

ANITA LORENC*, DANIEL KRÓL**,
ANDRZEJ SENDERSKI***, RENATA BOROWIECKA****

*Uniwersytet Warszawski, Wydział Polonistyki, Instytut Polonistyki Stosowanej,

**Akademia Nauk Stosowanych w Tarnowie, Katedra Informatyki,

***Oddział Otolaryngologii Dziecięcej i Pediatrii UCK WUM, APD Medical sp. z o.o,

****Centrum Edukacji, Diagnostyki i Terapii Psychologiczno-Pedagogicznej w Warszawie,
APD Medical sp. z o.o.

ORCID ID: <https://orcid.org/0000-0002-7614-0881>, <https://orcid.org/0000-0002-8611-0838>,
<https://orcid.org/0000-0003-0078-176X>, <https://orcid.org/0000-0001-5322-9632>

Percepcja \pm dźwięczności a akustyczny parametr VOT w terapii Neuroflow ATS

Perception of \pm voicing and Acoustic VOT Parameter in Neuroflow ATS Therapy

STRESZCZENIE

W artykule zaprezentowano nowy zestaw ćwiczeń w aktywnym treningu słuchowym Neuroflow przeznaczony dla osób z zaburzeniami słuchu fonemowego oraz z zaburzeniami przetwarzania słuchowego o podtypie fonologicznym. Są to ćwiczenia dyskryminacji i identyfikacji słuchowej cechy \pm dźwięczności w obrębie polskich spółgłosek zwarto-wybuchowych z wykorzystaniem modyfikacji akustycznego parametru VOT. Modyfikację długości VOT zrealizowano, opierając się na technice przepróbkowania sygnału cyfrowego.

Słowa kluczowe: \pm dźwięczność, VOT (czas rozpoczęcia dźwięczności), dyskryminacja i identyfikacja słuchowa, terapia Neuroflow ATS

SUMMARY

This paper presents a new set of exercises in Neuroflow active auditory training designed for people with phonemic hearing disorders and auditory processing disorders of the phonological subtype. These are exercises for the discrimination and identification of the auditory feature of \pm voicing within Polish plosive consonants using modification of the acoustic parameter VOT. The modification of VOT length was realised using a digital signal resampling technique.

Key words: ±voicing, VOT (Voice Onset Time), auditory discrimination and identification, Neuroflow ATS therapy

WPROWADZENIE – CHARAKTERYSTYKA DŹWIĘCZNOŚCI

Efekt dźwięczności należy rozumieć jako rezultat całego kompleksu mechanizmów aerodynamicznych i mięśniowych, podlegających kontroli ośrodkowego układu nerwowego, kierowanych przez korę ruchową przy współdziałaniu czuciowego i słuchowego mechanizmu sprzężenia zwrotnego. Aktywność krtaniowa musi być skoordynowana z poszczególnymi ruchami narządów nasady. Z fizjologicznego punktu widzenia podstawą elementów dźwięcznych są powtarzające się w krótkich odstępach czasu fazy zamknięcia i otwarcia głośni. Drgania te są konsekwencją dwóch przeciwstawnych mechanizmów: wzrastające ciśnienie podgłośniowe dąży do rozsunęcia zamkniętych fałdów głosowych, zaś skurcz mięśni krtani uniemożliwia przemieszczenie zgromadzonego powietrza. W efekcie różnicy ciśnień wprowadzone w wibracje fałdy głosowe stopniowo uwalniają nagromadzone powietrze podgłośniowe. Naprzemienne fazy zagęszczenia i rozrzedzenia powietrza przemieszczającego się przez wibrujące fałdy głosowe tworzą ton krtaniowy.

Zmienne aerodynamiczne są ściśle powiązane z kompleksami ruchów w obrębie krtani i jam rezonatorów. Eksperymentalne badania aktywności krtaniowej (Sawashima, Hirose 1983; Hirose 1997) pozwoliły wyróżnić cztery podstawowe typy ruchów. Dwa z nich są szczególnie istotne w procesie wytwarzania dźwięków mowy:

1. ruchy rozwierająco-zwierające – regulują stopień otwarcia fałdów głosowych w zależności od charakterystyki głoski i typu fonacji (dźwięczna, bezdźwięczna, szeptana i ich kombinacje);
2. ruchy napinająco-rozluźniające – regulują długość, powierzchnię i napięcie fałdów głosowych, w zależności od zmian częstotliwości podstawowej F_0 .

Dodatkowe typy ruchów to:

3. wznosząco-opadające – dzięki nim krtani porusza się w górę i w dół;
4. obkurczające górną część krtani – pełnią ważną rolę w realizacji zwarcia krtaniowego.

Efekt dźwięczności to nie tylko rezultat aktywności krtaniowej. Różnice pomiędzy spółgłoskami bezdźwięcznymi i dźwięcznymi tworzone są również w obszarze ponadkrtaniowym, co najlepiej oddaje fonetyczne rozróżnienie na spółgłoski mocne/napięte (bezdźwięczne) i słabe/nienapięte (dźwięczne) odnoszące się do siły artykulacji. A zatem w przypadku spółgłosek bezdźwięcz-

nych obserwuje się bardziej intensywny i obszerny kontakt artykulatorów niż w przypadku spółgłosek dźwięcznych (zob. Benni 1931), silniejsze zwarcie podniebienia miękkiego z tylną ścianą jamy gardłowej oraz większe wzniesienie nagłośni, kości gnykowej i krtani (zob. Koneczna, Zawadowski 1951). Bezdźwięczność wiąże się również z dłuższym iloczasem artykulacji (zob. Trochymiuk [Lorenc] 2008).

W ujęciu fonetyki artykulacyjnej fonacja definiowana jest jako użycie krtani z pomocą prądu powietrza wytwarzanego dzięki układowi oddechowemu w celu wygenerowania słyszalnego źródła energii akustycznej, która dzięki zmianom artykulacyjnym może być modyfikowana w pozostałych częściach toru głosowego (Laver 1994, 184; por. też: Laver 1980).

P. Ladefoged (1971) sugeruje, iż typy fonacji mogą stanowić kontinuum uzależnione od wielkości szczeliny pomiędzy fałdami głosowymi i chrząstkami nalewkowymi, gdzie największe otwarcie związane jest z fonacją bezdźwięczną, dalej znalazły się: fonacja oddechowa, fonacja modalna (prawidłowa), fonacja skrzecząca, aż po zwarcie krtaniowe charakteryzujące się całkowitym zamknięciem fałdów głosowych i chrząstek nalewkowych.

Z kolei J. Laver (1994, 184–201) dzieli typy fonacji na:

1. Bezdźwięczną (ang. *voicelessness*):
 - a) oddechową (ang. *breath phonation*) – fałdy głosowe nie drgają, są odwiedzione, a przepływające pomiędzy nimi z dużą prędkością powietrze ulega turbulencjom, które dają słuchowe wrażenie głosu chuchającego;
 - b) zerową (ang. *nil phonation*) – fałdy głosowe są odwiedzione, a powietrze przepływa pomiędzy nimi w niezakłócony sposób; odmianą fonacji zerowej jest zwarcie krtaniowe, podczas którego dochodzi do zablokowania przepływu powietrza z płuc przez zwarte fałdy głosowe.
2. Dźwięczną (ang. *voicing*):
 - a) modalną (ang. *modal voice*) – szpara głośni jest zamknięta na całej swej długości, a fałdy głosowe drgają w sposób quasi-periodyczny (prawie regularny); jest to fonacja neutralna;
 - b) skrzypiącą (ang. *creak*) – fałdy głosowe są słabo napięte, drgają wolniej i nieregularnie, w wyniku czego powstaje dźwięk o niskiej częstotliwości;
 - c) falset (ang. *falsetto*) – szpara głośni jest zamknięta, fałdy głosowe drgają z częstotliwością wyższą niż w przypadku fonacji modalnej, co daje słuchowe wrażenie głosu wysokiego;
3. Szept (ang. *whisper*).

Wszystkie pozostałe typy fonacji są ich kombinacjami, reprezentując tym samym fonacje złożone (np. głos skrzekliwy, ang. *creaky voice*, będący kombina-

cją fonacji modalnej i skrzeku, czy też głos aspirowany, ang. *breathy voice*, będący połączeniem fonacji modalnej i oddechowej itd.).

