Laboratoriach Haskinsa. Praktyczne wykorzystanie parametru VOT nastąpiło w specyficznym „syntezatorze mowy” (ang. *pattern-playback machine*), w którym metodami optyczno-mechanicznymi przetwarzano informację zawartą w obrazie spektrograficznym na odpowiadający jej przebieg akustyczny (Rottenberg, 2009). Upowszechnienie parametru VOT w badaniach nad sygnałem mowy nastąpiło wraz z opublikowaniem w 1964 r. przez L. Liskera i A. Abramsona prekursorskich wyników badań fonetyczno-akustycznych, w których *Voice Onset Time* zdefiniowano jako: „the interval between the release of the stop and the onset of glottal vibration, that is, voicing” (Lisker, Abramson, 1964, 389); „the time interval between the burst that marks release and the onset of periodicity that reflects laryngeal vibration” (Lisker, Abramson 1964, 422). Ideą podjętych badań było znalezienie akustycznych cech, które posłużą jako wskazówki do różnicowania spółgłosek zwartych poprzez ustalenie czasowych relacji pomiędzy zwolnieniem okluzji a początkiem dźwięczności. Wykonano zatem pomiary iloczynowych wartości VOT w nagłosowych realizacjach spółgłosek zwartych (wargowych, dentalnych, alweolarnych, retrofleksyjnych i welarnych) w 11 wybranych językach podzielonych na trzy grupy zależnie od liczby kategorii spółgłosek zwartych. Do grupy języków wykorzystujących dwie kategorie spółgłosek zakwalifikowano języki: holenderski, amerykański

⁷ Szersze omówienie zaledwie sygnalizowanych zjawisk wykracza poza ramy niniejszego opracowania.

⁸ Źródło: Wikipedia. http://en.wikipedia.org/wiki/Voice_onset_time (dostęp: 25 sierpnia 2013 r.).

angielski, kantoński, węgierski, hiszpański (Puerto Rico) i język tamil, do wykorzystujących trzy kategorie języki: koreański, wschodni armeński i tajski, a do wykorzystujących cztery kategorie – język hindi i język marathi. W opisanej przez Liskera i Abramsona procedurze wyznaczania odcinka pomiędzy zwolnieniem zwarcia realizowanej spółgłoski a początkiem wibracji fałdów głosowych przyjętym w domenie czasu punktem zerowym jest wybuch plosji. Na podstawie dokonanych pomiarów wartości parametru VOT stwierdzono, że w 11 badanych językach nagłosowe spółgłoski zwarte rozmieszczone są wzdłuż kontinuum⁹ VOT zasadniczo w dwóch z trzech wyszczególnionych kategorii VOT¹⁰ (Lisker, Abramson, 1964), które z czasem stały się podstawą do podziału języków na języki oparte na dźwięczności i na języki oparte na aspiracji (Keating, 1984). Wyszczególnione przez Liskera i Abramsona w oparciu o wartości, jakie przyjmuje parametr VOT w nagłosowych spółgłoskach zwarto-wybuchowych w pozycji przed samogłoską, trzy sposoby realizacji dźwięczności kształtują się następująco:

1) ujemne wartości parametru VOT – (ang.) *voicing lead*, dźwięczność wyprzedzająca, periodyczność pojawia się przed wybuchem spółgłoski w segmencie zwierania narządów artykulacyjnych, typowy zakres wartości VOT zawiera się w przedziale od (-125) [ms] do (-75) [ms], a średnia wartość wynosi około (-100) [ms] (dźwięczne nieaspirowane spółgłoski zwarto-wybuchowe);

2) dodatnie krótkie wartości parametru VOT – (ang.) *short voicing lag*, krótkie opóźnienie dźwięczności, periodyczność zaczyna się w momencie lub tuż po zwolnieniu zwarcia narządów artykulacyjnych, typowe wartości VOT zawierają się w przedziale od 0 [ms] do 25 [ms], a średnia wartość wynosi około 10 [ms] (bezdźwięczne nieaspirowane spółgłoski zwarto-wybuchowe);

3) dodatnie długie wartości parametru VOT – (ang.) *long voicing lag*, długie opóźnienie dźwięczności, periodyczność zaczyna się po zwolnieniu zwarcia narządów artykulacyjnych i następującej po zwolnieniu blokady aspiracji lub ciszy, typowe wartości VOT zawierają się w przedziale od 60 [ms] do 100 [ms], a średnia wartość wynosi około 75 [ms] (bezdźwięczne aspirowane spółgłoski zwarto-wybuchowe) (Lisker, Abramson, 1964).

W ujęciu Liskera i Abramsona VOT jest wymiarem krtaniowym, tj. krtaniową koordynacją czasową dźwięczności ze złożonym zestawem krzyżujących się, nachodzących na siebie lub nawet dyskretnie wyrażonych akustycznych wskaźników przejawiających się m.in. energią plosji, amplitudą szumu aspiracyjnego, skróceniem formantu F_1 (ang. *cutback*), niższą częstotliwością pierwszego

⁹ W ujęciu autorów, parametru VOT nie należy traktować jako kontinuum akustyczne, ale jako kontinuum artykulacyjne lub fizjologiczne (Abramson, 1977, 296).

¹⁰ Języki tajski, wschodni armeński i koreański wykorzystują trzy kategorie VOT (Lisker, Abramson, 1964).

formantu w początkowej fazie następującej samogłoski, zmianami częstotliwości podstawowej (Lisker, Abramson, 1964; Abramson, Lisker, 1965; Abramson, 1977). Z badań przeprowadzonych przez Liskera i Abramsona, a także z zapoczątkowanych przez ich publikację niezwykle licznych prac badawczych wynika, że wartości VOT (długie/krótkie trwanie w domenie czasu) zależą do pewnego stopnia od miejsca artykulacji i z reguły parametr VOT osiąga wyższe wartości wraz z postępującą w kierunku przednio-tylnym zmianą miejsca artykulacji. Relacje między miejscem artykulacji, masą artykulatorów i prędkością ruchów artykulacyjnych a wartościami VOT, określane też mianem uniwersaliów, kształtują się następująco:

1) im bardziej tylne zwarcie spółgłoski, tym wyższe wartości parametru (dłuższy) VOT, przy czym spółgłoski welarne mają zawsze wyższe wartości niż pozostałe spółgłoski zwarte,

2) im bardziej obszerny kontakt zwartych narządów artykulacyjnych, tym wyższe wartości parametru (dłuższy) VOT,

3) im szybszy ruch artykulatorów, tym niższe wartości parametru (krótszy) VOT,

4) im mniej czasu narządy artykulacyjne potrzebują na rekonfigurację układu (od miejsca zwarcia do układu masy języka właściwej dla następującej samogłoski), tym krótsze wartości VOT (Cho, Ladefoged, 1999).

Powszechnie znane są też inne, jakościowo odmienne, czynniki wpływające na długość VOT – kontekstowe i związane z tempem wypowiedzi. Wiele z tych czynników omówiono w odrębnym opracowaniu, poświęconym zagadnieniom iloczasu bezdźwięcznych spółgłosek trących (Konopska, 2013). Z badań nad uniwersaliami VOT wynika, że czas trwania VOT jest dłuższy przed wyższymi samogłoskami niż przed niższymi, w sylabie akcentowanej niż nieakcentowanej, w pojedynczych wyrazach niż w zdaniach, w wolniejszych niż w szybszych wypowiedziach (Lisker, Abramson, 1964; Lisker, Abramson, 1967; Lehiste, 1977; Flege, Brown, 1982; Docherty, 1992; Cho, Ladefoged, 1999; Mortensen, Tøndering, 2013).

