

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. I*

Słowa kluczowe: fonetyka, samogłoska, fonetyka akustyczna, fonetyka artykulacyjna, artykulacja, formant, samogłoski podstawowe

Keywords: phonetics, vowel, acoustic phonetics, articulatory phonetics, articulation, formant, cardinal vowels

Metoda wykresowa w badaniach samogłosek

Dostępność programów do analizy akustycznej pozwala bardzo dokładnie i obiektywnie badać samogłoski przy użyciu właśnie metod akustycznych¹. Do badań akustycznych wystarczy obecnie przeciętnej klasy komputer osobisty, a nie – jak dawniej – specjalistyczne analizatory akustyczne.

Stosowana powszechnie metoda akustycznej analizy samogłosek² – nazwijmy ją dla wygody „**metodą wykresową**” – składa się z następujących etapów:

* Druga część artykułu ukaże się w nrze następnym (1(19) 2015). Red.

1 Dostępne są wersje darmowe takich np. programów: Praat 5.3.60, <http://www.fon.hum.uva.nl/praat/>; Speech Analyzer 3.0.1, http://www-01.sil.org/computing/sa/sa_download.htm; SFS/WASP 1.54, <http://www.phon.ucl.ac.uk/resource/sfs/wasp.htm>; VoiceSauce 1.14, <http://www.seas.ucla.edu/spapl/voicesauce/>; Spectrogram 5.0, <http://www.electronics-lab.com/downloads/pc/003/>.

2 Zob. m.in. Garczyńska 2007a; Garczyńska 2007b; Orawski 2001.

- 1) Ustalenie częstotliwości formantowych badanych samogłosek (należących do zbioru, który będziemy dalej nazywać **zbiorem A**); wystarczy częstotliwości 1. i 2. formantu.
- 2) Umieszczenie na wykresie punktów (nazwijmy je skrótowo „**punktami samogłoskowymi**”), których współrzędnymi są ustalone wcześniej częstotliwości formantowe (zwykle oś OX reprezentuje wartości formantu 1., a oś OY – wartości formantu 2.³). Punkty samogłoskowe tworzą na wykresie **chmurę punktów**.
- 3) Porównanie położenia chmury punktów samogłoskowych z położeniem chmury punktów samogłosek będących odniesieniem (należących do zbioru B, np. ogólnej odmiany języka polskiego). Analiza uzyskanego wykresu pozwala sformułować jeden z poniższych wniosków:
 - a) Samogłoski badane (zbiór A) w pełni pokrywają się z punktami odniesienia (zbiorem B), są więc artykułowane dokładnie tak samo.
 - b) Samogłoski badane tylko częściowo pokrywają się z punktami odniesienia, ale są przesunięte w określonym kierunku (np. innych punktów odniesienia). Nie można więc mówić o identycznej, ale zbliżonej artykulacji samogłosek ze zbioru B w porównaniu do głosek ze zbioru A.
 - c) Głoski badane nie pokrywają się z punktami odniesienia, są więc artykułowane zupełnie inaczej niż samogłoski ze zbioru B. Tu możliwe są następujące przypadki:
 - samogłoski badane znajdują się pomiędzy dwiema chmurami punktów obrazujących cechy akustyczne głosek ze zbioru B (można więc mówić o artykulacji pośredniej);
 - chmura punktów samogłosek badanych nie leży pośrodku dwu innych chmur samogłosek odniesienia, ale jest przesunięta w kierunku jednej z nich (wówczas artykulacja samogłosek badanych nie jest dokładnie pośrednia);
 - punkty samogłosek badanych nie pokrywają się z żadnymi punktami samogłosek ze zbioru B ani nie leżą między chmurami samogłosek odniesienia (wówczas jedynie bliskość względem jakichś samogłosek ze zbioru B może stanowić podstawę do stwierdzenia podobieństwa między nimi a badanymi samogłoskami).

Zamiast punktów można zaznaczyć na wykresie prostokąt odzwierciedlający zakres występowania punktów samogłoskowych (samogłosek badanych oraz samogłosek odniesienia). Wówczas każdy bok prostokąta odnosi się do maksymalnej lub minimalnej wartości osiągananej przez częstotliwości formantowe danej grupy

3 By ułatwić odczytywanie takiego wykresu można odwrócić obie osie (taki zabieg zastosowano w przedstawionych niżej wykresach), dzięki czemu wykres punktów samogłoskowych przypomina artykulacyjny czworokąt samogłoskowy.

samogłosek. Praktycznym rozwiązaniem jest stosowanie chmury punktów w przypadku samogłosek badanych, a „**prostokątów zakresowych**” w przypadku samogłosek odniesienia⁴.

Do niedoskonałości metody wykresowej można zaliczyć następujące jej cechy:

- 1) Potrzebujemy głosek porównawczych (samogłosek odniesienia, elementów zbioru B) pochodzących od informatorów bardzo podobnych (zbliżonych wiekiem, pochodzeniem i o tej samej płci) do tych, którzy wymówili samogłoski badane.
- 2) Jesteśmy ograniczeni zestawem głosek porównawczych (zbiór B) i uzależnieni od sposobu ich opisu (który może być przecież niedokładny). Stwierdzić bowiem możemy jedynie tyle, że badana samogłoska: 1) jest identyczna z jakąś samogłoską odniesienia albo 2) jest nieco przesunięta względem samogłoski odniesienia, albo też, że samogłoska badana zupełnie nie pokrywa się z daną samogłoską odniesienia, lecz 3) zajmuje pozycję pośrednią (dokładnie pośrednią lub z przesunięciem w jakimś kierunku) między dwiema głoskami porównawczymi lub 4) nie zajmuje pozycji pośredniej między żadnymi dwoma modelami, lecz jest przesunięta względem jednego z nich.

Precyzyjne określenie artykulacji głoski badanej możliwe jest w przypadku nr 1; w kolejnych sytuacjach nasze wnioski są coraz mniej dokładne. W każdym przypadku uzależnieni jesteśmy w dużej mierze od opisu samogłosek zbioru B: jeśli jest mało precyzyjny, również nasze wnioski dotyczące samogłosek ze zbioru A będą ogólnikowe i będą takie już w pierwszym z wymienionych wyżej przypadków. Każdy następny – jak już powiedzieliśmy – wiąże się z coraz mniejszą precyzją. Tak więc uzyskany opis może być nie tylko mało precyzyjny, ale nawet błędny. Jeśli jeszcze uświadomimy sobie brak algorytmów jednoznacznego opisu artykulacji samogłosek⁵, trudno będzie znaleźć dokładny opis głosek zbioru B. A bez dokładnego opisu artykulacji głosek zbioru B nie jesteśmy w stanie dokładnie opisać głosek zbioru A.

- 3) Ograniczenie liczby analizowanych formantów do pierwszych dwóch jest jednocześnie rezygnacją z badania labializacji samogłosek. Uzasadnia się to tym, iż zaokrąglenie samogłosek nie jest fonologicznie relewantne (Garczyńska 2007a: 45), jest to jednak kolejny moment, w którym obniżamy poziom dokładności badania. Rezygnacja z analizy labializacji w badaniach nad innymi (np. gwarowymi) odmianami polszczyzny jest rozciągnięciem owej nierelewantności na badaną odmianę. Zakłada się więc z góry, że również w systemie, z którego pochodzą

4 Tak w: Garczyńska 2007a.

5 Przez „algorytmy jednoznacznego opisu” rozumiem tutaj taką metodę opisu, która użyta przez dwu różnych fonetyków pozwoliłaby na podstawie tego samego materiału – tego samego obrazu głoski – uzyskać identyczne rezultaty, tzn. takie same opisy artykulacji tejże głoski bez względu na poziom szczegółowości opisu – czy mówimy np. o 3, 5 lub większej ilości ułożeń języka w danej płaszczyźnie – oraz przyzwyczajenia lub doświadczenia danego badacza.

badane samogłoski, labializacja samogłosek także nie jest relewantna. Jednakże zaokrąglenie samogłosek powoduje obniżenie wszystkich formantów (badania własne, zob. też Jassem 1973: 191⁶; Fant 2004: 157), w związku z tym ignorowanie tego faktu może prowadzić do błędnego interpretowania niskich częstotliwości formantowych jako dowodu na cofnięcie masy języka zamiast zaokrąglenia warg. Tak więc pominięcie labializacji obniża dokładność wyników z jednej strony, a z drugiej zwiększa ryzyko błędnej interpretacji niskich częstotliwości formantowych.