Z akustycznego punktu widzenia artykulacjom dźwięcznym towarzyszy obecność tzw. tonu krtaniowego (por. Jassem 1973, Szczepankowski 1985; Pickett 1999). Ton krtaniowy to dźwięk złożony, o przebiegu zbliżonym do regularnego wielotonu harmonicznego, wytwarzany jest przez drgające fałdy głosowe. Składa się z tonu podstawowego o częstotliwości drgań bezpośrednich fałdów głosowych oraz tonów składowych o częstotliwościach, które są wielokrotnościami tonu podstawowego¹. Ton podstawowy (częstotliwość podstawowa, formant zerowy) to składowa częstotliwość tonu krtaniowego, dźwięku złożonego powstającego w krtani, odpowiadająca bezpośrednio częstotliwości drgań fałdów głosowych. W akustyce bywa określana również mianem formantu zerowego i oznaczana symbolem F_0 . Mierzona jest w cyklach na sekundę i wyrażana w hercach (Hz). Ton podstawowy przyjmuje różne wartości w zależności od rodzaju głosu. Jego średnia wartość podczas konwersacji wynosi 120 Hz dla mężczyzn, 220 Hz dla kobiet i około 330 Hz dla dzieci w wieku 10 lat (dane dla języków europejskich na podstawie: Laver 1994). Maksymalna rozpiętość częstotliwości podstawowej podczas przeciętnej rozmowy waha się w przedziale 50–250 Hz w przypadku głosów męskich i 120–480 Hz u kobiet. W mowie zmiany wysokości tonu podstawowego sięgają zazwyczaj jednej oktawy, natomiast w śpiewie (zwłaszcza w przypadku głosów szkolonych) jego rozpiętość jest znacznie większa i obejmuje dwie, a nawet trzy i więcej oktaw (np. przeciętny zakres szkolonego sopranu koloraturowego wynosi 245–1480 Hz).

Percepcyjnym odpowiednikiem zmian tonu podstawowego w dłuższej wypowiedzi jest subiektywne odczucie zmian wysokości głosu (melodii mowy, intonacji), a w przypadku pojedynczych segmentów obecność F_0 związana jest ze słuchowym wrażeniem dźwięczności.

AKUSTYCZNY PARAMETR VOT

Kontrast dotyczący dźwięczności, różnicujący na przykład klasę spółgłosek zwarto-wybuchowych, jest dosyć powszechny w językach świata i został zakodowany w alfabetycznych i sylabicznych skryptach od tysiącleci. Akustyczne realizacje tego kontrastu nie były jednak możliwe do zaobserwowania aż do czasów współczesnych. Prawie 60 lat temu Leigh Lisker i Arthur S. Abramson (1964)² na

¹ Ton krtaniowy stanowi jedynie bazę dla powstających dźwięków mowy, których ostateczny kształt zależy od wielu modyfikacji zachodzących w strukturach ponadkrtaniowych (jamach gardłowej, ustnej i nosowej). Niektóre częstotliwości składowe tonu krtaniowego ulegają wzmocnieniu w jamach nasady, co związane jest z powstaniem tzw. formantów dźwięków mowy.

² Warto podkreślić, iż redaktorzy „Journal of Phonetics” w 50-lecie badań nad akustycznym

podstawie analizy akustycznych danych pochodzących z 11 języków zaproponowali, iż obserwacja czasu pulsów krtaniowych w relacji do czasu artykulacji ponadkrtaniowej wyjaśniałaby większość homoorganicznych³ rozróżnień spółgłoskowych, o których zazwyczaj mówi się, że zależą od dźwięczności, aspiracji⁴, czy intensywności⁵. Analizowane zależności opisali za pomocą akustycznego parametru VOT (ang. *voice onset time* – czas rozpoczęcia dźwięczności). Zgodnie z pierwszą definicją, zaproponowaną w artykule założycielskim (Lisker, Abramson 1964), jest to odstęp czasowy, w jakim dźwięczność (wibracje fałdów głosowych) pojawia się w relacji do momentu zwolnienia zwarcia ustnego, czyli wybuchu spółgłoski zwartej. Stopień dźwięczności spółgłosek zwarto-wybuchowych możemy określić zatem jako iloczyn drgań periodycznych, które mogą wyprzedzać zwolnienie zwarcia lub następować po nim.

Badacze wyróżnili trzy sposoby realizacji dźwięczności:

1. *Short voicing lag* (krótkie opóźnienie dźwięczności) – drgania periodyczne pojawiają się tuż po zwolnieniu zwarcia (płozji) spółgłoski zwarto-wybuchowej (bezdźwięczne nieaspirowane), czemu odpowiadają dodatnie wartości VOT od 0 do +25 milisekund (dla języka polskiego zob. rycina 4).
2. *Long voicing lag* (długie opóźnienie dźwięczności) – dźwięczne drgania periodyczne występują po zwolnieniu zwarcia (płozji) spółgłoski i następującym po nim segmencie aspiracji (bezdźwięczne aspirowane), czemu odpowiadają dodatnie wartości VOT, mieszczące się w przedziale od +60 do +100 milisekund.
3. *Voicing lead* (dźwięczność wyprzedzająca) – dźwięczne drgania periodyczne pojawiają się już w segmencie odpowiadającym zwarcia narządów mowy (dźwięczne nieaspirowane spółgłoski zwarto-wybuchowe), czemu odpowiadają ujemne wartości VOT, mieszczące się w zakresie od –125 do –75 ms (dla języka polskiego zob. rycina 5).

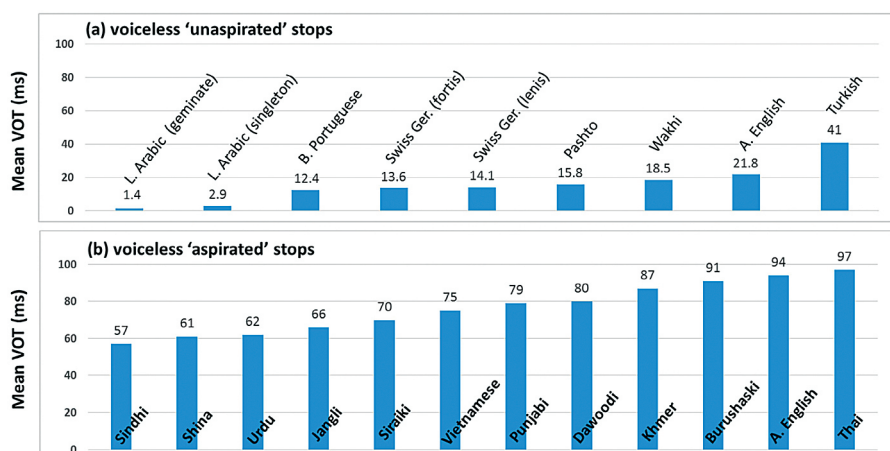
VOT poświęcili temu zagadnieniu całe specjalne wydanie czasopisma (Special Issue: Marking 50 Years of Research on Voice Onset Time po redakcją T. Cho, G. Docherty i D.H. Whalen).

³ W fonetyce spółgłoski homoorganiczne (homo – „pierwszy człon wyrazów złożonych mający znaczenie: taki sam, jednakowy, podobny” (SJP) i *organ* w znaczeniu „narząd mowy”) to takie, które są artykułowane w tym samym miejscu artykulacji. Przykładem spółgłosek homoorganicznych w języku polskim są [p b m], ponieważ wszystkie przyjmują dwuwargowe miejsce artykulacji.

⁴ Aspiracja powstaje po zwolnieniu zwarcia spółgłoski zwarto-wybuchowej. Posiada cechy bezdźwięcznej samogłoski, lecz krótszy iloczyn. Bywa określana jako tzw. przydech. Angielskie spółgłoski [p t k] są aspirowane na początku sylaby takich słów jak: *pea, tea, key* (zob. np. Jassem 1992).

⁵ Danych popierających to twierdzenie dostarczyły wczesne studia poświęcone zachowaniu krtani w realizacjach spółgłosek zwarto-wybuchowych (Lisker et al. 1969; Sawashima et al. 1970), akustyce (Lisker, Abramson 1964) i percepcji (Abramson, Lisker 1965; Lisker, Abramson 1970).

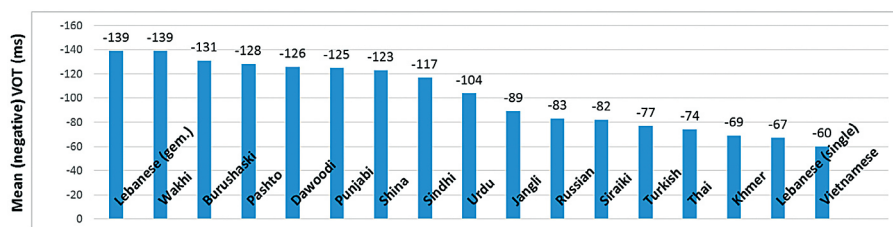
Rycina 1 poniżej przedstawia średnie pomiary VOT (zębowo-) dźwiękowych spółgłosek zwarto-wybuchowych w wybranych językach, w których występuje (a) krótkie opóźnienie dźwięczności (*short voicing lag* – na rysunku voiceless ‘unaspirated’ stops) oraz (b) długie opóźnienie dźwięczności (*long voicing lag* – na rysunku voiceless ‘aspirated’ stops).