Od czasu opublikowania pierwszej pracy na temat VOT mija właśnie pół wieku. W tym czasie z zastosowaniem parametru VOT zostały zbadane i opisane charakterystyki dźwięcznych i bezdźwięcznych spółgłosek zwarto-wybuchowych w wielu językach, m.in. w aspekcie rozwojowym, w aspekcie percepcji kategoryalnej, w aspekcie nauczania języków obcych, w aspekcie bilingwizmu (Caramazza i in., 1973; Macken, Barton, 1979; Macken, Barton, 1980; Keating i in., 1980; Davis, 1995; Kehoe i in., 2004; Hoonhorst i in., 2009; Ringen, Kulikov, 2012)¹¹. Parametr VOT znalazł również zastosowanie w badaniach

¹¹ Przytoczenie obszernego piśmiennictwa dotyczącego badań nad parametrem VOT wykracza poza ramy niniejszego opracowania.

klinicznych nad różnymi zaburzeniami mowy, m.in. jako wiarygodne narzędzie pomiaru i oceny percepcji i realizacji \pm dźwięczności w nagłosowych realizacjach spółgłosek zwarto-wybuchowych (np. Itoh i in., 1982; Auzou i in., 2000; Özsancak i in., 2001; Arnaut, Ávila, 2008; Collet i in., 2012). Wśród wielu prac badawczych z zastosowaniem parametru VOT tylko kilka dotyczy języka polskiego, a zaledwie trzy są źródłem opublikowanych danych na temat wartości VOT w sygnale mowy polskojęzycznych mówców. Polskie spółgłoski zwarto-wybuchowe lokują swoje wartości w dwóch kategoriach VOT. Dla głoskowych realizacji normatywnie dźwięcznych fonemów /b, d, g/ jest to dźwięczność wyprzedzająca (ang. *voicing lead*), a dla głoskowych realizacji normatywnie bezdźwięcznych fonemów /p, t, k/ jest to krótkie opóźnienie dźwięczności (ang. *short voicing lag*). Prawdopodobnie pierwszą polskojęzyczną pracą, w której zamieszczono wyniki pomiarów parametru VOT dla jednego dorosłego mówcy jest doniesienie A. Kopczyńskiego datowane na 1971 rok (Kopczyński, 1971)¹². Uzyskane dane autor przedstawił także w badaniach kontrastywnych języka polskiego i amerykańskiej odmiany języka angielskiego (Kopczyński, 1977, 72). W kolejnych pracach, opublikowanych w latach 1977–1981¹³ przez P. Keating i współautorów, omawiane są zagadnienia produkcji i percepcji VOT i stanowią one drugie w kolejności źródło danych o wielkościach parametru VOT w przypadku dorosłych mówców polskojęzycznych (Moslin, Keating, 1977; Mikoś i in., 1978; Keating, 1980; Keating i in., 1981).

Pomiaru parametru VOT w sygnale mowy polskojęzycznych dziecięcych mówców dokonała A. Trochymiuk (Trochymiuk, 2005; 2007; 2008). W pracy autorki, poświęconej analizie audytywnej i akustycznej wymowy dzieci niesłyszących korzystających z fonogestów, podjęto również zagadnienia związane z realizacją dźwięczności w nagłosowych spółgłoskach zwartych. Potrzeba rozwiązania problemu badawczego w grupie z zaburzoną wymową przyczyniła się do podjęcia tożsamyh badań w grupie kontrolnej. Uzyskano zatem dane o wartościach VOT także w sygnale mowy dziecięcych mówców z normatywną wymową w wieku 8–12 lat. W kontekście badań nad nauką drugiego języka i dydaktyką nauczania języków obcych należy także wymienić prace: W. Goneta (2001), A. Rojczyka (2009; 2010), E. Szalkowskiej-Kim (2010), a w kontekście bilingwizmu adnotację L. Newlin-Łukowicz (2010).

¹² Pomiary wykonano na Uniwersytecie Kalifornijskim w Los Angeles.

¹³ Dla /t/ i /d/ na podstawie nagrań 24 mówców, a dla wszystkich spółgłosek zwartych języka polskiego na podstawie nagrań sygnału mowy 5 dorosłych mówców.

BADANIA WŁASNE

W realizowanych badaniach własnych nad dyslalią desonoryzacyjną¹⁴ podjęto się rozwiązywania problemu badawczego dotyczącego jakości głoskowych realizacji opozycyjnych fonemów zwarto-wybuchowych, zwarto-trących i trących także z zastosowaniem metod fonetyki akustycznej. W przyjętym w badaniach 37-elemenowym systemie fonologicznym współczesnej polszczyzny autorstwa B. Ročławskiego zaburzenie to dotyczy fonemów dźwięcznych nienapiętych w 10 parach fonemów opozycyjnych (/b/ : /p/, /d/ : /t/, /g/ : /k/, /v/ : /f/, /z/ : /s/, /z/ : /ç/, /z/ : /ʃ/, /dz/ : /ts/, /dʒ/ : /tʃ/, /dʒ/ : /tʃ/ (Ročławski, 1986; 2005)¹⁵. W referowanych badaniach parametr VOT zastosowano celem oceny jakości realizacji kontrastu \pm dźwięczność w nagłosowych realizacjach dźwięcznych i bezdźwięcznych fonemów zwarto-wybuchowych – wargowych, dentalnych i welarnych – w sygnale mowy dzieci z dyslalią desonoryzacyjną i w sygnale mowy dzieci z prawidłową wymową (grupa porównawcza). Powołanie grupy porównawczej do celów analizy akustyczno-fonetycznej sygnału mowy było niezbędne, albowiem interpretacja sygnału mowy zaburzonej wymaga tożsamyh analiz sygnału mowy uzyskanego od mówców z prawidłową wymową. W niniejszym doniesieniu przedstawiono wyniki przeprowadzonych analiz akustycznych dotyczące pomiaru parametru VOT uzyskane w grupie porównawczej, która w tej pracy prezentowana jest jako główna grupa. Przedmiotem niniejszego doniesienia jest przedstawienie wyników badań, w których poszukiwano odpowiedzi na pytanie: Jakie wartości osiąga parametr VOT w nagłosowych realizacjach opozycyjnych fonemów zwarto-wybuchowych u osób z wymową prawidłową?

GRUPA BADAWCZA

Grupę badawczą stanowi 18 osób (9 płci żeńskiej i 9 płci męskiej) w wieku od 4,7 do 18,0 lat (średnia wieku 9,6 lat) zakwalifikowanych do uczestnictwa w sesjach nagraniowych spośród grupy ochotników, którzy wyrazili chęć wzięcia udziału w nagraniach. Wszyscy badani wychowują się w pełnych rodzinach, uczęszczają do powszechnych placówek oświatowych i mają powodzenia dydaktyczne. Nikt z omawianej grupy nie ma orzeczenia o niepełnosprawności, wadzie słuchu, nie był leczony neurologicznie i nie korzystał wcześniej z logopedycznej opieki, natomiast większość osób była poddawana logopedycznym badaniom

¹⁴ Dla wyodrębnionego w polskiej logopedii typu dyslalii określanego umownie „mową bezdźwięczną” przyjęto własną propozycję nazwy zaburzenia – dyslalia desonoryzacyjna (Konopska, Tarnowska, 2005).

¹⁵ Szczegółowe uzasadnienie B. Ročławskiego odnoszące się do wiązania dźwięczności z nienapiętością i bezdźwięczności z napiętością oraz stosowania połączenia obu cech w opisach fonologicznych zamieszczone są w pracach autora (Ročławski, 1986; 2005).

w trakcie badań przesiewowych wykonywanych w placówkach. Na podstawie przeprowadzonego z rodzicami wywiadu uzyskano dane o prawidłowym rozwoju fizycznym, psychicznym i prawidłowym rozwoju mowy. U wszystkich osób wykonano badanie logopedyczne, na podstawie którego wykluczono nieprawidłowości w zakresie wymowy, płynności mówienia, narządów artykulacyjnych, zgryzu i wędzidełka języka. Wszyscy badani w przednim odcinku zgryzu mieli pełne uzębienie i nie byli leczeni ortodontycznie. Udział wszystkich osób w sesjach nagraniowych był nieodpłatny.

MATERIAŁ LEKSYKALNY I SESJE NAGRANIOWE

W doborze materiału leksykalnego kierowano się głównie możliwością porównania głoskowych realizacji opozycyjnych fonemów obstruentalnych w takim samym samogłoskowym kontekście oraz potrzebą zarejestrowania u każdego z mówców kilku przykładowych realizacji poszczególnych fonemów. Kolejne uwarunkowania wiązały się z wymogiem ujednoczenia materiału wyrazowego dla wszystkich grup wiekowych i z założoną dla najmłodszych mówców przynajmniej bierną znajomością wyrazów, które planowano wykorzystać. Potrzeba ujednoczenia sposobu uzyskania od mówców materiału leksykalnego zdeterminowała wybór – uzyskiwanie izolowanych wyrazów za pomocą pokazywanych obrazków. Analizowany w badaniach akustycznych sygnał mowy zarejestrowano w trakcie sesji nagraniowych, które przeprowadzono w formie „gry językowej” z wykorzystaniem w teście nazywania serii obrazków („zabawa w nazywanie obrazków”) (Krajna, 2008, 6, 12). Zadaniem dziecka było dopowiedzenie brakującego słowa w zdaniu wypowiedzianym przez osobę prowadzącą badanie, np. „Pani ma parasolkę, bo deszcz bardzo mocno ... (*pada*)”. W przypadku trudności w uzyskaniu pożądanego wyrazu używano go w zdaniu, przechodzono do kolejnego wyrazu z listy, a następnie wracano do wcześniejszego zadania (Łobacz, 1996). Taki sposób uzyskiwania materiału leksykalnego określany też jest jako „kontrolowana mowa spontaniczna” (Kuijpers, 1991; za: Łobacz 1996, 99). Celem ułatwienia sytuacji zadaniowej i dyskrecjonalnego przymuszenia mówców do określonej reakcji słownej w obrazkowym teście nazywania wykorzystano pojedyncze, realne, kolorowe fotografie ludzi, zwierząt, przedmiotów, zjawisk i wykonywanych przez ludzi czynności, będące obrazowym przedstawieniem list wyrazowych.

