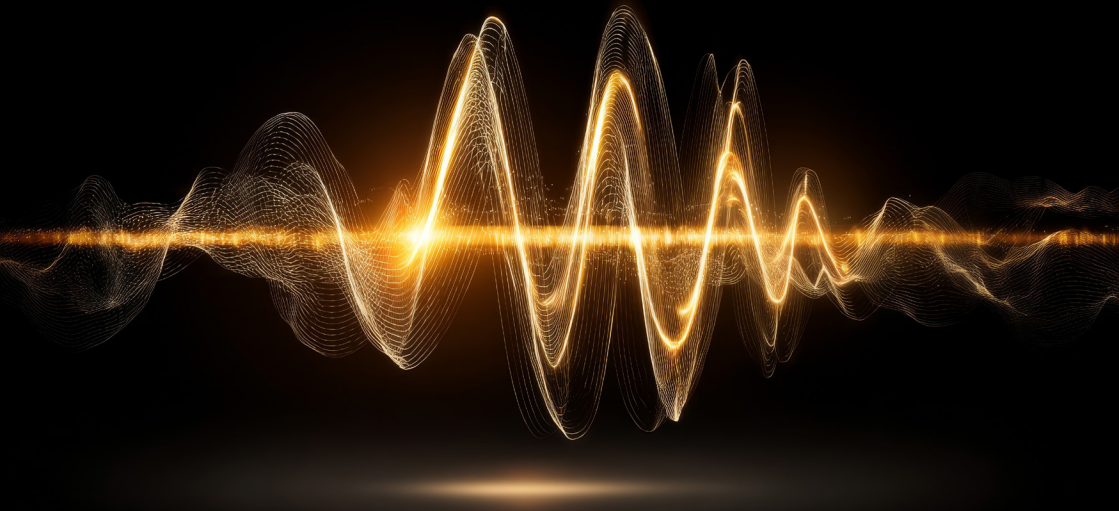


Ewa Piorun

Wybrane aspekty tonalności.
Badania porównawcze języków
chińskiego (mandaryńskiego) i polskiego



**Wybrane aspekty tonalności.
Badania porównawcze języków
chińskiego (mandaryńskiego) i polskiego**

Ewa Piorun

**Wybrane aspekty tonalności.
Badania porównawcze języków
chińskiego (mandaryńskiego) i polskiego**



Poznań 2025

Projekt okładki:
Wydawnictwo Rys

Recenzja:
dr hab. Paweł Nowakowski, prof. UAM

Copyright by:
Ewa Piorun

Copyright by:
Wydawnictwo Rys

Wydanie I
Poznań 2025

ISBN 978-83-68668-03-2

DOI 10.48226/978-83-68668-03-2

Wydanie:



Wydawnictwo Rys
ul. Kolejowa 41
62-070 Dąbrówka
tel. 600 44 55 80
e-mail: tomasz.paluszynski@wydawnictworys.com
www.wydawnictworys.com

Spis treści

Wstęp	9
Rozdział 1.	
Tonalność w języku mandaryńskim	13
1.1. Opisy i transkrypcje systemu tonalnego	14
1.2. Właściwości tonu i jego fonetyczne korelaty	19
1.2.1. Wysokość (<i>pitch</i>) i częstotliwość podstawowa (F_0)	19
1.2.2. Napięcie fałdów głosowych	21
1.2.3. „Szmer”	21
1.2.4. Temporalne właściwości tonów	21
1.3. Czynniki wpływające na realizację tonów	22
1.4. Fonologiczna reprezentacja tonu	23
1.4.1. Składniki tonu i ich poziomy kontrastywne	23
1.4.2. Tony konturowe (<i>contour tones</i>) i tony poziomowe (<i>level tones</i>) ...	23
1.4.3. Jednostka nacechowana tonem (TBU, <i>Tone Bearing Unit</i>)	28
1.5. Relacja ton – segment	30
1.5.1. Ton jako zjawisko suprasegmentalne	31
1.5.2. Ton jako zjawisko segmentalne	32
1.5.3. Ton w ujęciu fonologii autosegmentalnej	33
1.6. Ton, akcent i intonacja	35
1.6.1. Ton a akcent	35
1.6.2. Ton a intonacja	40
1.7. Akcent w języku polskim	45
Rozdział 2.	
Zjawisko tonalnego <i>sandhi</i>	61
2.1. <i>Sandhi</i> tonalne w języku chińskim (mandaryńskim)	62
2.1.1. <i>Sandhi</i> tonu 3	63
2.2. Domena <i>sandhi</i> tonu 3 w badaniach lingwistycznych	65
2.2.1. Podejście cykliczne (<i>The Cyclic Approach</i>)	65
2.2.2. Podejście Kaisse – Zasada gałęzi (<i>The Branch Condition</i>)	66
2.2.3. Podejście prozodyczne (<i>The Prosodic Approach</i>) – model Shih	68
2.2.4. Koncepcja Przypadku Specjalnego (<i>The Special Circumstance Approach</i>)	73
2.2.5. Koncepcja Jednostki Sensownej (<i>The Sense Unit Approach</i>)	73
2.2.6. Teoria Funktorów (<i>The Functor Approach</i>)	75
2.2.7. Teoria Efektu Semantyczno-Pragmatycznego (<i>The Semantic-Pragmatic Effect Approach</i>)	76

2.2.8. Teoria Klityk (<i>The Clitic Approach</i>)	77
2.2.9. Koncepcja Prozodycznej Hierarchii (<i>The Prosodic Hierarchy Proposal</i>)	77
2.2.10. Teoria Optymalności (<i>Optimality Theory</i>)	80
2.3. Fonetyczna natura tonu powstałego na skutek <i>sandhi</i>	92
Rozdział 3.	
Część eksperymentalna	97
3.1. Analiza tonograficzna wyrazów chińskich i polskich wypowiadanych przez rodzimych użytkowników języka mandaryńskiego	97
3.1.1. Materiał językowy	97
3.1.2. Uczestnicy	99
3.1.3. Przebieg eksperymentu	99
3.1.4. Edycja materiału dźwiękowego	100
3.1.5. Analiza	101
3.2. Analiza tonograficzna wyrazów chińskich i polskich wypowiadanych przez nierodzimych użytkowników języka mandaryńskiego	103
3.2.1. Materiał językowy	103
3.2.3. Przebieg eksperymentu, edycja oraz analiza danych	104
3.3. Badanie percepcji słuchowej tonów chińskich wśród nierodzimych użytkowników języka mandaryńskiego	105
3.3.1. Materiał językowy	105
3.3.2. Słuchacze	105
3.3.3. Przebieg eksperymentu	106
3.3.4. Analiza	107
Rozdział 4.	
Uzyskane rezultaty	109
4.1. Realizacja tonów chińskich przez rodzimych użytkowników języka mandaryńskiego	109
4.1.1. Część I – język mandaryński	109
4.1.2. Część II – język polski	122
4.1.3. Porównanie wyników obu części eksperymentu	137
4.2. Realizacja tonów chińskich przez nierodzimych użytkowników języka mandaryńskiego	139
4.2.1. Część I – język mandaryński	139
4.2.2. Część II – język polski	149
4.2.3. Porównanie wyników obu części eksperymentu	163
4.3. Porównanie wyników rodzimych i nierodzimych użytkowników języka mandaryńskiego	164

4.4. Test słuchowy	172
4.4.1. Kompetencja językowa oraz indywidualne predyspozycje słuchowe odbiorcy	175
4.4.2. Wpływ nadawcy	179
4.4.3. Zależność od rodzaju tonu	180
Konkluzje	191
Aneks	195
Bibliografia	197
Abstract	201
Spis tabel	203
Spis rycin	205

Wstęp

Języki polski i mandaryński niewątpliwie różnią się między sobą zasadniczo, począwszy od struktury morfologicznej jednostek leksykalnych, poprzez system pisma, składnię, sposób artykulacji głosek, na prozodii kończąc. Wzajemne zależności pomiędzy tymi elementami, które w procesie ewolucji każdego z języków ukształtowały się w odmienny sposób, spowodowały, że w jednym z systemów zaistniały podstawy do wykształcenia się tonów, podczas gdy w drugim takiej potrzeby nie było. Z uwagi na stałą liczbę sylab w języku mandaryńskim, która – w zależności od źródła – stanowi od 410 do 413 dozwolonych kombinacji nagłosów i wygłosów oraz fakt, że duża część wyrazów to monosylaby, konieczny był mechanizm umożliwiający dodatkowe różnicowanie homofonicznych sylab pod względem leksykalnym. Z kolei język polski, ze względu na zasady rządzące jego morfologią, ma praktycznie niewyczerpany potencjał słowotwórczy, umożliwiający generowanie nowych jednostek leksykalnych, którym przypisać można nowe znaczenia, z zestawu zaledwie około 40¹ pojedynczych segmentów. Właściwość ta przekłada się także na istnienie w języku polskim ogromnej, choć bliżej nieznannej liczby możliwych do wymówienia sylab. Przy tak szerokim inwentarzu sylab tony byłyby elementem zbędnym, mogącym wprowadzić niepotrzebny chaos w języku. Te i inne różnice, wyraźnie zarysowane praktycznie w każdym aspekcie lingwistycznym, sprawiają, że oba te języki stanowią niezbyt produktywny grunt dla porównawczych badań eksperymentalnych. Nie dziwi zatem fakt, że dotychczas nie próbowano zbadać, czy pewne cechy języka polskiego faktycznie stanowią barierę uniemożliwiającą zaistnienie w nim tonów, czy też jest to spowodowane względami czysto praktycznymi. Abstrahując od kwestii ich przydatności w polszczyźnie, interesujące jest jednak to, czy w języku nietonalnym, typowo spółgłoskowym, o strukturze tak dalece odbiegającej od struktury języka mandaryńskiego, realizacja tonów jest w ogóle możliwa. Zagadnienie to podejmuje niniejsza praca, której głównym celem jest sprawdzenie, czy osoby będące rodzimymi użytkownikami jednego z dwóch języków

¹ W. Jassem (2003) wyróżnia w języku polskim 38 fonemów (s. 103–107).

