LingVaria
X (2015) | 1 (19)
doi: 10.12797/LV.10.2015.19.01

Piotr Rybka
Uniwersytet Śląski, Katowice
piotr.rybka82@gmail.com

JAK BADAĆ SAMOGŁOSKI METODAMI AKUSTYCZNYMI? PROPOZYCJA METODY OPARTEJ NA WZGLĘDNYCH CZĘSTOTLIWOŚCIACH FORMANTOWYCH ORAZ MODELU SAMOGŁOSEK PODSTAWOWYCH. CZ. II[1]

Słowa klucze: fonetyka, samogłoska, fonetyka akustyczna, fonetyka artykulacyjna, artykulacja, formant, samogłoski podstawowe.
Keywords: phonetics, vowel, acoustic phonetics, articulatory phonetics, articulation, formant, cardinal vowels.

Propozycja częstotliwości względnych samogłosek modelowych

Oto proponowany zestaw częstotliwości względnych samogłosek modelu 98-elementowego:

1 Pierwszą część artykułu zob. „LingVaria" 2014, nr 2 (18).

Tabela 4. Względne częstotliwości formantów od I do IV dla samogłosek modelowych płaskich i zaokrąglonych

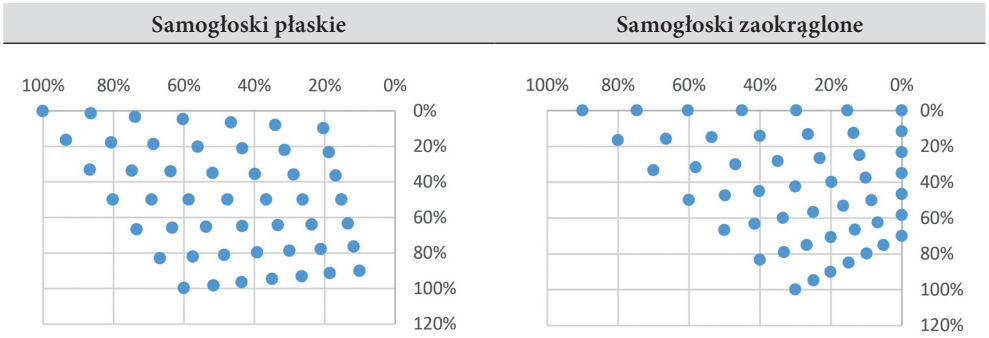
	f'_1	f'_2	f'_3	f'_4		f'_1	f'_2	f'_3	f'_4
	samogłoski płaskie					**samogłoski zaokrąglone**			
i̇	0,0000%	100,0000%	100,0000%	100,0000%	y	0,0000%	90,0000%	15,9810%	0,0000%
ï	1,7000%	86,4586%	91,4214%	93,2000%	ÿ	0,0000%	74,7000%	13,2642%	0,0000%
i̤	3,3000%	73,7137%	83,3475%	86,8000%	ü̟	0,0000%	60,3000%	10,7073%	0,0000%
i̞	5,0000%	60,1723%	74,7689%	80,0000%	u̞	0,0000%	45,0000%	7,9905%	0,0000%
i̱	6,7000%	46,6309%	66,1904%	73,2000%	u̱	0,0000%	29,7000%	5,2737%	0,0000%
ɯ̈	8,3000%	33,8861%	58,1164%	66,8000%	ü	0,0000%	15,3000%	2,7168%	0,0000%
ɯ	10,0000%	20,3447%	49,5379%	60,0000%	u	0,0000%	0,0000%	0,0000%	0,0000%
i̥	16,6667%	93,3333%	93,3333%	100,0000%	y̥	16,6667%	80,0000%	13,3175%	0,0000%
ɪ	17,8000%	80,6322%	87,3179%	94,0500%	ʏ	15,8167%	66,4000%	12,1869%	0,8500%
ɪ̈	18,8667%	68,6781%	81,6562%	88,4500%	ʏ̈	15,0167%	53,6000%	11,1227%	1,6500%
ï̥	20,0000%	55,9769%	75,6408%	82,5000%	ü̟̥	14,1667%	40,0000%	9,9921%	2,5000%
ɯ̽̈	21,1333%	43,2758%	69,6253%	76,5500%	ʊ̈	13,3167%	26,4000%	8,8614%	3,3500%
ɯ̽̈	22,2000%	31,3217%	63,9637%	70,9500%	ʊ	12,5167%	13,6000%	7,7973%	4,1500%
ɯ̥	23,3333%	18,6206%	57,9482%	65,0000%	u̥	11,6667%	0,0000%	6,6667%	5,0000%
e	33,3333%	86,6667%	86,6667%	100,0000%	ø	33,3333%	70,0000%	10,6540%	0,0000%
ë	33,9000%	74,8057%	83,2143%	94,9000%	ø̈	31,6333%	58,1000%	11,1095%	1,7000%
e̥	34,4333%	63,6425%	79,9650%	90,1000%	ø̥	30,0333%	46,9000%	11,5382%	3,3000%
ɘ	35,0000%	51,7816%	76,5126%	85,0000%	ɵ	28,3333%	35,0000%	11,9937%	5,0000%
e̱	35,5667%	39,9206%	73,0602%	79,9000%	ø̱	26,6333%	23,1000%	12,4492%	6,7000%
ɘ̈	36,1000%	28,7574%	69,8109%	75,1000%	ö	25,0333%	11,9000%	12,8778%	8,3000%
ɤ	36,6667%	16,8964%	66,3586%	70,0000%	o	23,3333%	0,0000%	13,3333%	10,0000%
ɐ	50,0000%	80,0000%	80,0000%	100,0000%	ø̞	50,0000%	60,0000%	7,9905%	0,0000%
ɐ̈	50,0000%	68,9793%	79,1107%	95,7500%	ø̞̈	47,4500%	49,8000%	10,0321%	2,5500%
ɐ̞	50,0000%	58,6069%	78,2737%	91,7500%	ɵ̞	45,0500%	40,2000%	11,9536%	4,9500%
ə	50,0000%	47,5862%	77,3845%	87,5000%	ɵ̞	42,5000%	30,0000%	13,9952%	7,5000%
ɐ̱	50,0000%	36,5655%	76,4952%	83,2500%	ɵ̱	39,9500%	19,8000%	16,0369%	10,0500%
ɤ̈	50,0000%	26,1930%	75,6582%	79,2500%	ö̞	37,5500%	10,2000%	17,9584%	12,4500%
ɣ	50,0000%	15,1723%	74,7689%	75,0000%	o̞	35,0000%	0,0000%	20,0000%	15,0000%

