

**Justyna Garczyńska**

Uniwersytet Warszawski

Instytut Języka Polskiego

ORCID: 0000-0002-5606-8990; e-mail: jgarczynska@uw.edu.pl

## Zastosowanie statystyki Pillai'a w badaniach nad samogłoskami

**Abstrakt:** W artykule przedstawiono możliwości zastosowania testu Pillai'a w badaniach nad samogłoskami na przykładzie akcentowanej samogłoski [a] występującej w mowie Polek posługujących się gwarą odmianą polszczyzny południowokresowej, deportowanych w 1936 r. z Ukrainy radzieckiej do północnego Kazachstanu.

**Słowa kluczowe:** test Pillai'a, samogłoska, fonetyka akustyczna, język polski, Kazachstan

**Abstract: The application of Pillai statistics in research on vowels.** The article shows the possibilities of using the Pillai test in research on vowels, using the example of the stressed vowel [a] in the speech of Polish women using the dialectal variety of the Southern Borderland Polish, and deported in 1936 from Soviet Ukraine to northern Kazakhstan.

**Keywords:** Pillai test, vowel, acoustic phonetics, Polish language, Kazakhstan.

### 1. Wstęp

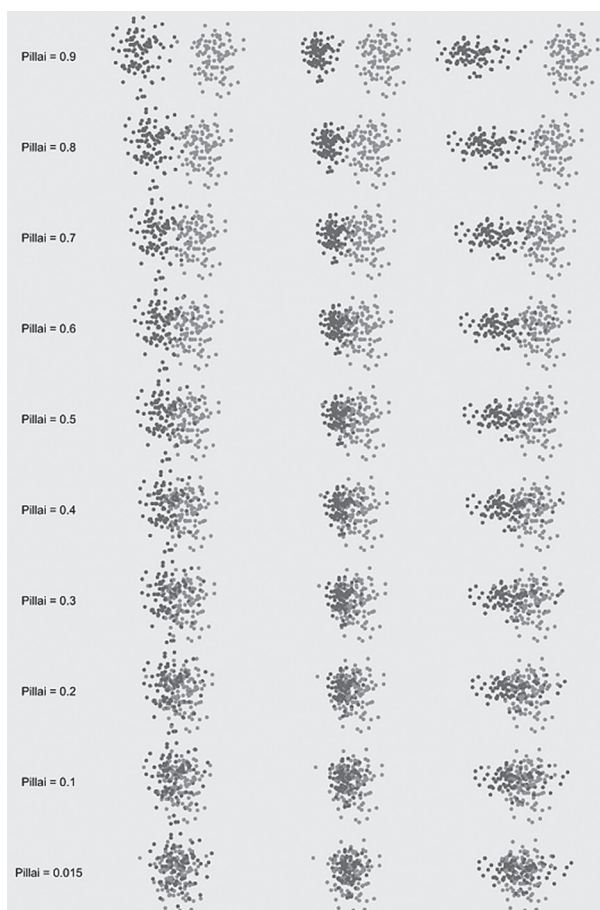
Powszechnie stosowaną miarą określającą stopień podobieństwa między samogłoskami jest odległość euklidesowa, obliczana najczęściej na podstawie średnich lub median wartości F1, F2, charakteryzujących samogłoski w przestrzeni akustycznej. Miara ta jednak nie uwzględnia stopnia i sposobu rozproszenia pomiarów F1 i F2. Jedną z propozycji alternatywnych<sup>1</sup>, uwzględniającą całe pola wyznaczone przez wartości dwóch pierwszych formantów pobrane dla każdej realizacji danej samogłoski, jest test Pillai'a (Pillai 1955), dostępny jako jeden z kilku testów w module MANOVA. MANOVA, czyli wielowymiarowa analiza wariancji jest używana w odniesieniu do więcej niż jednej zmiennej zależnej jednocześnie, np. F1 i F2. Warunkami do stosowania testu Pillai'a są: pomiar na skali interwałowej, normalność rozkładu danych oraz równość wariancji. Jednocześnie jest to test stosunkowo odporny w przypadku naruszenia tych założeń i preferowany przy

---

<sup>1</sup> Istnieje kilka metod mierzenia dystansu pomiędzy samogłoskami w przestrzeni akustycznej. Oprócz odległości euklidesowej oraz testu Pillai'a można tu wymienić SOAM (Wassink 2006), VOACH (Haynes, Taylor 2014) czy odległość Bhattacharyya (Warren 2017).

mało licznych próbach (Aranowska, Rytel 2010, 131; Olson 1974, 907). Jego wartości wahają się od 0 do 1, przy czym, w wypadku analiz akustycznych, wartości bliższe 0 wskazują na większe zbliżenie pól samogłosek, zaś bliższe 1 na większą ich separację. Wyniki testu Pillaia można zatem traktować jako pewnego rodzaju skalę podobieństwa między porównywanymi samogłoskami, np. jasnymi i pochylonymi w gwarach polskich. Mając wartości testu Pillaia dla par samogłosek w mowie różnych osób, możemy je porównywać, uzyskując informacje o stopniu zaawansowania nałożenia się pól tych głosek u każdej z nich.

Sposób przełożenia wyników testu Pillaia na dane akustyczne dotyczące samogłosek ukazuje wykres 1. Widać na nim trzy różne dystrybucje punktów wyznaczonych przez wartości F1, F2 samogłosek. Pierwsza kolumna ukazuje pola samogłosek o tym samym



Wykres 1. Wartości testu Pillaia wraz z odpowiadającym im stopniem nałożenia pól samogłosek

Źródło: <https://joestanley.com/blog/a-tutorial-in-calculating-vowel-overlap>.

rozproszeniu punktów, druga ukazuje pola o różnym rozproszeniu punktów (dla pierwszej samogłoski rozproszenie jest mniejsze zarówno pod względem F1, jak i F2), zaś ostatnia kolumna ukazuje pola samogłosek, które różnią się rozproszeniem pomiarów w zakresie jednego wymiaru (pierwszą z nich cechuje większe zróżnicowanie pomiarów F1, a drugą F2).

Te trzy przykłady pokazują także, że ten sam wynik testu Pillaia może oznaczać inną odległość euklidesową między punktami wyznaczonymi przez wartości średnie lub mediany F1, F2, por. także wykres 2. Najczęściej im większa odległość euklidesowa między punktami, tym mniejszy stopień nakładania się pól samogłosek (2a), ale może zdarzyć się także taka sytuacja, że dystans euklidesowy między samogłoskami będzie taki sam, a stopień pokrycia ich pól będzie inny (2b) lub odległość euklidesowa będzie inna, a stopień nałożenia się pól samogłosek taki sam (2c). Zależy to w dużej mierze od dystrybucji realizacji samogłoskowych.

Jednak statystyka Pillaia stosowana w badaniach nad samogłoskami ma również swoje wady. Pierwsza polega na tym, że nie istnieje określona skala wartości testu Pillaia, która mówiłaby, jak bardzo są nałożone na siebie pola samogłosek i kiedy na pewno nie ma już między nimi żadnej różnicy. Oczywiście, jak przy każdym teście, tak i przy wynikach Pillaia podawane są wartości  $p$ , ale pozwalają one określić tylko, czy wynik jest istotny statystycznie. Innymi słowy, uzyskujemy informację, że samogłoski różnią się, ale nie wiemy, jak bardzo się różnią. Poza tym wyniki istotne statystycznie w zakresie F1, F2 nie muszą oznaczać percepcyjnej różnicy między samogłoskami (Kewley-Port, Watson 1994), możemy je także uzyskać w sytuacji, kiedy pola samogłosek są niemal całkowicie nałożone (Stanley, Sneller 2023, 63–64). Dlatego niektórzy badacze nie podają wartości  $p$  przy wynikach tego testu, traktując je jako pewnego rodzaju skalę użyteczną przede wszystkim wówczas, kiedy chcemy porównywać stopień nałożenia się analizowanych samogłosek u różnych osób czy w różnych językach.