- 4) Gdybyśmy chcieli badane samogłoski porównać nie z samogłoskami ogólnopolskimi, ale z samogłoskami innej odmiany polszczyzny lub innego języka, musielibyśmy znaleźć wymówienia informatorów o podobnych (wiek) lub identycznych (płeć) cechach co mówcy pierwszej (A) grupy⁷. Poza tym uzyskalibyśmy znowu dane o charakterze relacyjnym: głoska badana zbioru A (głoski badane) jest albo identyczna z jakim elementem zbioru B (głoski stanowiące podstawę porównania), albo nieco przesunięta względem tego elementu, albo wreszcie pośrednia między 2 elementami zbioru B (tu możliwe są 2 warianty: względnie idealnie pośrednie położenie lub zauważalne przesunięcie w kierunku któregoś z elementów zbioru B).

Poza tym wypowiadamy się o wymowie mówców określonego typu (określonej płci i podobnego wieku, a więc takich, których aparaty mowy będą najbardziej podobne) i nie określamy artykulacji samogłosek w kategoriach, które można by uznać za uniwersalne (np. w odniesieniu do modelu samogłosek podstawowych).

Obliczanie odległości między punktami samogłoskowymi i zestaw samogłosek modelowych

Można jednak zaproponować inne podejście: zamiast porównywać cały zbiór głosek (zbiór A) do innego zbioru głosek (zbioru B) i stwierdzać, w jakim stopniu określone podzbiory zbiorów A i B się pokrywają, możemy obrać inny zbiór B i stosować go w każdym badaniu. Najlepiej, by był to zbiór powszechnie znany i jednoznacznie

6 W podanym miejscu mowa wprawdzie o F1 i F2, ale analiza danych na stronie 190 tej publikacji pokazuje, że (pomijając kilka wyjątków) F3 i F4 również są niższe w przypadku samogłosek zaokrąglonych. Wspomnianymi wyjątkami są częstotliwości następujących samogłosek: F1 dla [i y] (różnica 10 Hz), F3 dla [a ɒ] (różnica 130 Hz), F4 dla [ɤ ɔ] (różnica 100 Hz), F1, F3 i F4 dla [ɪ ʊ] (różnice, odpowiednio, 30, 100 i 50 Hz).

7 W tym miejscu można by zapytać, po co porównywać głoski jakiejś odmiany języka polskiego z zupełnie innym językiem. Jeśli oczywiście nie porównujemy bardzo odległych od siebie systemów (np. południowoafrykańskiego języka xhosa i gwary śląskiej), to oprócz czysto poznawczych celów można wymienić chęć stwierdzenia ewentualnego transferu z jednego systemu do drugiego (np. upodabniania się głosek jednego systemu do głosek drugiego). Poza tym wiedza o tego typu podobieństwach może ułatwić naukę wymowy głosek danego systemu. Są to więc co najmniej dwa praktyczne powody tego typu badań.

scharakteryzowany, choć – jak w przypadku każdego modelu – trudno oczekiwać pełnej doskonałości od takiego punktu odniesienia. Takim zbiorem mogą być samogłoski podstawowe, które można scharakteryzować przede wszystkim artykulacyjnie (skrajne ułożenia języka), w drugiej kolejności audytywnie (pośrednie ułożenia języka mają wytwarzać głoski słyszalnie różne). Jedną głoskę, mianowicie [ə], można też scharakteryzować akustycznie: ponieważ podczas artykulacji tej głoski aparat mowy przybiera postać względnie idealnej rury, to znaczy średnica kanału głosowego na każdym odcinku od krtani do szpary ustnej jest mniej więcej taka sama. Jak wynika z badań akustycznych, częstotliwości formantowe tej głoski powinny tworzyć ciąg typu 500, 1500, 2500, 3500⁸, lub – biorąc pod uwagę możliwe różnice osobnicze – n , $3 \cdot n$, $5 \cdot n$, $7 \cdot n$, gdzie n to częstotliwość pierwszego formantu samogłoski neutralnej [ə].

Taki model nie jest jednorodny. Zastosowanie kryterium audytywnego naraża go na subiektywność. Jest to jednak model o ściśle określonych elementach, pozwalający na dowolną dokładność opisu⁹, i co może najważniejsze, powszechnie stosowany, więc zapewniający swobodną wymianę informacji (na tym modelu jest oparty międzynarodowy alfabet fonetyczny) i porównywanie wyników badań.

Żeby jednak korzystać z pojedynczych modeli (samogłosek porównawczych, samogłosek odniesienia) zamiast ich zbiorów, musimy zmienić sposób porównania. Każdym elementem zbioru A jak w tradycyjnej, wykresowej metodzie będą punkty samogłoskowe opisane częstotliwościami formantowymi, ale nie będziemy analizować tego, w jakim stopniu dwie chmury punktów samogłoskowych (lub prostokątów zakresowych) pokrywają się wzajemnie. Trudno będzie bowiem zebrać dużą liczbę wymówień tej samej samogłoski podstawowej. Poza tym – jak zobaczymy w dalszej części tekstu – większa liczba realizacji tej samej samogłoski podstawowej może okazać się zupełnie nieprzydatna. Łatwiej będzie ustalić jedną, niejako prototypową artykulację danej samogłoski podstawowej i do niej porównywać wszystkie samogłoski badane. Porównywanie będzie trzeba powtarzać z wykorzystaniem wszystkich poszczególnych modeli samogłosek podstawowych, gdyż tylko dzięki temu będzie można stwierdzić największe podobieństwa między głoskami badanymi (zbiór A) a określonymi modelami (zbiór B).

Podobieństwo można najłatwiej zdefiniować jako odległość w kartezjańskim układzie współrzędnych. Takie postępowanie umożliwi nam także uwzględnienie

8 Zob. choćby Jassem 1973: 197.

9 Np. między ułożeniami języka typowymi dla głosek [i] a [u] możemy wyróżnić: jedną pośrednią pozycję języka: [i̯],
2 pozycje pośrednie: [i̯ ü̯],
3 pozycje pośrednie: [i̯ i̯ ü̯],
5 pozycji pośrednich: [i̯ i̯ i̯ i̯ ü̯],
9 pozycji pośrednich: [i̯ i̯ i̯ i̯ i̯ i̯ i̯ i̯ ü̯],
a nawet 13 pozycji pośrednich: [i̯ i̯ i̯ i̯ i̯ i̯ i̯ i̯ i̯ i̯ i̯ i̯ i̯ ü̯],
zależnie od stopnia dokładności opisu, jaki zakładamy i jaki jest nam potrzebny.

większej niż tylko 2 liczby formantów, a więc badanie labializacji. Wzór na odległość między punktami samogłoskowymi jest prostym przekształceniem wzoru z twierdzenia Pitagorasa:

$$\text{Wzór 1: } d = \sqrt{(f_{1A}-f_{1B})^2 + (f_{2A}-f_{2B})^2 + \dots + (f_{nA}-f_{nB})^2} = \sqrt{\sum_{i=1}^n (f_{iA}-f_{iB})^2}$$

gdzie d to odległość między 2 punktami samogłoskowymi, f_{iA} to częstotliwość i -tego formantu głoski badanej (zbiór A), a f_{iB} to częstotliwość i -tego formantu głoski modelowej (podstawowej, zbiór B). Liczba badanych formantów to n .

