

Antoni Grzanka^{1,2}, Radosław Pietruczuk³

¹ Instytut Systemów Elektronicznych Politechniki Warszawskiej

² Instytut Fizjologii i Patologii Słuchu, Warszawa

³ Instytut Telekomunikacji Politechniki Warszawskiej

Komputerowy test percepcji pseudoszeptu

Computer test of Pseudowhisper Perception

Słowa kluczowe: pseudoszept, listy zrównoważone fonetycznie i strukturalnie, test psychoakustyczny, test z zamkniętą listą wyboru.

Key words: pseudowhisper, phonetically and structurally balanced lists of words, psychoacoustic test, multiple choice test.

Streszczenie

W poniższej pracy przedstawiony jest komputerowy system do oceny zrozumiałości mowy bezkrtaniowców oparty na teście psychoakustycznym. Podkreślona zostanie sprawność, uniwersalizm oraz prostota obsługi systemu opartego na listach zrównoważonych strukturalnie i fonetycznie. Uzyskane wyniki wskazują na proces szybkiego uczenia się zrozumiałości pseudoszeptu przez wszystkie osoby biorące udział w testach.

Summary

A computer system for estimation of understanding pseudowhisper based on psychoacoustic test is presented in this paper. Efficiency, universality and simplicity of the system based on the phonetically and structurally balanced lists of words is emphasized. The results show the process of a very quick learning of understanding pseudowhisper by all the people taking part in the tests.

Istnieje wiele przyczyn zaburzeń mowy, z których operacyjne usunięcie krtani prowadzi do najpoważniejszego inwalidztwa. Mowa osób laryngotomowanych jest trudna do zrozumienia z dwóch powodów – braku dźwięczności i braku przepływu powietrza dla wytworzenia zwykłego szeptu. Pomocą

upośledzonym służyć może proces rehabilitacji lub stosowanie protez mechanicznych i elektronicznych, które dążą do zwiększenia komunikatywności.

Skuteczność technik operacyjnych rehabilitacji i protezowania powinna podlegać obiektywnej ocenie [Pastore (i in.) 1998]. Poprzez kontrolę stosowanych środków możliwe jest lepsze ich dopasowanie do indywidualnych predyspozycji. W przypadku opracowywania nowych protez ocena mowy wyznacza kierunki prac badawczych i konstrukcyjnych.

W literaturze przedmiotu można znaleźć dwa podejścia do zagadnienia. Pierwsze z nich to zastosowanie oceny cech akustycznych mowy i procesu wypowiedzi. Najczęściej stosowane do oceny parametry to: częstotliwość podstawowa, poziom i czas trwania [Hoasjoe (i in.) 1992].

Na drugim biegunie metod oceny mowy są testy psychoakustyczne, gdzie stopień zrozumiałości jest oceniany przez niezależnych słuchaczy. Dobrym przykładem takiej metody jest test przedstawiony w pracy [Pastore (i in.) 1998]. Test psychoakustyczny oparty był na wypowiedzi pięciu przedłużonych samogłosek i sześciu fonetycznie zrównoważonych zdań. Mikrofon znajdował się w odległości 30 cm od ust, a czas zapisu trwał 3 min. Oceny dokonywali wyuczeni słuchacze. Punktacja słuchaczy na specjalnych kartach oceny dotyczyła cech jakościowych mowy i charakterystyk perceptualnych. W ten sposób autorzy dowodzili skuteczności swojej techniki operacyjnej.

Ten autor, jak i inni łączyli obie techniki. Dobrym przykładem jest tu praca [Rydell (i in.) 1995], w której zastosowano cechy akustyczne do oceny jakości mowy (shimmer, jitter, breathiness, stosunek harmonicznego do szumu i częstotliwość podstawowa) obliczane programem Soundscope. Ponadto stosowano jako wskaźniki czas potrzebny na przeczytanie próbki mowy i liczbę oddechów. Ocena zrozumiałości dokonywana była przez dwie grupy słuchaczy – doświadczonych (A) i niedoświadczonych (B).