– mandaryńskiego lub polskiego – i znające drugi z nich na poziomie co najmniej podstawowym, potrafią wymówić polskie wyrazy jednosylabowe na czterech tonach mandaryńskich, zachowując prawidłowe kontury tonalne. Dodatkowo zbadano, czy tak zrealizowane tony są „czytelne” dla polskich słuchaczy mających zarówno wiedzę teoretyczną, jak i praktyczne umiejętności w zakresie posługiwania się językiem mandaryńskim. Istotne jest również to, w jakim stopniu ewentualna poprawna realizacja uwarunkowana jest poziomem znajomości danego języka obcego wśród mówców, a także jak dużą rolę odgrywają w tym procesie indywidualne możliwości bądź umiejętności naśladowcze poszczególnych osób. Dociekania te mają pomóc ustalić, czy z zadaniem tym lepiej radzą sobie *native speakerzy* języka tonalnego, którego element imitowano, czy też języka spółgłoskowego, o mało zróżnicowanej intonacji (za jaki powszechnie uważa się język polski), którego sylaby poddawano imitacji. Obszar zainteresowań niniejszej pracy stanowi również wpływ jakości realizowanych tonów na trafność ich identyfikacji oraz wszelkie inne czynniki mogące w znaczący sposób ułatwiać bądź utrudniać odbiór słuchowy. Uzyskanie odpowiedzi na poruszone tu kwestie mają umożliwić odpowiednio zaprojektowane eksperymenty, będące głównym przedmiotem tej pracy, badające zarówno produkcję, jak i percepcję tonów języka mandaryńskiego. Metodyka oraz wyniki przeprowadzonych badań własnych zostały zaprezentowane kolejno w rozdziałach 3 i 4.

Rozpatrując produkcję tonów w obu językach, warto przyjrzeć się różnicom, jakie występują pomiędzy nimi pod względem struktury. Choć wciąż trwają dyskusje nad pojemnością elementarnej jednostki nacechowanej tonem (por. § 1.4.3), różnice te najłatwiej zilustrować na przykładzie sylaby, która stanowi w tym wypadku dogodny wspólny mianownik. Z punktu widzenia fonetyki w języku mandaryńskim – podobnie jak i w wielu językach europejskich – wyróżnić można cztery podstawowe typy struktur sylabowych: V, VC, CV oraz CVC, jednakże istnieją pewne warunki, które odróżniają standardową odmianę języka chińskiego od wspomnianych języków europejskich. Zhang (1996) przedstawia je w postaci pięciu zasad:

- 1) W mówionym języku chińskim nie występują zbitki spółgłoskowe. Zarówno w pozycji nagłosu, jak i wygłosu dozwolone są co najwyżej pojedyncze spółgłoski.

- 2) W odmianie standardowej (mandaryńskiej) w wygłosie możliwe są jedynie spółgłoski nosowe /n/ [n] oraz /ng/ [ŋ].
- 3) Sylaby chińskie na ogół składają się z nie więcej niż czterech segmentów, przy czym ósrodek i wygłos mogą zawierać łącznie maksymalnie trzy segmenty.
- 4) Spółgłoska /ng/ [ŋ] nigdy nie występuje w nagłosie.
- 5) W języku mandaryńskim każda sylaba ma jeden z pięciu tonów: równy, rosnący, opadająco-wznoszący, opadający lub tzw. ton zerowy (neutralny), występujący jedynie na poziomie wyrazów.

Jak podaje Duanmu (2000: 79–81), struktura sylaby o maksymalnej pojemności ma postać CGVX, gdzie G jest wysoką półsamogłoską (ang. *glide*) zlokalizowaną przed ósrodkiem danej sylaby, natomiast VX może być dyftongiem, samogłoską długą lub też kombinacją samogłoski i spółgłoski.

Dla porównania język polski jest w tej kwestii mało restrykcyjny, bowiem „dopuszcza wyjątkowo dużą liczbę spółgłosek zarówno w nagłosie, jak i w wygłosie” (Śledziński 2008, za: Łobacz, Jassem 1974: 179–197). Nie stwarza on zatem warunków sprzyjających realizacji tonów, co może stanowić źródło potencjalnych trudności dla osób badanych (w szczególności rodzimych użytkowników języka mandaryńskiego), z którymi przyjdzie im się zmierzyć podczas uczestnictwa w jednym z eksperymentów, i które mogą okazać się kluczowe dla wyników badania.

Podstawę teoretyczną tematu przewodniego niniejszej pracy stanowi dokładne omówienie różnych aspektów tonalności języka mandaryńskiego, o czym traktuje rozdział 1. Rozdział 2 oferuje nieco bardziej szczegółowe ujęcie zjawiska tonalnego *sandhi*, które jest niezwykle istotnym, a zarazem ciekawym elementem tego zagadnienia. Choć – zgodnie z definicją – wystąpienia *sandhi* pierwotnie nie zakładano, możliwości takiej nie można było wykluczyć.

Podjmując problematykę tonu w kontekście porównania języka tonalnego (mandaryńskiego) z nietonalnym (polskim), należy zastanowić się, w jaki sposób ton, rozumiany jako określona modulacja częstotliwości podstawowej głosu, niosąca ze sobą konkretne znaczenie leksykalne, może być postrzegany przez rodzimego użytkownika języka, w którym informacja semantyczna wyrażana jest za pomocą innych środków. W języku polskim istotną rolę odgrywa intonacja, która – podobnie jak

ton – wykorzystuje zmiany wysokości głosu, jednak nie w celu rozróżniania znaczeń leksykalnych, lecz raczej w zakresie wyrażania emocji, ekspresji, modalności czy struktury informacji w wypowiedzi. Z intonacją wiąże się zjawisko akcentu, który wyróżnia sylaby pod względem ich rytmiczno-prozodycznego znaczenia w obrębie wyrazu lub frazy. Choć akcent, podobnie jak ton, bazuje na zmianach wysokości głosu i może być w pewnym stopniu postrzegany jako zjawisko akustycznie podobne, to jednak – jak podkreśla Majewicz (1989) – „ton w językach tonalnych nie jest akcentem” (s. 187), a ich funkcje w systemie językowym pozostają zasadniczo odmienne.

Dziękuję Pani Profesor Piotrze Łobacz za inspirację, nieocenione wsparcie mentorskie oraz konstruktywne sugestie, które znacząco przyczyniły się do powstania tej pracy.

Rozdział 1.

Tonalność w języku mandaryńskim

Choć tonalność jest zasadniczo terminem z dziedziny muzyki, używa się go także w kontekście lingwistycznym na określenie ogółu zjawisk towarzyszących językom tonalnym. Języki takie, do których należą wszelkie odmiany języka chińskiego² – w tym odmiana mandaryńska, o której niniejsza praca traktuje w sposób bardziej szczegółowy – według niektórych szacunków stanowią liczną grupę blisko 70% języków świata (Yip 2006). Mimo że różnią się one między sobą pod wieloma względami, istnieją pewne cechy, pozwalające na ujęcie ich wspólną definicją. Zdaniem Majewicza, autora obszernej klasyfikacji języków świata:

Języki tonalne charakteryzują się stałym, leksykalnie istotnym relatywnym tonem na każdej (lub prawie każdej) sylabie, będącym w opozycji do innego, bądź innych relatywnych tonów, które mogą wystąpić na takiej samej sylabie (1989: 190).

Badacz podkreśla pewien bardzo istotny fakt – mianowicie relatywną naturę tonu, która wymusza rozpatrywanie go w relacji do innych tonów. Kwestię tę rozwija Pike w pracy zatytułowanej *Phonetics: A Critical Analysis of Phonetic Theory and a Technic for the Practical Description of Sounds* (1966). Pisze on, że ucho rodzimego użytkownika języka tonalnego rejestruje na bieżąco dane dotyczące akustycznych parametrów sąsiadujących ze sobą dźwięków, które są poddawane analizie, a następnie oceniane jako „niskie” lub „wysokie” w stosunku do siebie, aniżeli jakiegokolwiek „bezwzględnej” wysokości; nawet w przypadku zmiany rejestru, powodującej zmianę wysokości poszczególnych dźwięków, tony pozostają niezmiennie, gdyż: „a single, level, isolated tone is not subject to classification – it must be put next to others to see if it is relatively low or high, and so on”(s. 28). MacKay (1987) umieszcza języki tonalne w opozycji do języków intonacyjnych, które, pomimo

² Według Majewicza: „Chiński może być rozpatrywany jednocześnie jako monolit etnolingwistyczny [...], jak i jako grupa oddzielnych języków »splicionych w nie kończąca się sieć dialektów«” (Voegelin-Voegelin 1977: 114)” (w: 1989: 66).

pewnego podobieństwa w nazewnictwie, różnią się od tych pierwszych diametralnie. Jak wyjaśnia, w językach takich jak np. angielski, pewne aspekty znaczeniowe są wyrażane za pomocą zmian intonacyjnych w obrębie zdania lub frazy, podczas gdy w wielu językach azjatyckich – w tym w mandaryńskim – cechą dystynktywną jest częstotliwość podstawowa (F_0), będąca nieodłącznym komponentem segmentów mniejszych – wyrazów; można zatem powiedzieć, że ton jest atrybutem natury fonemicznej³ (s. 170–171). W praktyce oznacza to, że w wyniku zmiany choćby jednego tonu danej sylaby sens wypowiedzi, której ta sylaba jest częścią, również się zmienia.