Na wartości testu Pillaia ma wpływ także wielkość próby. Badania Josepha Stanleya i Betsy Sneller (2023) wykazały, że większe próby dają niższy wynik Pillaia. Badacze wzięli pod uwagę wyniki tego testu, pochodzące z porównania prób równolicznych i różniących się liczebnością, zawierających od 5 do 100 wartości F1, F2 dla danej samogłoski. Okazało się, że nie ma znaczenia, czy grupy są równo-, czy nierównoliczne, ponieważ dla wyniku Pillaia ważna jest całkowita wielkość próby. Oznacza to, że jeśli próba dla jednej samogłoski zawiera 10 pomiarów, a dla drugiej 90, to wyniki testu Pillaia będą takie



Wykres 2. Zależność między odległością euklidesową a polami samogłosek wyznaczonymi przez wartości F1, F2

Źródło: Nycz, Hall-Lew 2014, 1.

same, jak przy próbach równolicznych, liczących 50 pomiarów lub w próbach liczących 30 i 70 pomiarów. Natomiast zawsze mniejsza próba będzie dawała wyższą wartość testu.

Z powyższych ustaleń wynika wiele konsekwencji dla interpretacji uzyskanych wyników, zwłaszcza w sytuacji, kiedy chcemy porównywać stopień nałożenia się samogłosek u różnych osób. Przy tym samym stopniu pokrycia pól samogłoskowych dla respondenta reprezentowanego przez większą liczbę danych otrzymamy niższy wynik Pillaia niż dla tego, dla którego uzyskano mniejszą liczbę pomiarów i co za tym idzie możemy błędnie zinterpretować zmiany zachodzące w ich samogłoskach. Problem ten jest szczególnie dotkliwy, kiedy chcemy porównać wyniki badań publikowanych w różnych pracach, dlatego tak ważne jest raportowanie liczebności próby<sup>2</sup>.

Statystykę Pillaia po raz pierwszy zastosowali Jennifer Hay, Paul Warren i Katie Drager (2006) do określenia stopnia nałożenia się pól samogłosek w języku angielskim Nowej Zelandii. Wartości testu Pillaia wykorzystano także w pracy doktorskiej Lauren Hall-Lew dotyczącej wariantywności fonetycznej w mowie mieszkańców okolic San Francisco (2009) oraz w jej artykule (2010), bazującym na tym samym materiale badawczym. Przykłady zastosowania statystyki Pillaia można także znaleźć w pracy magisterskiej Marianny Kennedy poświęconej zróżnicowaniu samogłosek w języku angielskim dzieci w Nowej Zelandii (2006). W 2013 r. ukazał się artykuł Jennifer Nycz i Lauren Hall-Lew, w którym autorki omówiły cztery metody przydatne w analizie stopnia pokrywania się pól samogłosek, w tym statystykę Pillaia. Kolejne porównanie takich metod zaprezentowali Matthew C. Kelley i Benjamin V. Tucker (2020). Z prac wykorzystujących wartości testu Pillaia do analizy zmian w obrębie samogłosek warto także wymienić artykuły Valerie Fridland, Tylera Kendalla oraz Charliego Farringtona (2014), w którym autorzy porównują charakterystykę spektralną oraz iloczynową samogłosek w dialektach Stanów Zjednoczonych, Marjoleine Sloos (2014), która przeprowadziła porównanie niektórych samogłosek języka niemieckiego na obszarze Austrii, Marka Amenguala i Pilar Chamorro (2015) dotyczący samogłosek języka galicyjskiego w Hiszpanii czy Marianny Nadeu i Margaret Renwick (2016) poświęcony akcentowanym samogłoskom szeregu średniego w języku katalońskim<sup>3</sup>.

## 2. Metodologia

Do ukazania zastosowania testu Pillaia wykorzystano akcentowaną samogłoskę [a] w obustronnym sąsiedztwie spółgłosek twardych, wyekscerpowaną z tekstów spontanicznych nagranych w 2000 r. wśród dziewięciu kobiet wywiezionych w 1936 r. z terenów Ukrainy do okręgu kokczetawskiego w północnym Kazachstanie. Charakterystykę informatek przedstawia tabela 1.

---

<sup>2</sup> J. Stanley i B. Sneller (2023, 61) próbują rozwiązać ten problem poprzez opracowanie wzoru pozwalającego obliczyć wartości testu Pillaia wskazujące na całkowite pokrycie się pól samogłosek w zależności od liczebności próby. Na przykład przy 20 pomiarach F1, F2 możemy uznać, że nie ma różnic między głoskami poniżej wartości Pillaia wynoszącej 0,27, zaś przy wielkości próby 80 przy wartości wynoszącej 0,05.

<sup>3</sup> Wymieniono tutaj tylko niektóre prace wykorzystujące statystykę Pillaia. Ich pełniejszą listę zawiera artykuł (Stanley, Sneller 2023).

Tabela 1. Wykaz informaterek wraz z charakterystyką socjalną

Osoba	Rok urodzenia	Miejsce pochodzenia (Ukraina)	Wykształcenie	Język używany w domu	
				rodzice	mąż/dzieci
HK	1909	Mała Radogoszcz	2 klasy szkoły polskiej	polski	pol/ukr
NK	1912	Pokostówka	8 klas szkoły ukraińskiej z lekcjami j. polskiego; 1 rok Instytutu Pedagogicznego w Kazachstanie	polski	pol/ros
WR	1914	Pokostówka	4 klasy szkoły polskiej	polski	pol/ukr
AZ	1919	Mała Radogoszcz	3 klasy szkoły polskiej; 1 klasa szkoły ukraińskiej	polski	pol/ukr
KB	1920	Dowbysz	1 klasa szkoły polskiej	polski	ukraiński
SG	1922	Słobódka Czernicka	4 klasy szkoły polskiej; 3 klasy szkoły rosyjskiej w Kazachstanie	polski	pol/ukr
ST	1925	Zawalijki	4 klasy szkoły ukraińskiej; 5 klas szkoły rosyjskiej, 3 lata rosyjskiej szkoły medycznej	ukraiński	rosyjski
AR	1932	Chutor Dobra	2 klasy szkoły rosyjskiej w Kazachstanie	polski	rosyjski
HT*	1939	Storonicze	4 klasy szkoły rosyjskiej w Kazachstanie	polski	ukraiński

\* Osoba HT urodziła się już w Kazachstanie. Storonicze to wieś, z której pochodzili rodzice informatorki.