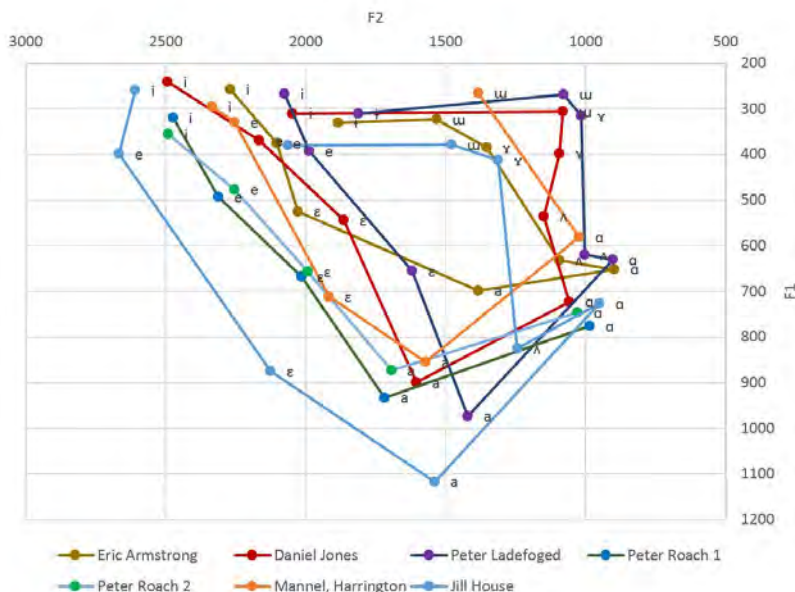
Każdorazowe obliczanie tych odległości będzie oczywiście czasochłonne, ale bardzo łatwo można skonstruować odpowiedni arkusz kalkulacyjny lub napisać prosty program komputerowy, który na podstawie wpisanych częstotliwości formantowych samogłosek zbioru A i B nie tylko błyskawicznie obliczy wartość d dla każdego porównania poszczególnych elementów zbioru A ze wszystkimi elementami zbioru B (będzie to konieczne, by mieć pewność, do którego elementu zbioru B badana samogłoska jest najbardziej podobna), ale także będzie mógł dopasować symbol fonetyczny tego elementu zbioru B, a więc tego modelu, który do badanej samogłoski jest najbardziej podobny, czyli automatycznie rozpoznać tę głoskę. Proponowane podejście jest więc krokiem ku automatycznemu rozpoznawaniu samogłosek ustnych.

Takie postępowanie napotyka jednak pewien dość istotny problem. Jeśli przyjrzymy się czworokątom samogłosek podstawowych określonym na podstawie częstotliwości formantowych, zauważymy, iż te – w przybliżeniu – czworokąty mają różne kształty i są względem siebie przesunięte. Szczególnie wyraźne różnice zauważymy, porównując wymówienia kobiet i mężczyzn (wykresy nr 1 i 2).

Porównajmy również czworokąty samogłosek polskich wymówionych przez kobietę i mężczyznę (wykres 3).

Jak widać, wymówienia kobiet (kolor jasnoniebieski na wykresach nr 1 i 2 oraz pomarańczowy na wykresie nr 3) są przesunięte względem wymówień mężczyzn w lewo i w dół, to znaczy częstotliwości formantowe analogicznych samogłosek mają w przypadku wymówień mówców płci żeńskiej wartości wyższe. Gdybyśmy więc za zestaw modeli przyjęli wymówienia mężczyzn, wymówienia kobiet byłyby błędnie rozpoznawane i odwrotnie: wymówienia mężczyzn byłyby błędnie interpretowane w odniesieniu do modeli opartych na wymówieniach kobiet. Tak samo byłoby w przypadku artykulacji dzieci, a także mężczyzn z wyjątkowo wysokimi głosami. Wynika to z indywidualnych różnic w budowie aparatu mowy, które przekładają się na inną częstotliwość tonu podstawowego (F0), który jest następnie filtrowany przez kolejne jamy nasady. I chociaż charakterystyka filtracyjna (sposób filtracji), tzn. wzmacnianie i tłumienie poszczególnych drgań składowych powinno być zbliżone dla każdego (względnie) identycznego ułożenia artykulatorów (pomijamy tutaj indywidualne różnice w budowie anatomicznej, jak np. zakrzywienie linii podniebienia twardego), to jednak na wyjściu uzyskuje się inne częstotliwości formantowe. Nie możemy

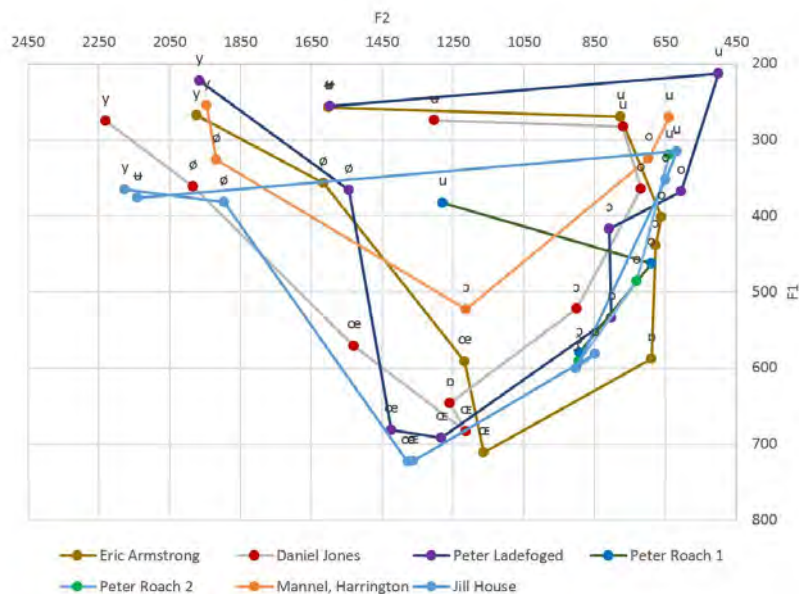
Wykres 1. Samogłoski podstawowe płaskie [i e ε a ʌ ɜ u ɪ] wymówione przez sześciu fonetyków



(Opracowanie własne na podstawie nagrań opublikowanych w różnych źródłach)¹⁰.

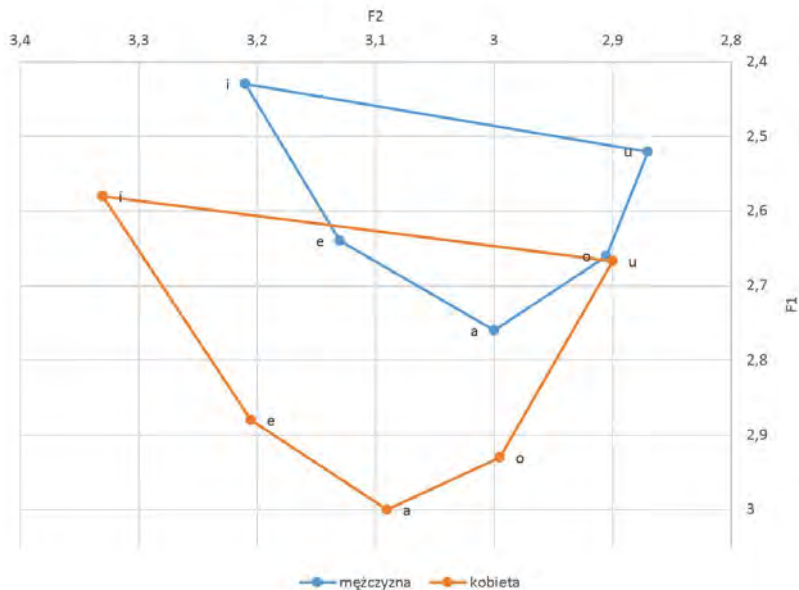
¹⁰ Eric Armstrong, associate profesor na Wydziale Sztuk Pięknych Uniwersytetu w Jorku, gdzie uczy wymowy studentów aktorstwa. Wymówienia samogłosek podstawowych pochodzą z jego strony *Eric Armstrong's voice & speech source*, <http://www.yorku.ca/earmstro/ipa/vowels.html>. Daniel Jones – jego wymówienia samogłosek podstawowych pochodzą ze strony *Audiufon: audio demonstrations, Utrech University, Phonetics*, www.let.uu.nl/~audiufon/data/e_root.html. Peter Ladefoged, nieżyjący już profesor lingwistyki na Uniwersytecie w Kalifornii, absolwent Uniwersytetu w Edynburgu. Nagrania pochodzą z jego istniejącej jeszcze strony *Vowels and Consonants*, www.phonetics.ucla.edu/vowels/contents.html. Peter Roach (2 serie wymówień), emerytowany profesor Uniwersytetu w Reading. Nagrania pochodzą z płyty dodanej do jego podręcznika *English Phonetics and Phonology. A Practical Course* (I wyd. z 1983 r.; IV wyd. z 2009 r.). Robert Mannel (dr, starszy wykładowca na Wydziale Nauk Humanistycznych Macquarie University w Sydney) i Jonathan Harrington (profesor fonetyki i cyfrowego przetwarzania mowy w Instytucie Fonetyki Uniwersytetu w Monachium, absolwent Uniwersytetu w Cambridge), autorzy materiałów dydaktycznych nt. fonetyki i fonologii opublikowanych na stronie Wydziału Lingwistyki Wydziału Nauk Humanistycznych (Macquarie University), clas.mq.edu.au/phonetics/transcription/ipa/ipa_vowel.html (na stronie internetowej nie podano nazwiska parlatora). Jill House (w pierwszej publikacji nagrań pojawiło się błędne nazwisko Susan Ramsaran), starszy wykładowca na University College w Londynie. Nagranie jej wymówień samogłosek podstawowych pojawia się na stronie <http://www.phonetics.ucla.edu/course/chapter1/wells/wells.html> obok wymówień Johna Wellsa i Petera Ladefogeda.