1.1. Opisy i transkrypcje systemu tonalnego

Wyraz 聲 *shēng* w znaczeniu „ton” funkcjonuje w języku chińskim już od V wieku n.e., co sugeruje świadomość różnic tonalnych wśród chińskich uczonych na wczesnym etapie rozwoju języka średniochińskiego, który przypada na okres od ok. VI do X wieku. Mniej więcej w tym samym czasie, w którym termin ten zaczęto stosować, poczyniono pierwsze udokumentowane próby opisu systemu tonalnego, co przypisuje się parze poetów-historyków Shen Yue i Zhou Yong. Jak głosi oficjalna historia Dynastii Południowych, wyróżnili oni cztery tony, znane do dziś pod nazwą tonów klasycznych: 平 *píng*, 上 *shàng*, 去 *qù* i 入 *rù*, którym nadali nieco poetyckie nazwy, zazwyczaj tłumaczone kolejno jako „równy” (*even*), „rosnący” (*rising*), „odchodzący” (*departing or going*) oraz „wchodzący” (*entering*, Sagart 1998). Ponadto, Shen Yue uważa się za twórcę zaginionej dziś pracy leksykograficznej 四聲譜 *Sì Shēng Pǔ*, w której używane ówczesznie znaki zostały sklasyfikowane pod względem tonów, co doskonale sprawdziło się jako – stosowana z powodzeniem także obecnie – technika układania słowników języka chińskiego. Pomimo zgodności w liczbie tonom klasycznym nie można przypisać konkretnych bezpośrednich odpowiedników wśród tych używanych współcześnie w wielu dialektach, m.in. w odmia-

³ Jak twierdzi MacKay (1987), fonemiczna natura tonu wynika z faktu, że jest on jedynym elementem różnicującym wyrazy (sylaby) pod względem semantycznym; mówiąc inaczej, wszelkie zmiany tonów widoczne są na poziomie wyrazów.

nie standardowej. W miarę ewolucji języka chińskiego ton „równy” uległ podziałowi na współczesny ton 1 i 2, w zależności od tego, czy w nagłosie sylaby znajdowała się spółgłoska bezdźwięczna, czy też dźwięczna. Ton „rosnący” w wyrazach zaczynających się od głosek bezdźwięcznych, płynnych lub nosowych odpowiada w przybliżeniu dzisiejszemu tonowi 3, natomiast zarówno wyrazy z dawnym tonem „odchodzącym”, jak i te opatrzone tonem „rosnącym” i zawierające w nagłosie dźwięczne spółgłoski zwarte lub szczelinowe otrzymałyby mandaryński ton 4. Pozostaje jeszcze grupa wyrazów z tonem „wchodzącym”, zakończonych w języku średniochińskim na [-p], [-t] lub [-k], które w pierwotnej formie zachowały się jedynie w języku kantońskim⁴, podczas gdy w odmianie standardowej i wielu dialektach używanych obecnie na północy Chin utraciły końcową spółgłoskę i otrzymały jeden z czterech współczesnych tonów (Chao 1970: 25, Pulleyblank 2003: 6).

W dzisiejszej odmianie standardowej każda akcentowana sylaba ma ton osadzony na jej dźwięcznej części, tak więc jeżeli w nagłosie znajduje się spółgłoska dźwięczna, ton rozciąga się na całą sylabę, z kolei w przypadku nagłosu bezdźwięcznego (lub jego braku) – obejmuje on swym zasięgiem ośrodek i ewentualny wygłos (Chao 1970: 19–20). Jak wspomniano, współczesną chińszczyznę charakteryzuje zaplecze tonalne w postaci czterech tonów o różnych wysokościach (*pitch*), najczęściej oznaczanych cyframi arabskimi, kolejno 1, 2, 3, 4. Ton 1 – wysoki równy (*high-level*) – charakteryzuje względnie równy przebieg w górnym rejestrze skali głosu mówcy; ton 2 – wysoki rosnący (*high-rising*) – jest tonem wznoszącym, osiagającym pod koniec rejestr wysoki; ton 3 – niski opadająco-wznoszący (*low-dipping*) – cechuje zmienna wysokość, która w momencie spadku osiąga rejestr dolny, po czym następuje zmiana kierunku na wznoszący, oraz ton 4 – wysoki opadający (*high-falling*) o szerokim zakresie, odznaczający się spadkiem wysokości od rejestru górnego do możliwie najniższego. Należy także zwrócić uwagę na obecność piątego „dodatkowego” tonu, który, choć jest elementem powszechnie występującym w języku mandaryńskim, często bywa przez badaczy pomijany, co zawdzięcza swej odmiennej naturze. Ton ten, zwany zerowym lub neutralnym (z chin. *qīng shēng* – dosł. „ton lekki”), wyróżnia m.in. Cheng (1973), umieszczając

⁴ Ton występujący na tego typu wyrazach zwany jest tonem zatrzymanym.

go w osobnej niż tony „podstawowe” kategorii. Ponieważ jednak dla niniejszej pracy ma on niewielkie znaczenie, zostanie jedynie pokrótce omówiony w jej dalszej części.

Oprócz przedstawionych nazw zwyczajowych w formie krótkich opisów każdy ton ma także poniekąd oficjalną nazwę chińską, zaczerpniętą od ich klasycznych pierwowzorów. Są to:

- 1) Ton 1 – 陰平 *Yīnpíng*
- 2) Ton 2 – 陽平 *Yángpíng*
- 3) Ton 3 – 上声 *Shàngshēng*
- 4) Ton 4 – 去声 *Qùshēng*

(Chao 1970: 26).

Aby móc dokładnie opisać i zanalizować jakikolwiek system tonalny, konieczne jest przyjęcie adekwatnej techniki zapisu symbolicznego. Przez lata specjaliści zajmujący się badaniem języków tonalnych wypracowali szereg metod transkrypcji, mniej lub bardziej czytelnych, odpowiadających potrzebom poszczególnych grup języków. Dotychczas nie znaleziono jednak optymalnej techniki, którą można by zastosować do każdego bez wyjątku systemu tonalnego, co nie dziwi z uwagi na wielość i ogromną różnorodność tego typu języków. O popularności danej metody transkrypcji zdaje się decydować kryterium geograficzne, co można wytłumaczyć podobieństwem wynikającym z bliskości terytorialnej pewnych języków. Jak zauważa Yip (2006), w kręgu ekspertów z dziedziny języków azjatyckich przeważa tendencja do stosowania cyfr, choć w użyciu są także notacje literowe oraz graficzne. Do opisu samego języka mandaryńskiego, w literaturze funkcjonuje aż siedem różnych metod transkrypcji, które zaprezentowano poniżej. Za ilustrację niech posłuży chętnie wykorzystywany przez lingwistów przykład, częściowo zaczerpnięty z *The Phonology of Standard Chinese*:

T1	T2	T3	T4
1) <i>ma</i> 1	<i>ma</i> 2	<i>ma</i> 3	<i>ma</i> 4
2) <i>ma</i> □	<i>ma</i> /	<i>ma</i> √	<i>ma</i> \
3) (ma)55	(ma)35	(ma)214	(ma)51
4) <i>mā</i>	<i>má</i>	<i>mǎ</i>	<i>mà</i>

5) <i>mha</i>	<i>mar</i>	<i>maa</i>	<i>mah</i>
6) (<i>ma</i>) HH	(<i>ma</i>) LH	(<i>ma</i>) LL	(<i>ma</i>) HL
7) (<i>ma</i>) HH	(<i>ma</i>) MH	(<i>ma</i>) MLH	(<i>ma</i>) HL
„matka”	„konopie”	„koń”	„łajać”

(Duanmu 2000: 210).

Przykład pierwszy przedstawia technikę szeroko stosowaną ze względów praktycznych. Użycie cyfr w zestawieniu z konkretnymi sylabami nie pozostawia wątpliwości co do tonów, na których powinny one być wymawiane; brak natomiast wskazówek co do wysokości i kierunku modulacji głosu, dlatego też taki sposób zapisu można spotkać tam, gdzie informacje tego typu nie są szczególnie istotne, np. w słownikach i niektórych podręcznikach do nauki języka chińskiego. Notacja zaprezentowana w przykładzie drugim, znana pod angielską nazwą *Chao letters*, została wprowadzona w roku 1930 przez ojca chińskiego językoznawstwa Yuen Ren Chao. System ten, w dużej mierze intuicyjny, ma na celu zobrazowanie przebiegu konturu tonalnego za pomocą specjalnych znaków graficznych. Punktem odniesienia jest pionowa linia symbolizująca skalę głosu mówcy. Jak zauważa Duanmu (2000), transkrypcja ta jest jednak dość mało precyzyjna oraz dopuszcza możliwość różnej interpretacji. Podobnie rzecz ma się z kolejną metodą zapisu, która jest notabene liczbowym odpowiednikiem grafów Chao. Tu poziomy toniczne wyrażane są za pomocą cyfr od 1 do 5, odpowiadających wysokości na skali głosu, przy czym 1 oznacza rejestr możliwie najniższy, natomiast 5 – najwyższy. Opis tonu przy użyciu tej notacji składa się z dwóch lub – w przypadku tonu opadająco-wznoszącego – trzech cyfr, reprezentujących jego wysokość początkową i końcową, oraz ewentualne pośrednie fluktuacje. Dodatkowy problem pojawia się przy próbie zaadaptowania tego systemu do wyrażenia tonów w zgodzie z koncepcją Yip, wedle której pojedynczy ton jest jednostką złożoną z szeregu mniejszych segmentów o równej, określonej wysokości (zagadnienie to zostanie omówione nieco dokładniej w dalszej części pracy).