Język respondentek można określić jako polskie gwary wyspowe należące niegdyś do zwartego obszaru polszczyzny południowokresowej. Polska ludność chłopska, która osiedliła się na Ukrainie i przyniosła tam swoją gwarę ludową, głównie małopolską i mazowiecką, wytworzyła gwarę mieszaną z ukraińskimi interferencjami lub uległa asymilacji z ukraińskim otoczeniem (Kurzowa 1997, 127, 151–152). Polskie wsie na Ukrainie, nawet leżące w bezpośrednim sąsiedztwie, mogły się między sobą różnić pod względem językowym, zakładali je bowiem osadnicy z różnych obszarów rdzennej Polski, w różnym czasie i różnym otoczeniu językowym (Rieger 2000, 109–110; 2001, 9). Niestety podczas nagrań nie udało się zdobyć informacji na temat czasu przybycia na Ukrainę oraz pochodzenia terytorialnego rodzin badanych osób. Po zesłaniu do Kazachstanu kontakt z polskim obszarem językowym został całkowicie zerwany. Polacy z różnych wsi na Ukrainie zostali przemieszani i osiedleni w nowych punktach (tzw. toczkach), w obrębie których utworzył się nowy układ językowy. Używanie gwary południowokresowej zostało ograniczone do kontaktów rodzinnych, językiem spotkań sąsiedzkich i pracy stał się przede wszystkim ukraiński, w szkole czy urzędzie posługiwano się rosyjskim (Garczyńska 2007, 20–42).

Do analizy porównawczej wykorzystano samogłoskę [a] języków ogólnopolskiego i ukraińskiego, wyekscerpowaną z nagrań starszych kobiet (50–80 lat) z wykształceniem wyższym, pochodzących z większych miast Mazowsza oraz z Kijowa i Lwowa. Materiał ten uzupełniono o dane dla [a] występującego w gwarach mazowieckich (Mazowsze

bliższe, Mazowsze dalsze, Kurpie, Podlasie) i małopolskich (Lubelszczyzna zachodnia),, uzyskane z tekstów nagrywanych wśród starszych osób, kobiet i mężczyzn. Do nagrań wykorzystano dyktafony MD SONY i OLYMPUS DM-650 rejestrujące dźwięk w formacie Linear PCM i częstotliwości próbkowania 44 kHz. Segmentację sygnału mowy przeprowadzono ręcznie, biorąc pod uwagę spektrogramy oraz obecność pełnej struktury formantowej w obrębie samogłoski (Machač, Skarnitzl 2009, 23–24).

Do analizy akustycznej samogłoski [a] wybrano częstotliwości dwóch pierwszych formantów. Pomiary wartości formantów wykonano ręcznie w środkowej części samogłoski, po odcięciu 20% początkowego i końcowego fragmentu jej trwania, aby zminimalizować wpływ sąsiedniego kontekstu. Ponieważ nie zawsze udało się wydzielić ustaloną część samogłoski, wartości formantów pobierano w punkcie, w którym F1 osiągało najwyższą wartość (Watson, Harrington 1999)<sup>4</sup>. Następnie wartości dwóch pierwszych formantów zostały znormalizowane z użyciem procedury z-score Lobanova dostępnej na stronie <http://lingtools.uoregon.edu/norm/norm1.php>. Uzyskane dane poddano analizom statystycznym opisanym w dalszej części artykułu<sup>5</sup>.

### 3. Wyniki

W ramach analizy wstępnej sprawdzono założenia normalności rozkładu oraz równości wariancji analizowanych zmiennych pomiędzy grupami. W tabeli 2. ukazano wyniki testu Shapiro-Wilka, który jest standardowym testem służącym do oceny, czy zebrane dane mają rozkład normalny. Hipoteza zerowa dla tego testu zakłada, że próba badawcza pochodzi z populacji o normalnym rozkładzie. Jeśli test Shapiro-Wilka osiąga istotność statystyczną ( $p \leq 0,05$ ), świadczy to o rozkładzie odbiegającym od normalnego. Na podstawie wyników ukazanych w tabeli 2 można stwierdzić, że rozkład F1, F2 w każdej z grup był normalny. Jednocześnie test Browna-Forsythe'a wykazał brak jednorodności wariancji w obrębie porównywanych grup, a zatem drugie z założeń testów w obrębie MANOVA nie zostało spełnione, por. tab. 3.

Mimo to zdecydowano się na przeprowadzenie testu Pillaia, ponieważ jest on uważany za odporny na umiarkowane odchylenia od założeń (por. Wstęp). Dodatkowo zastosowano test ANOVA Kruskala-Wallisa<sup>6</sup>, który jest nieparametryczną alternatywą jednozmienną analizy wariancji, oraz serię testów post-hoc z zastosowaniem poprawki Bonferroniego, osobno dla F1 i F2 każdego informatora. Poziom istotności skorygowano do  $p \leq 0,01$ .

---

<sup>4</sup> Zarówno segmentację, jak i pomiar wartości formantów przeprowadzono w programie PRAAT (wersja praat5400\_win64, pobrana ze strony: [http://www.fon.hum.uva.nl/praat/download\\_win.html](http://www.fon.hum.uva.nl/praat/download_win.html)).

<sup>5</sup> Wszystkie analizy statystyczne wykonano w programie STATISTICA 10 ([www.statsoft.pl](http://www.statsoft.pl)).

<sup>6</sup> Test ANOVA Kruskala-Wallisa stosujemy, gdy chcemy porównać przynajmniej trzy grupy pod względem zmiennej ilościowej – istotny statystycznie wynik tego testu mówi, że co najmniej jedna grupa z porównywanych różni się od innej. Aby dowiedzieć się, która z grup jest odpowiedzialna za różnice, należy przeprowadzić testy post-hoc. Przy występowaniu wielu zmiennych zaleca się korzystanie z tzw. poprawki Bonferroniego i korekcję poziomu istotności do  $p \leq 0,01$ , co chroni „przed wzrostem efektywnego poziomu istotności testu przy przeprowadzaniu wielu porównań” (Wątroba 2012, s. 83).

Tabela 2. Wyniki testu Shapiro-Wilka

Test Shapiro-Wilka, różnice są istotne z $p \leq 0,05$				
Grupa	F1 (z-score)		F2 (z-score)	
	W	Poziom – $p$	W	Poziom – $p$
HK	0,98	0,882	0,96	0,425
NK	0,94	0,278	0,95	0,374
WR	0,98	0,880	0,96	0,195
AZ	0,94	0,321	0,92	0,151
KB	0,96	0,541	0,95	0,301
SG	0,93	0,196	0,95	0,523
ST	0,97	0,705	0,96	0,407
AR	0,94	0,259	0,94	0,302
HT	0,98	0,987	0,97	0,588
j. polski	0,94	0,091	0,96	0,463
j. ukraiński	0,99	0,999	0,96	0,732
Kurpie	0,97	0,094	0,97	0,108
M. bliższe	0,95	0,129	0,97	0,442
M. dalsze	0,95	0,199	0,98	0,989
Podlasie	0,95	0,113	0,96	0,420
Lubelszczyzna	0,94	0,399	0,93	0,058

Tabela 3. Wyniki testu jednorodności wariancji Browna-Forsythe'a

Test jednorodności wariancji Browna-Forsythe'a Zaznaczone efekty są istotne z $p \leq 0,05$								
	SS – Efekt	df – Efekt	MS – Efekt	SS – Błąd	df – Błąd	MS – Błąd	F	p
F1 (z-score)	4,056	15	0,270	49,87	463	0,107	2,51	0,00*
F2 (z-score)	1,261	15	0,084	12,97	463	0,028	3,00	0,00*

W tabeli 4 przedstawiono podstawowe statystyki dla znormalizowanych wartości F1, F2 samogłoski [a] dla każdej z respondentek oraz dla języków ukraińskiego i polskiego, a także wybranych gwar mazowieckich i małopolskich. Wartości średnie z tabeli 4 ukazuje wykres 3.