Wykres 2. Samogłoski podstawowe zaokrąglone [y ø œ ɤ ɔ o u ʊ] wymówione przez sześciu fonetyków



(Opracowanie własne na podstawie nagrań opublikowanych w Internecie (zob. przyp. 10)).

Wykres 3. Czworokąty samogłosek polskich wymówionych przez kobietę i mężczyznę (częstotliwości formantowe zlogarytmowane)



(Opracowanie własne na podstawie danych z: Jassem 2003: 105).

więc porównywać dowolnego głosu (samogłosek ze zbioru A) z jednym zbiorem modeli B¹⁰, ale konieczne jest dopasowywanie modeli do badanego głosu. Musimy więc dysponować wymówieniami samogłosek podstawowych wykonanych przez przedstawicieli obu płci w różnym wieku.

Jest to jednak warunek trudny do spełnienia, ponieważ poprawne wymówienie samogłosek podstawowych wymaga przede wszystkim zrozumienia modelu samogłosek podstawowych, a więc wykształcenia fonetycznego. Pomijając fakt, iż szukanie modeli dopasowanych do badanego informatora jest niejako powrotem do metody wykresowej, już samo zebranie odpowiednich modeli samogłosek podstawowych w wykonaniu wielu różnych mówców jest niesamowicie trudne. Każde to szukać innego rozwiązania.

Względne częstotliwości formantowe

Sposobem usunięcia opisanego wyżej problemu może być zastosowanie częstotliwości względnych. Zauważmy bowiem, że na wykresach nr 1 i 2 każdy czworokąt samogłoskowy zajmuje określoną przestrzeń. Załóżmy na razie bezbłędne wymówienie ukazanych na wspomnianych wykresach samogłosek podstawowych, by nie przejmować się tym problemem przynajmniej w tym momencie. Można powiedzieć, że dla każdego informatora rozrzut punktów mieści się w pewnym niezmiennym zakresie. Jeśli dla danego mówcy X najbardziej przednia samogłoska [i] ma częstotliwość F₂ równą ok. 2500 Hz, nie trzeba się zastanawiać, co będzie, gdy odkryjemy u innej osoby samogłoskę o częstotliwości F₂ równej 2700 Hz. Analogicznie w przypadku innego parametru: jeśli maksymalnie niska samogłoska osiąga u informatora Y częstotliwość F₁ równą 700 Hz, to tę wartość powinniśmy przyjąć za maksimum częstotliwości dla tego jednego mówcy Y. Inne maksima i minima mogą być zauważalne, ale w wymówieniach innych badanych.

Potrzebny jest teraz sposób pozwalający bezwzględne maksima i minima, będące dla różnych informatorów różnymi wartościami, zamienić na takie same wartości, a więc żeby – ujmując rzecz obrazowo – rozciągnąć lub ścieśnić czworokąty samogłoskowe ukazane na wykresach nr 1 i 2 w taki sposób, by zajmowały taką samą przestrzeń, ale żeby jednocześnie zachować względne położenie samogłosek pośrednich mieszczących się wewnątrz każdego czworokąta, a dokładniej między jego skrajnymi punktami. Innymi słowy potrzebujemy uzyskać częstotliwości formantowe względne każdej samogłoski modelowej.

Jeśli dana częstotliwość formantowa f określonej samogłoski wymawianej przez tego samego mówcę mieści się w przedziale $f_{min} \leq f \leq f_{max}$, to wówczas f_{min} możemy wyzerować, tzn. przyjąć za tę, różną w przypadku poszczególnych informatorów,

¹⁰ Bez względu na to, czy uśrednimy, czy zmedianujemy, czy jeszcze w inny sposób uogólnimy uzyskane częstotliwości formantowe samogłosek podstawowych, czy też po prostu arbitralnie wybierzemy te realizacje, które wydają nam się najlepsze.

wartość 0. W tym celu odejmujemy od f i f_{max} minimalną częstotliwość formantową, jaką dany formant osiąga dla wszystkich samogłosek wymówionych przez tę samą osobę. A więc uzyskujemy przedziały:

$$\text{Wzór 2: } f_{min} - f_{min} \leq f - f_{min} \leq f_{max} - f_{min}$$

co daje:

$$\text{Wzór 3: } 0 \leq f - f_{min} \leq f_{max} - f_{min}$$

Jeżeli teraz wartości $f_{max} - f_{min}$ przypiszemy 100%, procentowy odpowiednik wartości $f - f_{min}$ bardzo łatwo obliczymy z proporcji:

$$\begin{aligned} \text{Wzór 4: } f_{max} - f_{min} &= 100\% \\ f - f_{min} &= x \end{aligned}$$

a więc:

$$\text{Wzór 5: } x = \frac{f - f_{min}}{f_{max} - f_{min}} \cdot 100\%$$

Częstotliwość względną formantu o numerze n możemy oznaczyć symbolem f'_n , a ogólny wzór takiej częstotliwości będzie wyglądać następująco:

$$\text{Wzór 6: } f'_n = \frac{f_n - f_{nmin}}{f_{nmax} - f_{nmin}} \cdot 100\% = \frac{f_n - f_{nmin}}{R_n} \cdot 100\%$$

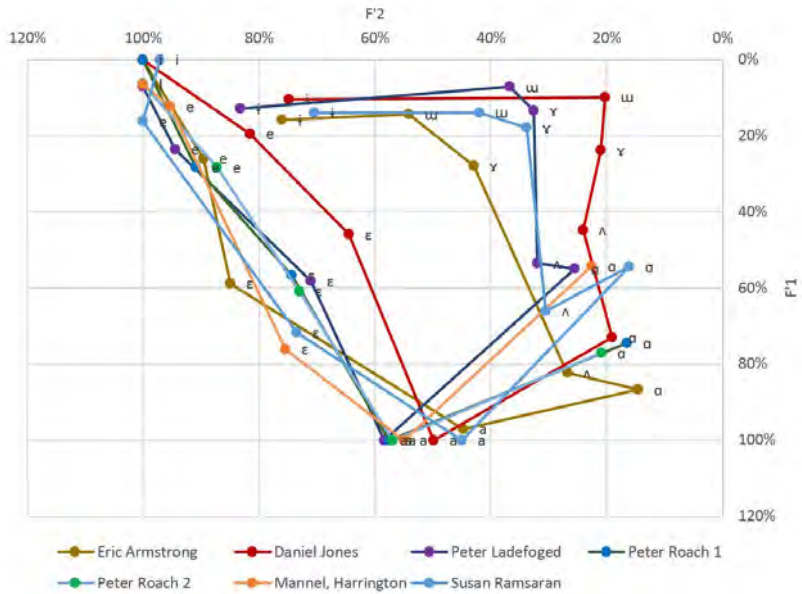
Różnicę $f_{nmax} - f_{nmin}$ możemy zastąpić rozstępem częstotliwości¹¹ dla formantu o numerze n . Mnożenie przez 100% nie jest konieczne, ale przydatne, by nie operować na ułamkach z przedziału [1; 0].