Kolejny zapis, będący w zasadzie uproszczoną wersją graficznej reprezentacji Chao (2), jest powszechnie używany w oficjalnej transkrypcji języka mandaryńskiego *Hanyu pinyin*. Znak diakrytyczny umieszczony nad nacechowanym tonem elementem danej sylaby obrazuje przebieg tonu wraz z kierunkiem i orientacyjną wysokością głosu. Notacja zastosowana w przykładzie nr 5, rzadko spotykana w literaturze, jest produktem działań nastawionych na romanizację języka chińskiego. Działania te mają na celu wyeliminowanie wszelkich znaków diakrytycznych z transkrypcji na rzecz liter jedynie alfabetu łacińskiego. Kodowanie tonów odbywa się tu poprzez zastosowanie różnej pisowni, co przekłada się także na pewne drobne różnice w wymowie. System ten, choć zdecydowanie mniej intuicyjny i niewątpliwie mało klarowny dla osób nieznających zasad kodowania, ma jednak pewne zalety. Stworzony głównie z myślą o uczących się języka chińskiego, ma za zadanie uprościć pisownię i wymowę, czyniąc go bardziej dostępnym dla odbiorcy z łacińskiego kręgu kulturowego. Ponadto wymusza niejako zapamiętywanie tonów poprzez integrację informacji semantycznej z fonetyczną, podkreślając w ten sposób nierozzerwalność obu tych elementów (Duanmu 2000: 210–211).

Dwie ostatnie transkrypcje do reprezentacji tonów wykorzystują litery alfabetu łacińskiego H, L i M, będące pierwszymi literami angielskich określeń poziomu wysokości dźwięku i oznaczającymi odpowiednio: *high* (wysoki), *low* (niski) oraz *middle* (średni). Pierwsza z nich, zaproponowana przez Woo (1969, za: McCawley 1978) i postulowana także przez Yip (1980, za: Duanmu 2000) zakłada zgoła inną koncepcję tonu, zgodnie z którą jest on realizowany jako sekwencja mniejszych segmentów, osadzonych w górnym (H) lub dolnym (L) rejestrze, reprezentowanym przez pierwszą literę, na jego wyższym (H) lub niższym (L) poziomie, na co wskazuje litera druga. Tymczasem notacja ostatnia jest inną, trzejelementową, a zatem nieco uboższą formą zapisu cyfrowego, w której litery oznaczają wysokość na skali głosu mówcy. Wprowadzony zostaje dodatkowy, średni poziom (M), w przybliżeniu odpowiadający poziomowi 3 w wersji „cyfrowej”, a tony opisywane są na zasadzie zestawiania oznaczeń, obrazując przebieg tonalny poprzez sygnalizowanie wzrostów i spadków wysokości.

1.2. Właściwości tonu i jego fonetyczne korelaty

Według Chao (1970: 19) ton jest przebiegiem zmian wysokości głosu zrealizowanym na dźwięcznej części sylaby. Jednak aby wyjaśnić, co jest przyczyną tych zmian, percypowanych przez ucho ludzkie jako ton rosnący bądź opadający, należy przywrócić się procesowi produkcji tonu oraz zdefiniować pewne parametry, w kategoriach których proces ten można opisać.

1.2.1. Wysokość (*pitch*) i częstotliwość podstawowa (F_0)

Jak wyjaśnia Francis Katamba (1996: 53), wysokość dźwięku jest wynikiem szybkości drgania fałdów głosowych, zjawiska cyklicznego będącego efektem współdziałania sił aerostatycznych i aerodynamicznych z przepływem powietrza regulowanym skurczami mięśni krtaniowych. Mięśnie te odpowiadają za zmiany napięcia fałdów głosowych, co z kolei wpływa na częstotliwość ich drgań. Tempo drgań, określane jako częstotliwość podstawowa (F_0), mierzone w hercach (Hz), rośnie wraz ze wzrostem napięcia i ciśnienia subglotalnego, a także zależy od właściwości anatomicznych: długości, masy i ogólnej wielkości fałdów głosowych. Dłuższe i cięższe fałdy wibrują wolniej, co skutkuje niższą F_0 , natomiast krótsze i lżejsze – szybciej, co prowadzi do wyższego tonu; stąd m.in. wyraźne różnice wysokości głosu między kobietami a mężczyznami. W językach takich jak mandaryński zmiany wysokości dźwięku są wykorzystywane do zmiany znaczenia danego segmentu, co określa się mianem tonu.

Jak podają Jongman, Wang, Moore i Sereno (2006), autorzy studium nad percepcją i produkcją tonów języka mandaryńskiego, przeciętny zakres F_0 rodzimych użytkowników wynosi: w przypadku kobiet 165–352 Hz, natomiast mężczyzn – 68–223 Hz⁵. Należy jednak zwrócić uwagę, że są to jedynie wartości przykładowe, uzyskane przez pojedynczych mówców w warunkach laboratoryjnych, mające głównie zarysować skalę różnic pomiędzy płciami, a zatem mogą, lecz nie muszą one odzwierciedlać rzeczywistych tendencji wśród populacji.

⁵ Wartości uzyskane przez pojedynczych, losowo wybranych rodzimych użytkowników odmiany standardowej obu płci.

Pewnego rodzaju szczególnym przypadkiem jest wspomniany wcześniej ton zerowy, lub inaczej neutralny, mający niezwykle interesujące fonetyczne implikacje. Otóż jak się okazuje, wbrew funkcjonującemu potocznie określeniu sylab z tonem zerowym jako sylab „bez tonu”, ton ten wykazuje w rzeczywistości zmienny przebieg tonalny, uzależniony od tonu bezpośrednio go poprzedzającego. Pierwszy wniosek w tej kwestii wysnuł Chao (1933, za: Cheng 1973: 55), stwierdzając, że ton ten brzmi jak 4, jeżeli następuje po tonie 1, 2 lub 4, natomiast gdy poprzedza go ton 3 – przyjmuje formę bardziej przypominającą ton 1. 15 lat później badacz dochodzi do bardziej precyzyjnej konkluzji, ściśle określając wysokość tonu zerowego po każdym z czterech tonów leksykalnych. Ustalenia są następujące:

- 1) po tonie 1 – średnio-niski,
 - 2) po tonie 2 – średni,
 - 3) po tonie 3 – średnio-wysoki,
 - 4) po tonie 4 – niski
- (Cheng 1973: 55-56).

Jak dodaje, zdarzają się także przypadki rozpoczynania frazy tonem zerowym; wówczas przyjmuje on z reguły wysokość średnią (Shen 1990: 32).

Dreher i Lee (za: Cheng 1973: 56) analizują kontury tonu zerowego, które następnie opisują przy użyciu cyfrowej notacji Chao – odpowiednio jw.:

- 1) 41,
- 2) 31,
- 3) 23,
- 4) 21.

Na podstawie powyższych danych można stwierdzić, że ton zerowy istotnie ma kontur opadający, jeśli występuje po tonie 1, 2 i 4, oraz kontur rosnący w przypadku, gdy poprzedza go ton 3. Co ciekawe, ton zerowy przyjmuje zupełnie przeciwne wartości w sytuacji, gdy stanowi drugi z sekwencji trzech tonów; wówczas odpowiada on wysokości końcowej części konturu tonu bezpośrednio go poprzedzającego (Shuh 1978: 238).

1.2.2. Napięcie fałdów głosowych

Stopień napięcia fałdów głosowych jest wymiarem niezwykle istotnym, gdyż – jak twierdzi para badaczy Halle i Stevens (1971, za: Duanmu 2000: 213) – stanowi on mechanizm bezpośrednio kontrolujący ton. Efektem zmian tego parametru są zmiany w częstotliwości podstawowej realizowanego tonu. Halle i Stevens wyróżniają trzy możliwości: sztywne struny głosowe, luźne struny głosowe oraz stadium pośrednie, które następnie odpowiadają za produkcję dźwięków na trzech poziomach tonicznych: wysokim, średnim i niskim. Jak wyjaśnia Anderson (1982), zeszywnieniu fałdów głosowych towarzyszy podwyższenie tonu na elementach nacechowanych tonem – samogłoskach, a także – w zależności od stopnia rozwarcia głośni – spółgłoskach bezdźwięcznych lub silnie aspirowanych. Natomiast „rozluźnienie strun głosowych zwiększa ich zdolność utrzymania drgań w warunkach, które skądinąd niezbyt sprzyjają dźwięczności” (s. 368–369).