Rycina 1. Średnie wartości VOT (zębowo-) dźwiękowych spółgłosek zwarto-wybuchowych w wybranych językach w kategorii: a) bezdźwięcznych nieaspirowanych (ang. *voiceless unaspirated stops*), b) bezdźwięcznych aspirowanych (ang. *voiceless aspirated stops*)

Źródło: Cho et al. (2019, 55).

Rycina 2 poniżej przedstawia średnie ujemne wartości VOT (zębowo-) dźwiękowych spółgłosek zwarto-wybuchowych. Dane opracowano dla 17 języków.

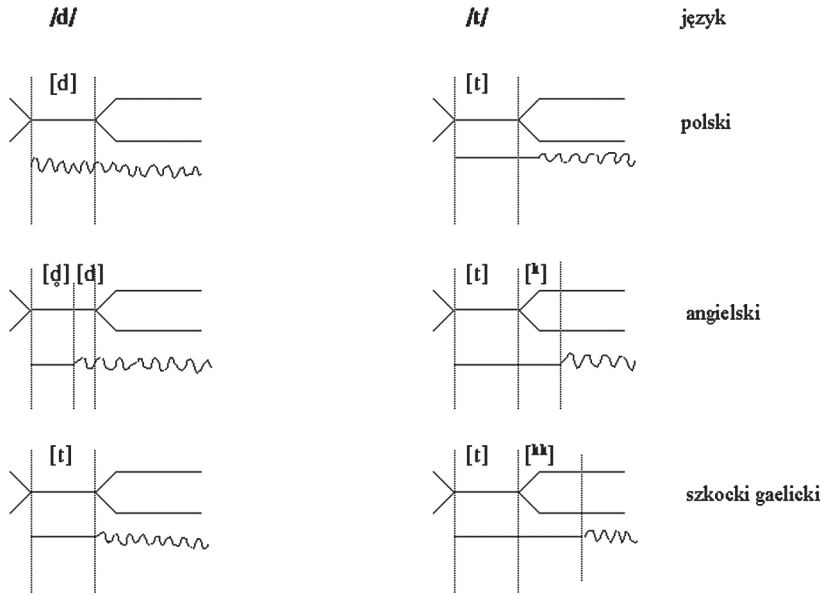


Rycina 2. Średnie ujemne wartości VOT (zębowo-) dźwiękowych spółgłosek zwarto-wybuchowych w wybranych językach

Źródło: Cho et al. (2019, 56).

Najczęściej w rozmaitych językach świata *continuum* VOT dzielone jest na dwie kategorie. W niektórych, tak jak w polskim, hiszpańskim czy węgierskim, kontrast fonemiczny \pm dźwięczna reprezentowany jest przez spółgłoski dźwięczne

o ujemnym VOT (*voicing lead*) i spółgłoski bezdźwięczne nieaspirowane o krótkim dodatnim VOT (*short voicing*). Ten sam kontrast fonemiczny ± dźwięczna w języku angielskim realizowany jest przez spółgłoski częściowo dźwięczne o krótkim ujemnym lub zerowym VOT (*voicing lead*) i bezdźwięczne aspirowane o długim dodatnim VOT (*long voicing*), zob. rycina 3 poniżej.

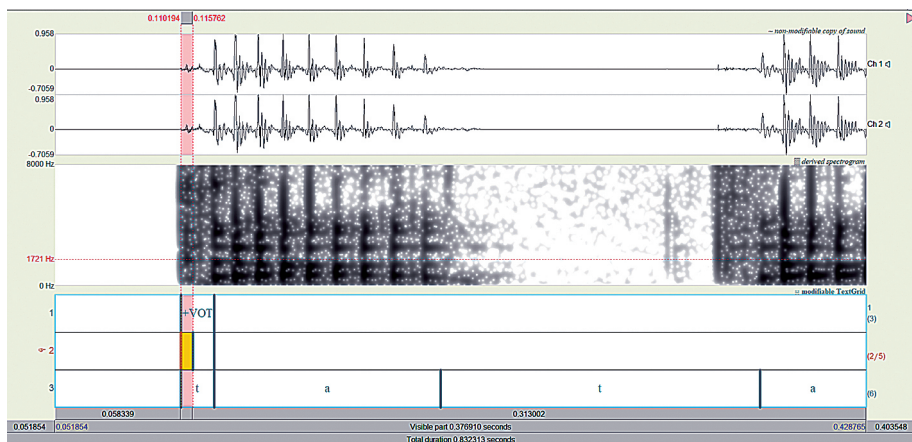


Rycina 3. VOT /t/ i /d/ w językach: polskim, angielskim i szkockim gaelickim

Źródło: Rogers 2000, s. 243.

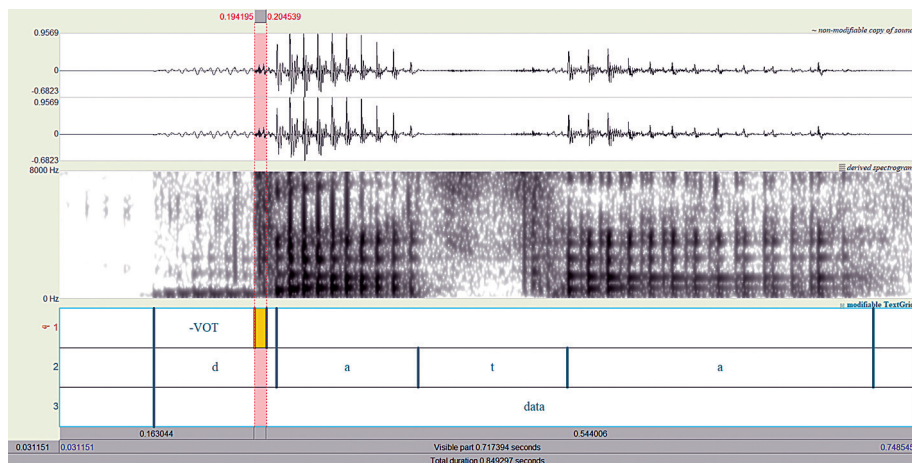
Część języków świata wykorzystuje trzystopniowy kontrast VOT, tak jak koreański, tajski lub hindi, a w przypadku języków indoaryjskich można mówić o kontraście więcej niż trzystopniowym (jangli i urdu z kontrastem czterostopniowym oraz sindhi i siraiki z kontrastem pięciostopniowym), dla uzyskania którego wykorzystywana jest nie tylko zmienna koordynacja dźwięczności, ale też kontrola kierunku przepływu powietrza, jak na przykład implozywność w języku sindhi. Więcej na ten temat można przeczytać w artykule Cho et al. (2019). Autorzy poddają w nim analizie 11 badań dotyczących kontrastu dźwięczności w 19 językach, opierając swoje rozważania na obszernym korpusie danych pochodzących od 270 użytkowników tychże języków.

Poniżej zamieszczono przykłady (zob. rycina 4, rycina 5) realizacji polskich zębowych spółgłosek zwarto-wybuchowych reprezentujących dwa typy realizacji dźwięczności.

Rycina 4. Oscylogramy i spektrogram wyrazu *tata*

Źródło: opracowanie własne z wykorzystaniem programu Praat (Boersma, Weenink 2014).

Na drugim poziomie anotacji wydzielono segment odpowiadający wybuchowi spółgłoski [t] (zaznaczony na kolor różowy) i następujący po nim krótki segment szumowy. Za punkt odniesienia w procedurze pomiarowej przyjęto początek realizacji ploszji spółgłoski zwartej, nadając jej czas zerowy. Jeżeli drgania periodyczne, charakterystyczne dla pracy fałdów głosowych, pojawiają się po wybuchu (tak jak na rycinie 4) parametr VOT przyjmuje wartości dodatnie. W analizowanym przykładzie VOT wyniósł +15 ms, a przykład ilustruje krótkie opóźnienie dźwięczności.

Rycina 5. Oscylogramy i spektrogram wyrazu *data*

Źródło: opracowanie własne z wykorzystaniem programu Praat (Boersma, Weenink 2014).