Niestety, to dalej nie rozwiązuje wszystkich problemów, ponieważ jeśli umieścimy wszystkie nasze modele na wykresie częstotliwości względnych, to zobaczymy, że pokryły nam się jedynie najbardziej skrajne samogłoski, przede wszystkim [i], [u], [a] (zob. wykres nr 4 i 5).

Nierównomiernie rozłożone są samogłoski przednie płaskie i tylne zaokrąglone. Dość duży rozrzut pojawia się także w przypadku samogłosek tylnych płaskich i przednich zaokrąglonych oraz centralnych (środkowych). Jest to zrozumiałe, ponieważ przyjmując za jednakową wartość dla każdego mówcy rozstęp częstotliwości ($f_{max} - f_{min}$), samogłoskami pośrednimi stały się te głoski, których częstotli-

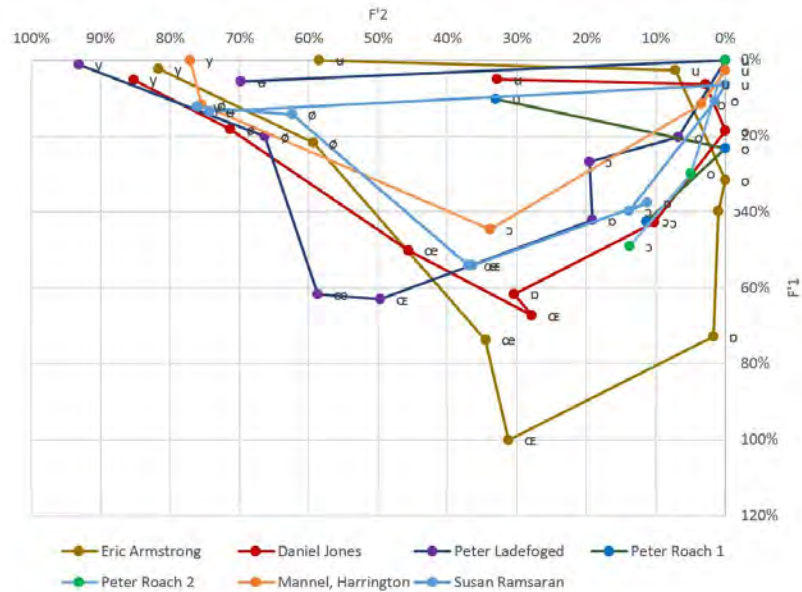
¹¹ Pojęcie wzorowane na rozstępie danych stosowanym w statystyce.

Wykres 4. Czworokąty samogłosek płaskich (wykres względnych częstotliwości formantowych)



(Opracowanie własne).

Wykres 5. Czworokąty samogłosek zaokrąglonych (wykres względnych częstotliwości formantowych)



(Opracowanie własne).

wości względne nie osiągają wartości skrajnych. Uzyskane rozbieżności wskazują także na różne wymawianie samogłosek podstawowych przez badanych informatorów (np. Eric Armstrong uprzednił samogłoskę [u], Peter Ladefoged upodobił do siebie samogłoski [ʌ] i [a]). To spostrzeżenie prowadzi z kolei do wniosku, iż uzyskane dane, tzn. częstotliwości względne samogłosek pośrednich są w niewielkim stopniu przydatne w celu opracowania uniwersalnego zestawu modeli do porównań. Należałoby bowiem arbitralnie odrzucić te samogłoski, które – jak się może wydawać – zostały wymówione błędnie (np. wspomniane wymówienia [u ʌ a]). To natomiast uszczupliłoby i tak już dość ograniczony zestaw modeli, a w przypadku samogłosek centralnych musielibyśmy zupełnie zrezygnować z niektórych modeli (np. wszystkie wymówienia samogłoski [ɜ] – nieuwzględnionej na wykresach, by zachować przejrzystość już i tak dosyć skomplikowanych wykresów – okazały się na tyle podobne do samogłoski [ɜ], że wydają się po prostu błędne i niewarte uwzględnienia w zestawie modeli).

Warto jednak w tym momencie przypomnieć, że mamy do czynienia z modelem samogłosek podstawowych (lub modelem czworokąta samogłoskowego Jonesa) zbudowanym według różnych kryteriów. Przywołajmy je w tym momencie:

- A. Samogłoski [i y a ɛ u ʊ ɔ] są scharakteryzowane artykulacyjnie jako skrajne ułożenia języka (choć w niektórych opisach czworokąta samogłoskowego [a ɛ], a nawet [u ʊ] uznaje się za pośrednie w odbiorze słuchowym). Te samogłoski możemy określić jako „**absolutnie skrajne artykulacyjnie**”.
- B. Samogłoski [e ø ε œ i ɪ ʊ o ʌ ɔ] są scharakteryzowane częściowo artykulacyjnie ([e ø ε œ] są skrajnie przednie, [i ɪ] są skrajnie wysokie itd.), a częściowo audytywnie – są to samogłoski wyraźnie różne od siebie i pośrednie między odpowiednimi skrajnymi samogłoskami wymienionymi w poprzednim punkcie. Ze względu na tę charakterystykę nazwijmy je „**częściowo skrajnymi artykulacyjnie**” samogłoskami, gdyż tylko jedna ich cecha dotyczy skrajnego położenia języka (nb. równie dobrze można by je nazwać częściowo pośrednimi, gdyż pośredniość w odbiorze słuchowym dotyczy tylko jednej z ich cech artykulacyjnych).
- C. Samogłoski średnio-środkowe [ə ɐ ɜ ɞ] mogą być określone tylko audytywnie – jako takie ułożenia pośrednie języka, które pozwalają wymówić samogłoski wyraźnie różne od odpowiednich samogłosek z grupy B. Skoro w przypadku tej grupy skrajne ułożenie języka nie ma zastosowania, ale jedynie słuchowe wrażenie pośredniości, nazwijmy te samogłoski „**absolutnie pośrednimi artykulacyjnie**”.
- D. Samogłoski [ɪ ʏ ʊ ɐ æ] są również scharakteryzowane audytywnie, choć można je również określić artykulacyjnie (np. [ɪ ʏ ʊ ɐ æ] to luźne warianty [i y u a]). Ponieważ wymówienia tych samogłosek okazały się nieprzydatne (pokrywały się z innymi samogłoskami), zostaną tutaj pominięte i zastąpione razem z samogłoskami poprzednich dwu grup samogłoskami o ekstrapolowanych¹² względnych częstotliwościach formantowych.

Przekładanie cech artykulacyjnych (ułożenie języka) na akustyczne (bezwzględne i względne częstotliwości formantowe)¹³ pozwala mówić o samogłoskach **akustycznie skrajnych i akustycznie pośrednich**, przy czym grupy samogłosek scharakteryzowanych akustycznie nie będą się pokrywać z analogicznymi grupami (samogłoskami skrajnymi i pośrednimi) określonymi artykulacyjnie.