W tabeli 1. zamieszczono listę wyrazów wykorzystanych do pomiaru parametru VOT w nagłosowych realizacjach fonemów zwarto-wybuchowych w zapisie ortograficznym i w transkrypcji szerokiej (ang. *broad transcription*) – fonematycznej z zastosowaniem symboli alfabetu IPA (Jassem, 2003; Trochymiuk, Świącicki, 2004). Transkrypcję fonematyczną materiału leksykalnego w tempie

lento, jednolitą z przedstawionymi w pracy W. Jassem (2003) dla języka polskiego symbolami IPA, przeprowadzono zgodnie z przyjętym w badaniach systemem fonemowym współczesnej polszczyzny B. Ročławskiego (2005).

Tab. 1. Lista wyrazowa obrazkowego testu nazywania z uwzględnieniem liczby przewidzianych do analizy akustycznej wystąpień realizacji bezdźwięcznych i dźwięcznych fonemów zwarto-wybuchowych w nagłosie wyrazów

Fonemy zwarto-wybuchowe			Zapis ortograficzny	Transkrypcja fonematyczna
liczba realizacji (51)	normatywnie bezdźwięczne (24)	normatywnie dźwięczne (27)		
/p/	8	–	pada, pasek, pije, piłka, pokój, półka, puszka, pudełko	/pada/, /pasek/, /pije/, /piwka/, /pokuj/, /puwka/, /pułka/, /pudewko/
/b/	–	11	bada, Basie, bazie, bije, bilet, bocian, bułka, buty, bucik, budy, budzik	/bada/, /bae/, /baze/, /bije/, /bilet/, /botean/, /buwka/, /buti/, /butaik/, /budi/, /budzik/
/t/	6	–	tata, talerz, Tomek, tory, tygrys, tyczka	/tata/, /taleʃ/, /tomek/, /tori/, /tigris/, /tiłka/
/d/	–	6	data, dach, domek, dom, dywan, dym	/data/, /dax/, /domek/, /dom/, /divan/, /dim/
/k/	10	–	kapa, kasa, kawa, kije, kino, kozy, kość, koziołek, kury, kubek	/kapa/, /kasa/, /kava/, /kije/, /kino/, /kozi/, /koete/, /kozowek/, /kuri/, /kubek/
/g/	–	10	gapa, gazeta, garaż, gitara, gips, gość, godzina, gotuje, guma, góry	/gapa/, /gazeta/, /garaʃ/, /gitara/, /gips/, /goete/, /godzina/, /gotuje/, /guma/, /guri/

Wszystkie sesje nagraniowe wykonano w studyjnym dźwiękoszczelnym pomieszczeniu Laboratorium Akustyki i Technologii Nagrań Dźwiękowych przy Katedrze Inżynierii Systemów, Sygnałów i Elektroniki Zachodniopomorskiego Uniwersytetu Technologicznego w Szczecinie. W każdej sesji brały udział dwie osoby: inżynier dźwięku odpowiedzialny za rejestrację sygnału mowy i nadzorującą nagrania od strony technicznej oraz osoba badająca przybywająca z dzieckiem w dźwiękoszczelnym pomieszczeniu studyjnym. Wszyscy uczestnicy nagrań zostali poinformowani o celu i sposobie ich przeprowadzenia, mieli możliwość zapoznania się z osobami realizującymi sesje nagraniowe, pomieszczeniem

studyjnym i sprzętem nagraniowym. Sesje nagraniowe przeprowadzono podczas jednorazowych spotkań, a czas trwania sesji nie przekraczał 20 minut.

REJESTRACJA SYGNAŁU MOWY

Do odbioru sygnału mowy zastosowano łącznie dwa różne przetworniki elektroakustyczne: szerokopasmowy mikrofon podstawowy AKG C-555L z ausznymi zaczepami (nagłowny) oraz dodatkowy podwójny mikrofon kontaktowy (krtaniowy, ang. *throat microphone*) LGF-24 firmy Navcomm, dokładnie rejestrujący dolny zakres częstotliwości sygnału, który umieszczano na szyi mówcy na wysokości krtani (powyżej chrząstki tarczowatej) (Flege, Brown, 1982; Acker-Mills i in., 2005)¹⁶. Sygnał mowy z obu przetworników zapisywano równoległe w dwóch kanałach z częstotliwością próbkowania 44,1 kHz z rozdzielczością 16 bitów na próbkę (standard CD-Audio) w profesjonalnym studyjnym rejestratorze cyfrowym Fostex FR-2 LE. Po zakończeniu nagrań archiwizowano cyfrowe pliki dźwiękowe na płytach CD-ROM w bezstratnym formacie WAV.

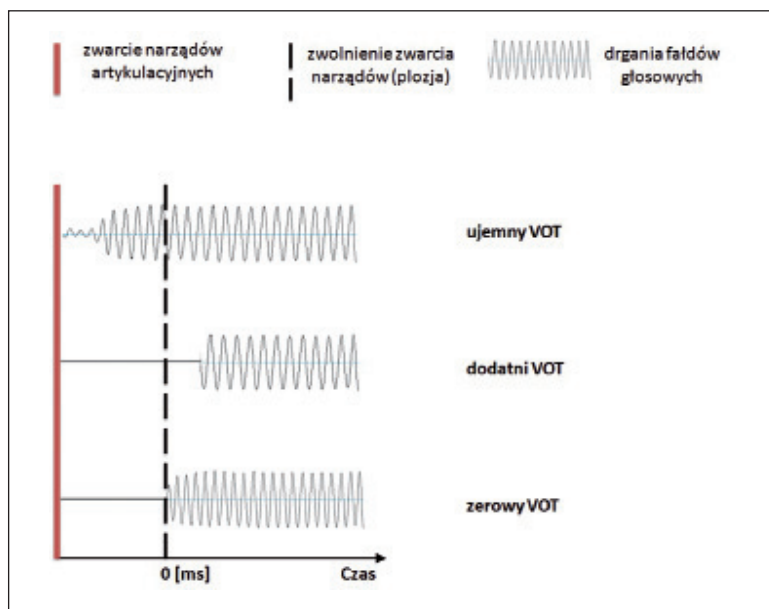
Zastosowanie w rejestracji sygnału mowy pierwszego przetwornika – mikrofonu nagłownego – zapewniło stałą odległość membrany mikrofonu od ust osoby mówiącej (w przybliżeniu 3–5 cm), co w przypadku mówców dziecięcych ma znaczenie, albowiem naturalna swoboda dziecięcych zachowań mogła w trakcie nagrań powodować niespodziewane i znaczne zmiany położenia mikrofonu względem ust, a co za tym idzie – zmiany wartości poziomu dźwięku i czasu propagacji sygnału akustycznego od ust do mikrofonu. Zastosowanie drugiego przetwornika podyktowane było tym, że mikrofon kontaktowy dokładnie rejestruje zakres niskich częstotliwości sygnału mowy, w tym sygnał tonu podstawowego nawet przy niskim poziomie amplitudy drgań krtaniowych, a zatem dostarcza pełnych danych o przebiegu czasowym pobudzenia krtaniowego, co w prowadzonych badaniach miało szczególne znaczenie. Czasowe przesunięcie sygnałów pochodzących od obu mikrofonów, wynikające z różnych czasów propagacji sygnału (mikrofony były umieszczone w różnych odległościach od źródła) dokładnie analizowano dla każdego mówcy w programie do edycji, analizy i syntezy mowy *Praat* (Boersma, Weenink, 2010). Dla każdego mówcy na podstawie 3–4 fragmentów nagrania ustalano średnie opóźnienie czasowe, a następnie synchronizowano sygnały z obu mikrofonów i powtórnie zapisywano w dwukanałowych plikach dźwiękowych. Niezależnie od pliku dźwiękowego ze zsynchronizowanymi sygnałami z obu mikrofonów, wydzielono dodatkowo plik dźwiękowy w formacie WAV z zapisem z mikrofonu podstawowego. Wszystkie uzyskane w badaniach wyniki pomiarów

¹⁶ Koncepcję i realizację technicznej strony rejestracji dźwięku opracował dr inż. Jerzy Sawicki z Katedry Inżynierii Systemów, Sygnałów i Elektroniki Zachodniopomorskiego Uniwersytetu Technologicznego w Szczecinie.

parametru VOT wyznaczono podczas ręcznej segmentacji sygnału mowy za pomocą programu *Praat* w wersji 5.2. (Boersma, Weenink, 2010).