Podobnie jak w poprzednim przykładzie za punkt odniesienia w procedurze pomiarowej VOT przyjęto początek realizacji płozi (zaznaczonej na kolor różowy) spółgłoski zwarto-wybuchowej [d], nadając jej czas zerowy. Drgania periodyczne tym razem poprzedzają wybuch, pojawiając się już w segmencie zwarcia spółgłoski [d], w związku z czym parametr VOT przyjmuje wartości ujemne. W analizowanym przykładzie VOT wyniósł -84 ms, a przykład ilustruje zjawisko dźwięczności wyprzedzającej.

Zamieszczone w poniższej tabeli (zob. tabela 1) średnie wartości VOT bezdźwięcznych spółgłosek zwarto-wybuchowych [p t k] nie tylko obrazują ich zmienność w różnych językach, ale również jedno z podstawowych uniwersaliów dotyczących tego parametru. Mianowicie już w pierwszym studium poświęconym VOT (Lisker, Abramson 1964) udowodniono, że jego poziom zależy od miejsca artykulacji i przyjmuje najwyższe wartości dla spółgłosek welarnych, najniższe dla labialnych, a średnie dla koronalnych. Zatem wraz z przesuwaniem się artykulacji ponadkraniowej do tyłu wzrasta długość VOT.

Tabela 1. Średnie wartości VOT [ms] w realizacjach fonemów /p/, /t/, /k/ w wybranych językach

Autor opracowania	język	/p/	/t/	/k/
Lisker, Abramson (1964)	węgierski	2	16	29
Lisker, Abramson (1964)	holenderski	10	15	25
Lisker, Abramson (1964)	koreański	18	25	47
Kollia (1993)	grecki	19	27	49
Keating i in. (1981)	polski	21	28	52
Raphael, Tobin (1983)	hebrajski	26	34	64

Źródło: opracowanie własne.

Kolejne badania (zob. np. Cho, Ladefoged 1999) potwierdziły opisaną powyżej zależność, pozwoliły też ustalić szereg innych uniwersaliów na temat VOT, a wśród nich między innymi:

- im większa jest płaszczyzna kontaktu pomiędzy artykulatorami, tym dłuższy jest VOT,
- im szybszy ruch artykulatorów, tym krótszy VOT,
- VOT jest zależny od następującej samogłoski – jego iloczyn jest dłuższy przed wysoką samogłoską, krótszy przed niską,

- VOT jest dłuższy dla spółgłosek w sylabie akcentowanej,
- VOT jest dłuższy w izolowanych słowach, w zdaniach ulega skróceniu.

W tabeli 2 poniżej zestawiono średnie wartości VOT polskich bezdźwięcznych [p t k] oraz dźwięcznych [b d g] spółgłosek zwarto-wybuchowych w ujęciu różnych autorów. Zaprezentowane wyniki odnoszą się do wymowy pięciu dorosłych osób (Keating et al. 1981), 10 dzieci w wieku 8–12 lat (Trochymiuk [Lorenc] 2008)⁶ oraz 18 osób w wieku od 5 do 18 lat (Konopska, Sawicki 2013). Zaprezentowane wyniki prowadzonych niezależnie badań ilustrują szereg omówionych wcześniej uniwersaliów dotyczących VOT oraz potwierdzają dwustopniowy kontrast fonemiczny \pm dźwięczności w języku polskim, reprezentowany przez zwarto-wybuchowe spółgłoski bezdźwięczne nieaspirowane o krótkim dodatnim VOT (*short voicing*) i zwarto-wybuchowe spółgłoski dźwięczne o ujemnym VOT (*voicing lead*).

Tabela 2. Średnie wartości (zaokrąglone do pełnych liczb) parametru VOT [ms] w realizacjach polskich spółgłosek zwarto-wybuchowych w ujęciu różnych autorów.

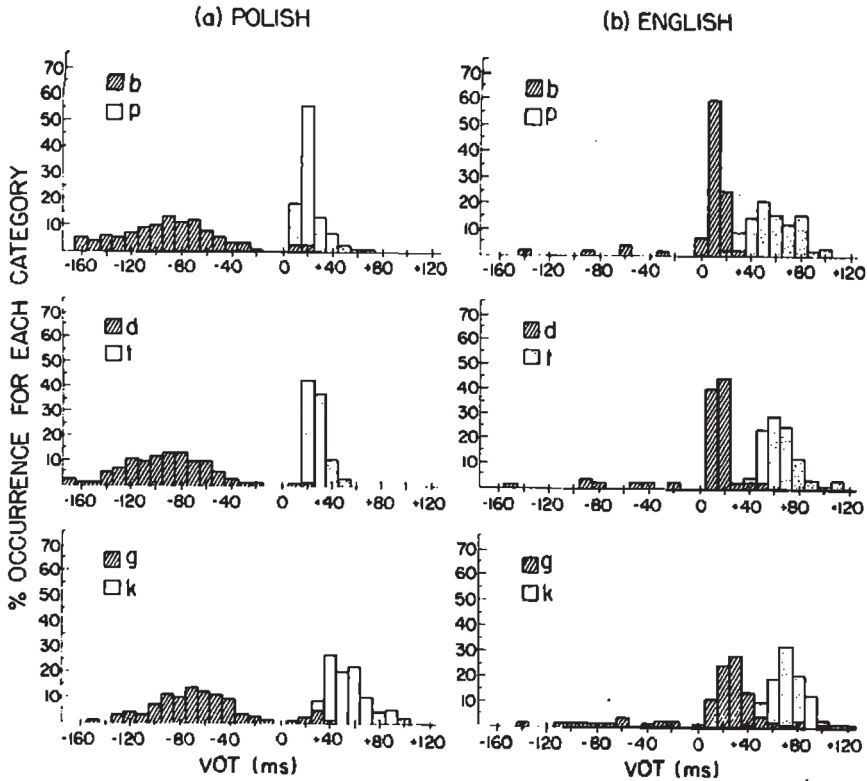
Spółgłoska	Keating et al. (1981)	Trochymiuk [Lorenc] (2008)	Konopska i Sawicki (2013)
	5 dorosłych osób	10 dzieci (8–12 lat)	18 osób (5–18 lat)
p	21	22	24
t	28	29	27
k	52	54	55
b	–88	–89	–102
d	–90	–107	–109
g	–66	–76	–82

Źródło: opracowanie własne.

Dane dotyczące realizacji spółgłosek zwarto-wybuchowych w języku polskim i odpowiadających im wartości parametru VOT wyraźnie wskazują, że kategoria \pm dźwięczności jest wyraźnie rozdzielona na dwie subkategorie tworzące

⁶ W monografii A. Trochymiuk [Lorenc] (2008) przedstawione są również wyniki pomiarów VOT w realizacjach 20 dzieci w wieku 8–12 lat z głębokimi uszkodzeniami słuchu. Podobne badania, dotyczące wymowy zaburzonej, przeprowadziła L. Konopska (2016). Wzięło w nich udział 30 osób w wieku od 5 do 18 lat z diagnozą wymowy bezdźwięcznej (tzw. dyslalią desonorizacyjną w ujęciu autorki).

kontrast dwustopniowy (zob. rycina 6 i rycina 7) w obrębie wartości ujemnych dla spółgłosek dźwięcznych i dodatnich dla spółgłosek bezdźwięcznych.



Rycina 6. Rozkład wartości VOT dla sześciu nagłosowych spółgłosek zwarto-wybuchowych w języku a) polskim, b) angielskim

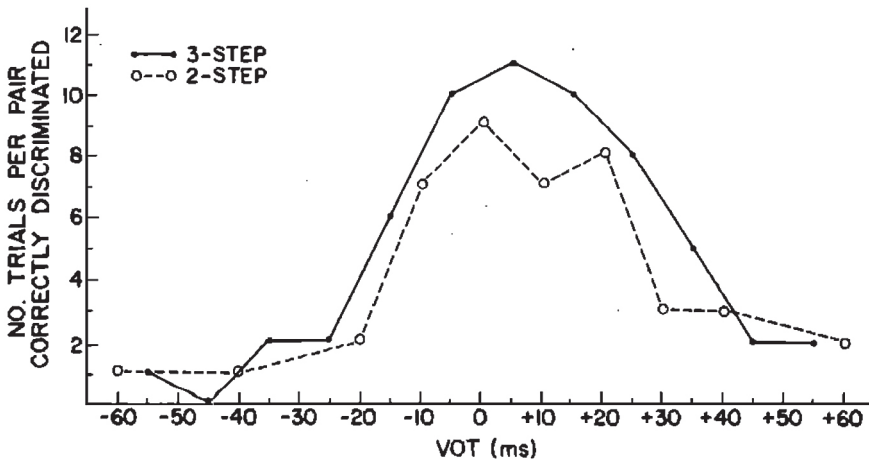
Źródło: Keating et al. 1981, s. 1263.