Niniejsza praca dotyczy testu psychoakustycznego. Szczególną wagę przywiązywano do właściwego doboru zestawu słów zrównoważonych fonematycznie i strukturalnie. Forma komputerowa testu została opracowana w ten sposób, że test może być przeprowadzany praktycznie na każdym komputerze osobistym. Baza słów jest łatwo rozszerzalna o dowolne reprezentatywne listy słów testowych. W artykule chcieliśmy podkreślić interesujące wyniki otrzymane w badaniach pilotowych zestawu testowego. Wyniki uzyskano po dokonaniu oceny przez 19 słuchaczy.

Zaobserwowano szczególnie wyraźnie zaznaczony proces szybkiego uczenia się zrozumiałości mowy bezkrtaniowców przez osoby, które nie miały wcześniej kontaktu z pseudoszeptem. Ważnym rezultatem jest też uzyskana klasyfikacja słów według stopnia trudności ich rozpoznania.

I. MATERIAŁ I METODA

Do eksperymentów wykorzystano próbki pseudoszeptu osoby po operacji usunięcia krtani. Był to mężczyzna w wieku 56 lat. Okres od operacji do momentu nagrania wynosił trzy miesiące. Pooperacyjny okres rehabilitacji nie wpłynął znacząco na podniesienie jakości wypowiedzi u nagrywanej osoby. Mężczyzna chętnie współpracował przy tworzeniu nagrań.

Podczas sesji nagraniowych osoba wypowiadała pojedyncze słowa z list słów zrównoważonych fonematycznie i strukturalnie według Zakrzewskiego, Pruszevicza i Kubzdeli. Każda zamknięta lista zawierała wyrazy jedno- i dwusylabowe. Struktura i budowa tych słów są reprezentatywne pod względem częstości występowania fonemów dla mówionego języka polskiego. Dla każdej z osób nagrano sześć list zrównoważonych, a każda lista zawierała 29 słów.

Nagrań dokonano w komorze bezchłowej Wydziału Elektroniki i Technik Informacyjnych Politechniki Warszawskiej. Takie warunki pozwoliły na wyeliminowanie zakłóceń i szumów akustycznych oraz uniezależniły jakość nagrań od specyficznych warunków pomieszczenia (brak pogłosu). Do nagrań zastosowano mikrofon dynamiczny (typ Md 286), który był podłączony bezpośrednio do wejścia karty dźwiękowej (SOUND-GALAXY PRO 16), zainstalowanej w komputerze osobistym. W ten sposób rejestrowane dźwięki były bezpośrednio przetwarzane na postać cyfrową bez etapów pośrednich, co zapewniło dużą wierność zapisu. Stosowano 16-bitowe przetwarzanie analogowo-cyfrowe w karcie muzycznej z częstotliwością próbkowania 22,05 kHz. Każde słowo było zapisywane na twardym dysku komputera w oddzielnym pliku zgodnym ze standardem WAV bez kompresji (kodowanie PCM – pulse code modulation).

Dla oceny stopnia zrozumiałości zebranych próbek mowy zastosowano test psychoakustyczny, w którym mowę oceniało 19 osób (15 mężczyzn i 4 kobiety) w wieku ok. 22 lat bez upośledzenia słuchu. Oceniający słuchacze byli studentami. Próbkę mowy były odtwarzane bezpośrednio z postaci cyfrowej za pomocą wymienionej wyżej karty dźwiękowej. Do odsłuchiwania dźwięków zastosowano słuchawki audiometryczne TDH 39P o paśmie do 8 kHz. Każdy ze słuchaczy oceniał zrozumiałość słów ze wszystkich sześciu list. Listy zrównoważone fonetycznie prezentowane były każdemu słuchaczowi kolejno od numeru 1 do numeru 6.