Na wykresie 3 u wszystkich badanych osób jest widoczna wyższa (spadek F1) niż w językach polskim i ukraińskim realizacja samogłoski [a]<sup>7</sup>. Biorąc pod uwagę wartości

<sup>7</sup> Charakteryzując samogłoski, posłużono się takimi terminami, jak *wysoka* w opozycji do *niska* oraz *przednia* w opozycji do *tylna*, które są tradycyjnie stosowane do opisu ich relacji na tzw. trójkącie samogłosek.

Tabela 4. Podstawowe statystyki dla F1, F2 (z-score) samogłoski [a] w mowie informaterek, językach polskim i ukraińskim oraz w gwarach polskich

Samogłoska [a]					
Grupa	Liczebność	Średnia		Odch. std	
		F1 (z-score)	F2 (z-score)	F1 (z-score)	F2 (z-score)
HK	25	1,11	-0,11	0,74	0,31
NK	25	1,45	-0,21	0,59	0,29
WR	50*	1,39	-0,34	0,40	0,24
AZ	25	1,52	-0,22	0,57	0,23
KB	26	1,22	-0,40	0,58	0,42
SG	25	1,42	-0,25	0,57	0,20
ST	27	1,19	-0,53	0,63	0,32
AR	25	1,50	-0,32	0,68	0,24
HT	29	1,31	-0,41	0,49	0,28
j. polski	30	1,61	-0,11	0,48	0,13
j. ukraiński	30	1,66	-0,56	0,73	0,36
Kurpie	33	0,92	0,09	0,43	0,20
M. bliższe	33	1,43	-0,03	0,61	0,29
M. dalsze	31	1,15	0,10	0,50	0,27
Podlasie	36	0,99	-0,04	0,66	0,23
Lubelszczyzna	29	1,56	-0,26	0,89	0,26

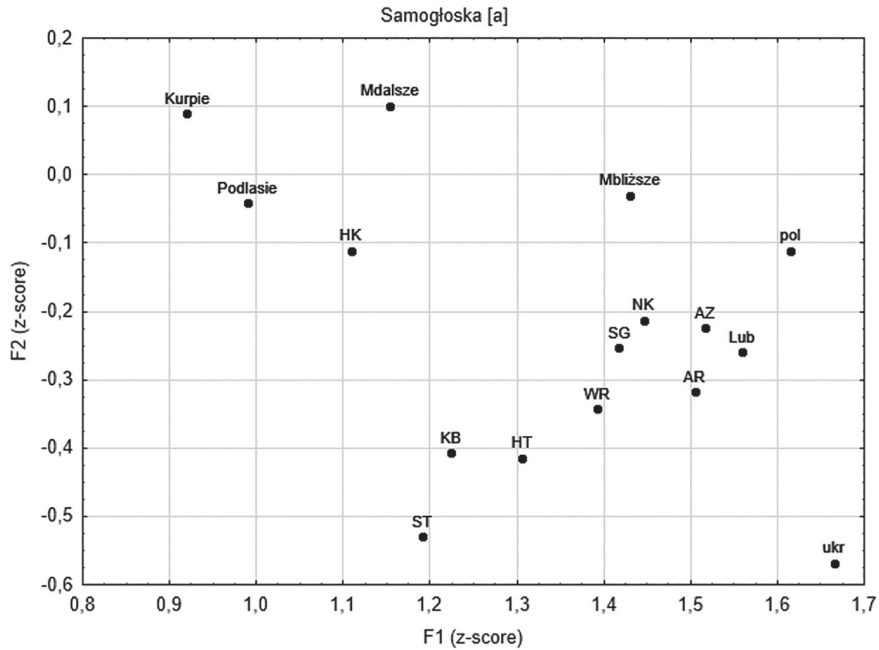
\* Ze względu na wpływ wielkości próby badawczej na wyniki testu Pillai'a zaleca się, aby porównywane grupy miały zbliżoną liczebność (Stanley, Sneller 2023, 62). Warunek ten został dla większości grup spełniony, jedynie dla osoby WR próba była zdecydowanie większa, co mogło wpłynąć na zaniżenie wartości testu Pillai'a.

F2, można analizowaną głoskę określić jako pośrednią między polską a ukraińską – cofniętą (spadek F2) w stosunku do odpowiednika polskiego i bardziej przednią w stosunku do głoski ukraińskiej.

Test Kruskala-Wallis'a dla większości osób nie wykazał istotnych statystycznie różnic między wartościami formantów samogłoski [a] w ich mowie oraz w językach polskim i ukraińskim, por. tab. 5. Jedynie u informatorki HK wystąpił istotny statystycznie wzrost wartości F2 w stosunku do głoski ukraińskiej, a u osoby ST istotny statystycznie spadek F2 w stosunku do odpowiednich wartości dla polszczyzny ogólnej. Zgodnie z otrzymanymi wynikami należałoby zatem uznać, że [a] u kobiety HK nie różni się pod względem F1, F2 od odpowiednika polskiego, u osoby ST od odpowiednika ukraińskiego, zaś dla pozostałych informaterek jest zgodne zarówno z [a] języka polskiego, jak i ukraińskiego.

W przedstawionym materiale terminy te są używane do określenia różnic między samogłoskami w przestrzeni akustycznej, a nie do opisu konkretnych pozycji języka (por. Ladefoged, Johnson 2011, 88–89, 197–198).





Wykres 3. Wartości średnie F1, F2 (z-score) samogłoski [a] w zależności od informatora na tle [a] języków polskiego i ukraińskiego oraz gwar mazowieckich i małopolskich (pol – j. polski, ukr – j. ukraiński, Lub – Lubelszczyzna)

Tabela 5. Wyniki testu Kruskala-Wallisa dla porównania F1, F2 samogłoski [a] w mowie informatorów i w językach polskim i ukraińskim

Test Kruskala-Wallisa					
H (15, N = 479) = 71,51, p = 0,000* dla F1					
H (15, N = 479) = 191,35, p = 0,000* dla F2					
Zaznaczone różnice są istotne z p < 0,01					
Osoba	Rok urodzenia	Porównywane [a] w mowie respondentów – [a] j.polski		Porównywane [a] w mowie respondentów – [a] j.ukraiński	
		Poziom – p F1	Poziom – p F2	Poziom – p F1	Poziom – p F2
HK	1909	0,374	1	1	0,004*
NK	1912	1	1	1	1
WR	1914	1	0,038	1	1
AZ	1919	1	1	1	1
KB	1920	1	0,740	1	1
SG	1922	1	1	1	1
ST	1925	1	0,001*	1	1
AR	1932	1	1	1	1
HT	1939	1	0,024	1	1

Ta trudność związana z określeniem, czy badana samogłoska należy do systemu wokalicznego ukraińskiego, czy polskiego, może być spowodowana m.in. pokrewieństwem obu systemów, a także mieszanym charakterem gwar, którymi posługiwały się informatorki. Wyniki tego samego testu wykazały istotne statystycznie różnice pomiędzy głoską [a] w językach polskim i ukraińskim jedynie w zakresie F2, por. tab. 6. W przestrzeni akustycznej samogłoska ukraińska jest głoską cofniętą w stosunku do odpowiednika polskiego, por. wykres 3.