Niektórzy badacze definiując VOT określają zakres tego parametru jako *continuum* w przestrzeni czasowej od -100 milisekund do około $+100$ milisekund. Wskazuje się również na uniwersalność VOT, wynikającą z możliwości ludzkiego toru głosowego, który może wyprodukować spółgłoski zwarte w jakimkolwiek punkcie podanego zakresu czasowego.

VOT A PERCEPCJA DŹWIĘCZNOŚCI

Nagłosowe spółgłoski zwarto-wybuchowe są również przedmiotem licznych badań percepcyjnych, zapoczątkowanych w latach 60. (zob. np. Yeni-Kom-

shian 2005). Materiałem bodźcowym jest w nich syntetyczne kontinuum, zazwyczaj sylab, różniących się długością VOT, którego wartość maleje lub wzrasta o stałą wartość (np. krok = 10 milisekund). Uczestnicy badań wskazują, czy prezentowane pary bodźców (np. [da da] lub [ta da] o zmodyfikowanej wartości VOT są identyczne („takie same”), czy różnią się od siebie („różne”). Jest to tzw. zadanie dyskryminacyjne. Natomiast w zadaniach identyfikacyjnych słuchacze wskazują, którą konkretnie sylabę słyszą ([ta], czy [da]). Wyniki tego typu badań pozwalają wyznaczyć tzw. bodziec graniczny oddzielający kategorię fonemiczną. Jest on jest niestabilny percepcyjnie, czyli nie jest odbierany jako głoska pośrednia, lecz raz jako [da], a innym razem jako [ta]. Taki schemat reakcji jest charakterystyczny dla zjawiska tzw. percepcji kategoryalnej, gdzie w obrębie zmieniającej się w sposób ciągły cechy fizycznej bodźca (iloczasu VOT skracanego lub wydłużanego co 10 milisekund) manifestuje się percepcyjna nieciągłość. Na rycinie 7 poniżej zaprezentowano wyniki zadania dyskryminacyjnego na podstawie parametru VOT przeprowadzonego w grupie ośmiu polskojęzycznych słuchaczy (zob. Keating 1981). Krzywa oznaczona linią przerywaną (tzw. 2-step) przedstawia reakcje słuchaczy na pary bodźców różniące się 20-milisekundową długością VOT (np. -70 ms/ -50 ms, -50 ms/ -30 ms, -30 ms/ -10 ms, -10 ms/ $+10$ ms, 0 ms/ $+20$ ms, $+10$ ms/ $+30$ ms, $+30$ ms/ $+50$), natomiast krzywa oznaczona linią ciągłą (tzw. 3-step) ilustruje wyniki zadania dyskryminacyjnego par różniących się parametrem VOT o długości 30 ms (np. -70 ms/ -40 ms, $+40$ ms/ $+70$ ms).



Rycina 7. Uśrednione wyniki zadania dyskryminacji VOT dla par różniących się parametrem VOT o długości 20 ms (2-step) oraz 30-ms (3-step)

Źródło: Keating 1981, s. 1263.

Na obu wykresach przedstawionych na rycinie 7 wyraźnie widać zarówno tendencję do rozpoznawania par bodźców jako ‘takich samych’, jak też wzrost

dyskryminacji w przedziale od -15 ms do $+25$ ms, ze szczytem dyskryminacji wynoszącym 0 ms dla opozycji 2-stopniowej oraz $+5$ ms dla opozycji 3-stopniowej. „Thus the data from production and discrimination independently point to a Polish category boundary at about 0 ms VOT”⁷ (Keating et al. 1981, 1264). Szczyt dyskryminacyjny znajduje się na granicy pomiędzy dwiema kategoriami fonemicznymi w obrębie kontrastu \pm dźwięczna.

CHARAKTERYSTYKA TRENINGU NEUROFLOW ATS

Aktywny trening słuchowy Neuroflow jest metodą rehabilitacji dzieci z trudnościami słuchowymi i prawidłową czułością słuchu, czyli dzieci z zaburzeniami wyższych funkcji słuchowych CAPD (ang. *Central Auditory Processing Disorder*). Aktywny trening słuchowy Neuroflow usprawnia procesy wstępujące i zstępujące, wykorzystując bodźce niejęzykowe: tony, szумы, jak i materiał językowy. Jest on realizowany w formie interaktywnych ćwiczeń on-line z wykorzystaniem procedur adaptacyjnych dostosowujących trudność ćwiczeń do możliwości dziecka. Neuroflow ATS jest przeznaczony głównie do pracy w domu z rodzicem, z wykorzystaniem indywidualnie dobranego dla dziecka zestawu ćwiczeń, ale może być również realizowany w gabinecie z terapeutą. Ćwiczenia wyższych funkcji słuchowych obejmują usprawnianie rozumienia mowy w szumie, ćwiczenia rozdzielnouszne poprawiające wymianę informacji pomiędzy półkulami mózgu, ćwiczenia usprawniające uwagę i pamięć słuchową oraz ćwiczenia różnicowania wysokości i długości dźwięków.

Wśród pacjentów z zaburzeniami przetwarzania słuchowego CAPD dużą grupę stanowią dzieci ze współistniejącymi problemami w zakresie słuchu fonemowego (tzw. podtyp fonologiczny CAPD). Nieprawidłowo funkcjonującym mechanizmem u tych dzieci są trudności w odkodowywaniu cech akustycznych dźwięków mowy. Manifestuje się to zaburzeniami w różnicowaniu i identyfikowaniu realizacji fonemów, w tym spółgłosek zwarto-wybuchowych, których akustyczną korelatą ze względu na cechę dźwięczności jest parametr VOT. U pacjentów z tego typu deficytami często wtórnie obserwuje się zaburzenia artykulacji – mowę bezdźwięczną oraz trudności w nabywaniu umiejętności czytania i poprawnego pisania.

Aktywny trening słuchowy Neuroflow poświęcony pacjentom z deficytem fonologicznym CAPD pozwala usprawniać wyższe funkcje słuchowe oraz funkcje fonologiczne. Interaktywna forma pracy dziecka z rodzicem wymaga skupienia uwagi, wykorzystania pamięci oraz zaangażowania emocjonalnego, co buduje wzajemnie relacje, pozwala wspólnie pokonywać trudności pojawiające się

⁷ Tak więc dane z realizacji i dyskryminacji niezależnie wskazują na granicę polskiej kategorii przy około 0 ms VOT [tłumaczenie własne].

w treningu, a w konsekwencji poprawiać funkcjonowanie dziecka w szkole. Trening Neuroflow jest autorską formą pracy z pacjentami z zaburzeniami przetwarzania słuchowego, ale może stanowić również element terapii logopedycznej i pedagogicznej⁸.

NOWE ĆWICZENIA W AKTYWNYM TRENINGU SŁUCHOWYM NEUROFLOW

Nowy zestaw ćwiczeń w terapii Neuroflow ATS przeznaczony jest dla dzieci z zaburzeniami słuchu fonematycznego oraz z zaburzeniami przetwarzania słuchowego o podtypie fonologicznym. Są to ćwiczenia dyskryminacji i identyfikacji słuchowej cechy ±dźwięczności w obrębie spółgłosek zwarto-wybuchowych z wykorzystaniem akustycznego parametru VOT. Opracowano je w interdyscyplinarnym zespole, w którego składzie znaleźli się:

- dr hab. Anita Lorenc, prof. ucz. – fonetyk, surdologopeda (Centrum Logopedyczne Uniwersytetu Warszawskiego, Pracownia Fonetyki Stosowanej IPS UW),
- dr inż. Daniel Król – specjalista ds. cyfrowego przetwarzania sygnałów (Katedra Informatyki Akademii Nauk Stosowanych w Tarnowie),
- dr n. med. Andrzej Senderski – otolaryngolog, audiolog, foniatra (Oddział Otolaryngologii Dziecięcej i Pediatrii UCK WUM, APD Medical sp. z o.o.),
- mgr Renata Borowiecka – pedagog specjalny i korekcyjny (APD Medical sp. z o.o.).

Nowe ćwiczenia wzbogaciły zestawy dotyczące funkcji fonologicznych realizowanych w czasie dziennej sesji terapeutycznej.

MATERIAŁ JĘZYKOWY

W ćwiczeniach z modyfikacją VOT zastosowano materiał językowy obejmujący 120 słów jedno- i dwusylabowych, różniących się realizacją jednego fonemu w obrębie par minimalnych /p:b/, /t:d/, /k:g/ w pozycji nagłosu wyrazowe w sylabach otwartych lub w grupie spółgłoskowej (zob. tabela 3).