1.2.3. „Szmer”

Duanmu (2000) wskazuje na obecność w niektórych językach kolejnego bezpośrednio skorelowanego z częstotliwością podstawową wymiaru, który nazywa „szmerem” (*murmur*). Tony, w których obecny jest „szmer”, zazwyczaj charakteryzuje niższa częstotliwość konturu, podczas gdy te „bezsmerowe” są na ogół osadzone w wyższym rejestrze (por. § 1.4.1). W standardowej odmianie języka chińskiego jedynym tonem, któremu towarzyszy „szmer”, jest ton 3 – niski, tymczasem każdy z pozostałych tonów jest przynajmniej częściowo realizowany w rejestrze wysokim.

1.2.4. Temporalne właściwości tonów

Jak zauważają Jongman et al. (2006), tony języka mandaryńskiego charakteryzują się różnymi właściwościami czasowymi, co zdaje się potwierdzać znaczna liczba danych empirycznych. Badania wykazują, że najdłuższe są na ogół tony 2 i 3, a najkrótszy przebieg czasowy

(spośród podstawowych tonów leksykalnych) ma z kolei ton 4. Woo (1969, za: Anderson 1982: 317) podaje, że sylaby z tonem 3 mogą być nawet do 50% dłuższe niż sylaby z pozostałymi tonami. Należy jednak podkreślić, że dane te dotyczą wyłącznie tonów w pozycji izolowanej, bowiem – jak się okazuje – względna długość każdego z tonów może ulec zmianie w zależności od pozycji jaką zajmuje on w zdaniu. I tak np. różnica długości pomiędzy tonami 1, 2 i 4 oraz tonem 3 w pozycji innej niż końcowa bądź izolowana spada zaledwie do kilku procent (Duanmu 2000: 232).

Ponownie poza klasyfikacją jest reprezentujący osobną kategorię ton neutralny. Jak podaje Cheng (1973), sylaby z tym tonem są niemal o połowę krótsze od sylab mających pełne tony leksykalne, bez względu na to, czy są one monofongami czy też nie. Co więcej, w złożeniach dwusylabowych, z których druga ma ton zerowy, sylaba pierwsza, pomimo znajdowania się w pozycji akcentowanej, także ulega redukcji (s. 55).

1.3. Czynniki wpływające na realizację tonów

Ogólnie rzecz biorąc, czynniki te można sklasyfikować jako świadome bądź niezamierzone. W grupie pierwszej znajdują się przede wszystkim działania powodowane chęcią zakomunikowania przez nadawcę pewnych treści poprzez: emocje, intonację, informacje pragmatyczne, a także świadomą manipulację strukturą składniową. Spośród czynników niezamierzonych wymienić można pewne indywidualne predyspozycje oraz różnice związane z budową toru głosowego i funkcjonowaniem narządów biorących udział w produkcji poszczególnych dźwięków (Zuo 2002).

Osobną kwestię stanowią zmiany fonetyczne i fonologiczne, obejmujące przyległe tony występujące w odpowiednich konfiguracjach. Ponieważ zjawisko to, znane pod nazwą *sandhi* tonalnego, które stało się przedmiotem wielu sporów natury lingwistycznej na rozmaitych płaszczyznach ma dla niniejszej pracy istotne znaczenie, poświęcono mu w całości kolejny rozdział.

1.4. Fonologiczna reprezentacja tonu

Jak sugeruje Duanmu (2000), ton jako jednostka natury fonologicznej powinien być rozpatrywany pod czterema względami: liczby jego elementów składowych, liczby poziomów kontrastywnych w ramach każdego elementu, formy przebiegu tonalnego, a także postaci jednostki nacechowanej tonem (TBU) (*tone-bearing unit*).

1.4.1. Składniki tonu i ich poziomy kontrastywne

Duanmu wyróżnia dwa składniki, które nazywa wysokością (*pitch*) (nie należy mylić z wysokością rozumianą jako wartość F_0) i rejestrem (*register*). Pierwszy z nich odpowiada za zmiany częstotliwości głosu (por. § 1.2.2), podczas gdy drugi wpływa na ewentualną obecność „szmeru” w realizacji danego tonu. Ponieważ zarówno wysokość, jak i rejestr są wymiarami dwubiegunowymi, ich interakcja daje w rezultacie cztery możliwe kombinacje:

<u>Rejestr</u>	<u>Wysokość</u>	
(stopień napięcia	(grubość	
fałdów głosowych)	fałdów głosowych)	
szytwe	cienkie	– H bez „szmeru”
szytwe	grube	– L bez „szmeru”
luźne	cienkie	– H z „szmerem”
luźne	grube	– L z „szmerem”

(Duanmu 2000: 213–216).

1.4.2. Tony konturowe (*contour tones*) i tony poziome (*level tones*)

Aby opisać system tonalny języka mandaryńskiego z jego specyfiką, konieczne jest przedstawienie typologii, na której opiera się klasyfikacja języków tonalnych. W zależności od kształtu przebiegu tonalnego wyróżnia się dwa typy: języki rejestrowe (*register tone languages*) oraz języki konturowe (*contour tone languages*). W „czystym” systemie rejestrowym tony mają niemal równą, osadzoną w konkretnym

rejestrze wysokość, która nie ulega zauważalnym zmianom w czasie ich trwania. Najczęściej wśród języków tego typu spotyka się dwa lub trzy, natomiast prawdopodobnie nie więcej niż cztery dystynktywne rejestry, gdyż – jak pisze Majewicz: „choć w literaturze cytuje się języki o większej liczbie rejestrów, wydaje się mało prawdopodobne, aby takie subtelne różnice były w ogóle percepowalne” (1989: 191). Zuo (2002) zwraca uwagę, że rejestry te nie mają stałej, ściśle określonej wysokości; jest ona raczej relatywna, rozpatrywana względem wysokości innych poziomów, które są na tej podstawie określane jako wysokie, średnie lub niskie, w zależności od ich liczby.

W odróżnieniu od języków rejestrowych tony w systemach konturowych charakteryzuje zmienna wysokość, która podlega różnego rodzaju fluktuacjom. Elementem dystynktywnym jest tu kierunek, w jakim zmiany te zachodzą, natomiast nie sama wartość wysokości czy choćby jej względny poziom. Jak zauważa Francis Katamba (1996: 188), systemy te w „czystszej” postaci należą jednak do rzadkości, bowiem w znacznej części języków tonalnych występuje tendencja do łączenia cech obu typów. Jednym z takich języków jest właśnie *pǔtōnghuà* (dosł. *wspólny język*), jak brzmi oficjalna nazwa standardowej odmiany chińskiego. W języku tym jedynym tonem rejestrowym jest ton 1; reszta to tony konturowe, realizowane w obrębie trzech rejestrów.

O ile koncepcja konturu tonalnego jest kwestią niebudzącą sprzeciwu lingwistów, przynajmniej w sferze nazewnictwa, o tyle jego integralność została przez wielu poddana w wątpliwość. Podział, jaki można zaobserwować wśród badaczy zajmujących się tym zagadnieniem, jest dychotomiczny: jedni opowiadają się za niepodzielnością tonu jako jednostki, natomiast według innych jest on sekwencją kilku tonów równych. Spór ten, dotychczas nierozwiązany, pociąga za sobą także pytanie o jednostkę nacechowaną tonem (por. § 1.4.3), co w dalszej kolejności prowadzi do kwestionowania tonu jako zjawiska segmentalnego bądź suprasegmentalnego (por. § 1.5).

Jednym ze znanych badaczy, którzy postulowali traktowanie konturu tonalnego jako całości, jest Pike (1948, za: Anderson 1978). Pomimo że w ramach szerokiej klasyfikacji, obejmującej także języki tonalne z obszaru Afryki i Ameryki, dopuszcza on możliwość istnienia konieczności rozgraniczenia pomiędzy językami, w których kontury tonalne są wynikiem występowania więcej niż jednego poziomu tonalnego

w ramach jednego tonu a tymi, w których elementem dystynktywnym jest ruch jako taki o pewnym określonym przebiegu, w kwestii języków z rodziny sino-tybetańskiej możliwość ta zostaje przez badacza wykluczona. Pike uważa, że języki azjatyckie w przeważającej liczbie znacznie różnią się pod tym względem od pozostałych, dlatego też wysuwa tezę o niepodzielności ich konturów tonalnych, którą przedstawia w postaci czterech argumentów traktujących o specyfice systemów konturowych. Brzmiały one następująco:

- 1) Podstawową jednostkę tonemiczną charakteryzuje płynne przejście pomiędzy punktami o skrajnych wysokościach, natomiast nie poziom równy.
 - 2) Jednostkowych glissand⁶ nie mogą dzielić żadne granice, w odróżnieniu od niefonemicznych jednostek złożonych, charakterystycznych dla systemów rejestrowych.
 - 3) Skrajne punkty glissand nie mogą pokrywać się z tonami równymi w ramach tego samego systemu konturowego.
 - 4) W systemach konturowych na jedną sylabę przypada jeden ton, podczas gdy w niektórych systemach rejestrowych, jak np. Mazateco, możliwe są dwa lub więcej tonów
- (Pike 1948: 8, za: Anderson 1978: 151–153).

W odniesieniu do punktu 1) należy zaznaczyć, że Pike miał na myśli jednostki najbardziej powszechne w danych systemach, czyli w pewnym sensie dla nich charakterystyczne, gdyż – jak sam stwierdza – nieznane są języki, w których nie występuje żaden ton równy. Z kolei w argumencie drugim zwraca uwagę na istnienie pewnych języków rejestrowych, w których początek i koniec jednego tonu są realizowane w obrębie różnych morfemów, co nie może mieć miejsca w językach typu konturowego (Anderson 1978: 151–152).