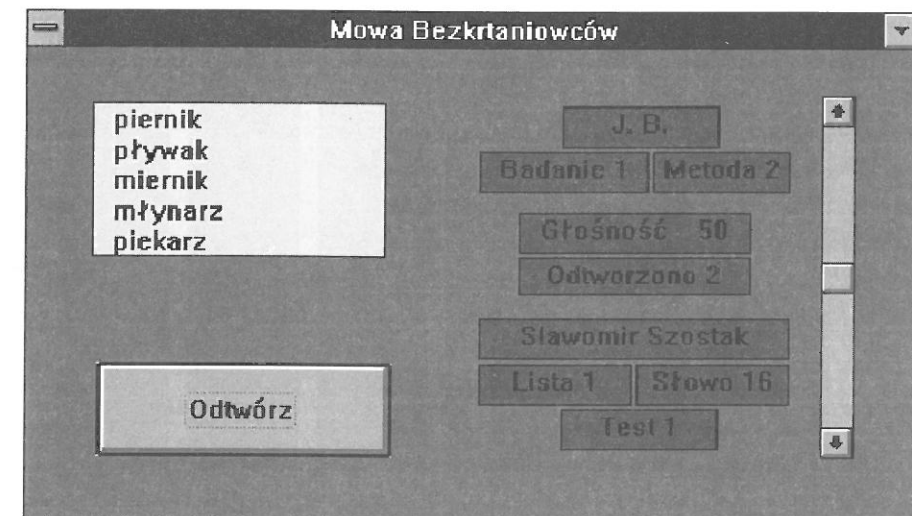
Z powodu znacznej trudności w identyfikacji słów zastosowano testy z zamkniętą listą wyboru (*multiple choice test*). Do każdego z testowanych słów dobrano cztery słowa mylące. Słowa mylące dobrano kierując się następującymi zasadami: słowa te musiały mieć tę samą liczbę sylab i charakteryzować się tym samym rytmem wypowiedzi, musiały też zdecydowanie różnić się wzajemnie znaczeniowo. Starano się przy tym dobierać słowa potoczne. Jedną z list wraz ze słowami mylącymi przedstawiono w tab. 1.

Tab. 1. Pierwsza z sześciu list słów zrównoważonych stosowana w testach zrozumiałości (właściwe słowa listy przedstawione są w drugiej kolumnie; dalsze kolumny zawierają słowa mylące w zestawach zamkniętych; w dwóch ostatnich kolumnach przedstawiono wyniki testu)

Lp.	Słowo	Słowo mylące 1	Słowo Mylące 2	Słowo mylące 3	Słowo mylące 4	Liczba trafień (19 prób)	% trafień
1	no	ma	na	ci	mi	3	15,8
2	suw	sok	tok	tuz	rów	5	26,3
3	kilo	kipi	halo	halka	pralka	16	84,2
4	hušta	prośba	próba	huba	luba	13	68,4
5	ze	źle	swe	trwa	twa	0	0,0
6	owe	mowa	chowa	sowa	rowek	1	5,3
7	różny	późny	próżny	puka	półka	7	36,8
8	pralka	lalka	gałka	lampa	rampa	9	47,4
9	ty	by	my	me	mnie	4	21,1
10	pas	las	para	mara	larwa	6	31,6
11	myje	gnije	pije	pijar	bojar	5	26,3
12	dawno	darmo	karma	karny	marny	14	73,7
13	my	ty	by	me	mnie	8	42,1
14	biec	móc	wlec	bies	pies	11	57,9
15	sadził	radził	chodził	rama	sama	10	52,6
16	młynarz	piekarz	piernik	plywak	miernik	11	57,9
17	co	to	ta	za	zoo	11	57,9
18	jem	je	Jan	cham	hen	8	42,1
19	tuba	gruba	luba	tuzin	Gruzin	10	52,6
20	warkocz	smarkacz	skoczyć	gonić	skurcz	11	57,9
21	za	co	to	ta	zoo	7	36,8
22	wre	gdzie	kra	gra	brać	8	42,1
23	gonić	warkocz	smarkacz	skoczyć	skurcz	2	10,5
24	gnije	myje	pije	pijar	bojar	4	21,1
25	je	jem	Jan	cham	hen	1	5,3
26	pot	pod	kot	kąt	bat	3	15,8
27	niesie	prosię	Promień	Niemiec	szaniec	4	21,1
28	szkoda	broda	brokuł	szkopał	szkoła	7	36,8
29	te	wre	kra	bo	co	3	15,8

Podczas prezentacji dźwiękowej osoba oceniająca widziała na ekranie komputera listę słów proponowanych ułożoną w przypadkowej kolejności. Wśród tych słów znajdowało się słowo właściwe. Dobór poziomu odtwarzania dźwięków należał do oceniającego i był on pouczony o możliwości dostosowania tego

poziomu do własnych preferencji. Istniała możliwość trzykrotnego powtórzenia prezentacji danego słowa. Swoją decyzję oceniający sygnalizował najeżdżając wskaźnikiem myszy na wybrane słowo i naciskając przycisk myszy.