Skoro rozstęp częstotliwości (wartość $f_{n\max} - f_{n\min}$ dla formantu o numerze n) określany jest przez samogłoski o skrajnych bezwzględnych częstotliwościach formantowych, to jedynie w ich przypadku można mówić o skrajności akustycznej. Nie zawsze będzie to skrajność absolutna: absolutnie skrajnymi akustycznie są samogłoski [i u], natomiast artykulacyjnie absolutnie skrajne [a ɶ y] okazały się jedynie częściowo skrajnie akustycznie: tylko jedna częstotliwość tych samogłosek przybrała wartość 0% lub 100%.

Z kolei absolutnie skrajne pod względem artykulacyjnym [y u a ɶ] okazały się **absolutnie pośrednie akustycznie**: wartości względnych częstotliwości formantowych tych samogłosek mieszczą się w zakresach obustronnie otwartych (0%; 100%). Oczywiście, pozostałe samogłoski pośrednie artykulacyjnie (grupy C i D w powyższym wyliczeniu) są także absolutnie pośrednie akustycznie.

Pośredniość audytywną można rozumieć jako pośredniość częstotliwości formantowych, gdyż właśnie głoski o różnej budowie akustycznej mogą być percypowane przez słuchacza jako różne. Jeśli dwie samogłoski mają prawie identyczną budowę akustyczną (a więc zbliżone częstotliwości formantowe), słuchacz nie powinien rozróżnić tych głosek. Możemy zatem pośredniość audytywną postrzegać w kategoriach akustycznych, a więc jako pośredniość częstotliwości formantowych. Tak więc [e] i [ɛ], scharakteryzowane już w modelu samogłosek podstawowych jako pośrednie ułożenia języka w kierunku pionowym, możemy zamienić na proporcjonalne zmiany względnej częstotliwości formantu pierwszego. Ponieważ skrajne (akustycznie) samogłoski [i a] przyjęły prawie zawsze skrajne wartości f'_1 (0% dla [i], 100% dla [a]), wówczas samogłosce [e] możemy przypisać $\frac{1}{3}$ rozstępu częstotliwości względnej F_1 , a samogłosce [ɛ] $\frac{2}{3}$ tej wartości.

Nieco inaczej przedstawia się kwestia częstotliwości względnych formantu drugiego. Analiza wykresu nr 4 pokazuje, że zmienia się ona liniowo w przypadku samogłosek [e ɛ] – zmniejsza się wraz z obniżaniem artykulacji. Położenie języka podczas artykulacji [e] i [ɛ] jest w modelu samogłosek podstawowych pośrednie także w kierunku poziomym. Możemy więc różnicę między f'_2 samogłosek [i] oraz [a] podzielić na 3 części tak jak w przypadku różnicy między f'_1 samogłoski [i] oraz [a]:

- samogłoska [e]: $f'_2 = \frac{1}{3} \cdot (|f'_{2[i]} - f'_{2[a]}|)$
- samogłoska [ɛ]: $f'_2 = \frac{2}{3} \cdot (|f'_{2[i]} - f'_{2[a]}|)$

12 Tzn. obliczonych na podstawie skrajnych wartości f' samogłosek wymówionych przez badanych informatorów (sposób obliczania częstotliwości względnych ekstrapolowanych opisano w dalszej części).

13 A taką operację wykonaliśmy w przypadku samogłosek [i y u a ɶ a ɶ] – skrajne ułożenie języka zamieniliśmy na względne częstotliwości formantowe.

Analogicznie możemy ustalić względne częstotliwości formantowe wszystkich samogłosek podstawowych, dysponując najbardziej skrajnymi wartościami względnych częstotliwości formantowych uzyskanych z zamiany bezwzględnych częstotliwości formantowych na względne. Ciągłe punktem wyjścia są wymówienia wspomnianych wcześniej fonetyków, które stanowią wzór, na podstawie którego stwierdzamy, które względne częstotliwości formantowe osiągają wartości skrajne. Bezwzględne częstotliwości formantowe będą różne dla różnych mówców, ale stosunki zachodzące między nimi powinny być (takie jest przyjmowane tu założenie) takie same dla takiego samego ułożenia narządów mowy¹⁴.

Samogłoskę neutralną [ə], jak powiedzieliśmy wcześniej, można scharakteryzować akustycznie, a więc w odniesieniu do częstotliwości bezwzględnych, a co za tym idzie także i względnych. Zanim jednak spróbujemy ustalić te parametry, przypomnijmy charakterystykę artykulacyjną głoski [ə]. Otóż jest ona określana jako średnia, środkowa, płaska i ustna. A więc znajduje się dokładnie w centrum pola artykulacyjnego samogłosek. Taka definicja odniesiona do względnych częstotliwości formantowych wskazywałaby na zbliżone wartości względnych częstotliwości formantowych tej samogłoski i równe (w przybliżeniu) w każdym przypadku 50%. Sprawdźmy więc, czy rzeczywiście dla samogłoski [ə] zachodzi poniższa zależność:

Wzór 7: $f'_1 \approx f'_2 \approx f'_3 \approx f'_4 \approx 50\%$

Przypomnijmy wzór na częstotliwość względną formantu n (wzór nr 6): $f'_n = \frac{f_n - f_{nmin}}{R_n}$ i relacje między częstotliwościami bezwzględnymi dla samogłoski neutralnej:

Wzór 8: $f_1 = \frac{f_2}{3} = \frac{f_3}{5} = \frac{f_4}{7} \dots$

Ze wzoru na częstotliwość względną f'_n (wzór nr 6) obliczymy częstotliwość bezwzględną:

Wzór 9: $f_n = f'_n \cdot R_n + f_{nmin}$

a po wstawieniu do relacji między częstotliwościami bezwzględnymi samogłoski neutralnej (wzór nr 8):

Wzór 10: $f'_1 \cdot R_1 + f_{1min} = \frac{1}{3}(f'_2 \cdot R_2 + f_{2min}) = \frac{1}{5}(f'_3 \cdot R_3 + f_{3min}) \dots$

Podstawmy za f' wartość 50%, którą sprawdzamy:

Wzór 11: $50\% R_1 + f_{1min} = \frac{1}{3}(50\% R_2 + f_{2min}) = \frac{1}{5}(50\% R_3 + f_{3min}) \dots$

¹⁴ Niewielkie różnice między samogłoskami rozpoznanymi a docelowymi (tzn. jakie starali się wymówić badani informatorzy) uzyskane w przeprowadzonym teście metody częstotliwości względnych (opis testu w dalszej części artykułu) zdają się potwierdzać przyjęte tu założenie: samogłoski podstawowe wymawiane przez badanych informatorów (trzech fonetyków) charakteryzowały się bardzo podobnymi względnymi częstotliwościami formantowymi.

Przypomnijmy, że rozstęp częstotliwości R_n dla formantu o numerze n to: $f_{n\max} - f_{n\min}$, czyli po wstawieniu za R_n te same różnice uzyskujemy:

$$\begin{aligned} \text{Wzór 12: } 50\% f_{1\max} - 50\% f_{1\min} + 100\% f_{1\min} &= \frac{1}{3} (50\% f_{2\max} - 50\% f_{2\min} + 100\% f_{2\min}) = \\ &= \frac{1}{5} (50\% f_{3\max} - 50\% f_{3\min} + 100\% f_{3\min}) \dots \end{aligned}$$

Dodatkowo zamieniliśmy f_{\min} na $100\% f_{\min}$, co ułatwi dalsze obliczenia:

$$\text{Wzór 13: } 50\% (f_{1\max} + f_{1\min}) = \frac{50\%}{3} (f_{2\max} + f_{2\min}) = \frac{50\%}{5} (f_{3\max} + f_{3\min}) \dots$$

Po skróceniu 50% otrzymujemy:

$$\text{Wzór 14: } f_{1\max} + f_{1\min} = \frac{1}{3} (f_{2\max} + f_{2\min}) = \frac{1}{5} (f_{3\max} + f_{3\min}) \dots$$

Maksymalną wartość częstotliwości dla każdego formantu obliczymy ze wzoru na rozstęp częstotliwości: skoro $R_n = f_{n\max} - f_{n\min}$, to $f_{n\max} = R_n + f_{n\min}$, a to pozwoli skrócić uzyskaną relację, gdyż po podstawieniu za $f_{n\max}$ sumy $R_n + f_{n\min}$ uzyskamy:

$$\text{Wzór 15: } R_1 + f_{1\min} + f_{1\min} = \frac{1}{3} (R_2 + f_{2\min} + f_{2\min}) = \frac{1}{5} (R_3 + f_{3\min} + f_{3\min}) \dots$$

czyli:

$$\text{Wzór 16: } R_1 + 2f_{1\min} = \frac{1}{3} (R_2 + 2f_{2\min}) = \frac{1}{5} (R_3 + 2f_{3\min}) \dots$$

Tak więc częstotliwości względne poszczególnych formantów samogłoski [ə] będą równe 50%, jeśli tylko suma zakresów (rozstępów) częstotliwości bezwzględnych i podwojonej częstotliwości minimalnej będzie wielokrotnością kolejnych nieparzystych liczb naturalnych – 1, 3, 5 i 7. Tabela nr 1 zawiera zestawienie rozstępów częstotliwości badanych samogłosek podstawowych oraz ilorazy tych wartości i liczb 1, 3, 5, 7.