⁸ Informacje o terapii Neuroflow dla dzieci z CAPD dostępne są na stronie www.neuroflow.pl

Tabela 3. Liczba par minimalnych w realizacjach polskich fonemów zwarto-wybuchowych w ćwiczeniach z modyfikacją VOT w terapii Neuroflow ATS

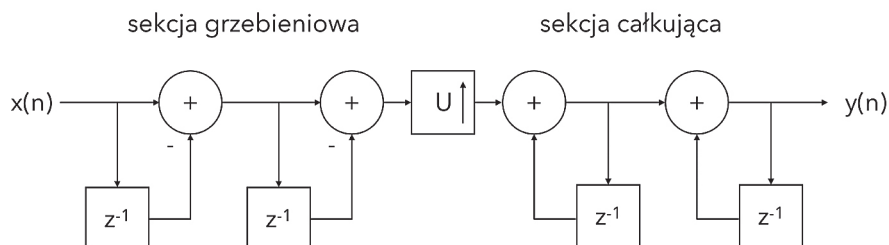
NAGŁOS	/p:b/	/t:d/	/k:g/
w sylabie otwartej	18 (np. pada: bada)	14 (np. tata: data)	15 (np. kary: gary)
w zbitce spółgłoskowej	6 (np. pranie: branie)	3 (np. tratwa: dratwa)	7 (np. kładki: gładki)
RAZEM	24	17	22

Źródło: opracowanie własne.

Nagrania zrealizowano w profesjonalnym studiu. Głosu udzielił znany lektor filmowy – Piotr Borowiec, członek Stowarzyszenia Lektorów Rzeczypospolitej Polskiej.

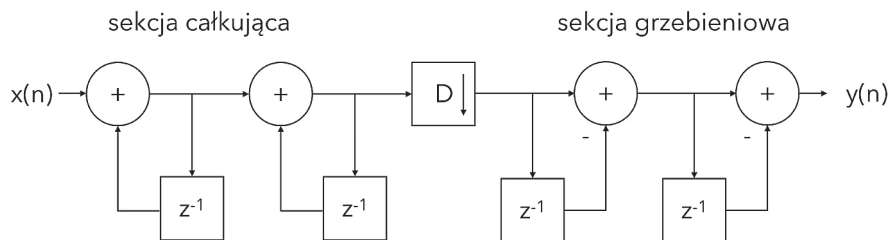
MODYFIKACJA DŁUGOŚCI VOT

Modyfikację długości VOT zrealizowano posługując się techniką przepróbkowania (ang. *resampling*) sygnału cyfrowego. Ze względu na potrzebę uzyskiwania ułamkowego współczynnika przepróbkowania, w modyfikowanych fragmentach sygnału, zastosowano połączenie algorytmów cyfrowego interpolatora (zob. rycina 8) oraz decymatora (zob. rycina 9) kaskadowego.



Rycina 8. Schemat blokowy interpolatora cyfrowego

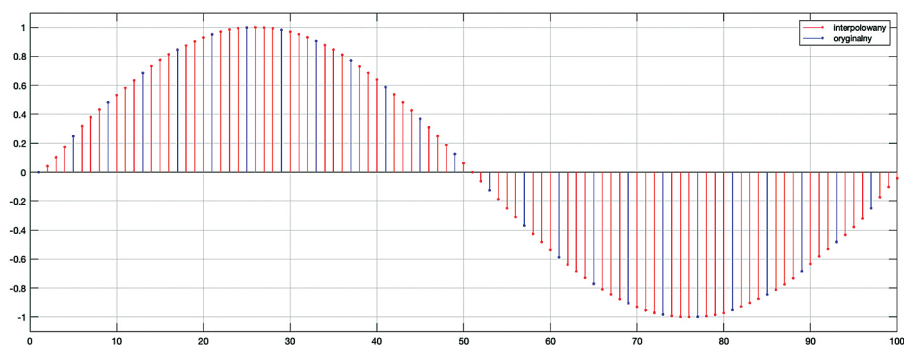
Źródło: opracowanie własne.



Rycina 9. Schemat blokowy decymatora cyfrowego

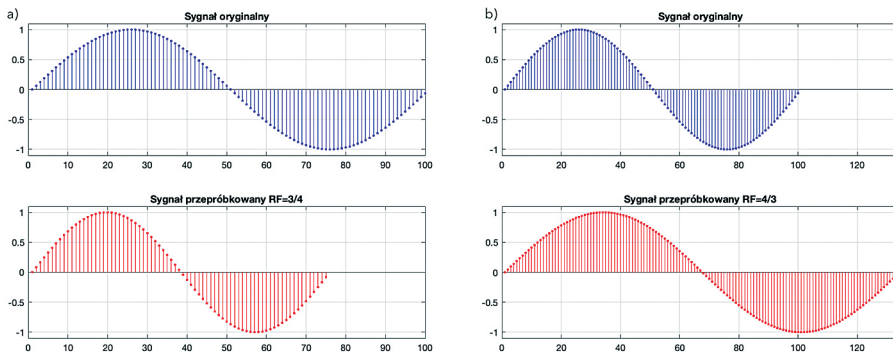
Źródło: opracowanie własne.

Interpolator cyfrowy generuje U dodatkowych próbek pośrednich w przetwarzanym sygnale (zob. rycina 10) natomiast decymator umożliwi pozostawienie co D -tej próbki oraz usunięcie pozostałych.

Rycina 10. Przykład interpolacji sygnału dla $U=4$

Źródło: opracowanie własne.

Działanie algorytmu zmieniającego liczbę próbek sygnału, z ułamkowym współczynnikiem $RF = U/D$, polega na jego interpolacji ze współczynnikiem U a następnie decymacji ze współczynnikiem D . Proces przepróbowania, w zależności od współczynnika, umożliwił zatem usunięcie lub wstawienie dodatkowych próbek pośrednich (zob. rycina 11).



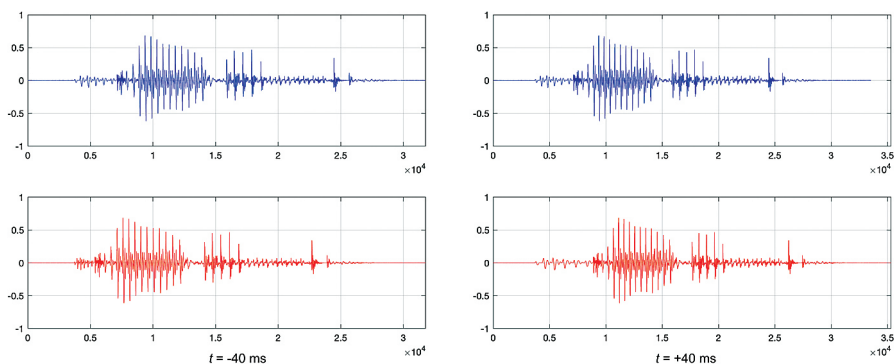
Rycina 11. Przykład zmiany długości sygnału z wykorzystaniem przepróbkowania: a) skrócenie – współczynnik przepróbkowania $RF=3/4$ oraz b) wydłużenie – współczynnik przepróbkowania $RF=4/3$

Źródło: opracowanie własne.

Konsekwencją przepróbkowania fragmentu sygnału zawierającego VOT jest jego precyzyjne skrócenie lub wydłużenie w odtwarzanym słowie. Liczba usuniętych lub wstawionych dodatkowych próbek pośrednich jest obliczana na podstawie długości fragmentu VOT w każdym (referencyjnym) przetwarzanym słowie w stosunku do zadanego wydłużenia/skrócenia w generowanych (nowych) słowach. W badaniach wykorzystano zmianę długości VOT z krokiem 10 ms. Z modyfikacji wyłączono segment wybuchu⁹.