	f'_1	f'_2	f'_3	f'_4		f'_1	f'_2	f'_3	f'_4
	samogłoski płaskie					**samogłoski zaokrąglone**			
ɛ	66,6667%	73,3333%	73,3333%	100,0000%	œ	66,6667%	50,0000%	5,3270%	0,0000%
ɛ̈	66,1000%	63,1529%	75,0071%	96,6000%	œ̈	63,2667%	41,5000%	8,9547%	3,4000%
ɜ̟	65,5667%	53,5712%	76,5825%	93,4000%	ɞ̟	60,0667%	33,5000%	12,3691%	6,6000%
ɜ	65,0000%	43,3908%	78,2563%	90,0000%	ɞ	56,6667%	25,0000%	15,9968%	10,0000%
ɜ̠	64,4333%	33,2103%	79,9301%	86,6000%	ɞ̠	53,2667%	16,5000%	19,6246%	13,4000%
ʌ̈	63,9000%	23,6287%	81,5055%	83,4000%	ɔ	50,0667%	8,5000%	23,0389%	16,6000%
ʌ	63,3333%	13,4482%	83,1793%	80,0000%	ɔ̈	46,6667%	0,0000%	26,6667%	20,0000%
æ	83,3333%	66,6667%	66,6667%	100,0000%	ɶ̟	83,3333%	40,0000%	2,6635%	0,0000%
ă̟	82,2000%	57,3264%	70,9036%	97,4500%	ɶ̆	79,0833%	33,2000%	7,8774%	4,2500%
ɐ̟	81,1333%	48,5356%	74,8912%	95,0500%	ɐ̟˕	75,0833%	26,8000%	12,7845%	8,2500%
ɐ	80,0000%	39,1954%	79,1282%	92,5000%	ɐ̟	70,8333%	20,0000%	17,9984%	12,5000%
ɐ̠	78,8667%	29,8552%	83,3651%	89,9500%	ɐ̠˕	66,5833%	13,2000%	23,2123%	16,7500%
ɑ̆	77,8000%	21,0643%	87,3527%	87,5500%	ɒ̆	62,5833%	6,8000%	28,1195%	20,7500%
ɑ̟	76,6667%	11,7241%	91,5896%	85,0000%	ɒ̟	58,3333%	0,0000%	33,3333%	25,0000%
a	100,0000%	60,0000%	60,0000%	100,0000%	ɶ	100,0000%	30,0000%	0,0000%	0,0000%
ä	98,3000%	51,5000%	66,8000%	98,3000%	ɶ̈	94,9000%	24,9000%	6,8000%	5,1000%
ɐ̟˔	96,7000%	43,5000%	73,2000%	96,7000%	ɐ̟˔˔	90,1000%	20,1000%	13,2000%	9,9000%
ɐ̟	95,0000%	35,0000%	80,0000%	95,0000%	ɐ̟	85,0000%	15,0000%	20,0000%	15,0000%
ɐ̠˔	93,3000%	26,5000%	86,8000%	93,3000%	ɐ̠˔˔	79,9000%	9,9000%	26,8000%	20,1000%
ɑ̈	91,7000%	18,5000%	93,2000%	91,7000%	ɒ̈	75,1000%	5,1000%	33,2000%	24,9000%
ɑ	90,0000%	10,0000%	100,0000%	90,0000%	ɒ	70,0000%	0,0000%	40,0000%	30,0000%

Opracowanie własne.

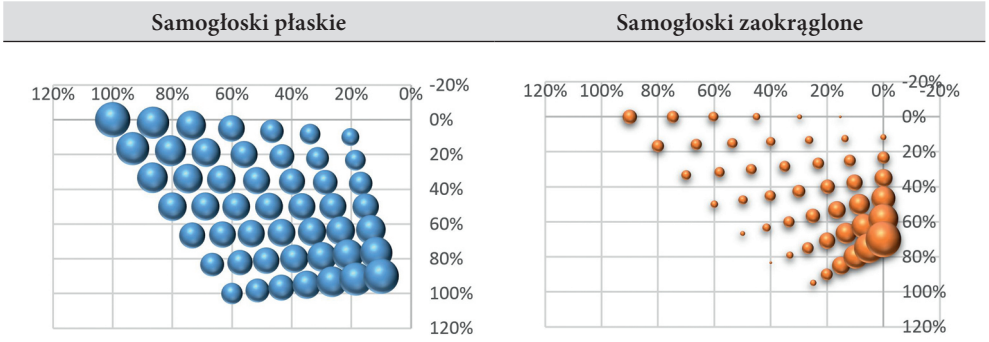
Poniższe wykresy przedstawiają regularne rozłożenie częstotliwości względnych powyższych modeli:

Wykres 6. Rozłożenie punktów samogłoskowych modeli



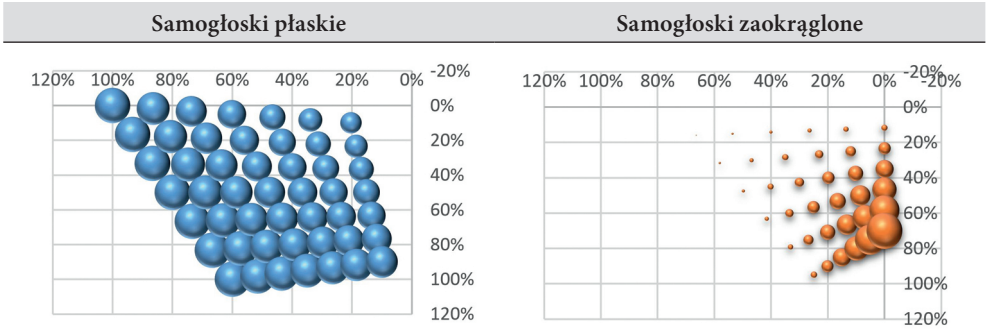
Opracowanie własne. Osie OX – względne częstotliwości F2, osie OY – względne częstotliwości F1.

Wykres 7. Rozłożenie punktów samogłoskowych modeli z uwzględnieniem F3



Opracowanie własne. Osie OX – względne częstotliwości F2, osie OY – względne częstotliwości F1, wielkość bąbli: względna częstotliwość F3 w skali 1:2 (brak bąbla oznacza wartość 0%).

Wykres 8. Rozłożenie punktów samogłoskowych modeli z uwzględnieniem F4



Opracowanie własne. Osie OX – względne częstotliwości F2, osie OY – względne częstotliwości F1, wielkość bąbli: względna częstotliwość F4 w skali 1:2 (brak bąbla oznacza wartość 0%).

Test metody częstotliwości względnych

Trudno przeprowadzić test proponowanej tutaj metody, ponieważ brakuje odpowiedniego punktu odniesienia, to znaczy takiego sposobu opisu samogłosek, który umożliwiłby równie precyzyjne rozpoznanie[2]. Ponieważ zamieniamy dane akustyczne na opis artykulacji, wydawać by się mogło, że najlepszym sposobem sprawdzenia prezentowanej tu metody będzie porównanie uzyskanych za jej pomocą wyników z rezultatem analizy artykulacji. Przedstawiona tutaj metoda częstotliwości względnych pozwala na obiektywne przypisanie danemu zbiorowi wektorów względnych częstotliwości formantowych samogłoski badanej wektora względnych częstotliwości formantowych określonej samogłoski modelowej. Można powiedzieć, że w istocie wykonujemy pomiar samogłoski, ponieważ mierzymy odległość między punktem tej samogłoski opisanym czterema względnymi częstotliwościami formantowymi a punktami, których współrzędne są także względnymi częstotliwościami formantowymi ustalonych wcześniej samogłosek modelowych. Ustalenie najmniejszej odległości między tymi punktami można rozumieć jako rozpoznanie najmniejszej różnicy między samogłoską badaną a danym modelem, co pozwala rozpoznać tę samogłoskę jako najbardziej zbliżoną do danego modelu. W dalszej kolejności można przypisać tej samogłosce symbol najbliższego modelu, co jednocześnie będzie oznaczało przypisanie tej samogłosce zbioru etykiet określających artykulację tej samogłoski.

2 Porównywalną dokładność może zagwarantować użycie pomiarów antropometrycznych, których zastosowanie proponuję w osobnym artykule (Rybka w druku), złożonym do druku już po opublikowaniu pierwszej części niniejszego tekstu.