Pomiar parametru VOT w realizacjach bezdźwięcznych i dźwięcznych fonemów zwarto-wybuchowych

Procedurę pomiaru parametru VOT przeprowadzono zgodnie z opisem zawartym w pracy L. Liskera i A. Abramsona (1964, 389). Wartość VOT ustalano przez wyznaczenie odcinka czasu pomiędzy zwolnieniem zwarcia realizowanej spółgłoski a początkiem wibracji fałdów głosowych. Za zerowy w wymiarze czasowym punkt odniesienia przyjęto moment realizacji płozi spółgłoski zwartej. Parametr VOT przyjmuje ujemne wartości liczbowe, jeżeli periodyczność pojawia się przed momentem zwolnienia zwarcia, dodatnie, gdy periodyczność pojawia się po zwolnieniu zwarcia, i wartość zerową, jeśli periodyczność pojawia się w momencie zwolnienia zwarcia narządów artykulacyjnych (ryc. 2.).



Ryc. 2. Układy czasowe trzech możliwych wariantów pomiaru parametru VOT¹⁷

Za początek realizacji płozi (zwolnienie zwarcia) bezdźwięcznych i dźwięcznych spółgłosek zwarto-wybuchowych przyjęto moment, w którym na oscylogramie uwidacznia się impuls płozi. Jeśli to było tylko możliwe¹⁸, pierwszy punkt

¹⁷ Opracowano na podstawie ryciny na stronie: <http://www.phon.ucl.ac.uk/>

¹⁸ Zapis oscylograficzny nie zawsze przechodzi przez zero w momencie zwolnienia płozi. Procedurę postępowania w takich przypadkach omówiono w dalszej części opisu wyznaczania parametru VOT.

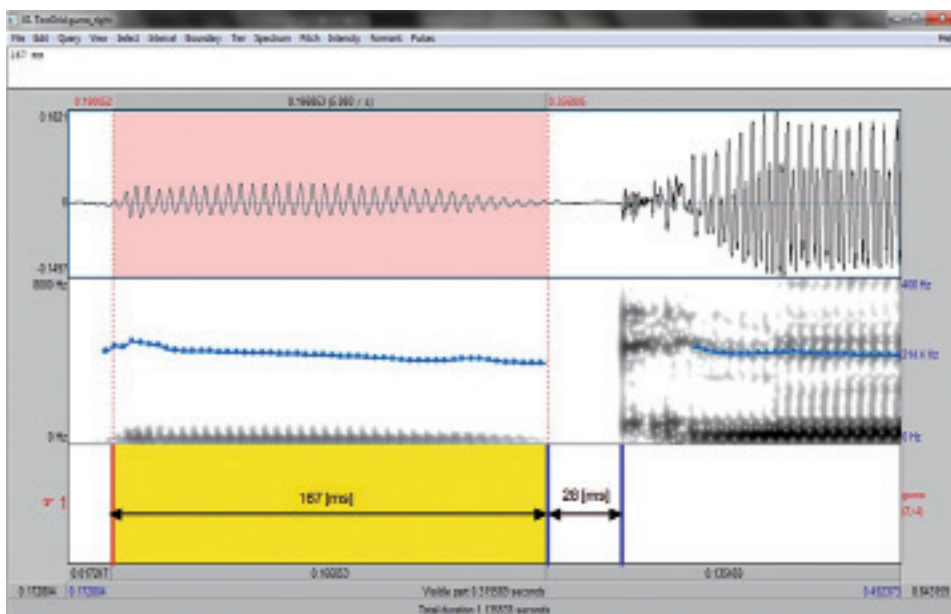
pomiarowy wyznaczano w momencie przejścia przez zero przebiegu akustycznego sygnału mowy o charakterze impulsu, gdy jego wychylenie ze stanu równowagi zmieniało się rosnąco w kierunku wartości dodatnich lub ujemnych. W przypadku dźwięczności wyprzedzającej, gdy impuls plozji nakładał się na periodyczność zapisu i występował np. na szczytach maksimum fali dźwiękowej, punkt graniczny wyznaczano w miejscu początku wystąpienia impulsu. W celu jak najdokładniejszego pomiaru przy słabo wyrażonych plozjach bezdźwięcznej spółgłoski [p] dostosowywano powiększenie graficzne oscylogramu oraz zmieniano zakres dynamiki spektrogramu dla lepszego zobrazowania szczegółów widma wybuchu. Zwolnienie zwarcia [p] wyraża się zazwyczaj znacznie mniejszą amplitudą i bardziej rozproszonym widmowo wybuchem niż w pozostałych spółgłoskach zwartych (Stevens, 1980, 190; za: Henton i in., 1992), co sprawia niekiedy trudności w wyznaczaniu wartości VOT¹⁹. W przypadku występowania w przebiegu plozji kilku powtarzających się impulsów punkt pomiarowy ustanawiano przy pierwszym impulsie (Davis, 1995; za: Özsancak i in., 2001; Trochymiuk, 2008). Drugi punkt pomiarowy parametru VOT wyznaczano na oscylogramie: a) w dźwięczności wyprzedzającej (ang. *voicing lead*) – w momencie przejścia przez zero w kierunku wartości dodatnich pierwszego okresu nieprzerwanej do momentu pojawienia się plozji serii periodycznych drgań fałdów głosowych; b) w krótkim opóźnieniu dźwięczności (ang. *short voicing lag*) – w momencie przejścia przez zero w kierunku wartości dodatnich pierwszego okresu serii periodycznych drgań fałdów głosowych w następującej po spółgłosce zwarto-wybuchowej samogłosce. Określenie położenia granicznych punktów w analizowanym przebiegu pozwalało wyznaczyć poszukiwany czas trwania odcinka pomiędzy punktami granicznymi i jego zwrot, i na tej podstawie odczytać wartość parametru VOT.

W wyznaczaniu VOT wiodącą rolę pełnił zapis oscylograficzny z mikrofonu podstawowego, natomiast trójwymiarowy, szerokopasmowy zapis spektrograficzny wspomagał analizę. W niejednoznacznych zapisach sygnału mowy, np. w przypadku opóźnienia w przejściu przez oś oscylogramu impulsu czy periodycznej fali dźwiękowej²⁰, wspierano się spektrogramem sygnału z mikrofonu podstawowego i/lub zapisem oscylograficznym z mikrofonu kontaktowego, na

¹⁹ W niektórych pracach badawczych zwraca się uwagę na niemożność wyznaczenia parametru VOT, np. z powodu braku widocznego wybuchu, co tłumaczy się niedostatecznym zwarciem narządów artykulacyjnych (najczęściej podczas realizacji /p/, rzadziej /k/, najrzadziej /t/). Z dostępnych danych wynika, że procent niemierzalnych próbek sygnału mowy przy wyznaczaniu VOT waha się od około 4% do 15%–20%. W badaniach francuskojęzycznych odsetek niemierzalnych próbek dla mówców z dyzartrią hypokinetyczną wynosił 20%, a dla mówców z grupy kontrolnej 6% (Özsancak i in., 2001; Auzou i in., 2000). O problemach w wyznaczaniu parametru VOT w nagłosowych realizacjach dźwięcznej spółgłoski [d] u polskojęzycznych mówców pisze P. Keating (1980, 36–38).

²⁰ Zob. np. w P. Keating (1980, 37).

którym czasowy zapis ciśnienia akustycznego niejednokrotnie przystępniej obrazował złożoność przebiegu nadkrtaniowo-krtaniowych pozycji i ruchów oddechowo-fonacyjno-artykulacyjnych. W referowanych badaniach do odrębnej analizy realizacji dźwięczności w nagłosowych realizacjach fonemów spółgłoskowych /b/, /d/, /g/ wyłączono wszystkie te realizacje, w których dźwięczność występowała przed zwolnieniem zwarcia spółgłoski, ale jeszcze przed wybuchem pojawiało się także segment ciszy²¹. W badaniach własnych dla tego typu realizacji przyjęto określenie „dźwięczne z segmentem bezdźwięcznym przed płożą”. Opisany sposób realizacji dźwięczności, niespełniający kryterium pomiaru parametru VOT, obrazuje rycina 3.