Anderson przytacza także spostrzeżenia dwóch innych badaczy, którzy opowiadają się za podobnym ujęciem problemu konturów tonalnych – mianowicie Grubera (1964) oraz Wang (1967). Obaj skupili się na językach azjatyckich, w szczególności tych należących do rodziny sino-tybetańskiej, przyjmując wstępne założenie, że domeną tonu jest sylaba. Zaproponowane przez nich analizy konturów tonalnych opie-

⁶ Z uwagi na brak polskiego ekwiwalentu proponowany jest termin z dziedziny muzyki, używany na określenie analogicznego zjawiska.

rają się na modelu, zgodnie z którym każdy ton należy rozpatrywać pod kątem dwóch wymiarów: rejestru, w jakim jest on realizowany (u Grubera „wysoki” lub „wysoki2”; u Wanga – „wysoki”, „centralny” lub „średni”), oraz kierunku, w którym odbywa się ruch, czyli „rosnący” i/lub „opadający”. Wang wyszczególnia ponadto dwie dodatkowe cechy: „kontur”, właściwą dla tonów rosnących i opadających, oraz „wypukły”, charakteryzującą tony złożone, łączące oba kierunki. Badacze ci różnią się także podejściem do kwestii reguł przypisywania konturom tonalnym ich rejestru. Koncepcja Grubera zakłada istnienie czterech dystynktywnych poziomów, spośród których danemu tonowi – bez względu na to, czy jest on rosnący, opadający czy też złożony – przypisuje się dokładnie jeden. I tak przyjmuje się, że każdy ton, którego kontur nie wzrasta ponad drugi poziom od dołu, przynależy do rejestru najniższego oraz – analogicznie – każdy kontur, który nie opada poniżej poziomu drugiego od góry, jest powiązany z rejestrem najwyższym. Natomiast wszystkie inne tony, których kontury przekraczają te granice, przynależą do jednego z pozostałych dwóch rejestrów, w zależności od tego, w którym z nich znajduje się dominująca część ich przebiegu. U Wanga, pomimo podziału skali głosu na pięć poziomów, tony mogą być zlokalizowane jedynie w rejestrach skrajnych: najwyższym lub najniższym, gdyż, jak twierdzi, w jednym języku możliwe są jedynie dwa kontury o tym samym kształcie – jeden w niższej, drugi w wyższej części skali. Zasada głosi, że kontur, którego przebieg w jakimkolwiek punkcie osiągnął maksymalną wysokość, należy do rejestru wysokiego; w przeciwnym razie musiał osiągnąć wartość minimalną, w związku z czym znajduje się w rejestrze niskim. Ponieważ zarówno w koncepcji Grubera, jak i Wanga o kwestii przynależności do danego rejestru nie decydują żadne szczególne punkty w czasowym przebiegu tonu, można wywnioskować, że wszystkie one mają jednakowy status, co wspiera tezę o jednostkowości konturów tonalnych (Anderson 1978: 147–148).

W opozycji do opisanego wyżej stanowiska znajduje się m.in. koncepcja Woo (1969, za: Anderson 1978), która jako pierwsza zakwestionowała panujący w owym czasie pogląd o integralności tonu, sugerując, że wszystkie tony konturowe są w zasadzie kombinacjami tonów równych. Teoria ta wyrosła głównie na bazie krytyki koncepcji Wanga (za: Anderson 1978), która, według badaczki, nie tylko nie dostarczyła przekonujących dowodów na poparcie własnej tezy, ale

także ignorowała przypadki jej zdaniem świadczące na rzecz stanowiska przeciwnego. Swoją koncepcję Woo oparła na założeniu, że „im bardziej złożony jest (kontrastujący) wzór toniczny, tym obszerniejszy jest segmentalny materiał objęty jego realizacją” (Anderson 1982: 316–317), co następnie usiłowała poprzeć materiałem dowodowym, uzyskanym z badań nad długością trwania sylab o różnych tonach. Jeden z takich pomiarów wykazał, że w języku mandaryńskim realizowane w izolacji sylaby z tonem 3, którego przebieg można opisać przy użyciu notacji Chao jako [315] lub też [213], są z reguły blisko 50% dłuższe od sylab opatrzonych pozostałymi tonami. Na tej podstawie wysunęła hipotezę, że jednostkami TBU nie są sylaby, jak twierdzili Gruber i Wang, lecz jednostki zwane morami, oraz że zestaw możliwych do przypisania im tonów składa się wyłącznie z tonów równych o określonych rejestrach. Zgodnie z tym każda sylaba z tonem złożonym musiałaby zawierać przynajmniej dwie mory, jak ma to miejsce w przypadku mandaryńskich tonów 1, 2 i 4, z których każdy obejmuje dwa dystynktywne poziomy toniczne. Na tej samej zasadzie sylaby z tonem 3, o trzech poziomach tonicznym powinny zawierać odpowiednio trzy mory (Anderson 1978: 148–149).

Należy w tym miejscu zaznaczyć, że swoim nowatorskim ujęciem zagadnienia konturów tonalnych Woo zaproponowała alternatywę dla obowiązującego w tamtym czasie sposobu postrzegania tonu jako zjawiska segmentalnego bądź suprasegmentalnego, wpisując się tym samym w zapoczątkowany niedługo potem nurt fonologii autosegmentalnej (*autosegmental phonology*).

Ustosunkowując się do koncepcji Woo, Anderson (1978) przyjmuje nieco ostrożne stanowisko, zgadzając się z badaczką jedynie po części. W świetle odkryć płynących z prac innych współczesnych mu lingwistów ostatecznie stwierdza, że przynajmniej w niektórych przypadkach – jak ma to miejsce m.in. w standardowej odmianie języka chińskiego – kontury tonalne powinny być traktowane jako sekwencje elementów o równej wysokości. Natomiast nie zgadza się z poprzedniczką co do przypisywania statusu jednostek TBU pojedynczym morom, gdyż według niego jest to teza nie znajdująca pokrycia w niektórych językach. W zamian oferuje wyjaśnienie, że w pewnych przypadkach możliwe jest przyporządkowanie jednej samogłosce krótkiej (lub innej domenie) więcej niż jednej jednostki TBU, co ilustruje na przykładzie

cechy nosowości czy też zjawiska dyftongizacji w cytowanej pracy Andersona (1972, za: Anderson 1978: 150–151).

W ciekawy sposób zagadnienie jedności konturów tonalnych ujmuje Alexis Michaud (2012), który słuszność swojej tezy argumentuje za pomocą wyników najnowszych badań nad pomyłkami językowymi. Otóż – wyjaśnia – gdyby kontur tonalny składał się z sekwencji tonów równych, wówczas zdarzałyby się błędy spowodowane pominięciem jednego z poziomów danej sekwencji, co często ma miejsce w systemach rejestrowych, a – jak dowodzą wnikliwe analizy – nie występuje w języku mandaryńskim. Badacz zwraca jednak uwagę, że wyników tych nie należy traktować jako podstawy do jakichkolwiek generalizacji, bowiem istnieją języki, w których reprezentacja ta jest jak najbardziej prawidłowa, podczas gdy z powodu braku dowodów świadczących przeciw, została ona niegdyś błędnie zaadaptowana jako alternatywna dla wielu odmian języka chińskiego.

1.4.3. Jednostka nacechowana tonem (TBU, *Tone Bearing Unit*)

Choć rozważania nad charakterem jednostki TBU pojawiły się już przy okazji omawiania struktury konturów tonalnych, był to jedynie niewielki wycinek szerszej dyskusji na ten temat, która toczy się w środowisku językoznawczym od drugiej połowy lat 60., przyjmując rozmaite formy w koncepcjach różnych badaczy. San Duanmu (2000: 218) przedstawia pięć prób odpowiedzi na pytanie, czym tak naprawdę jest TBU. Zagadnienie to ilustruje w kontekście języka mandaryńskiego, za pomocą sylaby *mai* (*sprzedawać*) z tonem opadającym (HL):

<i>mai</i>	<i>mai</i>	<i>mai</i>	<i>mai</i>	HL
∨	∨	/	/	/
S	O R	m m	m m	<i>mai</i>
∧	∧			∖
HL	HL	H L	H L	mm
(1)	(2)	(3)	(4)	(5)

Przykład pierwszy ukazuje tradycyjne ujęcie jednostki TBU zgodnie z koncepcją Wanga (1967: 95, za: Duanmu 2000: 218–219) oraz Chao (1968: 19, 25, za: Duanmu 2000), który reprezentację tonu przedstawia jako krzywą wysokości, mającą przebieg w danym czasie od punktu początkowego do punktu końcowego dźwięcznej części sylaby, czyli – w powyższym przypadku – od dźwięcznego [m] o wysokości H do [j] o wysokości L, a zatem na jedną sylabę przypadają dwa tony. Ton jest w tym ujęciu cechą suprasegmentalną, właściwą jednostkom większym niż segmenty. W punkcie drugim zaprezentowano pogląd reprezentowany m.in. przez Kratochvila (za: Duanmu 2000: 218–219) oraz Howiego (1976: 218, za: Duanmu 2000: 218–219), którzy zakładają podział sylaby na dwie części: nagłos oraz rym (*rhyme*), złożony z jądra i kody. Zgodnie z tą interpretacją nagłos nie jest nośnikiem żadnego tonu, natomiast rym – aż dwóch. Wniosek ten wynika z badań z dziedziny fonetyki, przeprowadzonych niezależnie przez obu badaczy. Wykryli oni bowiem, że kontur danego tonu przybiera pożądaną kształt dopiero z początkiem rymu, tak więc w przypadku sylaby z tonem rozciągającym się pomiędzy dwoma skrajnymi rejestrami, jej rym będzie składał się z dwóch jednostek TBU.