Tabela 6. Wyniki testu Kruskala-Wallisa dla porównania F1, F2 samogłoski [a] w językach polskim i ukraińskim

Test Kruskala-Wallisa	
H (15, N = 479) = 71,51, p = 0,000* dla F1	
H (15, N = 479) = 191,35, p = 0,000* dla F2	
Zaznaczone różnice są istotne z p < 0,01	
Porównywane [a] j. ukraiński – [a] j. polski	
Poziom – p F1	Poziom – p F2
1	0,008*

Przeprowadzona analiza statystyczna wykazała także brak istotnych statystycznie różnic w zakresie wartości F1, F2 [a] w mowie respondentek i w gwarze Lubelszczyzny. Poza tym u kobiet SG i AR wartości te nie różniły się istotnie od F1, F2 [a] w gwarze Mazowska bliższego, u NK i AZ Mazowska bliższego, Mazowska dalszego i Podlasia, zaś u HK od wszystkich gwar uwzględnionych w porównaniach, por. tab. 7.

Tabela 7. Wyniki testu Kruskala-Wallisa dla porównania F1, F2 samogłoski [a] w mowie informatorów i w gwarach polskich.

Test Kruskala-Wallisa											
H (15, N = 479) = 71,51, p = 0,000* dla F1											
H (15, N = 479) = 191,35, p = 0,000* dla F2											
Zaznaczone różnice są istotne z p < 0,01											
Porównywane [a] w mowie respondentów – [a] w gwarach polskich											
Osoba	Rok urodzenia	Kurpie		M. bliższe		M. dalsze		Podlasie		Lubelszczyzna	
		Poziom – p		Poziom – p		Poziom – p		Poziom – p		Poziom – p	
		F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
HK	1909	1	0,570	1	1	1	1	1	1	1	1
NK	1912	0,076	0,006*	1	1	1	0,100	1	1	1	1
WR	1914	0,000*	0,000*	1	0,000*	1	0,000*	0,119	0,000*	1	1
AZ	1919	0,015	0,008*	1	1	1	0,108	0,367	1	1	1
KB	1920	1	0,000*	1	0,007*	1	0,000*	1	0,004*	1	1
SG	1922	0,146	0,000*	1	0,819	1	0,000*	1	0,000*	1	1
ST	1925	1	0,000*	1	0,000*	1	0,000*	1	0,000*	1	1
AR	1932	0,015	0,000*	1	0,062	1	0,000*	0,478	0,006*	1	1
HT	1939	0,133	0,000*	1	0,000*	1	0,000*	1	0,000*	1	1

Podsumowując przeprowadzone wyżej analizy, można stwierdzić, że u części osób nieparametryczny test Kruskala-Wallisa nie pozwolił na jednoznaczne zakwalifikowanie samogłoski [a] do któregoś z uwzględnionych w porównaniach języków czy dialektów. Zdecydowano się zatem na przeprowadzenie testu Pillaia, aby zobaczyć, czy procedura ta pozwoli na uzupełnienie dotychczasowych wyników.

Wartości testu Pillaia potwierdziły brak istotnych statystycznie różnic w zakresie F1, F2 [a] w mowie HK i w języku polskim oraz w mowie ST i w języku ukraińskim. Jednocześnie ukazały nieco inne wyniki dla pozostałych kobiet. I tak NK oraz AZ reprezentowały realizacje nieróżniące się istotnie od polskich, zaś KB, AR i HT od ukraińskich. Dla [a] osób WR i SG test wykazał istotne statystycznie różnice w zakresie F1, F2 w stosunku do obu uwzględnionych w analizie porównawczej języków, por. tab. 8.

Tabela 8. Wartości testu Pillaia dla porównania [a] w mowie respondentów oraz w językach polskim i ukraińskim

Test Pillaia					
Zaznaczone różnice są istotne z $p < 0,01$					
Osoba	Rok urodzenia	Porównywane [a] w mowie respondentów – [a] j. polski		Porównywane [a] w mowie respondentów – [a] j. ukraiński	
		Wartość	Poziom – $p$	Wartość	Poziom – $p$
HK	1909	0,100	0,073	0,369	0,000*
NK	1912	0,074	0,161	0,247	0,009*
WR	1914	0,237	0,000*	0,169	0,002*
AZ	1919	0,095	0,103	0,294	0,004*
KB	1920	0,240	0,000*	0,145	0,046
SG	1922	0,200	0,005*	0,252	0,009*
ST	1925	0,443	0,000*	0,118	0,079
AR	1932	0,290	0,000*	0,170	0,035
HT	1939	0,369	0,000*	0,110	0,079

Test Pillaia przeprowadzono także dla porównania [a] w mowie badanych osób i w gwarach polskich. Częściowo potwierdził on wyniki testu Kruskala-Wallisa. U wszystkich kobiet wartości F1, F2 nadal nie wykazały istotnych statystycznie różnic w stosunku do [a] w gwarze Lubelszczyzny. Takie same wyniki jak we wcześniejszej analizie otrzymano także dla osób WR, KB, SG, ST i HT. Jednak dla respondentek HK, AR, NK i AZ były one inne: u HK test Pillaia wykazał różnice w stosunku do [a] Kurpiów, u NK w stosunku do [a] Mazowsza dalszego, u AZ w stosunku do [a] Mazowsza dalszego i Podlasia, zaś u AR w stosunku do [a] Mazowsza bliźszego, por. tab. 9.

Tabela 9. Wartości testu Pillaia dla porównania [a] w mowie respondentów oraz w gwarach polskich

Test Pillaia										
Zaznaczone różnice są istotne z $p < 0,01$										
Porównywane [a] w mowie respondentów – [a] w gwarach polskich										
Osoba	Kurpie		M. bliższe		M. dalsze		Podlasie		Lubelszczyzna	
	Wartość	Poziom – p	Wartość	Poziom – p	Wartość	Poziom – p	Wartość	Poziom – p	Wartość	Poziom – p
HK	0,12	0,001*	0,08	0,080	0,13	0,037	0,02	0,590	0,11	0,173
NK	0,30	0,000*	0,08	0,073	0,25	0,001*	0,14	0,016	0,01	0,822
WR	0,56	0,000*	0,26	0,000*	0,45	0,000*	0,33	0,000*	0,02	0,470
AZ	0,40	0,000*	0,10	0,046	0,37	0,000*	0,20	0,002*	0,01	0,846
KB	0,41	0,000*	0,24	0,000*	0,35	0,000*	0,25	0,000*	0,07	0,299
SG	0,36	0,000*	0,14	0,012	0,35	0,000*	0,19	0,003*	0,01	0,858
ST	0,58	0,000*	0,40	0,000*	0,53	0,000*	0,44	0,000*	0,15	0,062
AR	0,52	0,000*	0,21	0,000*	0,49	0,000*	0,33	0,000*	0,00	0,866
HT	0,53	0,000*	0,32	0,000*	0,48	0,000*	0,36	0,000*	0,08	0,205

Dla większej przejrzystości uzyskane na podstawie obu testów wyniki ukazano w postaci tabeli, w której plus oznacza brak istotnych statystycznie różnic w zakresie F1, F2 [a] w mowie badanych osób i w porównywanych językach/gwarach, zaś minus wskazuje na istnienie takich różnic, por. tab. 10. Jak można zauważyć, nadal nie otrzymano odpowiedzi, jaką głoskę realizują badane osoby – ogólnopolską, ukraińską czy może zgodną z jedną z gwar mazowieckich lub małopolskich (por. np. osoby KB, ST, AR, HT, których [a], zgodnie z wynikami testu Pillaia, nie różni się od głoski właściwej gwarze Lubelszczyzny oraz ukraińskiej).