Ryc. 3. Wyznaczone granice i czasy trwania przebiegu *+periodyczności* (167 milisekund) i *-periodyczności* (28 milisekund) w realizacji fonemu /g/ w nagłosie wyrazu *guma* wypowiedzianego przez mówcę płci męskiej w wieku 10 lat i odebranego jako /guma/

Liczba wszystkich nagłosowych realizacji spółgłoskowych fonemów zwarto-wybuchowych potencjalnie możliwych do analizy akustycznej dla 18 osób wynosiła 918. Na podstawie wykonanych odsłuchów, a także już w trakcie segmentacji sygnału mowy z analizy akustycznej wyłączono 29 realizacji, co stanowi 3,16%

²¹ P. Łobacz określa ten typ realizacji dźwięczności jako „wymowę częściowo bezdźwięczną” – „(...) początkowa część zwarcia jest dźwięczna, po czym następuje przynajmniej 30 ms segment ciszy przed płożą i aspiracją” (1996, 181). W podobny sposób ten typ realizacji dźwięczności spółgłosek zwarto-wybuchowych opisuje A. Trochymiuk (2008, 151–152).

wykluczeń. Najliczniejszą grupę wykluczeń stanowią realizacje dźwięczne z segmentem bezdźwięcznym przed płożą (72,4%), a następnie realizacje, w których po spółgłosce wybuchowej segment samogłoskowy zrealizowano bezdźwięcznie (13,8%). W pozostałych przypadkach (substytucja spółgłoski 3,45%, pomiar niemożliwy 6,90%, dźwięk protetyczny 6,90%) rozkład wykluczonych z analizy akustycznej realizacji fonemów zwarto-wybuchowych jest zbliżony. Łącznie wykonano analizę akustyczną 889 realizacji fonemów zwarto-wybuchowych w nagłosie wyrazów, w tym 424 głoskowych realizacji fonemów bezdźwięcznych /p/, /t/, /k/ oraz 465 głoskowych realizacji fonemów dźwięcznych /b/, /d/, /g/.

METODY STATYSTYCZNE

Spśród miar opisu statystycznego w pracy zastosowano: średnią arytmetyczną, do stopnia rozproszenia (zmienności) wyników – miarę stopnia zmienności (wartości minimalne i maksymalne oraz odchylenie standardowe), błąd standardowy (SE) i 95% przedział ufności (CI). We wnioskowaniu statystycznym zastosowano test t-Studenta, analizę wariancji. Do analizy stosowano test Chi-kwadrat Pearsona, dokładny test Fishera, korelację rang Spearmana. Za wyniki istotne statystycznie we wszystkich badanych testach przyjęto wyniki, dla których $p \leq 0,05$ (w tabelach pogrubiony druk), za wyniki na poziomie tendencji statystycznej przyjęto wyniki, dla których prawdopodobieństwo p zawiera się w przedziale $p > 0,05$ i $p \leq 0,10$ (w tabelach kursywa), a za nieistotne statystycznie, dla których $p > 0,10$ (Brzeziński, 1996; Ferguson, Takane, 2002).

Wyniki pomiarów parametru VOT w realizacjach bezdźwięcznych fonemów zwarto-wybuchowych /p/, /t/, /k/

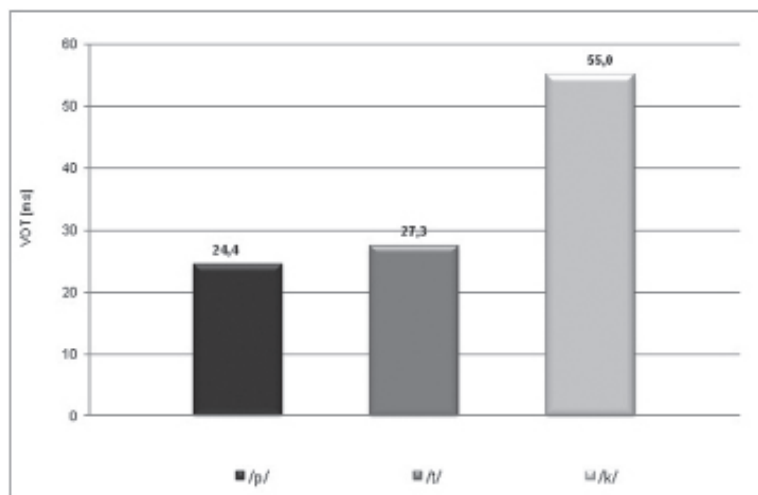
Wszystkie wyznaczone wartości parametru VOT w realizacjach bezdźwięcznych fonemów zwarto-wybuchowych są dodatnie i zawierają się w przedziale od 1 do 165 milisekund, w tym dla fonemu /p/ w przedziale od 1 [ms] do 67 [ms] – przy średniej wartości VOT 24,4 [ms], dla fonemu /t/ w przedziale od 9 [ms] do 79 [ms] – przy średniej wartości VOT 27,3 [ms], a dla fonemu /k/ w przedziale od 11 [ms] do 165 [ms] – przy średniej wartości VOT 55,0 [ms] (tab. 2.).

Wartość VOT jest istotnie wyższa w głoskowych realizacjach fonemu /k/ niż w głoskowych realizacjach fonemów /p/ i /t/ ($p=0,00001$). Natomiast różnica średnich wartości VOT w realizacjach fonemów /p/ i /t/ jest na granicy istotności statystycznej ($p=0,07087$). Z analizy uzyskanych w badaniach wartości parametru VOT w realizacjach bezdźwięcznych fonemów zwarto-wybuchowych /p/, /t/, /k/ wynika, że w omawianej grupie z prawidłową wymową:

– średnie wartości VOT są dodatnie i lokują się dla języka polskiego wzdłuż kontinuum VOT w krótkim opóźnieniu dźwięczności (ang. *short voicing lag*),

Tab. 2. Wyniki pomiarów parametru VOT [ms] w realizacjach bezdźwięcznych fonemów /p/, /t/, /k/ z uwzględnieniem (\pm SD), (SE), (CI), wartości granicznych oraz analizy statystycznej

Wartość VOT w realizacjach fonemów /p/, /t/, /k/								
Fonem	Liczba realizacji (424)	Średnia arytmetyczna [ms]	Odchylenie standardowe (\pm SD)	Błąd standardowy (SE)	95% przedziały ufności (CI)		Wartości graniczne [ms]	
					(-95% CI)	(+95% CI)	min.	maks.
/p/	137	24,4	13,2	1,1	22,2	26,6	1,0	67,0
/t/	108	27,3	11,7	1,1	25,1	29,6	9,0	79,0
/k/	179	55,0	23,4	1,7	51,6	58,5	11,0	165,0
Analiza statystyczna								
Porównywane			/p/ : /t/		/p/ : /k/		/t/ : /k/	
test t-Studenta			0,07087		0,00001		0,00001	
ANOVA (rozproszenie wyników)			0,18394		0,00001		0,00001	



Ryc. 4. Średnie wartości parametru VOT dla normalywie bezdźwięcznych realizacji fonemów zwarto-wybuchowych /p/, /t/, /k/

– średnie wartości parametru VOT realizacji spółgłoski welarnej są istotnie wyższe niż średnie wartości parametru VOT spółgłosek zwarto-wybuchowych z przednim miejscem artykulacji (labialnej, dentalnej),

– wartość parametru VOT zmienia się wraz z postępującą w kierunku przednio-tylnym zmianą miejsca artykulacji i jest to tendencja narastająca od spółgło-

ski labialnej, poprzez spółgłoskę dentalną, do spółgłoski welarnej ([p] → [t] → [k]) (ryc. 4.).

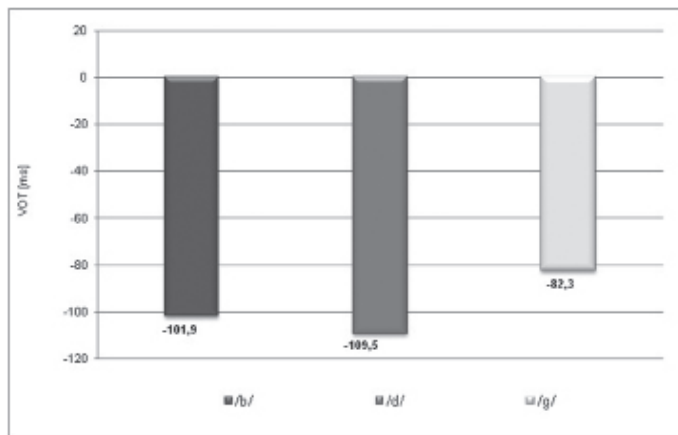
Wyniki pomiarów parametru VOT w realizacjach dźwięcznych fonemów zwarto-wybuchowych /b/, /d/, /g/

Zarejestrowane w sygnale mowy realizacje dźwięcznych fonemów spółgłoskowych /b/, /d/, /g/ mają taki sam prawostronny kontekst samogłoskowy jak omówione realizacje bezdźwięcznych fonemów spółgłoskowych /p/, /t/, /k/. Uzyskane wyniki średnich wartości parametru VOT w nagłosowych realizacjach dźwięcznych zwarto-wybuchowych fonemów spółgłoskowych zamieszczono w tabeli 3.