Miarodajne sprawdzenie tej metody wymagałoby zastosowania analogicznej metody umożliwiającej pomiar samej artykulacji. O ile jesteśmy w stanie dokładnie zarejestrować artykulację samogłoski (najlepszą metodą wydaje się artykulografia i rentgenografia), o tyle nie istnieje sposób tak dokładnego i obiektywnego opisu ułożenia artykulatorów, który pozwalałby na jednoznaczne przypisanie mu któregoś z 98 symboli fonetycznych z założonego tu modelu. Metoda takiego opisu musiałaby uwzględniać różnice w budowie anatomicznej (a w przypadku zdjęć rentgenowskich także różnice w ułożeniu całej głowy podczas wykonywania zdjęcia). Naniesienie na układ artykulatorów siatki punktów oznaczających artykulacje każdej z 49[3] samogłosek modelowych wymagałoby także ustalenia pewnych względnych (ustalanych tak samo przy indywidualnym aparacie mowy) punktów odniesienia, oznaczających skrajne położenie najwyższego punktu na powierzchni języka. Określanie tego punktu również wymagałoby zaproponowania jakiejś metody, ponieważ na wielu obrazach samogłosek (i spółgłosek) artykułowanych przy niskim ułożeniu języka powierzchnia tego organu jest w części grzbietowej bardzo płaska. Nie można więc, bez przyjęcia pewnych założeń[4], określić położenia najwyższego punktu na powierzchni języka w przypadku tak artykułowanych głosek.

Mimo powyższych trudności można jednak zaproponować sposób szacunkowej oceny poprawności uzyskiwanych wyników przy zastosowaniu metody częstotliwości względnych. Można bowiem, stosując wspomnianą metodę, przeanalizować zbadane już artykulacyjnie samogłoski lub samogłoski dobrze opisane pod względem artykulacyjnym (np. samogłoski podstawowe inne od tych wykorzystanych do ustalenia samogłosek modelowych). Uzyskane wyniki nie powinny być identyczne z wynikami uzyskanymi przy użyciu innych metod badawczych lub opisami podanymi przez samych informatorów, ale nie powinny być również skrajnie różne.

By wykonać tak obmyślane porównanie, musimy ustalić, co rozumiemy przez „skrajnie różne" wyniki. Albo inaczej: gdzie leży granica, poza którą dane rozpoznanie wykonane metodą częstotliwości względnych uznamy za wysoce wątpliwe. Na pewno z taką sytuacją będziemy mieli do czynienia wtedy, gdy w wyniku analizy metodą częstotliwości względnych uzyskalibyśmy np. samogłoskę [ɒ], podczas gdy badaną samogłoską była głoska [i], a przynajmniej taką głoskę starał się wymówić dany informator. Jeśli bowiem wykorzystamy do testów wymówienia wykształconych fonetyków, trudno będzie przypuszczać, że któryś z nich pomylił samogłoskę wysoką, przednią, płaską z niską, tylną, zaokrągloną. Z drugiej jednak strony nie możemy oczekiwać idealnych wymówień, skoro już wcześniej stwierdziliśmy zna-

3 Liczba wszystkich modeli jest tu zredukowana o połowę, gdyż można oddzielnie analizować artykulację wargową, a oddzielnie ułożenie języka.

4 Przykładowe założenie może polegać na dzieleniu na połowę, tzw. bisekcji, odcinka leżącego na stycznej do powierzchni grzbietowej języka i mającej najwięcej punktów wspólnych z tą powierzchnią.

czące różnice w artykulacji samogłosek podstawowych wykonanych przez fonetyków brytyjskich. Konieczne jest więc zachowanie pewnego marginesu, powyżej którego rozpoznanie zostanie uznane za błędne, a poniżej – za dopuszczalne.

Żeby jednak nie porównywać za każdym razem symboli fonetycznych (zwłaszcza że przy zastosowaniu 98-elementowego zbioru modeli „ręczne" porównywanie wyników byłoby utrudnione), konieczne jest zaproponowanie jakiegoś sposobu pomiaru różnicy między samogłoskami. Porównywane będą w istocie dwa symbole fonetyczne, a dokładniej dwa zestawy terminów fonetycznych opisujących artykulację danej samogłoski, których skrótowym zapisem jest właśnie dany symbol. Możemy więc każdy taki termin zamienić na liczbę naturalną według określonej reguły, którą niech będzie nasilenie lub osłabienie pewnej cechy fonetycznej. Na przykład dla obniżenia języka w modelu 98-elementowym można przypisać liczby od 0 do 6 dla każdej pozycji języka w kierunku pionowym (0 dla maksymalnie wysokiej, 6 dla maksymalnie niskiej, 3 dla dokładnie pośredniej). Podobnie dla cofnięcia języka: 0 – minimalnie tylne ułożenie języka (a więc maksymalnie przednie), 6 – maksymalnie tylne, 3 – dokładnie pośrednie ułożenie. Zaokrągleniu również można przypisać wiele wartości pośrednich, ale w modelu przyjęliśmy dwie wartości: obecność i brak zaokrąglenia. Niech więc 0 oznacza brak labializacji, a 1 jej obecność[5].

Uzyskaliśmy w ten sposób następującą matrycę cech artykulacyjnych dla modelu 98-elementowego (pierwsza liczba kodu oznacza położenie języka w pionie, druga – w poziomie, trzecia – labializację):

Tabela 5. Kody określające cechy artykulacyjne samogłosek modelowych

Samogłoski płaskie							Kody samogłosek						
i	ï	i̟	ɨ	i̠	ɯ̈	ɯ	000	010	020	030	040	050	060
i̟	ɪ	ï	i̟	ɯ̟̈	ɯ̟̈	ɯ̟	100	110	120	130	140	150	160
e	ë	e̞	ɘ	ɘ̠	ɤ̈	ɤ	200	210	220	230	240	250	260
e̟	ë̟	e̞	ɘ	ɘ̠	ɤ̞̈	ɤ̞	300	310	320	330	340	350	360
ɛ	ë	ɜ̞	ɜ	ɜ̠	ʌ̈	ʌ	400	410	420	430	440	450	460
æ	ä̟	a̞	ɐ	a̞̠	ɑ̟̈	ʌ̟	500	510	520	530	540	550	560
a	ä	a̞˕	a̞	a̞̠˕	ɑ̈	ɑ	600	610	620	630	640	650	660

5 Podobne rozwiązanie zaproponowano w (Łobacz et al. 2003: 16).

Samogłoski zaokrąglone							Kody samogłosek						
y	ÿ	ʉ̟	ʉ	ʉ̠	ü	u	001	011	021	031	041	051	061
y̟	ʏ	ʏ̈	ʉ̟	ʊ̈	ʊ	ʊ̠	101	111	121	131	141	151	161
ø	ø̈	ø̟	ɵ	ø̠	ö	o	201	211	221	231	241	251	261
ø̟	ø̟̈	ø̟ˠ	ø̟	ø̠ˠ	ö̟	o̟	301	311	321	331	341	351	361
œ	œ̈	œ̟	ɞ	ɞ̠	ɔ	ɔ	401	411	421	431	441	451	461
œ̟	œ̆	œ̟ˠ	ɞ̟	ɞ̠ˠ	ɔ̆	ɔ̟	501	511	521	531	541	551	561
ɶ	ɶ̈	ɶ̟ˠˠ	ɞ̟̟	ɞ̟ˠˠ	ɔ̈	ɒ	601	611	621	631	641	651	661

Opracowanie własne.