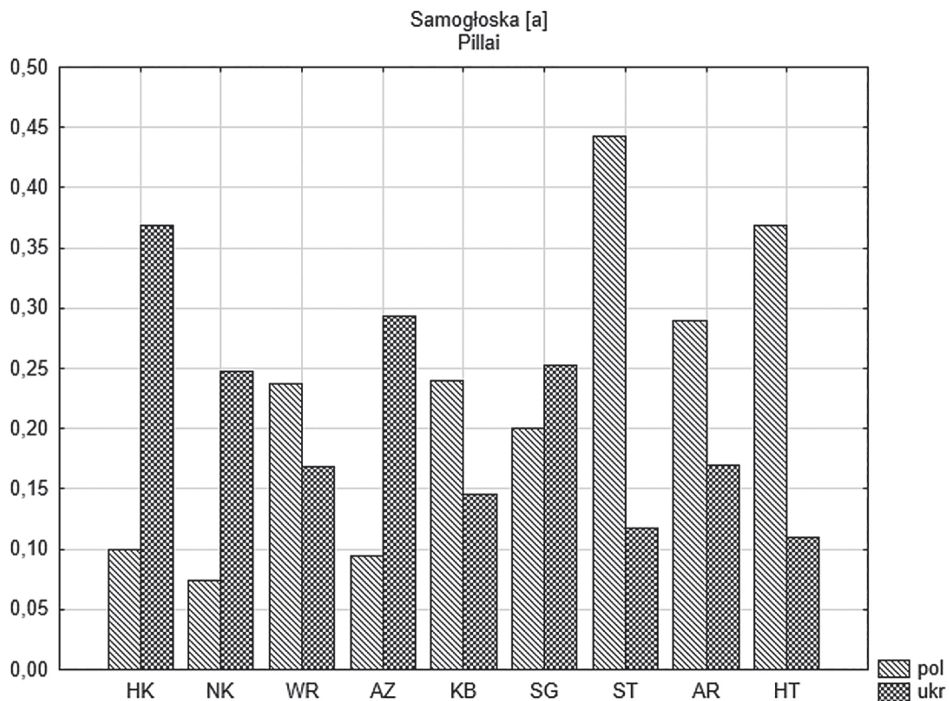
Tabela 10. Podsumowanie wyników testów Kruskala-Wallisa (K-W) oraz Pillaia (P)

Osoba	j. polski		j. ukraiński		M. bliższe		M. dalsze		Kurpie		Podlasie		Lubelszczyzna	
	K-W	P	K-W	P	K-W	P	K-W	P	K-W	P	K-W	P	K-W	P
HK	+	+	–	–	+	+	+	+	+	–	+	+	+	+
NK	+	+	+	–	+	+	+	–	–	–	+	+	+	+
WR	+	–	+	–	–	–	–	–	–	–	–	–	+	+
AZ	+	+	+	–	+	+	+	–	–	–	+	–	+	+
KB	+	–	+	+	–	–	–	–	–	–	–	–	+	+
SG	+	–	+	–	+	+	–	–	–	–	–	–	+	+
ST	–	–	+	+	–	–	–	–	–	–	–	–	+	+
AR	+	–	+	+	+	–	–	–	–	–	–	–	+	+
HT	+	–	+	+	–	–	–	–	–	–	–	–	+	+

(plus – brak różnic, minus – istnienie różnic; szarym kolorem zaznaczono różnice w wynikach obu testów)

W rozwiązaniu tego problemu może być pomocne potraktowanie wartości testu Pillai'a jako skali ukazującej stopień nałożenia się pól samogłosek i co za tym idzie, stopień podobieństwa między nimi. Wykres 4 przedstawia wyniki tego testu z tabeli 8 uporządkowane według wieku respondentek. Wyraźnie widać, że grupa kobiet KB, ST, HT i AR, u których stwierdzono brak istotnych statystycznie różnic w zakresie wartości F1, F2 [a] w stosunku do języka ukraińskiego, nie jest jednorodna – u HT [a] wykazuje największe podobieństwo do odpowiednika ukraińskiego, zaś u AR najmniejsze. Także u takich osób, jak SG i WR, dla których wyniki obu testów były niejednoznaczne, potraktowanie wartości Pillai'a jako skali pozwala na stwierdzenie, że u WR analizowana głoska była bliższa ukraińskiej, zaś u SG polskiej.

Dzięki ukazaniu wartości testu Pillai'a w zależności od wieku kobiet, widać też wyraźnie, że najstarsze z nich, przede wszystkim HK, NK oraz AZ, reprezentują realizacje najbliższe polszczyźnie. Są to osoby, które młodość spędziły na Ukrainie, posługiwały się w domu gwarą południowokresową oraz uczyły się języka polskiego w szkołach podstawowych. Respondentka NK pracowała także trzy lata jako nauczycielka języka polskiego w szkole podstawowej na Ukrainie.



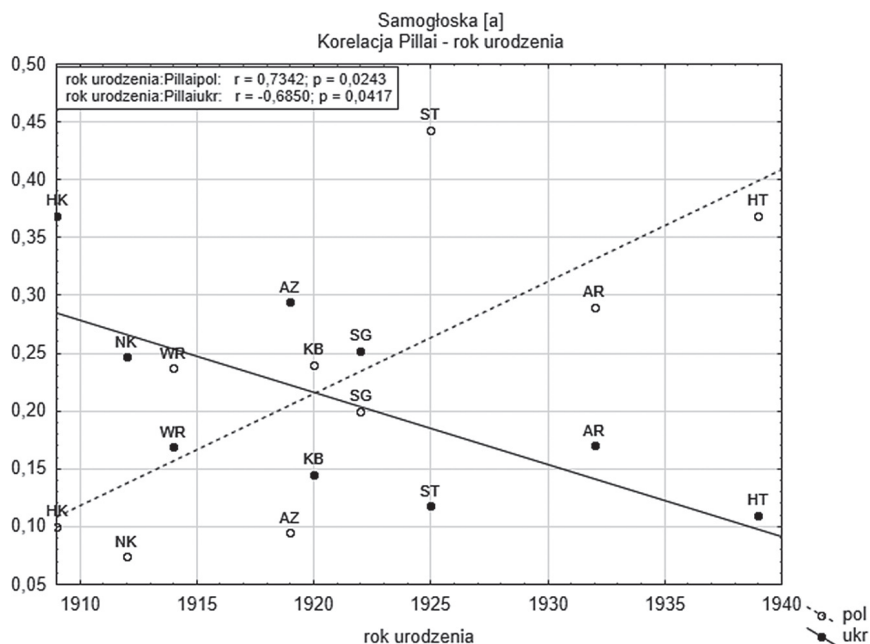
Wykres 4. Wartości testu Pillai'a dla porównania [a] w mowie respondentek oraz w językach polskim i ukraińskim

(słupki oznaczone skrótem pol i ukr ukazują stopień zbliżenia pól samogłoski [a] w mowie informaterek i odpowiednio w języku polskim i ukraińskim)