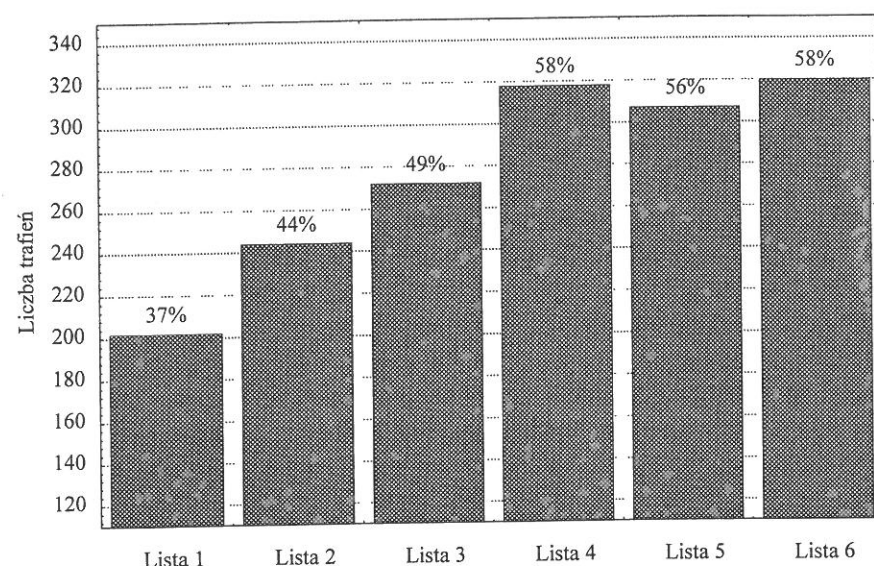


Ryc. 1. Okno dialogowe programu komputerowego nadzorującego prowadzenie testów (zamknięta lista słów pokazana jest w panelu w lewym górnym rogu; suwak po prawej stronie umożliwia dostosowanie poziomu głośności do preferencji słuchacza; przycisk „Odtwórz” umożliwia powtórzną prezentację słowa – do trzech razy)

II. WYNIKI

Najbardziej charakterystycznym wynikiem badań jest zależność, jaką prezentuje wykres na ryc. 2. Wykres ten przedstawia zmianę poziomów trafień w funkcji kolejno prezentowanych list. Lista 1 była prezentowana 19 słuchaczom, którzy pierwszy raz stykali się z pseudoszeptem – liczba trafień była równa 202 przy 551 łącznych prezentacjach, co stanowiło 36,7% poprawnych odgadnięć. Lista numer 6 była oceniana na końcu sesji, gdy słuchacze już przez dłuższy czas odsłuchiwali próbki mowy bezkraniowców. W poprzednich pięciu listach słuchacz próbował odgadnąć $10 \times 29 = 290$ słów, ponieważ każda lista była prezentowana dwukrotnie, przy czym wyniki analizowano dla pierwszej prezentacji. Po takim czasie kontaktu z tą formą mowy procent poprawnych trafień wynosił 58,1%. Średnia liczba trafień rosła wraz ze wzrostem numeru prezentowanej listy, przy czym listy były podawane zawsze w kolejności rosnącej, tj. od numeru 1 do numeru 6. Jak wynika z wykresu, średni wzrost poziomu trafień dla ostatniej listy w stosunku do listy pierwszej wyniósł 21,4%.

Pełny czas ekspozycji słuchacza na dźwięki pseudoszeptu nie może być ustalony, słuchacz bowiem miał prawo trzykrotnie odtworzyć każde słowo.