Różnicę artykulacyjną między dwiema samogłoskami można dzięki powyższemu rozwiązaniu przedstawić jako sumę modułów różnic poszczególnych elementów kodu każdej samogłoski. Na przykład różnica między [i] a [y] wynosi 1, ponieważ:

Tabela 6. Sposób obliczania różnicy artykulacyjnej $r = 1$ między samogłoskami [i y]

	[i]		[y]	**Wynik (moduł różnicy)**
Różnica między wysokością (niskością) samogłosek	0	–	0	$\|0 - 0\| = 0$
Różnica między przedniością (tylnością) samogłosek	0	–	0	$\|0 - 0\| = 0$
Różnica między labializacją samogłosek	0	–	1	$\|0 - 1\| = \|-1\| = 1$
Suma różnic:				$0 + 0 + 1 = 1$

Opracowanie własne.

Różnica między [i] a [ə] to 6:

Tabela 7. Sposób obliczania różnicy artykulacyjnej $r = 6$ między samogłoskami [i ə]

	[i]		[ə]	**Wynik (moduł różnicy)**
Różnica między wysokością (niskością) samogłosek	0	–	3	$\|0 - 3\| = \|-3\| = 3$
Różnica między przedniością (tylnością) samogłosek	0	–	3	$\|0 - 3\| = \|-3\| = 3$
Różnica między labializacją samogłosek	0	–	0	$\|0 - 0\| = 0$
Suma modułów różnic:				$3 + 3 + 0 = 6$

Opracowanie własne.

Maksymalna różnica zachodzi między [i] a [ɒ] (i odwrotnie: między [y] a [ɑ]):

Tabela 8. Sposób obliczania różnicy artykulacyjnej $r = 13$ między samogłoskami [i ɒ]

	[i]		[ɒ]	**Wynik (moduł różnicy)**
Różnica między wysokością (niskością) samogłosek	0	–	6	$\|0 - 6\| = \|-6\| = 6$
Różnica między przedniością (tylnością) samogłosek	0	–	6	$\|0 - 6\| = \|-6\| = 6$
Różnica między labializacją samogłosek	0	–	1	$\|0 - 1\| = \|-1\| = 1$
Suma modułów różnic:				$6 + 6 + 1 = 13$

Opracowanie własne.

Tabela 9. Sposób obliczania różnicy artykulacyjnej $r = 13$ między samogłoskami [y ɑ]

	[y]		[ɑ]	**Wynik (moduł różnicy)**
Różnica między wysokością (niskością) samogłosek	0	–	6	$\|0 - 6\| = \|-6\| = 6$
Różnica między przedniością (tylnością) samogłosek	0	–	6	$\|0 - 6\| = \|-6\| = 6$
Różnica między labializacją samogłosek	1	–	0	$\|0 - 1\| = \|-1\| = 1$
Suma modułów różnic:				$6 + 6 + 1 = 13$

Opracowanie własne.

Znając maksymalną różnicę artykulacyjną wyrażoną liczbami całkowitymi, możemy określić procentową różnicę między dwiema samogłoskami: będzie to iloraz sumy modułów różnic między poszczególnymi elementami kodu każdej samogłoski (cechami artykulacyjnymi zamienionymi na liczby naturalne) i liczby 13 pomnożony przez 100%:

Wzór 17:

$$r_{\%} = \left({}^{1}/_{13} \cdot \sum \left| c_{1n} - c_{2n} \right| \right) \cdot 100\%$$

gdzie $r_{\%}$ to różnica artykulacyjna między samogłoskami wyrażona procentowo, c_{1n} to n-ta cecha głoski 1. zamieniona na liczbę naturalną z przedziału [0; 6], a c_{2n} to n-ta cecha głoski 2. także zamieniona na liczbę naturalną z podanego wcześniej przedziału.

Obliczone wyżej różnice artykulacyjne odpowiadają więc następującym wartościom procentowym:

Tabela 10. Porównanie bezwzględnych (r) i względnych ($r_{\%}$) wartości różnic artykulacyjnych między samogłoskami [i y], [i ə] oraz [i ɒ] i [y ɑ]

Porównywane głoski:	[i y]	[i ə]	[i ɒ], [y ɑ]
r	1	6	13
$r_{\%}$	≈ 7,69%	≈ 46,15%	100%

Opracowanie własne.

Zdefiniowanie różnicy artykulacyjnej wyrażonej procentowo ułatwia nam ustalenie wspomnianego wcześniej progu akceptowalności rozpoznania samogłoski. Powyższe obliczenia dla pary samogłosek [i] i [ə] pokazują, że gdyby parlator pomylił samogłoskę wysoką, przednią ze średnią, środkową, co można byłoby już uznać za dość znaczny błąd, wówczas wartość $r_{\%}$ byłaby większa od 40%. Wartość $r_{\%}$ = 40% nigdy nie zostanie uzyskana w przypadku modelu 98-elementowego, gdyż nawet niewielkie zbliżenie [ə] do [i], spowoduje zmianę $r_{\%}$ o niecałe 8%, tak więc różnica między [i] a [ᵩ] oraz między [i] a [ɘ] będzie wynosiła 46,15% – 7,69%, czyli 38,46%. Możemy więc właśnie 40% przyjąć jako granicę akceptowalności wyników rozpoznania samogłosek, gdyż dzięki temu nie będziemy musieli rozstrzygać kwestii, czy rozpoznanie, w przypadku którego $r_{\%}$ = 40% , uznać za poprawne czy błędne.

W teście omawianej tu metody wykorzystano wymówienia samogłosek podstawowych w wykonaniu fonetyków. W jednym przypadku (Wiktor Jassem) skorzystano z bezpośrednich danych: częstotliwości formantowych (bezwzględnych) podanych przez badacza. W pozostałych przypadkach bezwzględne częstotliwości formantowe ustalano przy wykorzystaniu programu Praat. Następnie zamieniano częstotliwości bezwzględne na względne, korzystając z podanego wcześniej wzoru (wzór 6). W dalszej kolejności obliczano odległość między punktem badanej samogłoski opisanej częstotliwościami względnymi a najbardziej podobnym modelem, jaki został ustalony w porównaniu wykonanym opisywaną tu metodą. Po ustaleniu najmniejszej odległości, a więc zarazem największego podobieństwa, przypisywano badanej samogłosce symbol danego modelu.

Następnym krokiem była zamiana symboli badanych samogłosek przypisanych im w materiale źródłowym lub przez mówcę oraz przypisanych im podczas rozpoznawania symboli – na kody artykulacyjne, a w końcu obliczenie różnic między samogłoską, jaka miała być wymówiona, a jaka została rozpoznana. Wyniki zawarto w poniższej tabeli:

Tabela 11. Wyniki testu metody częstotliwości względnych dla wymówień samogłosek podstawowych trzech informatorów