W teoriach zilustrowanych trzema kolejnymi przykładami jednostki te utożsamiane są z morami (m). Przykłady (3) i (4), choć z pozoru bardzo zbliżone, różnią się zasięgiem pierwszej morę; u Hymana (1985, za: Duanmu 2000: 218–219), autora koncepcji z przykładu (3), obejmuje ona spółgłoskę w nagłosie, podczas gdy u Hayesa (1989, za: Duanmu 2000: 218–219) nie. Ostatni z przedstawionych modeli, proponowany przez Woo (1969, za: Duanmu 2000: 218–219) i poparty przez Duanmu (1990, za: Duanmu 2000: 218–219), w dużej mierze przypomina model (4), gdyż w istocie jest jego drobną modyfikacją. W obu z nich każdy element rymu zawiera jedną morę, z tą jednak różnicą, że w (5) tony łączą się nie z morami, lecz bezpośrednio z segmentami rymu. Jak pisze Duanmu w komentarzu, najbardziej adekwatne są modele (2) i (5), choć w świetle późniejszych badań stosowniejszy okazuje się (5). Zasadnicza różnica pomiędzy nimi polega na tym, że model (5), w odróżnieniu od modelu (2), określa zależność pomiędzy długością sylaby a liczbą jednostek TBU; i tak: sylaby krótkie zawierają jedną jednostkę TBU, natomiast sylaby długie mogą zawierać dwie jednostki. Jest to widoczne w dialektach języka chińskiego, w których większość

sylab – tzw. sylab pełnych (*full syllables*) – ma tony równe oraz proste tony konturowe. Ponieważ sylaby te są segmentami dwumorowymi, zawierają odpowiednio dwie jednostki TBU. Z kolei znacznie mniej powszechne sylaby o złożonych konturach tonalnych występują z reguły jedynie w pozycjach końcowych, dzięki czemu zachowują pełną formę, a co za tym idzie – są wydłużone. Jak potwierdzają badania, istnieje większe prawdopodobieństwo, że kontury złożone, które przez wzgląd na swoją nieco bardziej skomplikowaną budowę w istocie wymagają trzech mor, zachowają niezmienną formę w pozycji końcowej, gdyż w pozostałych niezmiernie często ulegają redukcji (2000: 218–219).

Jak konkluduje Duanmu, hipoteza, zgodnie z którą jednostką TBU jest sylaba bądź rym, nie znalazła dotychczas potwierdzenia w dowodach. Stwierdzono natomiast, że może nią być mora, oraz że mieści ona nie więcej niż jeden ton.

Zagadnienie natury jednostki TBU podejmują w swoich pracach także autosegmentaliści, i choć ujęcie to nie różni się w tej kwestii diametralnie od wcześniej proponowanych koncepcji, zmienia się perspektywa podejścia do pewnych zjawisk fonologicznych. Jednak zanim zaprezentowane zostaną istotne z punktu widzenia niniejszej pracy zagadnienia fonologii autosegmentalnej, należy przedstawić tradycyjny sposób ujmowania zjawiska tonu, którego krytyka stanowiła grunt dla alternatywnej teorii.

1.5. Relacja ton – segment

Tradycyjne podejście do zjawisk o charakterze fonologicznym, do których niewątpliwie należy ton, opiera się na postrzeganiu języka jako hierarchii poziomów. Zakłada ono, że pewne zjawiska zachodzą w obrębie danych segmentów⁷, podczas gdy inne – ponad poziomem segmentu. Tego typu procesy są określane mianem suprasegmentalnych (MacKay 1987: 154). W kontekście tonu odpowiedź na pytanie, do której z tych grup zjawisko to należy bynajmniej nie jest oczywista, gdyż – jak pokazuje San Duanmu (2000: 233–234) – istnieją argumenty

⁷ W językoznawstwie: „any discrete unit that can be identified, either physically or auditorily, in the stream of speech” (David Crystal, *A Dictionary of Linguistics & Phonetics*, Blackwell, 2003).

świadczące na korzyść obu stanowisk. Jednakże znalezienie rozwiązania tej kwestii rodzi kolejne pytanie o domenę tonu, a ponieważ możliwości jest kilka – jak zazwyczaj w takich sytuacjach zdania są podzielone.

Przyjrzyjmy się zatem bliżej obecnym w literaturze koncepcjom.

1.5.1. Ton jako zjawisko suprasegmentalne

Aby w pełni zrozumieć, na czym polega podejście suprasegmentalne, należałoby wyjaśnić, czym tak naprawdę są te „suprasegmenty” (*suprasegmentals*), co okazuje się zadaniem niełatwym. Genezę tego terminu w niezwykle zwięzły sposób przedstawia John Goldsmith (1979) w swojej pracy *Autosegmental Phonology*:

Ever since there have been segments in phonology, there have been phenomena that evaded segmental classification, and so there have been suprasegmentals (s. 5).

Choć próby zdefiniowania pojęcia „suprasegment” podejmowano wielokrotnie, jego znaczenie wciąż pozostaje uchwytnie w dużej mierze intuicyjnie. Jedną z definicji, którą autor przytacza, jest propozycja Haugena (1949):

Some sounds are thought of as occurring one after another, like bricks in a wall, while others occur simultaneously with these and usually span a number of the individual bricks at a time (s. 278, za: Goldsmith 1979: 7).

Pomimo nieco rozmytego sensu tego terminu wśród językoznawców istnieje powszechny konsensus co do prozodycznej natury zjawisk suprasegmentalnych.

Koncepcja tonu jako cechy prozodycznej pojawiła się w chińskiej lingwistyce za sprawą Chao w latach 1930–1933. W kolejnych latach stopniowo zyskiwała ona poparcie coraz szerszego grona lingwistów, na czele z Pikiem, Firthem oraz Wangiem (odpowiednio: 1948, 1957, 1967, za: Duanmu 2000: 233–234). Wszyscy oni postulowali związek

tonu z jednostkami większymi niż segmenty. Leben (1973:117, za: Schuh 1978: 222) wskazuje na kilka form, jakie proponowano w ramach ujęcia suprasegmentalnego, w zależności od domeny tonu, mianowicie – ton przedstawiano jako cechę wyrazów, morfemów lub sylab, co w przypadku języka mandaryńskiego redukuje liczbę możliwości do dwóch, gdyż w języku tym każdy morfem składa się z pojedynczej sylaby. W *The Phonology of Standard Chinese*, Duanmu (2000: 233) analizuje związek mandaryńskich tonów z segmentami, przedstawiając m.in. argumenty przemawiające za koncepcją suprasegmentalną. Po pierwsze stwierdza, że przebieg tonu zwykle pozostaje niezmienny, niezależnie od jednostki, która jest jego nośnikiem. Po drugie, nawet w przypadku redukcji sylaby o samogłoskę będącą ośrodkiem danego tonu, ton ten może zachować pierwotną formę, która zostaje następnie nałożona na inną samogłoskę. I wreszcie: ton, w odróżnieniu od cech segmentalnych, jest elementem ruchomym, gdyż wykazuje tendencję do częstej migracji na sąsiednie sylaby, co zostaje poparte przykładami przez Chao i Yanga (1962, za: Schuh 1978: 231).

Według Jongmana (et al. 2006: 13) jedna z fundamentalnych różnic pomiędzy cechami suprasegmentalnymi a segmentalnymi leży w ich związku z otaczającym kontekstem. Ponieważ tony określane są kategoriami relatywnymi takimi, jak np. „wysoki” lub „niski”, wykazują one bardzo ścisły związek z otoczeniem, w przeciwieństwie do segmentów, których cechy są ich wewnętrznymi właściwościami.

1.5.2. Ton jako zjawisko segmentalne

Jak zauważa Leben (1973: 117, za: Schuh 1978: 222), jednym z powodów istnienia wielu różnych, często przeciwstawnych wyjaśnień zasad działania procesów tonalnych jest fakt istnienia języków, w których procesy te w niektórych przypadkach wykazują związek z informacją segmentalną, podczas gdy w innych zdają się zachodzić niezależnie od niej. Zgodnie z ujęciem segmentalnym ton jest jedną z cech segmentu, przedstawioną jako element pewnego rodzaju matrycy. Matryca ta, czyli zbiór cech, opisujących np. wysokość tonu lub sposób artykulacji, stanowi swoistą charakterystykę danego segmentu. Spostrzeżenia mogące świadczyć o słuszności podejścia segmentalnego przedstawia Duanmu

(2000: 233–234), nawiązując w ten sposób polemikę z argumentacją zwolenników suprasegmentalizmu. Jak pisze, pomimo że ton może rozciągać się na dwie lub nawet więcej sylab, każda mora, stanowiąca niepodzielny segment, może nosić tylko jeden poziom toniczny – H lub L, co zostało potwierdzone w badaniach (por. § 1.4.3). W świetle tej koncepcji ton, funkcjonujący jak cecha segmentalna, zachowuje się w analogiczny sposób, a zatem nie może wystąpić dwukrotnie w ramach tego samego segmentu. Poza tym, wedle praw tonogenezy, cechy tonalne i cechy spółgłosek wzajemnie na siebie oddziałują. Autor podaje przykład właściwości określonej jako [luźne struny głosowe], która będąc jednocześnie cechą dźwięczności spółgłoski oraz cechą rejestru samogłoski, może rozciągać się z pierwszej na drugą, co dowodzi jej mobilności. Stwierdza natomiast, że nie jest to jedyna ruchoma cecha; znane są bowiem przypadki, w których cechy samogłosek takie, jak [tylna] i [zaokrąglona], oraz spółgłosek, jak [nosowa] czy [retrofleksyjna] również podlegały migracji. Wyjaśnienie tego zjawiska oferuje teoria cech (*feature theory*), która głosi, że każda cecha zajmuje osobny poziom i jest zdolna do swobodnej migracji wówczas, gdy nic nie stoi jej na przeszkodzie, tj. nie jest ona także cechą segmentów pośrednich. Zgodnie z tym ton jest ruchomy, ponieważ może rozciągać się tylko na sylaby nieakcentowane, które cechy tej w swoim zestawie nie mają. Ton nie jest natomiast jedyną cechą, która pozostaje pomimo redukcji segmentalnej; podobne właściwości wykazują również cechy [nosowa], [tylna] oraz [zaokrąglona].