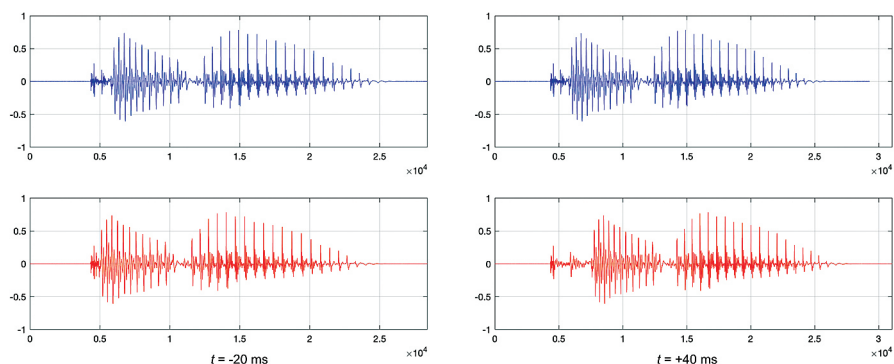
Przykładowe efekty zastosowanej modyfikacji długości VOT można zaobserwować na rycinie 12 i rycinie 13 poniżej. Rycina 12 przedstawia oscylogramy wyrazu *góra*, gdzie kolor niebieski (oscylogramy na górze) odnosi się do sygnału niezmodyfikowanego, natomiast kolor czerwony (na dole) ilustruje skrócony o 40 ms (po lewej) lub wydłużony o 40 ms (po prawej) iloczas VOT nagłosowej spółgłoski [g] w wyrazie *góra*.

⁹ Modyfikację VOT wykorzystwały autorki *Testu do badania rozumienia mowy u dzieci i dorosłych. Nowe spojrzenie na zegar mózgowy* (Szelağ, Szymaszek 2014). Zastosowano ją w odniesieniu tylko do jednego wyrazu *Tomek*, na bazie którego uzyskano kolejne bodźce o ujemnej lub dodatniej wartości VOT na kontinuum od – 100 ms do + 90 ms, z krokiem wynoszącym 10 ms. Niestety, autorki nie opisują dokładnie, w jaki sposób doprowadziły do modyfikacji VOT, trudno zatem odnieść się do ich badania w bardziej szczegółowy sposób.



Ryc. 12. Oscylogramy niezmodyfikowanego sygnału akustycznego wyrazu *góra* (kolor niebieski) oraz ze skróconym (kolor czerwony po lewej) lub wydłużonym (kolor czerwony po prawej) o 40 milisekund iloczasm VOT.

Z kolei na rycinie 13 przedstawiono oscylogramy wyrazu *kura*. Podobnie jak w przykładzie powyżej kolorem niebieskim na górze oznaczono sygnał niezmodyfikowany, natomiast kolor czerwony (na dole) ilustruje skrócony o 20 ms (po lewej) lub wydłużony o 40 ms (po prawej) iloczasm VOT nagłosowej spółgłoski [k] w wyrazie *kura*. W związku z tym, że w przypadku polskich spółgłosek bezdźwięcznych mamy do czynienia z tzw. krótkim opóźnieniem dźwięczności (ang. *short voicing lag*; zob. fragment *Akustyczny parametr VOT* w niniejszym artykule), skracanie VOT miało zakres ograniczony do około 20–30 ms w zależności od charakterystyki sygnału wyjściowego.



Rycina 13. Oscylogramy niezmodyfikowanego sygnału akustycznego wyrazu *kura* (kolor niebieski) oraz ze skróconym o 20 ms (kolor czerwony po lewej) lub wydłużonym o 40 ms (kolor czerwony po prawej) iloczasm VOT

Charakterystyka ćwiczeń słuchowych wykorzystujących modyfikację iloczasu VOT w aktywnym treningu słuchowym Neuroflow

Dzięki modyfikacji iloczasu VOT dla każdej pary opozycyjnych fonemów udało się uzyskać 9 poziomów trudności w ćwiczeniach dyskryminacji i identyfikacji słuchowej. Bazowy poziom trudności to mowa niezmodyfikowana – poziom trudności 0.

Modyfikując VOT opracowano 4 poziomy łatwiejsze niż mowa naturalna (wydłużony iloczas VOT, poziom trudności od najłatwiejszego do bliskiego mowie naturalnej: -4, -3, -2, -1) oraz 4 poziomy trudniejsze niż percepcja mowy naturalnej (skrócony iloczas VOT, poziom trudności od najbliższego mowie naturalnej do coraz łatwiejszych percepcyjnie: +1, +2, +3, +4).

Wyrazy ze spółgłoskami [p t k] oraz [b d k] są łatwiejsze do dyskryminacji i identyfikacji słuchowej, jeżeli będziemy wydłużać ich VOT. Dla spółgłosek [p t k] przy wydłużaniu VOT (podobnie jak przy skracaniu) został zastosowany skok 5 ms, a dla spółgłosek [b d g] skok wynosił 20 ms.

Wyrazy ze spółgłoskami [p t k] oraz [b d k] są trudniejsze do dyskryminacji i identyfikacji słuchowej, jeżeli będziemy skracać ich VOT. Dla głosek [p t k] przy skracaniu VOT zastosowano skok 5 ms dla głosek [b d g] skok wynosił 20 ms.

Każdy poziom trudności zawiera 40 słów, z których program losowo wybiera jedno słowo lub parę słów do kolejnych prezentacji. Ćwiczenie kończy się po dwudziestu prezentacjach o różnym poziomie trudności w zależności od poprawności odpowiedzi dziecka. W ćwiczeniach zastosowano algorytmy adaptacyjne (zob. Stropahl 2020), dzięki którym po dwóch dobrych odpowiedziach następne zadanie było trudniejsze, a po jednej nieprawidłowej odpowiedzi następne zadanie było o stopień łatwiejsze. Zastosowanie algorytmów adaptacyjnych oraz modyfikacji VOT pozwoliło na opracowanie ćwiczeń o różnym poziomie trudności dostosowanych do indywidualnego poziomu rozwoju funkcji słuchowych dziecka.

W przypadku występowania zaburzeń o podtypie fonologicznym u pacjentów obserwuje się wzrost ich nasilenia w przypadku obecności dodatkowych bodźców zakłócających (szumu/hałasu). Ten wzrost nasilenia objawów jest niewspółmiernie wysoki w stosunku do głośności zakłócającego szumu. Z tego też powodu badanie słuchu fonemowego w ciszy niekiedy nie wykazuje zaburzeń, a w realnych warunkach akustycznych panujących w przedszkolu, domu czy szkole, gdzie poziom hałasu jest podwyższony, dziecko prezentuje cechy zaburzeń słuchu fonemowego (zob. White-Schwoch et al. 2015; Song et al. 2011). W treningu słuchowym Neuroflow ATS każde ćwiczenie ma trzy poziomy trudności różniące się obecnością i głośnością dodatkowego szumu „bale noise” (tzn. nakładających się na siebie głosów wielu mówców): 1. ćwiczenia bez dodatkowych bodźców

zakłócających w tle, 2. ćwiczenia z dodatkowymi bodźcami zakłócającymi w tle, o stosunku sygnału do szumu (SNR) + 9dB i 3. ćwiczenia z dodatkowymi bodźcami zakłócającymi w tle, o stosunku sygnału do szumu (SNR) + 5dB. Na początku trwania terapii stosuje się ćwiczenia bez towarzyszącego szumu, a w jej trakcie dołączane są ćwiczenia z dodatkowymi bodźcami zakłócającymi w tle o zróżnicowanej głośności.

Polecenie skierowane do dziecka w zależności od rodzaju ćwiczenia brzmi: „Czy słowa, które usłyszysz są takie same czy różne?”, „Powtórz usłyszane słowo”, „Powtórz usłyszane dwa słowa”.

W trakcie wykonywania ćwiczenia rodzic/terapeuta przekazuje dziecku informację zwrotną o poprawności wykonania zadania, a także wzmacnia motywację dziecka, stosując pochwały werbalne. Dodatkowo dziecko po wykonaniu całego ćwiczenia otrzymuje motywacyjny puchar widoczny na ekranie, w zależności od osiągniętego poziomu trudności: brązowy, srebrny lub złoty. Poniżej przedstawiono kilka przykładowych zestawów ćwiczeń dyskryminacji i identyfikacji słuchowej z wykorzystaniem akustycznego parametru VOT w terapii Neuroflow ATS.

Zestaw 1

Poziomy trudności VOT: od -4 do 0 (łatwiejsze niż mowa naturalna).

Polecenie dla dziecka: Czy słowa, które usłyszysz są takie same czy różne?

Prezentacja bodźców dźwiękowych: dwa słowa prezentowane obustronnie, jedno po drugim.

Szum: bez dodatkowego szumu zakłócającego.

Zestaw 2

Poziomy trudności VOT: od -4 do 0 (łatwiejsze niż mowa naturalna).

Polecenie dla dziecka: Czy słowa, które usłyszysz są takie same czy różne?

Prezentacja bodźców dźwiękowych: dwa słowa prezentowane obustronnie, jedno po drugim.

Szum: +9 dB SNR.

Zestaw 3

Poziomy trudności: VOT od -4 do 0 (łatwiejsze niż mowa naturalna).