Mówca (źródło)	Częstotliwości bezwzględne				Częstotliwości względne				Samogłoska badana		Wynik rozpoznania		$r_{\%}$	P_r
	f1	f2	f3	f4	f'1	f'2	f'3	f'4	Symbol	Kod	Symbol	Kod		
W.J.	210,00	2750,00	3500,00	4200,00	0,0%	100,0%	100,0%	100,0%	i	000	i	000	0,00%	100,0%
W.J.	380,00	2630,00	3050,00	3650,00	25,8%	94,4%	65,4%	50,0%	e	200	e	200	0,00%	24,7%
W.J.	590,00	2280,00	2700,00	3600,00	57,6%	78,0%	38,5%	45,5%	ɛ	400		–	–	0,8%
W.J.	870,00	1750,00	2700,00	3650,00	100,0%	53,2%	38,5%	50,0%	a	600	a	600	0,00%	43,0%
W.J.	800,00	1050,00	2720,00	3500,00	89,4%	20,4%	40,0%	36,4%	ɑ	660	ɐ˕˕	641	23,08%	17,9%
W.J.	570,00	940,00	2700,00	3300,00	54,5%	15,2%	38,5%	18,2%	ʌ	460	ɒ̆	551	23,08%	11,6%
W.J.	450,00	850,00	2500,00	3100,00	36,4%	11,0%	23,1%	0,0%	ɤ	260	ɵ	351	23,08%	91,3%
W.J.	280,00	850,00	2250,00	3150,00	10,6%	11,0%	3,8%	4,5%	ɯ	060	ʊ	151	23,08%	92,3%
W.J.	240,00	1550,00	2400,00	3300,00	4,5%	43,8%	15,4%	18,2%	ɨ	030	ʉ	031	7,69%	51,7%
W.J.	220,00	2550,00	3100,00	3900,00	1,5%	90,6%	69,2%	72,7%	y	001	ɻ	020	23,08%	4,2%
W.J.	350,00	2320,00	2600,00	3380,00	21,2%	79,9%	30,8%	25,5%	ø	201	ɣ̞	101	7,69%	49,0%
W.J.	520,00	1950,00	2500,00	3500,00	47,0%	62,5%	23,1%	36,4%	œ	401	ø̞	301	7,69%	30,3%
W.J.	790,00	1650,00	2600,00	3640,00	87,9%	48,5%	30,8%	49,1%	œ	601	ɶ̆	511	15,38%	5,5%
W.J.	710,00	900,00	2850,00	3350,00	75,8%	13,3%	50,0%	22,7%	ɒ	661	ɒ	661	0,00%	11,9%
W.J.	550,00	820,00	2500,00	3200,00	51,5%	9,6%	23,1%	9,1%	ɔ	461	ɔ	451	7,69%	99,4%
W.J.	400,00	730,00	2300,00	3200,00	28,8%	5,4%	7,7%	9,1%	o	261	ö	251	7,69%	8,2%
W.J.	270,00	615,00	2200,00	3150,00	9,1%	0,0%	0,0%	4,5%	u	061	ụ	161	7,69%	44,7%
W.J.	270,00	1370,00	2500,00	3350,00	9,1%	35,4%	23,1%	22,7%	ʉ	031	ʉ̞	131	7,69%	30,8%
L.P.A.	272,04	2383,25	3074,27	3533,53	1,9%	100,0%	100,0%	61,9%	i	000	i	000	0,00%	100,0%

Mówca (źródło)	Częstotliwości bezwzględne				Częstotliwości względne				Samogłoska badana		Wynik rozpoznania		$r_{\%}$	P_r
	f1	f2	f3	f4	f'1	f'2	f'3	f'4	Symbol	Kod	Symbol	Kod		
L.P.A.	398,99	2161,51	2963,77	3874,69	20,3%	87,5%	84,9%	100,0%	e	200	ɪ	110	15,38%	60,6%
L.P.A.	604,26	1776,59	2634,88	3352,81	50,2%	65,7%	39,9%	41,7%	ɛ	400	ø	301	15,38%	9,2%
L.P.A.	946,86	1618,87	2774,67	3323,14	100,0%	56,8%	59,0%	38,4%	a	600	a	600	0,00%	97,0%
L.P.A.	667,21	898,29	2965,70	3316,46	59,3%	16,1%	85,2%	37,7%	ɑ	660	ʌ	460	15,38%	84,2%
L.P.A.	508,81	1023,79	2861,91	3342,00	36,3%	23,2%	71,0%	40,5%	ɔ	461	ɤ̞	250	30,77%	55,0%
L.P.A.	380,02	759,03	2531,91	3134,12	17,6%	8,2%	25,8%	17,3%	o	261	ö	251	7,69%	7,9%
L.P.A.	259,07	613,99	2772,68	3152,21	0,0%	0,0%	58,8%	19,4%	u	061	ɯ	060	7,69%	37,8%
L.P.A.	265,57	2012,95	2795,45	3486,22	0,9%	79,1%	61,9%	56,6%	y	001	ɨ	020	23,08%	8,6%
L.P.A.	373,57	1860,80	2614,95	3362,45	16,6%	70,5%	37,2%	42,8%	ø	201	ʏ	111	15,38%	2,8%
L.P.A.	530,96	1630,09	2342,85	2978,79	39,5%	57,4%	0,0%	0,0%	œ	401	ø	301	7,69%	3,4%
L.P.A.	910,24	1480,82	2673,69	3327,83	94,7%	49,0%	45,2%	39,0%	œ	601	a	600	7,69%	27,0%
L.P.A.	711,36	909,99	2910,00	3234,74	65,8%	16,7%	77,5%	28,6%	ɒ	661	ʌ	460	23,08%	31,0%
L.P.A.	718,94	864,08	2914,52	3325,83	66,9%	14,1%	78,2%	38,7%	ʌ	460	ʌ	460	0,00%	78,6%
L.P.A.	396,66	948,69	2738,10	3220,01	20,0%	18,9%	54,0%	26,9%	ɤ	260	ɯ̟	160	7,69%	91,1%
L.P.A.	272,18	989,24	2667,59	3150,86	1,9%	21,2%	44,4%	19,2%	ɯ	060	ɯ	060	0,00%	89,3%
P.R.	202,81	2203,04	3009,41	3502,57	0,0%	96,5%	93,8%	43,2%	i	000	i	000	0,00%	64,8%
P.R.	263,60	1793,97	2027,66	2926,97	13,3%	72,0%	32,8%	12,7%	y	001	y̟	101	7,69%	1,9%
P.R.	261,52	2163,01	2869,09	3320,54	12,8%	94,1%	85,1%	33,6%	i	000	ɨ	100	7,69%	73,2%
P.R.	361,70	794,56	2303,98	3135,93	34,8%	12,3%	50,0%	23,8%	ɯ	060	ɯ̟	160	7,69%	22,1%
P.R.	587,53	884,72	1876,61	2903,57	84,2%	17,7%	23,4%	11,5%	a	600	ɐ̟	631	30,77%	91,5%