Ryc. 2. Zależność liczby poprawnych odpowiedzi od kolejnego numeru ocenianej listy (wyniki zsumowane dla 19 słuchaczy)

U poszczególnych słuchaczy wyznaczono metodą regresji liniowej stopień dynamiki przyrostu poprawnych odpowiedzi w miarę osłuchiwania się z mową. Przyrost ten był dodatni u wszystkich słuchaczy. Współczynnik przyrostu liczby trafień na listę wahał się w granicach od 0,46 do 1,8 słowa na listę, przy średnim wzroście 1,24 słowa na listę.

W dwóch ostatnich kolumnach tab. 1 przedstawiono wyniki testu u 19 osób dla pierwszej listy w postaci wartości liczbowych i procentowych.

III. WNIOSKI

Przede wszystkim wyraźnie widoczny jest proces uczenia się rozpoznawania pseudoszeptu przez słuchaczy, co pokazuje rysunek na ryc. 2. W miarę trwania testu, w wyniku osłuchiwania się z mową bezkrtaniowców, liczba poprawnych odpowiedzi zdecydowanie rosła. Zaobserwowane zjawisko jest procesem uczenia się rozumienia nowej formy mowy. W zastosowanym teście uczenie to zachodzi bez nauczyciela, czyli bez sprzężenia zwrotnego, które wskazywałoby właściwe odpowiedzi uczącemu się słuchaczowi. Na tak szybkie

tempo uczenia się fonetycznego na pewno duży wpływ miał młody wiek słuchaczy i ich poziom intelektualny.

Analiza poziomów zrozumiałości poszczególnych słów prowadzi również do interesujących wniosków. Na listach znajdowały się słowa, które niezależnie od metody badania były, statystycznie rzecz biorąc, zupełnie niezrozumiałe. Są to m.in. następujące wyrazy: „ze”, „owe”, „je”, „kije”, „śni”, „ani”. Te w większości jednosylabowce składają się z głosek dźwięcznych, których wymowa sprawia bezkrtaniowcom olbrzymią trudność. Czas trwania wypowiedzi jest krótki, intensywność i energia akustyczna jest stosunkowo niska, duża zaś część informacji tracona.

Na drugim biegunie znajdują się wyrazy, których zrozumienie nie przysporzyło słuchaczom żadnych trudności. Do grupy tej należą m.in.: „kilo”, „siedzi”, „torba”, „harcerz”, „dzień”, „mleko”, „reszta”, „mocno”, „wnuczka”, „dać”, „grzech”. Są to w większości wyrazy dwusylabowe, a więc o dłuższym czasie trwania, większej intensywności oraz energii akustycznej.

W podsumowaniu należy podkreślić, że zaproponowany w pracy system do testowania zrozumiałości mowy bezkrtaniowców okazał się sprawnym, efektywnym i łatwym w użyciu narzędziem. Umożliwia on prowadzenie prac badawczych nad percepcją pseudoszeptu, a również nad formami przetwarzania elektrokustycznego tej mowy dla zwiększenia stopnia jej zrozumiałości.

Bibliografia

- Hoasjoe D. K., Martin G. F., Doyle P. C., Wong F. S. (1992): A comparative acoustic analysis of voice production by near-total laryngectomy and normal laryngeal speakers. „J. Otolaryngol.” 21, 1, 39-43.
- Pastore A., Yuceturk A. V., Trevisi P. (1998): Evaluation of voice and speech following subtotal reconstructive laryngectomy. „Eur. Arch. Otorhinolaryngol.” 255, 7, 371-374.
- Rydell R., Schalen L., Fex S., Elner A. (1995): Voice evaluation before and after laser excision vs. radiotherapy of T1A glottic carcinoma. „Acta Otolaryngol.” (Stockh.) 115, 4, 560-565.
- Stewart M. G., Chen A. Y., Stach C. B. (1998): Outcomes analysis of voice and quality of life in patients with laryngeal cancer. „Arch. Otolaryngol. Head Neck Surg.” 124, 2, 143-148.