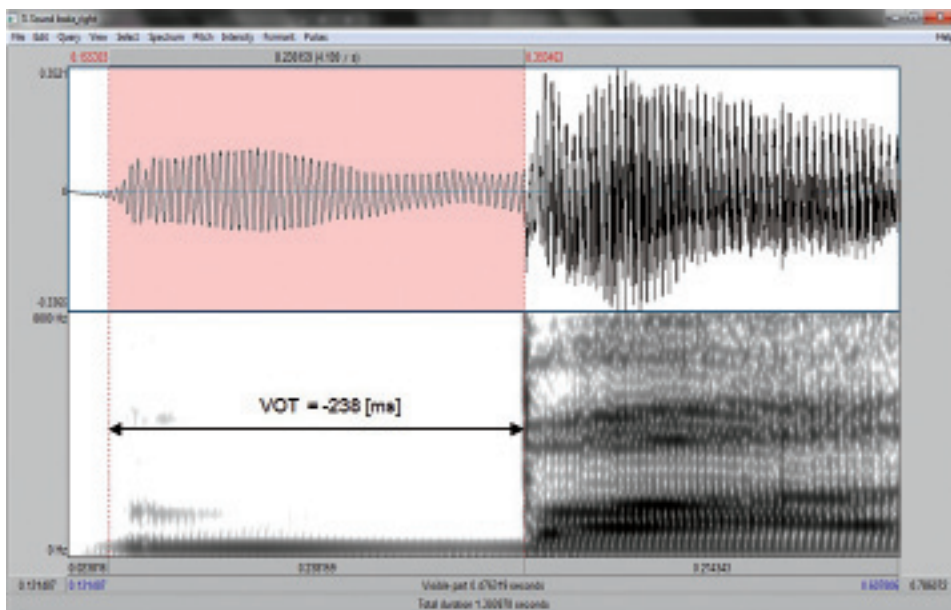
W badanej grupie graniczne wartości minimalne i maksymalne parametru VOT przybierają wartości dodatnie i ujemne, zawierając się w przedziale od 8 [ms] do (-238) [ms]. Średnie wartości parametru VOT są ujemne i wynoszą: (-101,9) [ms] dla głoskowych realizacji fonemu /b/, przy wartościach granicznych zawierających się w przedziale od 17 [ms] do (-238) [ms]; (-109,5) [ms] dla głoskowych realizacji fonemu /d/, przy wartościach granicznych zawierających się w przedziale od 8 [ms] do (-210) [ms]; (-82,3) [ms] dla głoskowych realizacji fonemu /g/, przy wartościach minimalnych i maksymalnych zawierających się w przedziale od 38 [ms] do (-192) [ms]. Średnie ujemne wartości VOT dla /g/ są istotnie niższe niż uzyskane średnie ujemne wartości VOT dla /b/ i /d/ ($p = 0,00001$). Natomiast różnica w średnich wartościach VOT w realizacjach /b/ i /d/ jest na granicy istotności statystycznej ($p = 0,07347$) (tab. 3., ryc. 5.).

Tab. 3. Wyniki pomiarów parametru VOT w realizacjach dźwięcznych fonemów /b/, /d/, /g/ z uwzględnieniem (\pm SD), (SE), (CI), wartości granicznych i analizy statystycznej

Wartości VOT w realizacjach fonemów /b/, /d/, /g/								
Fonem	Liczba realizacji (465)	Średnia arytmetyczna [ms]	Odchylenie standardowe (\pm SD)	Błąd standardowy (SE)	95% przedziały ufności (CI)		Wartości graniczne [ms]	
					(-95% CI)	(+95% CI)	min.	maks.
/b/	196	-101,9	34,7	2,5	-97,0	-106,8	17,0	-238,0
/d/	103	-109,5	34,7	3,4	-102,7	-116,3	8,0	-210,0
/g/	166	-82,3	46,2	3,6	-75,2	-89,4	38,0	-192,0
Analiza statystyczna								
Porównywane			/b/ : /d/		/b/ : /g/		/d/ : /g/	
test t-Studenta			0,07347		0,00001		0,00001	
ANOVA (rozproszenie wyników)			0,97101		0,00013		0,00198	



Ryc. 5. Średnie wartości parametru VOT w realizacjach normatywnie dźwięcznych fonemów zwarto-wybuchowych



Ryc. 6. Oscylogram i spektrogram realizacji fonemu spółgłoskowego /b/ w wyrazie *bada* wypowiedzianego przez siedmioletniego mówcę płci męskiej

Rycina 6. obrazuje wyjątkowo dużą wartość VOT wynoszącą (-238) [ms] w realizacji fonemu /b/ w wyrazie *bada* wypowiedzianym przez siedmioletniego chłopca.

Z analizy uzyskanych średnich wartości parametru VOT w realizacjach normatywnie dźwięcznych fonemów zwarto-wybuchowych /b/, /d/, /g/ w grupie z prawidłową wymową wynika, że:

- wartości minimalne i maksymalne parametru VOT są dodatnie i ujemne,
- średnie wartości parametru VOT dla głoskowych realizacji normatywnie dźwięcznych fonemów zwarto-wybuchowych /b/, /d/, /g/ są ujemne i lokują się wzdłuż kontinuum VOT w dźwięczności wyprzedzającej (ang. *voicing lead*),
- średnia, bezwzględna wartość parametru VOT zmniejsza się wraz z postępującą w kierunku przednio-tylnym zmianą miejsca artykulacji: od spółgłoski dentalnej, poprzez spółgłoskę labialną, do spółgłoski welarnej ([d] → [b] → [g]).

DYSKUSJA

Zbiorcze zestawienie średnich wartości VOT uzyskanych w dotychczasowych badaniach w realizacjach polskich opozycyjnych zwarto-wybuchowych fonemów obstruentalnych w sygnale mowy dorosłych i dziecięcych polskojęzycznych mówców z wymową prawidłową zamieszczono w tabeli 4.

Z analizy zebranych danych wynika, że wartości parametru VOT dla głoskowych realizacji bezdźwięcznych fonemów zwarto-wybuchowych otrzymane w badaniach własnych są zbliżone do wyników uzyskanych przez A. Trochymiuk oraz P. Keating i współautorów (Keating i in., 1981; Trochymiuk, 2008). Wartości VOT podane przez A. Kopczyńskiego dla /p/ i /t/ nieznacznie odbiegają od pozostałych danych, co zapewne wynika z faktu, że analizowany sygnał mowy pochodzi od jednego mówcy (Kopczyński, 1971). Średnie wartości parametru VOT dla bezdźwięcznej spółgłoski labialnej zawierają się w przedziale od 21,5 [ms] do 37,5 [ms], dla dentalnej w przedziale od 27,3 [ms] do 33 [ms], dla welarnej od 49 [ms] do 55 [ms]. Natomiast średnie wartości parametru VOT dla głoskowych realizacji polskich bezdźwięcznych fonemów zwarto-wybuchowych /p, t, k/ w badaniach przeprowadzonych przez P. Keating i współautorów, A. Trochymiuk i w badaniach własnych zawierają się w przedziale 34,1 [ms]–35,6 [ms] (tab. 4.).

W badaniach przeprowadzonych dla polskojęzycznych mówców dziecięcych w wieku 8–12 lat wartości minimalne i maksymalne parametru VOT w realizacjach fonemów dźwięcznych /b, d, g/ przybierają, podobnie jak w badaniach własnych, wartości dodatnie i ujemne, natomiast średnie wartości VOT w realizacjach /b/ wynoszą (-88,9) [ms], w realizacjach /d/ (-107,2) [ms], w realizacjach /g/ (-76,2) [ms] (Trochymiuk, 2008, 117–125). Uzyskane wyniki pomiaru VOT w badaniach własnych są zbliżone do pomiarów uzyskanych przez autorkę w przypadku spółgłoski welarnej i dentalnej, różnią się natomiast średnimi uzyskanymi w odniesieniu do spółgłoski labialnej. Różnice te należy objaśniać sąsiedztwem samogło-

Tab. 4. Analiza porównawcza uzyskanych w dotychczasowych badaniach średnich wartości parametru VOT w realizacjach polskich fonemów zwarto-wybuchowych: bezdźwięcznych /p/, /t/, /k/ i dźwięcznych /b/, /d/, /g/

Średnie wartości parametru VOT [ms] w realizacjach polskich fonemów zwarto-wybuchowych								
Autor, rok, grupa badawcza	A. Kopczyński (1971) 1 dorosła osoba (wiek nieznan)		P. Keating i in. (1981) 5 dorosłych osób (osoby studiujące)		A. Trochymiuk (2008) 10 dzieci w wieku 8–12 lat		L. Konopska, J. Sawicki (2013) 18 osób w wieku 5–18 lat	
Fonemy	Dźwięczne	Bezdźwięczne	Dźwięczne	Bezdźwięczne	Dźwięczne	Bezdźwięczne	Dźwięczne	Bezdźwięczne
/p, b/	-78	37,5	-88,2	21,5	-88,9	21,6	-101,9	24,4
/t, d/	-72	33	-89,9	27,9	-107,2	28,8	-109,5	27,3
/k, g/	-61	49	-66,1	52,7	-76,2	54,2	-82,3	55,0
Średni VOT	-70,3	39,8	-81,4	34,1	-90,8	34,9	-97,9	35,6

Źródło: Kopczyński A., 1971, *Degree of voicing in initial stops in educated Polish and American English*, „Studia Anglica Posnaniensia”, 3, s. 75–79; Keating, P. A., Mikoś, M. J., Gannon W. F., 1981, *A cross-language study of range of voice onset time in the perception of initial stop voicing*, „Journal of the Acoustical Society of America”, 70, s. 1261–1271; Trochymiuk A., 2008, *Wymowa dzieci niesłyszących. Analiza audytywna i akustyczna*, Wyd. UMCS, Lublin; Konopska L.: badania własne.

skowym. W badaniach własnych parametr VOT w realizacjach /b/ analizowano w kontekście czterech samogłosek [a, o, u, i] (niskiej, średniej i dwóch wysokich), natomiast w przytoczonej pracy analizowano VOT w kontekście jednej niskiej ([a]) i jednej wysokiej ([u]) samogłoski (Trochymiuk, 2008, 117–121). Ponadto, w badaniach własnych uzyskano wyższe wartości VOT w realizacjach /b/ przed samogłoską [a]²² niż w badaniach A. Trochymiuk, co mogło także wpłynąć na stwierdzone różnice.