Mówca (źródło)	Częstotliwości bezwzględne				Częstotliwości względne				Samogłoska badana		Wynik rozpoznania		$r_{\%}$	P_r
	f1	f2	f3	f4	f'1	f'2	f'3	f'4	Symbol	Kod	Symbol	Kod		
P.R.	602,95	1163,39	2910,09	3676,72	87,6%	34,4%	87,7%	52,5%	ɑ	600	ɐ̟˞	640	30,77%	18,0%
P.R.	458,05	1550,83	2160,80	2822,79	55,9%	57,5%	41,0%	7,2%	ə	330	ɵ̞	311	23,08%	3,7%
P.R.	245,54	2261,96	3108,52	4573,63	9,4%	100,0%	100,0%	100,0%	i	000	i	000	0,00%	45,2%
P.R.	280,06	1979,33	2633,02	2950,81	16,9%	83,1%	70,4%	14,0%	i	000	ɪ	110	15,38%	14,5%
P.R.	289,21	1882,36	2206,70	2699,42	18,9%	77,3%	43,9%	0,7%	y	001	ʏ̞	101	7,69%	17,8%
P.R.	277,24	1600,90	1942,51	2686,72	16,3%	60,5%	27,5%	0,0%	y	001	ʏ	111	15,38%	16,2%
P.R.	359,23	795,85	2257,08	3074,35	34,2%	12,4%	47,0%	20,5%	ɯ	060	ɯ̟	160	7,69%	35,2%
P.R.	321,56	588,51	2255,00	3000,00	26,0%	0,0%	46,9%	16,6%	u	061	ɯ̟	160	15,38%	33,8%
P.R.	657,81	1419,26	2556,97	3931,91	99,6%	49,6%	65,7%	66,0%	a	600	ä	610	7,69%	99,2%
P.R.	615,62	1297,33	2254,37	3579,38	90,3%	42,4%	46,9%	47,3%	a	600	ä	610	7,69%	5,8%
P.R.	580,36	1119,55	2435,82	3349,85	82,6%	31,7%	58,2%	35,1%	œ	601	ɐ	530	38,46%	11,3%
P.R.	579,08	1304,16	2509,15	3466,84	82,4%	42,8%	62,7%	41,3%	œ	601	ɐ̟	520	30,77%	39,8%
P.R.	659,72	897,05	2889,58	3468,46	100,0%	18,4%	86,4%	41,4%	ɑ	660	ɐ̟˞	640	15,38%	4,7%
P.R.	634,18	829,97	2863,49	3653,14	94,4%	14,4%	84,8%	51,2%	ɑ	660	ɑ̈	650	7,69%	43,2%
P.R.	593,12	804,52	2859,92	3445,20	85,4%	12,9%	84,5%	40,2%	ɒ	661	ʌ̟	560	15,38%	3,6%
P.R.	470,94	674,40	2924,91	3614,29	58,7%	5,1%	88,6%	49,2%	ɒ	661	ʌ	460	23,08%	80,7%
P.R.	445,20	1601,00	2327,25	3163,68	53,0%	60,5%	51,4%	25,3%	ə	330	–	–	–	0,1%
P.R.	443,34	1530,33	2215,16	2888,87	52,6%	56,3%	44,4%	10,7%	ə	330	–	–	–	0,5%
P.R.	434,86	1462,40	2128,25	2851,95	50,8%	52,2%	39,0%	8,8%	ə	330	ɵ̞	311	23,08%	6,0%
P.R.	388,16	1414,29	2082,20	2793,79	40,6%	49,3%	36,2%	5,7%	ə	330	ɐ̟˞	321	15,38%	4,7%

Mówca (źródło)	Częstotliwości bezwzględne				Częstotliwości względne				Samogłoska badana		Wynik rozpoznania		$r_{\%}$	P_r
	f1	f2	f3	f4	f'1	f'2	f'3	f'4	Symbol	Kod	Symbol	Kod		
P.R.	393,15	759,35	2024,73	2863,14	41,7%	10,2%	32,6%	9,3%	ɯ	060	–	–	–	0,2%
P.R.	605,76	1288,61	2336,19	3603,81	88,2%	41,8%	52,0%	48,6%	a	600	ä	610	7,69%	23,1%
P.R.	526,32	756,58	1501,04	2811,53	70,8%	10,0%	0,0%	6,6%	ɑ	660	ɐ̟	531	38,46%	8,9%
P.R.	540,85	1492,57	2036,73	2816,63	74,0%	54,0%	33,3%	6,9%	ə	330	œ	401	38,46%	1,4%

Opracowanie własne na podstawie własnych wymówień (P.R.), danych zawartych w literaturze (W.J. – Wiktor Jassem, za: Jassem 1973: 190) oraz opublikowanych w Internecie (L.P.A. – dr Leong Ping Alvin pracujący w Nanyang Technological University; nagrania pochodzą z jego prywatnej strony http://www.alvinleong.info/sounds/cardinals.html).

Do badania wprowadzono jeszcze jedną wartość, którą nazwano roboczo „**pewnością rozpoznania**", a która opisuje położenie punktu samogłoski badanej między dwoma najbliższymi modelami. Teoretycznie możliwe są bowiem następujące sytuacje:

a) punkt samogłoski badanej pokrywa się z najbliższym modelem;
b) punkt samogłoski badanej nie pokrywa się z najbliższym modelem, ale:
   1) leży bardzo blisko niego;
   2) znajduje się gdzieś między najbliższym modelem a połową odległości między najbliższym modelem a drugim w kolejności najbliższym modelem;
   3) znajduje się mniej więcej w połowie odległości między dwoma modelami;
   4) znajduje się dokładnie w połowie odległości między dwoma modelami.

Dokładne pokrywanie się punktu z modelem (sytuacja *a*) jest praktycznie niemożliwe w przypadku częstotliwości bezwzględnych: oznaczałoby to, iż danemu mówcy udało się nie tylko odtworzyć takie samo ułożenie artykulatorów, jakie charakteryzowało innego informatora, którego wymówienia porównujemy, ale także iż obaj mają aparaty mowy o identycznych właściwościach filtracyjnych. A więc jest to sytuacja być może równie prawdopodobna, co posiadanie identycznych linii papilarnych przez dwie niespokrewnione osoby. Operujemy jednak na częstotliwościach względnych, więc nie można takiej sytuacji wykluczyć (dane w tabeli nr 11 pokazują, że sytuacja *a* zdarzyła się kilka razy).

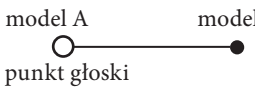
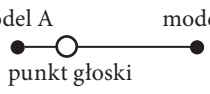
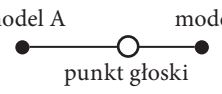
Częstsze powinny być sytuacje typu *b*, wśród których typ 4. jest również mało prawdopodobny: oznacza idealne trafienie w punkt między dwoma modelami wszystkimi czterema częstotliwościami. Taki wynik, choć teoretyczny, nie pozwala na przypisanie badanej samogłosce symbolu żadnego z modeli: jest ona tak samo podobna do dwu różnych samogłosek modelowych, jak i różna od nich.

Najbardziej prawdopodobne są sytuacje *b1*, *b2* i *b3*. Przypadek *b1* jest najbardziej pożądany: oznacza największe podobieństwo między głoską badaną a modelem. Sytuacja *b2* jest nieco gorsza, ale dalej wskazuje na podobieństwo badanej samogłoski do najbliższego modelu.

Sytuacje typu *b2*, *b3* i *b4* wymagają osobnego rozważenia, ponieważ operując na liczbach rzeczywistych, musimy znać konkretną, wymierną granicę oddzielającą sytuacje, w których rozpoznanie możemy uznać za pewne. Innymi słowy, musimy odpowiedzieć na pytanie, co to znaczy, że punkt samogłoski badanej znajduje się „mniej więcej" w połowie odległości między dwoma modelami, a więc gdzie leży granica między sytuacjami *b2*, *b3* i *b4*.

Zilustrujmy sytuacje *a*, *b2* i *b4* w sposób ukazany na rysunku nr 2 (pominięto dla większej przejrzystości przypadki *b1* i *b3*, ponieważ sytuacja *b1* jest bardzo podobna do *a*, natomiast przypadek *b3* jest podobny do *b4*):

Rysunek 2. Możliwe położenia punktu samogłoski badanej między dwoma modelami

Sytuacja *a*	Sytuacja *b2*	Sytuacja *b4*
		

Opracowanie własne.

Obierzmy następujące wartości:

x – odległość między punktem głoski a najbliższym modelem (w przykładzie na rysunku nr 2 – model A);

y – odległość między punktem głoski a drugim najbliższym modelem (w przykładzie powyżej model B);

l – odległość między modelami A i B obliczana według wzoru nr 1 lub przez zsumowanie x i y;

P_r – wyrażoną w procentach bliskość punktu głoski do modelu A, a zarazem odległość od połowy odległości między modelami A i B, tj. $\frac{1}{2}l$

W sytuacji *a* wartość x będzie równa 0; w sytuacjach *b2* i *b4* będzie to jakaś wartość różna od zera, spróbujmy jednak w wyniku obliczeń uzyskać zero w sytuacji *b4*, ponieważ ta sytuacja, choć jedynie teoretycznie możliwa, z całą pewnością nie upoważnia do przypisania badanej głosce symbolu czy to modelu A, czy to modelu B. Zero w sytuacji *b4* uzyskamy jako wynik różnicy: $y - x$ (bo wówczas $y = x$).