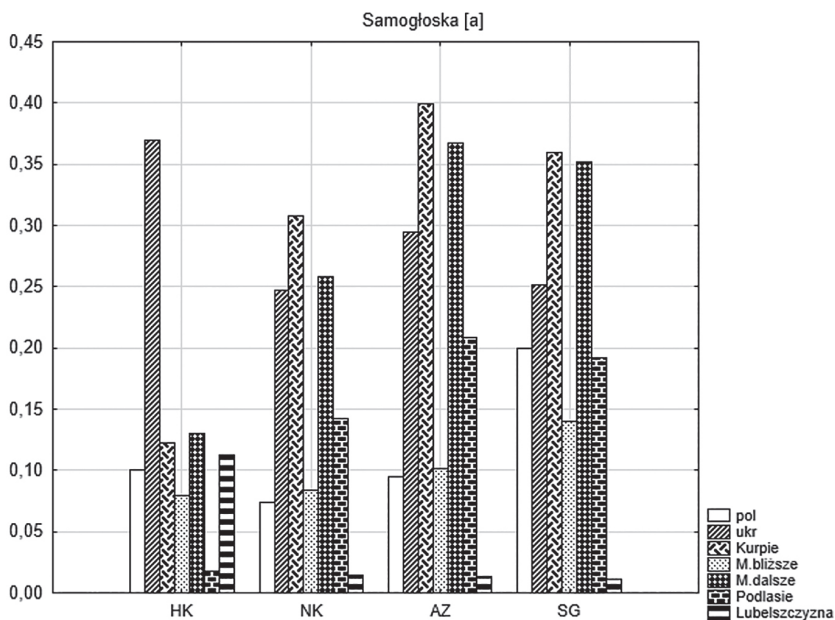
U kobiet młodszych (KB, ST, AR i HT) relacje się odwracają – ich realizacje zbliżają się do ukraińskich i oddalają od polskich, co szczególnie wyraźnie widać u osoby ST, która kończyła szkoły ukraińskie i rosyjskie, a gwarą południowokresową posługiwała się rzadko, preferując, także w domu, języki ukraiński i rosyjski, por. tab. 1. Zależność między rokiem urodzenia respondentek a wartościami testu Pillaia z tabeli 8 widać wyraźnie na wykresie 5. Na podstawie współczynników korelacji (odpowiednio  $r = 0,73$  i  $r = -0,68$ ) możemy ją określić jako silną oraz istotną statystycznie.

Powyzszą analizę, uwzględniającą jedynie języki ogólnopolski i ukraiński, uzupełniają wartości testu Pillaia, określające stopień nałożenia się pól samogłoski [a] w mowie respondentek i w gwarach polskich, por. wykresy 6 i 7. Na wykresie 6 ukazano dane dla osób HK, NK, AZ oraz SG, dla których test ten nie potwierdził różnic w zakresie F1, F2 [a] zarówno w stosunku do gwary Lubelszczyzny, jak i niektórych gwar mazowieckich. Wynika z niego, że [a] w mowie HK była najbliższa odpowiednikowi w gwarze podlaskiej, a następnie głosce ogólnopolskiej, u AZ i NK samogłosce w gwarze Lubelszczyzny, a następnie ogólnopolskiej, zaś u SG [a] w gwarze Lubelszczyzny i w drugiej kolejności Mazowsza bliższego.

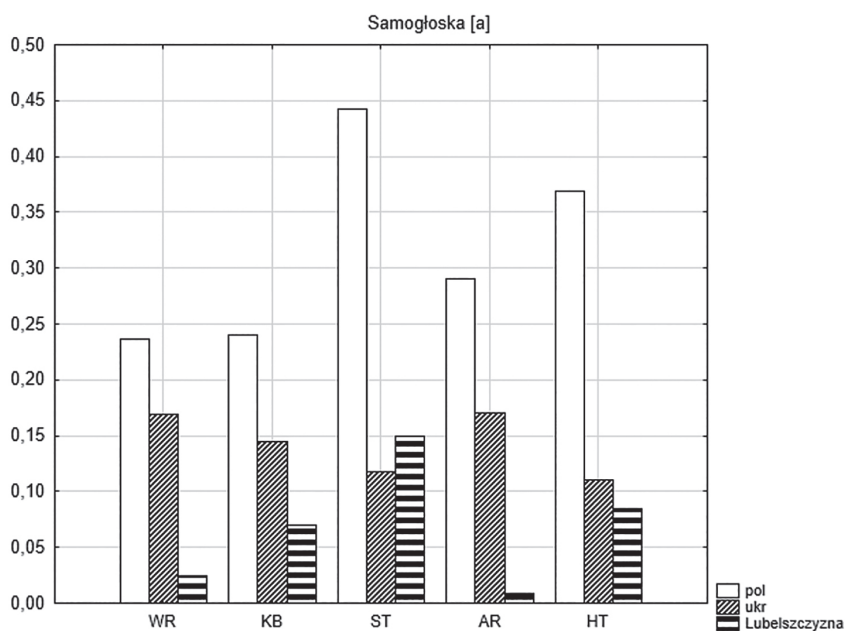
Wykres 7 ukazuje wartości testu Pillaia dla pozostałych kobiet, u których stwierdzono realizacje [a] różne od ogólnopolskich i mazowieckich. U osób WR, KB, AR i HT analizowana samogłoska była najbliższa odpowiedniej głosce w gwarze Lubelszczyzny i dopiero w drugiej kolejności samogłosce ukraińskiej. Natomiast u ST była najbliższa odpowiedniej głosce ukraińskiej, a następnie głosce w gwarze Lubelszczyzny.



Wykres 5. Korelacja między rokiem urodzenia respondentek a wartościami testu Pillaia dla porównania [a] w mowie respondentek oraz w językach polskim i ukraińskim



Wykres 6. Wartości testu Pillaia dla porównania [a] w mowie respondentek oraz w językach polskim, ukraińskim i w gwarach polskich



Wykres 7. Wartości testu Pillaia dla porównania [a] w mowie respondentek oraz w językach polskim, ukraińskim i w gwarze Lubelszczyzny

#### 4. Podsumowanie

Artykuł miał na celu ukazanie możliwości zastosowania statystyki Pillaia w badaniach nad samogłoskami. Jako przykładowy materiał badawczy wykorzystano samogłoskę [a] wyekscerpowaną z tekstów nagranych wśród Polek posługujących się gwarą południowokresową, wywiezionych w 1936 r. z Ukrainy radzieckiej do Kazachstanu. Zastosowanie testu Pillaia, zarówno jako tradycyjnego testu statystycznego, jak i skali określającej stopień pokrywania się pól samogłosek, pomogło w określeniu, do którego z porównywanych języków i dialektów<sup>8</sup> zbliża się samogłoska [a] w mowie badanych osób. Ukazana metodologia jest szczególnie cenna na obszarach językowo mieszanych, zwłaszcza tam, gdzie występują języki blisko ze sobą spokrewnione i metody opisu głosek oparte na odsłuchu mogą okazać się zawodne.

Ukazane w artykule zastosowanie statystyki Pillaia nie jest jedynym możliwym. Test ten może być pomocny m.in. w badaniach dialektologicznych, np. na gruncie języka polskiego w określaniu stopnia zaniku samogłosek pochyłonych (Garczyńska 2022a) czy w określaniu stopnia opanowania języka obcego (Mairano i in. 2020). Pojawiają się też próby jego wykorzystywania do badania spółgłosek (Regan 2020).

Chociaż ustalony zakres punktacji testu Pillaia od 0 do 1 jest przydatny do porównań pomiędzy mówcami, nie jest on łatwy w interpretacji, np. nie określa, przy jakich dokładnych wartościach możemy mówić o całkowitym braku różnicy między porównywanymi samogłoskami<sup>9</sup>. Test ten ma też swoje ograniczenia, związane z rodzajem rozkładu danych, jednorodnością wariancji czy liczebnością próby. Mimo coraz szerszego zastosowania w pracach dotyczących fonetyki, zwłaszcza socjolingwistycznych, nadal jest metodą, której zalety i ograniczenia dopiero są poznawane i z tego względu powinien być używany razem z innymi miarami odległości między samogłoskami, np. odległością euklidesową.