Uzyskane w prezentowanych badaniach wyniki pomiaru parametru VOT pozwalają na obliczenie i porównanie uogólnionych średnich wartości VOT. Dla spółgłosek dźwięcznych średnie te zawierają się w przedziale od (-70,3) [ms] do (-97,9) [ms] lub – z pominięciem danych od jednego mówcy – od (-81,4) [ms] do (-97,9) [ms], a dla spółgłosek bezdźwięcznych w przedziale od 34,1 [ms] do 39,8 [ms] lub – z pominięciem danych od jednego mówcy – od 34,1 [ms] do 35,6 [ms]. Przedstawione dane z analiz fonetyczno-akustycznych sygnału mowy dorosłych

²² Analiza parametru VOT w kontekście następującej po spółgłosce zwarto-wybuchowej samogłoski jest przedmiotem odrębnego opracowania (w przygotowaniu).

i dziecięcych mówców potwierdzają, że w języku polskim bezdźwięczne spółgłoski zwarte lokują swoje wartości w krótkim opóźnieniu dźwięczności, a dźwięczne spółgłoski zwarte w dźwięczności wyprzedzającej.

W trzech z przedstawionych czterech niezależnie wykonanych badań parametru VOT u polskojęzycznych użytkowników języka w realizacjach bezdźwięcznych spółgłosek zwarto-wybuchowych VOT wzrasta wraz z postępującą w kierunku przednio-tylnym zmianą miejsca artykulacji: od spółgłoski labialnej, poprzez spółgłoskę dentalną, do spółgłoski welarnej ([p] → [t] → [k]). W realizacjach dźwięcznych spółgłosek zwarto-wybuchowych VOT zmniejsza się wraz z postępującą w kierunku przednio-tylnym zmianą miejsca artykulacji: od spółgłoski dentalnej, poprzez spółgłoskę labialną, do spółgłoski welarnej ([d] → [b] → [g]). Dane te w całej rozciągłości potwierdzają przytoczone uniwersalia na temat VOT, że im bardziej tylne zwanie spółgłoski oraz im bardziej obszerny kontakt zwartych narządów artykulacyjnych, tym większa wartość VOT, a także – im szybszy ruch artykulatorów, tym mniejsza wartość VOT. Różnice dotyczące czasowego aspektu parametru VOT w obrębie spółgłosek zwartych tłumaczone są zasadami wynikającymi z praw aerodynamiki (różnice ciśnień związane z objętością jam rezonacyjnych za i przed miejscem zwania artykulatorów), prędkością ruchów artykulacyjnych poszczególnych artykulatorów i wielkością powierzchni zwania podczas kontaktu artykulatorów, a także zmiennym stopniem otwarcia szpary głośni (Cho, Ladefoged, 1999; Ohala, 1997). Współcześnie, różnice w prędkościach ruchu przedniej (w przypadku [t]), środkowej oraz tylnej (w przypadku [k]) części języka, oprócz odmiennej pod względem wielkości masy mięśniowej i zróżnicowanego udziału poszczególnych części języka, znajdują swoje wyjaśnienie w badaniach miologicznych. Prace badawcze nad występowaniem poszczególnych typów włókien mięśniowych (szybkokurczliwe, wolnokurczliwe i włókna pośrednie między szybko- i wolnokurczliwymi) w obrębie mięśni wewnętrznych języka (mięsień pionowy, mięsień poprzeczny, mięsień podłużny) wykazują, że w mięśniach wewnętrznych języka dominują głównie włókna szybkokurczliwe (60%), ale ich rozmieszczenie jest zróżnicowane. W przedniej części języka przeważają mniejsze włókna szybkokurczliwe – 71%, w tylnej części języka większe włókna wolnokurczliwe i włókna pośrednie – 66%. Zdaniem autorów, taki skład włókien prawdopodobnie odzwierciedla zarówno genotypową, jak i fenotypową specjalizację funkcji jamy ustnej. Przewaga włókien typu II (szybkokurczliwe) pozwala na szybkie i elastyczne działania w pozycjonowaniu i kształtowaniu ruchów języka podczas żucia, połykania, oddychania i mówienia (Stål i in., 2003).

Przeprowadzone badania uzupełniają dotychczasową wiedzę na temat wartości parametru VOT u mówców polskojęzycznych²³. Jednakże dokładne poznanie

²³ Współautorka opracowania składa podziękowania dr. inż. Jerzemu Sawickiemu z Katedry Inżynierii Systemów, Sygnałów i Elektroniki Zachodniopomorskiego Uniwersytetu Technologicz-

procesu doskonalenia realizacji dźwięczności w dziecięcych produkcjach wymaga podjęcia badań przekrojowych na większej liczbiej populacji. Kształtowanie się poszczególnych typów realizacji dźwięczności w spółgłoskach zwarto-wybuchowych, a zwłaszcza w wymowie dzieci przedszkolnych, powinno się stać przedmiotem odrębnych badań fonetyczno-akustycznych z zastosowaniem parametru VOT.

BIBLIOGRAFIA

- Abramson A., 1977, *Laryngeal timing in consonant distinctions*, „Phonetica”, 34, s. 295–303.
- Abramson A., Lisker L., 1965, *Voice onset time in stop consonants. Acoustic analysis and synthesis*, [w:] *Proceedings of the 5th International Congress of Acoustics*, Liège, A51.
- Acker-Mills B., Houtsma A., Ahroon W., 2005, *Speech Intelligibility with Acoustic and Contact Microphones*, [in:] *New Directions for Improving Audio Effectiveness* (pp. 7–1 – 7–14). Meeting Proceedings RTO-MP-HFM-123, Paper 7. Neuilly-sur-Seine, France: RTO. (Pobrano 19 września 2010 r. ze strony: <http://ftp.rta.nato.int/Public/PubFullText/RTO/MP/RTO-MP-HFM-123/MP-HFM-123-07.pdf>).
- Arnaut M., Ávila C., 2008, *Devoicing of plosive phonemes in dysphonic children's speech*, „Rev. Soc. Bras. Fonoaudiol.”, 13 (1), s. 37–44.
- Auzou P., Özsancak C., Morris R., Jan M., Eustache F., Hannequin D., 2000, *Voice onset time in aphasia, apraxia of speech and dysarthria. A review*, „Clinical Linguistics and Phonetics”, 14, s. 131–150.
- Boersma P., Weenink D., 2010, *Praat: doing phonetics by computer (Version 5.2)* [Computer program]. Pobrano 30 października 2010 r. ze strony: <http://www.praat.org/>.
- Brzeziński J., 1996, *Metodologia badań psychologicznych*, PWN, Warszawa.
- Caramazza A., Yeni-Komshian G., Zurif E., Carbone E., 1973, *The acquisition of a new phonological contrast. The case of stop consonants in French-English bilinguals*, „Journal of the Acoustical Society of America”, 54, s. 421.
- Cho T., Ladefoged P., 1999, *Variations and universals in VOT. Evidence from 18 languages*, „Journal of Phonetics”, 27, s. 207–222.
- Collet G., Colin C., Serniclaes W., Hoonhorst I., Markessis E., Deltenre P., Leybaert J., 2012, *Effect of phonological training in French children with SLI: perspectives on voicing identification, discrimination and categorical perception*, „Research in Developmental Disabilities”, 33(6), s. 1805–1818.
- Davis K., 1995, *Phonetic and phonological contrasts in the acquisition of voicing. Voice onset time production in Hindi and English*, „Journal of Child Language”, 22, s. 275–305.
- Docherty G., 1992, *The Timing of Voicing in British English Obstruents*, s. 9.
- Ferguson G., Takane Y., 2002, *Analiza statystyczna w psychologii i pedagogice*, PWN, Warszawa.
- Flege J., Brown W., 1982, *The voicing contrast between English /p/ and /b/ as a function of stress and position-in-utterance*, „Journal of Phonetics”, 10, s. 335–345.
- Gonet W., 2001, *Obstruent Voicing in English and Polish. A Pedagogical Perspective*, „International Journal of English Studies”, 1 (1), s. 73–92.

nego w Szczecinie za wprowadzenie w złożone zagadnienia teorii sygnałów w aspekcie fonetyki akustycznej, a dr Anicie Lorenz (wcześniej Trochymiuk) z Zakładu Logopedii i Językoznawstwa Stosowanego UMCS w Lublinie za pomoc w rozstrzygnięciu niejednoznacznych przypadków realizacji w zarejestrowanym sygnale mowy.