Sytuacji *a* spróbujmy przypisać maksymalną możliwą wartość w danej sytuacji. Może to być wartość $l = y = y - x$ (skoro $x = 0$).

Odległości między punktami samogłosek modelowych są różne (większe w przypadku samogłosek wysokich, mniejsze w przypadku niskich), dlatego uzasadnione jest operowanie nie na wartościach bezwzględnych, ale względnych, niezależnych od rzeczywistej odległości l między modelami. Może je stanowić wartość procentowa P_r, którą proponuje się obliczać według wzoru:

Wzór 18:

$$P_r = \left(\frac{y - x}{y + x}\right) \cdot 100\% = \left(\frac{y - x}{l}\right) \cdot 100\%$$

Możemy więc sytuacji *a* przypisać $P_r = 100\%$, a sytuacji *b4*: $P_r = 0\%$. Nie wiemy dalej, gdzie leży wymierna granica między sytuacjami *b3* i *b4*, ale mamy narzędzie pozwalające ustalić granicę między rozpoznaniami akceptowanymi (lub pewnymi) a nieakceptowanymi (niepewnymi).

Przyjrzyjmy się uzyskanym wartościom P_r w przeprowadzonym teście opisywanej tu metody (tabela nr 12):

Tabela 12. Liczby i odsetki rozpoznań przy różnych minimalnych wartościach P_r pozwalających uwzględnić rozpoznanie

Zakres P_r:	**[100; 90%]**	**(90%; 80%]**	**(80%; 70%]**	**(70% 60%]**	**(60%; 50%]**	**(50%; 40%]**	**(40%; 30%]**	**(30%; 20%]**	**(20%; 10%]**	**(10%; 1%]**	**Σ rozpoznań:**
n rozpoznań dla danego zakresu P_r:	9	3	2	2	2	5	7	4	8	21	N = 63
% rozpoznań dla danego zakresu P_r:	14%	5%	3%	3%	3%	8%	11%	6%	13%	33%	100%

Opracowanie własne.

Jak widać, najwięcej rozpoznań (n = 21, 1/3 wyników) znajduje się między punktami modeli (choć nie dokładnie między nimi, gdyż $P_r > 0$ przy 21 rozpoznaniach). Dużą liczbę rozpoznań bliskich danemu modelowi ($P_r \geq 90\%$) można wytłumaczyć budową metody: punktem odniesienia są wymówienia skrajne, które uznaje się za takie same u różnych parlatorów, wobec tego odnalezienie skrajnych częstotliwości formantowych jest nieuniknione. Jeśli pominąć te wyniki, okaże się, iż większość rozpoznań charakteryzuje się wartością P_r mniejszą niż 50%. *Notabene* średnia wartość P_r wyniosła 34,9%. Trudno zresztą oczekiwać, by badane samogłoski akustycznie pośrednie (a tych jest przecież najwięcej) były bliskie modelom, z których przecież większość to modele określone ekstrapolowanymi[6] częstotliwościami względnymi.

Pozostaje jednak pytanie, czy maksymalne P_r, przy którym rozpoznawanie jest nieakceptowane, ma wynosić 0, czy też należy tę granicę podnieść, a jeśli tak, to do jakiej wartości.

Przeanalizujmy kolejne zestawienie (tabela 13). Zawiera ono średnią wartość $\bar{r}_{\%}$, najczęściej pojawiającą się wartość $r_{\%}$ (tzw. dominantę, symbol: m_0), minimalne i maksymalne $r_{\%}$ oraz liczbę i procent rozpoznań dla danego minimalnego P_r, od którego rozpoznanie było uznane, a samogłosce badanej przypisano symbol najbliższego modelu.

6 Ustalonymi na podstawie obliczania równych odległości między punktami skrajnych akustycznie samogłosek.

Tabela 13. Średni odsetek różnic artykulacyjnych ($\bar{r}_{\%}$) między samogłoskami, które starali się wymówić badani mówcy a rozpoznanymi w wyniku badania akustycznego

średnie $\bar{r}_{\%}$	0,0%	11,1%	11,5%	10,4%	10,1%	11,1%	9,7%	10,8%	10,2%	11,9%	13,6%
m_0	---	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	7,7%	7,7%	7,7%	7,7%
min $r_{\%}$	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
max $r_{\%}$	0,0%	30,8%	30,8%	30,8%	30,8%	30,8%	30,8%	30,8%	30,8%	38,5%	38,5%
n rozpoznań	1	9	12	14	16	18	23	30	34	42	59
% rozpoznań	1,6%	14,3%	19,0%	22,2%	25,4%	28,6%	36,5%	47,6%	54,0%	66,7%	93,7%
min P_r	100,0%	90,0%	80,0%	70,0%	60,0%	50,0%	40,0%	30,0%	20,0%	10,0%	1,0%

Opracowanie własne.

Można zauważyć, że procent rozpoznań rośnie niemal wykładniczo wraz ze spadkiem wartości P_r. Jednocześnie średnia procentowych różnic $r_{\%}$ nie zmienia się diametralnie, lecz oscyluje wokół 10%, by przy $P_r = 1\%$ osiągnąć ok. 14%.

Fluktuacje wartości $r_{\%}$ mogą dziwić, lecz wynikają one z włączania do analiz nowych rozpoznań, które wcześniej były ignorowane (por. wzrost liczby i procentu rozpoznań), a które to rozpoznania okazują się nie zawsze zbieżne z wartościami spodziewanymi (symbolami fonetycznymi przypisanymi badanym samogłoskom w materiale źródłowym).

Warto również zauważyć, iż zmienność maksymalnej wartości $r_{\%}$ jest niewielka: do $P_r = 20\%$ wynosi ok. 30%, a przy $P_r < 20\%$ wzrasta do niecałych 40%. Tę wartość uznaliśmy wprawdzie za decydującą o poprawności całej metody, lecz zauważmy, że $r_{\%}$ bliskie 40% to maksymalna, a nie średnia wartość, i że pojawiła się tylko trzy razy na 63 rozpoznania. 5% błędnych wymówień, za które może być odpowiedzialny mówca, można chyba uznać za dopuszczalną wartość (*notabene* wszystkie trzy były wykonane przez autora tego artykułu, co chyba w większym stopniu świadczy na korzyść prezentowanej metody niż poprawności wymówionych przezeń samogłosek). Najważniejsze jest to, że średnie odchylenie od wartości oczekiwanych wynosi niewiele ponad 10%, a nawet gdybyśmy zaokrąglili je do 15%, będzie to w istocie różnica dwóch cech artykulacyjnych (zob. wcześniejsze obliczenia, zgodnie z którymi $r = 1$ odpowiadało $r_{\%} \approx 7{,}69\%$ dla modelu 98-elementowego). Jest to więc przesunięcie od szeregu tylnych do środkowych cofniętych, od przednich do środkowych uprzednionych lub przesunięcie o jeden szereg przy jednoczesnej zmianie artykulacji wargowej. Takie niedokładności w wymowie parlatorów, nawet jeśli są wykształconymi fonetykami, można chyba dopuścić.