Jak widać, uzyskane wartości nie są równe, tak więc względne częstotliwości formantowe samogłoski [ə] nie powinny być równe 50%. Jest to zrozumiałe, jeśli spróbujemy sobie wyobrazić, gdzie na wykresach nr 4 i 5 znajdzie się samogłoska [ə]. Jej położenie będzie można ustalić w odniesieniu do kilku również względnych (ustalonych na podstawie innych) punktów. Sposób określenia tego położenia przedstawia rysunek nr 1.

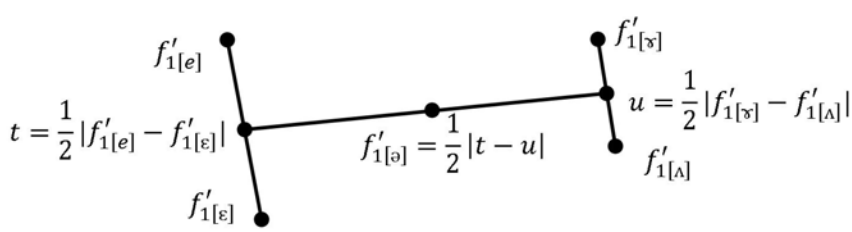
Wiemy już, że wartości $f'_{1[e]}$, $f'_{1[ε]}$, $f'_{1[ɜ]}$, $f'_{1[ə]}$ nie są skrajne (samogłoski [e ε ɜ ə] określiliśmy jako absolutnie pośrednie akustycznie), trudno więc oczekiwać, że $f'_{1[ə]}$ będzie miało dokładnie wartość równą połowie rozstępu częstotliwości R_1 .

Tabela 1. Sumy rozstępów częstotliwości bezwzględnych (R_n) i podwojonych minimów poszczególnych częstotliwości formantowych oraz te same wartości podzielone przez kolejne nieparzyste liczby naturalne

Suma zakresów bezwzględnych i podwojonych minimów częstotliwości							
Parlator	Eric Armstrong	Daniel Jones	Peter Ladefoged	Peter Roach	Peter Roach 2	Mannel, Harrington	Jill House
$R_1 + 2f_{1min}$	969,4	1141,0	1186,1	1252,6	1193,9	1110,2	1376,3
$R_2 + 2f_{2min}$	2931,2	3213,0	2575,1	3164,2	3129,3	2972,8	3285,3
$R_3 + 2f_{3min}$	4336,3	5554,0	5132,6	5620,2	5507,6	5588,7	6229,7
$R_4 + 2f_{4min}$	6453,3	6950,1	7057,5	7159,3	6921,0	7187,7	8727,7
Wyniki dzielenia							
$\frac{R_1 + 2f_{1min}}{1}$	969,4	1141,0	1186,1	1252,6	1193,9	1110,2	1376,3
$\frac{R_2 + 2f_{2min}}{3}$	977,1	1071,0	858,4	1054,7	1043,1	990,9	1095,1
$\frac{R_3 + 2f_{3min}}{5}$	867,3	1110,8	1026,5	1124,0	1101,5	1117,7	1245,9
$\frac{R_4 + 2f_{4min}}{7}$	921,9	992,9	1008,2	1022,8	988,7	1026,8	1246,8

(Opracowanie własne).

Rysunek 1. Sposób określenia względnej częstotliwości formantu pierwszego samogłoski neutralnej



(Opracowanie własne).

Obliczanie względnych częstotliwości formantowych samogłosek modelowych

Na podstawie dotychczasowych ustaleń możemy zbudować macierz częstotliwości względnych samogłosek podstawowych płaskich i zaokrąglonych, a także uzupełnić tę macierz o samogłoski pośrednie względem tych, które pojawiły się nie tylko

w czworokącie samogłoskowym Jonesa, ale także w jego wersji rozszerzonej przyjętej przez Międzynarodowe Towarzystwo Fonetyczne w 1989 r.¹⁵:

Tabela 2. Symbole samogłosek płaskich i zaokrąglonych rozszerzonej wersji czworokąta samogłoskowego przyjętej przez Międzynarodowe Towarzystwo Fonetyczne w 1989 r.

Samogłoski płaskie			Samogłoski zaokrąglone		
i	ɪ	ʉ	y	ʏ	u
ɪ			ʏ		ʊ
e	ə	ɘ	ø	ɵ	o
	ə				
ɛ	ɜ	ʌ	œ	ɞ	ɔ
æ	ɐ				
a		ɑ	ɶ		ɒ

(Opracowanie własne na podstawie Nolan i in. 1999).

Czworokąty samogłoskowe ukazane na wykresach nr 4 i 5 pozwalają ustalić względne częstotliwości samogłosek skrajnych akustycznie (absolutnie i częściowo), czyli [i a u y ɛ ɐ u] – będą to skrajne częstotliwości względne osiągnięte przez te samogłoski. Położenie pozostałych samogłosek (czyli ich względne częstotliwości formantowe) obliczymy na podstawie częstotliwości samogłosek skrajnych (według metody opisanej wyżej dla [e ɛ] oraz [ə]). Tym sposobem będziemy mogli uzupełnić wszystkie puste miejsca w tabeli nr 2.

Uzasadnieniem dla uzupełniania zestawu samogłosek modelowych o samogłoski nieobecne w zestawie zastosowanym w międzynarodowym alfabecie fonetycznym jest z jednej strony zauważona wcześniej nieprzydatność parametrów samogłosek [ɪ ʏ ʊ ɐ æ], a z drugiej nierównomierne rozłożenie samogłosek w rozszerzonym czworokącie samogłoskowym użytym w IPA (np. brak zaokrąglonych odpowiedników [ɐ æ ə] oraz brak płaskiego wariantu [ʊ]).

Poza tym uszczegółowienie zestawu modeli zwiększy precyzję rozpoznawania samogłosek i (co ważniejsze) zmniejszy ryzyko zbyt automatycznego przypisywania im etykiet artykulacyjnych. Jeśli bowiem korzystalibyśmy z ograniczonego zestawu modeli (np. takiego jak w tabeli nr 2, ale bez samogłosek [ɪ ʏ ʊ æ ɐ]) i badalibyśmy niezbyt staranne wymówienie samogłoski nieśrodkowej (np. przedniej), wówczas zaimplementowany w arkuszu kalkulacyjnym algorytm rozpoznawania samogłosek na podstawie ich częstotliwości względnych mógłby takie wymówienie przyporządkować do szeregu samogłosek środkowych. Dysponując zestawem uzupełnionym

¹⁵ Zob. Nolan i in. 1999.