#### Literatura

- Amengual M., Chamorro P. (2015), *The Effects of Language Dominance in the Perception and Production of the Galician Mid Vowel Contrasts*, „Phonetica” 72, s. 207–236.
- Aranowska E., Rytel J. (2010), *Wielowymiarowa analiza wariancji – MANOVA*, „Psychologia Społeczna”, nr 2–3 (14), s. 117–141.
- Fridland V., Kendall T., Farrington Ch. (2014), *Durational and Spectral Differences in American English Vowels: Dialect Variation within and across Regions*, „The Journal of the Acoustical Society of America” 136 (1), s. 341–349.

---

<sup>8</sup> Przeprowadzone analizy miały jedynie charakter przykładowy, a ich celem nie było określenie ani pochodzenia badanych osób, ani stopnia wpływów w ich mowie uwzględnionych w analizie języków. Do tego typu analiz potrzebna jest zdecydowanie większa próba badawcza oraz uwzględnienie gwar ukraińskich, a także języka rosyjskiego/gwar rosyjskich, z którymi na Ukrainie, a przede wszystkim na terenie Kazachstanu badane osoby także miały kontakt.

<sup>9</sup> Por. przypis 2.



- Garczyńska J. (2007), *Analiza fonetyczna akcentowanych samogłosek ustnych w mowie Polek z Kazachstanu*, Warszawa.
- Garczyńska J. (2022), *Wpływ wieku i płci na realizację samogłosek [i], [y] w gwarze Mazowska bliższego*, „LingVaria”, vol. XVII, nr 1 (33), s. 175–193.
- Garczyńska J. (2022a), *Samogłoski [a] oraz [ã] w gwarze Mazowska bliższego*, [w:] *Polszczyzna w różnych jej aspektach. Język polski dawniej i dziś*, red. A.S. Dyszak, A. Paluszak-Bronka, Bydgoszcz, s. 47–68.
- Hall-Lew L. (2009), *Ethnicity and Fhonetice Variation in a San Francisco Neighborhood*, Stanford University, San Francisco, <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=daed7ef856a51643f58889e4ff5d8e40e43ba9b2>.
- Hall-Lew L. (2010), *Improved Representation of Variance in Measures of Vowel Merger*, „Proceedings of Meetings on Acoustics”, vol. 9, <https://asa.scitation.org/doi/pdf/10.1121/1.3460625>.
- Hay J., Warren P., Drager K. (2006), *Factors Influencing Speech Perception in the Context of a Merger-in-Progress*, „Journal of Phonetics” 34, s. 458–484.
- Haynes E. F., Taylor M. (2014), *An Assessment of Acoustic Contrast between Long and Short Vowels Using Convex Hulls*, „The Journal of the Acoustical Society of America” 136 (2), s. 883–891.
- Kennedy M. (2006), *Variation in the Pronunciation of English by New Zealand School Children*, Victoria University of Wellington, Wellington, <https://core.ac.uk/download/pdf/41335595.pdf>
- Kelley M.C., Tucker B.V. (2020), *A Comparison of Four Vowel Overlap Measures*, „The Journal of the Acoustical Society of America” 147 (1), s. 137–145.
- Kewley-Port D., Watson C.S. (1994), *Formant-frequency Discrimination for Isolated English Vowels*, „The Journal of the Acoustical Society of America” 95 (1), s. 485–496.
- Kurzowa Z. (1997), *Historia i współczesność języka polskiego na Kresach południowo-wschodnich*, [w:] *Historia i współczesność języka polskiego na Kresach wschodnich*, red. I. Grek-Pabisowa, Warszawa, s. 111–165.
- Ladefoged P., Johnson K. (2011), *A Course in Phonetics*, Boston.
- Machač P., Skarnitzl R. (2009), *Principles of Phonetic Segmentation*, Praha.
- Mairano P., Bouzon C., Capliez M., de Iacovo V. (2019), *Acoustic Distances, Pillai Scores and LDA Classification Scores as Metrics of L2 Comprehensibility and Nativelikeness*, [w:] *Proceedings of the 19th International Congress of Phonetic Sciences, sierpień 2019, Melbourne*, red. S. Calhoun, P. Escudero, M. Tabain, P. Warren, s. 1104–1108.
- Nadeu M., Renwick M. (2016), *Variation in the Lexical Distribution and Implementation of Phonetically Similar Phonemes in Catalan*, „Journal of Phonetics” 58, s. 22–47.
- Nycz J., Hall-Lew L. (2013), *Best Practices in Measuring Vowel Merger*, [w:] *Proceedings of Meetings on Acoustics* 20 (1), <https://asa.scitation.org/doi/pdf/10.1121/1.4894063>.
- Olson Ch. (1974), *Comparative Robustness of Six Tests in Multivariate Analysis of Variances*, „The Journal of the American Statistical Association”, vol. 69, nr 348, s. 894–908.
- Pillai K.C.S. (1955), *Some New Test Criteria in Multivariate Analysis*, „The Annals of Mathematical Statistics” 26 (1), s. 117–121.
- Regan B. (2020), *Extending Pillai Scores to Fricative Mergers: Advancing a Gradient Analysis of a Split-in-Progress in Andalusian Spanish*, „University of Pennsylvania Working Papers in Linguistics”, vol. 26, s. 109–118.
- Rieger J. (2000), *Językowe i pozajęzykowe aspekty badania polszczyzny Kresów*, [w:] *Język – człowiek – kultura. Rozprawy i artykuły. Księga pamiątkowa poświęcona profesorowi Marianowi Jaurkowskiemu*, red. B. Czopek-Kopciuch, Piotrków Trybunalski, s. 107–114.

- Sloos M. (2014), *The Reversal of the BAREN-BEEREN Merger in Austrian Standard German*, „Mental Lexicon” 8 (3), s. 353–371.
- Stanley J.A. (2021), *A Tutorial in Measuring Vowel Overlap in R*, <https://joestanley.com/blog/a-tutorial-in-calculating-vowel-overlap> (dostęp: 15.03.2023).
- Stanley J.A., Sneller B. (2023), *Sample Size Matters in Calculating Pillai Scores*, „The Journal of the Acoustical Society of America” 153 (54), s. 54–67.
- Warren P. (2018), *Quality and Quantity in New Zealand English Vowel Contrasts*, „The Journal of the International Phonetic Association” 48 (3), s. 305–330.
- Wassink A.B. (2006), *A Geometric Representation of Spectral and Temporal Vowel Features: Quantification of Vowel Overlap in Three Linguistic Varieties*, „The Journal of the Acoustical Society of America” 119 (4), s. 2334–2350.
- Watson C.I., Harrington J. (1999), *Acoustic Evidence for Dynamic Formant Trajectories in Australian English Vowels*, „The Journal of the Acoustical Society of America” 106 (1), s. 458–468.
- Wątroba J. (2012), *Jak poprawnie zastosować wielowymiarową analizę wariancji (MANOVA) – pokonać trudności i osiągnąć korzyści*, StatSoft Polska, s. 79–90, [https://media.statsoft.pl/\\_old\\_dnn/update/downloads/jak\\_poprawnie\\_zastosowac\\_wielowymiarowa\\_analize\\_wariancji.pdf](https://media.statsoft.pl/_old_dnn/update/downloads/jak_poprawnie_zastosowac_wielowymiarowa_analize_wariancji.pdf).