Polecenie dla dziecka: Czy słowa, które usłyszysz są takie same czy różne?

Prezentacja bodźców dźwiękowych: dwa słowa prezentowane obustronnie jedno po drugim.

Szum: +5 dB SNR.

Zestaw 4

Poziomy trudności: VOT od 0 do 4 (trudniejsze niż mowa naturalna).

Polecenie dla dziecka: Czy słowa, które usłyszysz są takie same czy różne?

Prezentacja bodźców dźwiękowych: dwa słowa prezentowane obustronnie, jedno po drugim.

Szum: bez dodatkowego szumu zakłócającego.

Zestaw 7

Poziomy trudności: VOT – 4 do 0 (łatwiejsze niż mowa naturalna).

Polecenie dla dziecka: Powtórz dwa usłyszane słowa.

Prezentacja bodźców dźwiękowych: dwa słowa prezentowane obustronnie, jedno po drugim

Szum: bez szumu.

Zestaw 14

Poziomy trudności: VOT – 4 do 0 (łatwiejsze niż mowa naturalna).

Polecenie dla dziecka: Powtórz usłyszane słowo.

Prezentacja bodźców dźwiękowych: jedno słowo prezentowane obustronnie.

Szum: +9 dB SNR.

Pełny zestaw ćwiczeń dla par par minimalnych /p:b/, /t:d/, /k:g/ zaprezentowano w tabeli 4 poniżej.

Tabela 4. Zestaw ćwiczeń dyskryminacji i identyfikacji słuchowej z wykorzystaniem akustycznego parametru VOT w terapii Neuroflow ATS

Nr zestawu ćwiczeń	VOT adaptacyjny	Polecenie	Szum w dB SNR
1	-4-0	Czy słowa, które usłyszysz są takie same czy różne?	Bez szumu
2	-4-0	Czy słowa, które usłyszysz są takie same czy różne?	+9
3	-4-0	Czy słowa, które usłyszysz są takie same czy różne?	+5
4	0-4	Czy słowa, które usłyszysz są takie same czy różne?	Bez szumu
5	0-4	Czy słowa, które usłyszysz są takie same czy różne?	+9
6	0-4	Czy słowa, które usłyszysz są takie same czy różne ?	+5
7	-4-0	Powtórz dwa słyszane słowa.	Bez szumu
8	-4-0	Powtórz dwa słyszane słowa.	+9

Tabela 4. Ciąg dalszy

9	-4-0	Powtórz dwa słyszane słowa.	+5
10	0-4	Powtórz dwa słyszane słowa.	Bez szumu
11	0-4	Powtórz dwa słyszane słowa.	+9
12	0-4	Powtórz dwa słyszane słowa.	+5
13	-4-0	Powtórz usłyszane słowo.	Bez szumu
14	-4-0	Powtórz usłyszane słowo.	+9
15	-4-0	Powtórz usłyszane słowo.	+5
16	0-4	Powtórz usłyszane słowo.	Bez szumu
17	0-4	Powtórz usłyszane słowo.	+9
18	0-4	Powtórz usłyszane słowo.	+5

PODSUMOWANIE

Nowe ćwiczenia w terapii Neuroflow ATS wykorzystują modyfikację akustycznego parametru VOT w nauce dyskryminacji i identyfikacji słuchowej cechy \pm dźwięczności. Zaletą tego typu ćwiczeń jest możliwość wydłużania iloczasu VOT w celu ułatwienia percepcji słuchowej oraz z drugiej strony skracania iloczasu VOT na potrzeby zwiększenia stopnia trudności ćwiczeń. Przedstawiona propozycja uzupełnia tradycyjne wychowanie słuchowe dzięki modyfikacji sygnału akustycznego oraz zastosowaniu dodatkowych bodźców zakłócających.

BIBLIOGRAFIA

- Abramson A.S., Lisker L., 1965, *Voice onset time in stop consonants: Acoustic analysis and synthesis*. Rapports du 5^e Congrès International d'Acoustique, vol. 1a, s. A51.
- Benni T., 1931, *Palatogramy polskie*, Kraków.
- Boersma P., Weenink D., 2014, *Praat: Doing phonetics by computer* [program komputerowy]. Version, 5(3), 57 <http://www.praat.org/>.
- Cho T., Ladefoged P., 1999, *Variations and Universals in VOT: Evidence from 18 Languages*, „Journal of Phonetics”, 27, s. 207–222.
- Cho T., Whalen D.H., Docherty G., 2019, *Voice onset time and beyond: Exploring laryngeal contrast in 19 languages*, „Journal of Phonetics”, 72, s. 52–65.
- Hirose H., 1997, *Investigating the physiology of laryngeal structures*, [w:] *The Handbook of Phonetic Sciences*, red. W.J. Hardcastle, J. Laver, Cambridge, s. 116–136.

- Jassem W., 1973, *Podstawy fonetyki akustycznej*, Warszawa.
- Jassem W., 1992, *Podręcznik wymowy angielskiej*, Warszawa.
- Keating P.A., Mikoś M.J., Ganong W.F. III., 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.
- Kollia H.B., 1993, *Segmental duration changes due to variations in stress, vowel, place of articulation and voicing of stop consonants in Greek*, "Journal of the Acoustical Society of America", 93, s. 2298.
- Koneczna H., Zawadowski W., 1951, *Przekroje rentgenograficzne głosek polskich*, Warszawa.
- Konopska L., Sawicki J., 2013, *Badania parametry VOT w realizacjach polskich bezdźwięcznych i dźwięcznych fonemów zwarto-wybuchowych*, „Logopedia”, 42, s. 101–124.
- Ladefoged P., 1971, *Preliminaries to Linguistic Phonetics*, Chicago.
- Laver J., 1980, *The phonetic description of voice quality*, Cambridge.
- Laver J., 1994, *Principles of Phonetics*, Cambridge.
- Lisker L., Abramson A.S., 1964, *A Cross-Language Study of Voicing in Initial Stops. Acoustical Measurements*, „Word” 20, s. 384–422.
- Lisker L., Abramson A.S., 1970, *The voicing dimension: Some experiments in comparative phonetics*. Proceedings of the 6th International Congress of Phonetic Sciences, Prague.
- Lisker L., Abramson A.S., Cooper F.S., Schvey M.H., 1969, *Transillumination of the larynx in running speech*, „Journal of the Acoustical Society of America”, 45, s. 1544–1546.
- Picett J.M., 1999, *The Acoustics of Speech Communication*, Boston.
- Raphael L.J., Tobin Y., 1983, *Perceptual and acoustic studies of voice onset time in Hebrew*, [w:] *Abstracts of the Tenth International Congress of Phonetic Sciences*, red. A. Cohen, M.P.R. van der Broeke, Dordrecht, s. 516.
- Rogers H., 2000, *The Sounds of Language. An Introduction to Phonetics*, London.
- Sawashima M., Abramson A.S., Cooper F.S., Lisker L., 1970, *Observing laryngeal adjustments during running speech by use of a fiberoptics system.*, *Phonetica*, 22, s. 193–201.
- Sawashima M., Hirose H., 1983, *Laryngeal gestures in speech production*, [w:] *The Production of Speech*, red. P. MacNeilage, New York, s. 11–38.
- Song J.H., Skoe E., Banai K., Kraus N., 2011, *Perception of speech in noise: neural correlates*, „Journal of Cognitive Neuroscience”, 23, s. 2268–79.
- Stropahl M., Besser J., Launer S., 2020, *Auditory Training Supports Auditory Rehabilitation: A State-of-the-Art Review*, „Ear and Hearing”, 41(4), s. 697–704.
- Szeląg E., Szymaszek A., 2014, *Test do badania rozumienia mowy u dzieci i dorosłych. Nowe spojrzenie na zegar mózgowy*, Sopot.
- Szczepankowski B., 1985, *Fonetyka akustyczna, audytywna i wizualna*, Warszawa.
- Trochymiuk [Lorenc] A., 2008, *Wymowa dzieci niesłyszących. Analiza audytywna i akustyczna*, seria: Komunikacja Językowa i Jej Zaburzenia, t. 22, Lublin.
- White-Schwoch T., Woodruff Carr K., Thompson E.C., Anderson S., Nicol T., Bradlow A.R., Zecker S.G., Kraus N., 2015, *Auditory processing in noise: A preschool biomarker for literacy*, „Plos Biology” (dostępny w internecie: <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002196>).
- Yeni-Komshian G.H., 2005, *Percepcja mowy*, [w:] *Psycholingwistyka*, red. J. Berko Gleason, N. Bernstein Ratner, Gdańsk.