Tabela 3. Rozszerzony zestaw modelowych samogłosek (razem z samogłoskami ekstrapolowanymi [o ekstrapolowanych wszystkich częstotliwościach formantowych względnych] zaznaczonymi jaśniejszą czcionką)

Samogłoski płaskie							Samogłoski zaokrąglone						
i	ĩ	ᵢ	ɪ	ᵢ̃	ũ	u	y	ÿ	ʏ	ʘ	ʘ̃	ü	u
ɨ	ɨ̃	ᵢ̃	ɨ̃	ũ̃	ũ̃	ʉ	ɤ	ɤ̃	ʏ̃	ʘ̃	ʘ̃	ʉ̃	ʉ̃
e	ẽ	ɘ	ɘ̃	ɘ̃	ʏ	ʏ̃	ø	ø̃	ɘ̃	ɘ̃	ɘ̃	ö	o
ɛ	ẽ̃	ɘ̃	ɘ̃	ɘ̃	ʏ̃	ʏ̃	ø̃	ø̃̃	ɘ̃̃	ɘ̃̃	ɘ̃̃	ö̃	õ
ɛ̃	ẽ̃̃	ɘ̃̃	ɘ̃̃	ɘ̃̃	ʏ̃̃	ʏ̃̃	ø̃̃	ø̃̃̃	ɘ̃̃̃	ɘ̃̃̃	ɘ̃̃̃	ö̃̃	õ̃
æ	ǣ	ɶ	ɶ̃	ɶ̃	ǣ̃	ʌ	œ	œ̃	ɶ̃̃	ɶ̃̃	ɶ̃̃	ǣ̃̃	ʌ̃̃
a	ǣ̃̃	ɶ̃̃̃	ɶ̃̃̃	ɶ̃̃̃	ǣ̃̃̃	ʌ̃̃̃	œ̃̃̃	œ̃̃̃̃	ɶ̃̃̃̃	ɶ̃̃̃̃	ɶ̃̃̃̃	ǣ̃̃̃̃	ʌ̃̃̃̃

(Opracowanie własne).

o większą liczbę samogłosek (zob. tabela nr 3), można by uzyskać dokładniejszy wynik: samogłoskę przednią cofniętą lub środkową przesuniętą do przodu.

Można wreszcie podać argument jednolitości zestawu samogłosek modelowych: korzystamy bowiem z rozszerzonego czworokąta samogłoskowego (tabela nr 2), w którym wypełnienie przestrzeni artykulacyjnej samogłosek jest nierównomierne. Oczywiście, wynika to z przyczyn praktycznych: wprowadzenie symboli samogłosek [ɪ ʏ ʊ æ ɶ ɘ] okazało się przydatne w fonetyce i fonologii, spowodowało jednak powstanie niejednolitego zestawu modeli¹⁶. Proponowany model (tabela nr 3) w pełni równomiernie wypełnia tę przestrzeń.

Można przytoczyć jeszcze dwa argumenty za wyborem dość dużego, bo 98-elementowego zbioru samogłosek modelowych (49 płaskich i 49 zaokrąglonych; zob. przypis 17). Pierwszy dotyczy badań kontrastywnych i dydaktyki języków obcych: Rezygnacja z wielu samogłosek ukazanych w tabeli nr 3 mogłaby prowadzić do nieco mylnych rozpoznań. Na przykład ogólnopolska samogłoska w wyrazach typu *byk*, *syk* mogłaby być – zależnie od otoczenia fonetycznego lub pochodzenia informatora – rozpoznana jako bliska [i] lub [ɨ], a nawet [e] bądź [ɘ], a wiemy przecież, że niezależnie od odmiany polszczyzny (za wyjątkiem może gwar północnej Polski, gdzie omawiana

16 Zwróćmy uwagę, że w płaszczyźnie pionowej mamy siedem stopni zwężenia (lub rozwarcia): samogłoski wysokie typu [i ʊ], wysokie obniżone (typ [ɪ ʊ]), średniowysokie (typ [e ɔ]), średnie (tu tylko [ə]), średnioniskie (typ [ɛ ɔ]), niskie zwężone (typ [æ ɐ]), niskie (typ [a ɑ]). W płaszczyźnie poziomej co najmniej pięć: przednie, przednie cofnięte – [ɪ ʏ], środkowe (centralne), tylne przesunięte do przodu – [ʊ], tylne. Jednakże samogłoski [ɪ ʏ ʊ] nie leżą w modelu IPA dokładnie w połowie odległości między szeregiem przednich (tylnych dla [ʊ]) i środkowych samogłosek, lecz są nieco przesunięte ku przodowi ([ɪ ʏ]) lub tyłowi ([ʊ]), a to z kolei pozwala mówić o jeszcze dwóch szeregach pośrednich: samogłoski środkowe (centralne) nieco uprzednione, samogłoski środkowe (centralne) nieco cofnięte. Mamy więc 49 (7 × 7) modeli samogłosek płaskich i tylu zaokrąglonych.

tu głoska może być mieszana z [i]) jest to samogłoska znajdująca się gdzieś pomiędzy samogłoskami [i i e ɐ]. Niezupełnie satysfakcjonujące rozpoznania mogłyby się pojawiać również w przypadku samogłosek niskich. Na przykład polska samogłoska w wyrazach typu *bat*, *bak* mogłaby być w odniesieniu do uproszczonego modelu uznana za [a] lub [a]. Pierwsze rozpoznanie zrówna ją z angielską [æ] (która z kolei mogła by być zaklasyfikowana do [a] lub [ɛ]), a drugie z angielską [ɑː].

Poza tym wreszcie model 98-elementowy jest tutaj proponowany z myślą o automatycznym (z wykorzystaniem komputera) rozpoznawaniu samogłosek i przypisywaniu im symboli fonetycznych, a także przetwarzaniu w dalszej kolejności tych symboli (po zamianie na kody oznaczające cechy artykulacyjne, o czym będzie mowa w dalszej kolejności). Tak więc proponowany model wydaje się optymalny, szczególnie gdy bierze się pod uwagę wykorzystanie prezentowanej tu metody w badaniach dialektalnych (a przecież polskie gwary posiadają więcej samogłosek o artykulacji pośredniej względem samogłosek ogólnopolskich) i kontrastywnych.

Literatura

- FANT G., 2004: *Speech Acoustics and Phonetics*, London.
- GARCZYŃSKA J., 2007a, *Analiza fonetyczna akcentowanych samogłosek ustnych w mowie Polek z Kazachstanu*, Warszawa.
- GARCZYŃSKA J., 2007b, *Zastosowanie metod fonetyki akustycznej w badaniach dialektologicznych (na przykładzie akcentowanej samogłoski [a]), [w:] J. Sierociuk (red.), Gwary dziś*, t. 4, Poznań, s. 237–247.
- JASSEM W., 1973, *Podstawy fonetyki akustycznej*, Warszawa.
- JASSEM W., 2003, *Polish*, „Journal of the International Phonetic Association” 33/1, s. 103–107.
- NOLAN F. i in., 1999, *Handbook of the International Phonetic Association. A Guide to the Use of the International Phonetic Alphabet*, Cambridge.
- ORAWSKI M., 2001, *Analiza akustyczna samogłosek pochylonych na podstawie przykładów z Orawy i Śląska*, [w:] J. Sierociuk (red.), *Gwary dziś*, t. 1, Poznań, s. 147–152.

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

Summary

The first part of the paper proposes a method of studying (or, more precisely, of recognizing) oral vowels, using the so-called relative formant frequencies.

First, we describe the traditional method of acoustic study of oral vowels, together with its limitations. Next, we introduce the assumptions of the proposed method where, unlike in the traditional approach, a set of fixed models is used. The comparison, instead of being reduced to establishing whether two sets of vowels match, requires a distance to be calculated between each vowel and each model.

The paper introduces the notion of relative formant frequency (denoted f') which was used in the comparisons to minimize the impact of individual articulatory variation on the outcome of the analysis. It was defined as the proportion between, on the one hand, the difference between the absolute frequency of the given formant, and its minimal value observed in actual pronunciations, and on the other hand, the difference between the minimal and the maximal frequency of the formant, as observed in pronunciations.

In order to establish the f' of model vowels, we used the cardinal vowels pronounced by several British phoneticians (including Daniel Jones).