Uzyskane wyniki pozwalają brać pod uwagę następujące graniczne (minimalne) wartości P_r:

a) min. $P_r = 20\%$ – przy tej wartości uzyskano blisko połowę rozpoznań, z czego żadne nie różniły się od wartości spodziewanych o więcej niż ok. 30%, a średnia różnic między rozpoznaniami a wartościami spodziewanymi była niska – ok. 10%;
b) min. $P_r = 10\%$ – liczba rozpoznań wzrasta do 2/3, lecz maksymalna różnica wobec wartości oczekiwanych wzrasta do ok. 40%, choć średnia zmienia się o zaledwie 2%;
c) min. $P_r = 1\%$ – uzyskuje się blisko 100% rozpoznań, a różnica między rozpoznaniami a wartościami oczekiwanymi nie przekracza 40% ($\max r_{\%} \approx 40\%$, $\bar{r}_{\%} \approx 14\%$).

Należy tutaj zaznaczyć, iż podwyższanie progu P_r powodujące, jak mogłoby się wydawać, pożądane zmniejszanie wartości $\bar{r}_{\%}$ oraz $\max r_{\%}$ polega w istocie na ignorowaniu rozpoznań dalekich od założonych modeli i tylko nieznacznie (por. duży wzrost rozpoznań wobec niewielkiego wzrostu $\bar{r}_{\%}$ i $\max r_{\%}$ przy obniżeniu progu P_r do 1%) różniących się od wartości oczekiwanych. Można więc powiedzieć, że jeśli zależy nam na szybkim, automatycznym i obiektywnym rozpoznawaniu samogłosek (a to właśnie umożliwia prezentowana metoda), to wówczas nie powinniśmy

ignorować żadnych wyników, zwłaszcza że większa część rozpoznań skupiona jest przy niskich wartościach P_r, a więc znajduje się między założonymi modelami. Poza tym uzyskane rezultaty nie wskazują na częste mylenie przez informatorów samogłosek odległych od siebie artykulacyjnie (choć zastanawiające mogą być przypadki przeciwnego rozpoznania labializacji samogłosek wysokich).

Wobec powyższego przyjęto za minimalne P_r pozwalające uwzględnić dane rozpoznanie wartość równą 1%. Można wprawdzie przyjąć dowolny ułamek jednego procentu, lecz P_r = 1% wydawało się praktyczne, gdyż z jednej strony przyjmowanie mniejszej wartości nie zwiększyłoby już znacznie liczby rozpoznań, a utrudniłoby wizualną analizę uzyskanych wyników (na przykład sprawdzanie, czy zaimplementowany w arkuszu kalkulacyjnym algorytm poprawnie rozpoznaje samogłoski, ignorując te, dla których P_r wynosi mniej niż założona wartość). To właśnie (założenie minimalnego P_r = 1%, powyżej którego rozpoznanie będzie uznane, a samogłosce zostanie przypisany symbol fonetycznych) jest przyczyną pojawienia się pustych miejsc w tabeli nr 11. W tych przypadkach wartość P_r była mniejsza niż 1%.

Uzyskane wartości $\tilde{r}_{\%}$ oraz max$r_{\%}$ pokazują także, że zastosowana metoda nie przynosi rezultatów zupełnie nieoczekiwanych (okazuje się, iż badani wymówili głoski bardzo podobne).

Podsumowanie

Zalety ukazanej metody są oczywiste: szybkie (automatyczne), obiektywne i bardzo precyzyjne rozpoznawanie samogłosek na podstawie modelu stosowanego powszechnie w opisie różnych języków, zatem ułatwiającego porównywanie rezultatów badań, a więc i języków i ich odmian. Zastosowany model został scharakteryzowany w zasadzie tak samo jak pierwotny model Jonesa, przy czym różnicę audytywną zastąpiono arytmetycznie równymi zmianami częstotliwości względnych w zakresie wyznaczonym przez samogłoski skrajne akustycznie, tj. [i u a].

Nie oznacza to jednak, że uzyskana metoda nie ma ograniczeń. Wśród dostrzegalnych wad można wymienić:

– uzależnienie od wymówień skrajnych (jeśli analizujemy wymowę z częstymi redukcjami i bez spółgłosek [j w], które mogłyby wskazywać najbardziej przednie, tylne i wysokie ułożenie języka – uzyskane wyniki mogą być błędne);
– niejednolity podział przestrzeni artykulacyjnej: w zakresie pionowych i poziomych ruchów języka wyróżniono aż siedem położeń, podczas gdy w zakresie labializacji wyróżniono tylko dwa ułożenia: płaskie i zaokrąglone;
– zastanawiające jest pojawiające się dość często rozpoznawanie innej artykulacji wargowej (samogłosce określonej jako płaska został przypisany symbol głoski zaokrąglonej i odwrotnie); być może przygotowanie dokładniejszego zestawu wymówień modelowych (zweryfikowanych badaniami artykulacyjnymi) pozwoli skorygować zestaw częstotliwości względnych proponowanych tu modeli.

Poza tym w dalszym ciągu prezentowany model jest oderwany od sposobu artykulacji spółgłosek (z wyjątkiem labializacji), a przecież wiąże płaszczyznę akustyczną z artykulacyjną. Rzecz jest tym bardziej zasadna, jeśli zapytamy o sytuację, gdy punkty samogłoskowe znajdą się poza polem zakreślanym przez skrajne punkty uzyskanej matrycy. Co, jeśli parlator spróbuje wymówić na przykład tak skrajnie uprzednione [ɛ], iż czubek jego języka znajdzie się poza linią siekaczy? Albo gdzie w zaprezentowanym modelu znajdą się półsamogłoski [j ɥ w ɯ̯]? Pozostaje wreszcie problem retrofleksyjności i nosowości samogłosek. Kwestie te, tak samo jak wspomniane wyżej zagadnienie stopni labializacji, wymagają jednak dostępności do precyzyjnie opisanych pod względem artykulacyjnym nagrań samogłosek podstawowych, gdyż – jak wykazały przedstawione analizy wymówień samogłosek podstawowych – staranna i wyćwiczona artykulacja jest niewystarczająca, by mieć absolutną pewność, iż mamy do czynienia z samogłoską o określonych cechach artykulacyjnych.

Literatura

Jassem W., 1973, *Podstawy fonetyki akustycznej*, „Biblioteka Mechaniki Stosowanej", Warszawa.

Łobacz P., Grygiel W., Baranowska E., Francuzik K., 2003, *Klasyfikacja samogłosek polskich za pomocą sieci neuronowych w wymowie dzieci niesłyszących*, „Audiofonologia" XXIII, s. 8–31.

Rybka P., 2015, w druku, *Wykorzystanie pomiarów antropometrycznych w badaniach artykulacji samogłosek (na materiale polskim)*, „Linguarum Silva" t. 3.

How to study vowels using acoustic methods? A proposal of a method based on relative formant frequencies and cardinal vowels model. Part II. Summary

The second part of the paper consists of a list of relative formant frequencies of model vowels (the method of calculation was discussed in the first part).

In this part, a test of the proposed method is designed using several new devices, such as the articulatory difference between vowels (expressed both as an integer and as a percentage), and the percentage distance between the analysed vowels and the particular model. In addition, a few simple statistics such as median, mode and arithmetic mean are used. The test itself consists of recognizing new vowels as pronounced by three different phoneticians.

The calculations lead to the conclusion that recognition of vowels carried out with the use of the proposed method produces very accurate results which, most importantly, do not deviate in most instances from the expected values (the difference between the recognized vowel and the vowel described by the speaker did not exceed 40%).

In conclusion, this part of the paper describes the advantages and limitations of the proposed method, and suggests possible solutions which may help improve the results in the future.