- Henton C., Ladefoged P., Maddieson I., 1992, *Stops in the World's Languages*, „Phonetica”, s. 49.
- Hoonhorst I., Colin C., Markessis E., Radeau M., Deltenre P., Serniclaes W., 2009, *French native speakers in the making. From language-general to language-specific voicing boundaries*, „Journal of Experimental Child Psychology”, 104(4), s. 353–366.
- Itoh M., Sasanuma S., Tatsumi I., Murakami S., Fukusako Y., Suzuki T., 1982, *Voice onset time characteristics in apraxia of speech*, „Brain and Language”, 17(2), s. 193–210.
- Jassem W., 2003, *Illustrations of the IPA. Polish*, „Journal of the International Phonetic Association”, 33 (1), s. 103–107.
- Keating P., 1980, *A Phonetic Study of a Voicing Contrast in Polish*. Unpublished Ph.D. dissertation, Brown University. (Pobrano 29 września 2010 r. ze strony: <http://www.linguistics.ucla.edu/people/keating/thesis1-55.pdf>)
- Keating P., 1984, *Phonetic and phonological representation of stop consonant voicing*, „Language”, 60, s. 286–319.
- Keating P., Mikoś M., Ganong W., 1981, *A cross-language study of range of voice onset time in the perception of initial stop voicing*, „Journal of the Acoustical Society of America”, 70, s. 1261–1271.
- Keating P., Westbury J., Stevens K., 1980, *Mechanisms of stop-consonant release for different places of articulation*, „Journal Acoust. Soc. Am.”, 67, s. 93. Pobrano: 21 września 2010 r. ze strony: <http://www.linguistics.ucla.edu/people/keating/KeatingWestburyStevensASA1980.pdf>
- Kehoe M., Lleó C., Rakow M., 2004, *Voice onset time in bilingual German-Spanish children*, „Bilingualism. Language and Cognition”, 7(01), s. 71–88.
- Konopska L., 2013, *Iloczas głoskowych realizacji bezdźwięcznych fonemów trących /f/, /s/, /ɸ/ i /ʃ/*, [w:] *Język. Człowiek. Społeczeństwo*, red. J. Panasiuk, T. Woźniak, Wyd. UMCS, Lublin, s. 441–477.
- Konopska L., Tarnowska Cz., 2005, *Dyslalia desonoryzacyjno-fonacyjna, czyli o tzw. mowie bezdźwięcznej – inaczej*, [w:] *Logopedia. Teoria i praktyka*, red. M. Młynarska, T. Smereka, Ag. Wyd. A linea, Wrocław, s. 307–315.
- Kopczyński A., 1971, *Degree of voicing in initial stops in educated Polish and American English*. „Studia Anglica Posnaniensia”, 3, s. 75–79.
- Kopczyński A., 1977, *Polish and American English Consonant Phonemes. A Contrastive Study*. PWN, Warszawa.
- Krajna E., 2008, *100-wyrazowy test artykulacyjny. Założenia teoretyczne. Podręcznik*, Wyd. Komlogo, Gliwice.
- Ladefoged P., Maddieson I., 2008, *The Sounds of the World's Languages*, Blackwell Publishing.
- Lehiste I., 1977, *Suprasegmentals*, MIT Press, Cambridge–Massachusetts–London.
- Lisker L., Abramson A., 1964, *A cross language study of voicing in initial stops. Acoustic measurements*, „Word”, 20, s. 384–422.
- Lisker L., Abramson A., 1967, *Some effects of context on voice onset time in English stops*, „Language and Speech”, 10, s. 1–28.
- Łobacz P., 1996, *Polska fonologia dziecięca. Studia fonetyczno-akustyczne*, Wyd. Energeia, Warszawa.
- Macken M., Barton D., 1979, *The acquisition of the voicing contrast in Spanish. A phonetic and phonological study of word-initial stop consonants*, „Journal of Child Language”, 7(03), s. 433–458.
- Macken M., Barton D., 1980, *A longitudinal study of the acquisition of the voicing contrast in American-English word-initial stops, as measured by voice onset time*, „Journal of Child Lang.”, 7, s. 41–74.
- Mikoś M., Keating P., Moslin B., 1978, *The perception of voice onset time in Polish*, „Journal of the Acoustical Society of America” (S1), 63, S19.

- Mortensen J., Tøndering J., 2013, *The effect of vowel height on Voice Onset Time in stop consonants in CV sequences in spontaneous Danish. Proceedings of Fonetik 2013 The XXVIth Annual Phonetics Meeting 12–13 June 2013*, „Sweden Studies in Language and Culture”, Linköping University, Linköping, s. 21.
- Moslin B., Keating P., 1977, *Voicing distinction in Polish word-initial stop consonants*, „J. Acoust. Soc. Am.”, 62, S1, S27–S27.
- Newlin-Lukowicz L., 2010, *L1 phonetic interference in the VOTs produced by Polish-English bilinguals*. Dostęp na stronie: http://wa.amu.edu.pl/plm/2013/files/Abstracts/Newlin-Lukowicz_Luiza_PLM2013b.pdf
- Ohala J., 1997, *Aerodynamics of phonology. Proceedings of the Seoul International Conference on Linguistics*, Linguistic Society of Korea, Seoul, s. 92–97.
- Özsancak C., Auzou P., Jan M., Hannequin D., 2001, *Measurement of voice onset time in dysarthric patients. Methodological considerations*, „Folia Phoniatrica at Logopédica”, 53, s. 48–57.
- Ringen C., Kulikov V., 2012, *Voicing in Russian stops. Cross-linguistic implications*, „Journal of Slavic Linguistics”.
- Rocławski B., 1986, *Zarys fonologii, fonetyki, fonotaktyki i fonostatystyki współczesnego języka polskiego*, Wyd. Uniw. Gdańskiego, Gdańsk.
- Rocławski B., 2005, *Podstawy wiedzy o języku polskim dla glottodydaktyków, pedagogów, psychologów i logopedów*, Glottispol, Gdańsk.
- Rojczyk A., 2009, *Parametr VOT w języku polskim i angielskim. Badanie percepcji*, „LingVaria” 4, s. 29–47.
- Rojczyk A., 2010, *Temporal and spectral parameters in perception of the voicing contrast in English and Polish*, Wyd. Uniw. Śląskiego, Katowice.
- Rottenberg M., 2009, *Voice onset time vs. articulatory modeling for stop consonants*, „Logopedics Phoniatrics Vocology”, 34, s. 171–180.
- Stål P., Marklund S., Thornell L., De Paul R., Eriksson P., 2003, *Fibre composition of human intrinsic tongue muscles*, „Cells Tissues Organs”, 173, s. 147–61.
- Szalkowska-Kim E., 2010, *Charakterystyka fonetyczna kontrastu homoorganicznych spółgłosek zwartych w językach: polskim i koreańskim*, „Postscriptum Polonistyczne”, (2 (6), s. 31–57.
- Trochymiuk A., 2005, *Wymowa bezdźwięczna w świetle analizy akustycznej*, [w:] *Logopedia. Teoria i praktyka*, red. M. Młynarska, T. Smereka, Ag. Wyd. A linea, Wrocław, s. 499–519.
- Trochymiuk A., 2007, *Wykorzystanie akustycznego parametru VOT w analizie mowy bezdźwięcznej*, [w:] *Język. Interakcja. Zaburzenia mowy. Metodologia badań*, t. 2, red. T. Woźniak, A. Domańska, Wyd. UMCS, Lublin, s. 212–240.
- Trochymiuk A., 2008, *Wymowa dzieci niesłyszących. Analiza audytywna i akustyczna*, Wyd. UMCS, Lublin.
- Trochymiuk A., Świącicki R., 2004, *Symbole podstawowej transkrypcji Międzynarodowego Towarzystwa Fonetycznego (IPA) i jej rozszerzenia (ExtIPA)*, „Audiofonologia”, XXV, s. 97–114.

Źródła internetowe

http://en.wikipedia.org/wiki/Voice_onset_time

<http://www.phon.ucl.ac.uk/>