1.5.3. Ton w ujęciu fonologii autosegmentalnej

Zanim omówiony zostanie sposób, w jaki autosegmentaliści postrzegali zjawisko tonu, należy pokrótce przedstawić zasadę konceptualizacji wszelkich zjawisk fonetyczno-fonologicznych według fonologii autosegmentalnej. Jak pisze Goldsmith (1979: 16), twórca podejścia autosegmentalnego, fonologia autosegmentalna jest pewnym spojrzeniem na geometrię reprezentacji fonetycznych, zgodnie z którym ukazywane są one jako zestaw paralelnych sekwencji cech akustyczno-artykulacyjnych, osadzonych na izolowanych poziomach i charakteryzujących ustawione liniowo poszczególne segmenty mowy oraz podlegające pewnym zasadom, określającym sposób, w jaki poziomy

te mogą być ze sobą powiązane. A ponieważ – wyjaśnia dalej – żaden z tych segmentów nie podlega ani też nie dominuje nad innymi, lecz każdy z nich jest niezależny sam w sobie, relacje te najlepiej oddaje przymiotnik „autosegmentalny”. Początek fonologii autosegmentalnej sięga połowy lat 70. i – jak przyznaje sam Goldsmith – w dużej mierze bazuje na wkładzie J.R. Firtha w rozwój dziedziny zajmującej się prozodią (1992: 151–152).

Do postrzegania tonu z perspektywy autosegmentalnej znacznie przyczyniła się krytyka nurtu suprasegmentalnego autorstwa Lebena (1973, za: Ladd 1996: 255). Stwierdzono wówczas, że „tone is notorious for its independence from the segments on which it is realized” (Yip 2006: 7), co spowodowało odejście od uznawanego do tamtego czasu zasadniczego dychotomicznego podziału na rzecz zgoła innej reprezentacji, która przedstawiała ton jako cechę umieszczoną na oddzielnym poziomie, lecz połączoną z odpowiednimi jednostkami TBU za pomocą specjalnych wiązań (*association lines*, Yip 2006: 7). Wiązania te są tworzone w oparciu o zbiór reguł, zaczerpniętych poniżej z pracy Yip (Yip 2006: 7) i stanowiących modyfikację pierwotnej wersji Warunku Dobrego Dopasowania (*Well-formedness Condition*) Goldsmitha (1979: 27). Brzmia one następująco:

- 1) Każda jednostka TBU musi być nacechowana tonem.
- 2) Każdy ton musi być związany z jakąś jednostką TBU.
- 3) Wiązanie odbywa się jeden-do-jednego i przebiega od lewej do prawej.
- 4) Wiązania nie mogą się przecinać.

Zgodnie z tymi warunkami tony mają możliwość rozszerzania swojego zasięgu poza segment, na których są zlokalizowane. Na podobnej zasadzie pojedyncze segmenty mogą łączyć się z więcej niż jednym tonem, formując w ten sposób kontury tonalne. W kwestii tej podejście autosegmentalne przypomina koncepcję Woo (por. § 1.4.2), postulującą traktowanie konturów jako sekwencji tonów równych, którą Goldsmith przyjmuje bezdyskusyjnie. Działanie powyższych reguł ma jeszcze jedną konsekwencję, uwzględnioną także w podejściu suprasegmentalnym – mianowicie w sytuacji, gdy sylaba zostaje zredukowana o element, na którym realizowany jest ton, tj. samogłoskę, ton ten nie zanika, lecz lokuje się na innej samogłosce. Właściwość ta, znana pod nazwą stabilności tonu (*tone stability*, Goldsmith 1979: 30), dowodzi wzajemnej niezależności obu poziomów – pomimo zmian

w strukturze sylabowej, forma tonu zostaje zachowana. Przejawia się ona także w sytuacji, gdy dany segment podwaja się, natomiast ton – nie. Koncepcja autosegmentalna zgadza się z suprasegmentalną także w kwestii mobilności tonów, gdyż, jak pisze Yip (2006: 7), mogą one przenosić się na inne segmenty.

1.6. Ton, akcent i intonacja

W *The Prosody of Mandarin Chinese* (1990) Susan Shen podsumowuje wyniki swoich badań stwierdzeniem, że ton, akcent i intonacja są ze sobą bardzo ściśle związane (s. 64). Choć badacze różnią się opiniami co do prymatu jednej z tych cech prozodycznych nad innymi, jednak bez względu na przyjętą hierarchię ważności oczywiste jest, że wszystkie one tworzą sieć wzajemnych zależności, wpływając na siebie w mniejszym lub większym stopniu.

1.6.1. Ton a akcent

Według Wierzchowskiej (1980) akcent to „uwydatnienie pewnych elementów ciągu mownego, np. pewnych sylab w wyrazach albo wyrazów w zdaniach” (s. 133). Na tej podstawie wyróżniamy akcent wyrazowy⁸ oraz akcent zdaniowy, związany z intonacją. Wierzchowska kontynuuje:

Uwydatnienie sylaby akcentowanej może polegać na silniejszym, tj. wykonanym z większą energią artykulacyjną, bardziej precyzyjnym, a zarazem głośniejszym jej wypowiedzeniu, na przedłużeniu tej sylaby, na podwyższeniu (czasem na obniżeniu) wysokości tonu krtańowego. W zależności, który z tych czynników przeważa, akcent określa się jako dynamiczny (ekspiracyjny), rytmiczny (kwantytatywny) lub muzyczny (melodyczny) (s. 133–134).

⁸ Wyraz jest tu rozumiany jako pewna całość pod względem prozodycznym, natomiast niekoniecznie ortograficznym.

Należy przy tym zaznaczyć, że kwestia nazewnictwa ma związek z dominacją jednego z wymienionych czynników, co nie oznacza, że musi on być jedynym, gdyż – jak pisze Majewicz (1989) – istnieją systemy o akcentach stanowiących typy pośrednie bądź też wykazujące cechy charakterystyczne dla więcej niż jednego spośród wskazanych sposobów akcentowania.

Jak zauważa Duanmu (2000: 125), w niektórych językach akcent jest zjawiskiem niewątpliwym, uchwytnym percepcyjnie i w dużej mierze realizowanym świadomie, a także z reguły łatwo poddającym się klasyfikacji, natomiast w innych nie jest on kwestią oczywistą. Jednym z takich języków jest właśnie standardowa odmiana języka chińskiego, w której akcent wywołuje opinie skrajne do tego stopnia, że kwestionowana jest jego obecność w ogóle, choć pogląd ten reprezentuje stosunkowo niewielu badaczy. Niestety samo stwierdzenie, że akcent faktycznie istnieje w języku mandaryńskim, zdaje się nie rozwiązywać problemu, lecz – jak pokazują badania – nastęrczać dodatkowych trudności, począwszy od wyróżnienia pewnych podtypów w ramach danego rodzaju akcentu do zasad jego dystrybucji w strukturach różnej długości. Jedną z podstaw, na której opierają się wątpliwości dotyczące obecności akcentu, jest udokumentowana niezgodność rodzimych użytkowników języka mandaryńskiego w testach percepcyjnych, polegających na wskazywaniu wyrazów objętych akcentem o różnym stopniu natężenia (Chao 1970: 38). Jednakże w obliczu dużej liczby trudnych do zignorowania danych uzyskanych z licznych badań akustycznych, które dowodzą wymiernych różnic pomiędzy sylabami akcentowanymi a nieakcentowanymi, można z dużą dozą prawdopodobieństwa stwierdzić, że akcent istotnie występuje w języku mandaryńskim, niemniej jest zjawiskiem mało wyrazistym pod względem fonetycznym. Duanmu (2000: 143–144) sugeruje, że przyczyna tego stanu rzeczy leży w specyfice języków tonalnych: ponieważ w językach tych zmiany częstotliwości podstawowej, będącej głównym wyznacznikiem akcentu, służą różnicowaniu wyrazów pod względem leksykalnym, w związku z tym nie może ona być dowolnie modyfikowana w celu sygnalizacji akcentu. Należy w tym miejscu zwrócić uwagę na pewne istotne rozgraniczenie pomiędzy akcentem tonalnym (zwanym również muzycznym lub melodycznym – por. s. 35) a tonem, który, choć realizowany z użyciem charakterystycznych dla tego rodzaju akcentu środków fonicznych,

notabene akcentem nie jest. Z tego względu języki tonalne musiały wypracować swoisty zestaw właściwości, jakimi powinny odznaczać się sylaby zajmujące pewne prominentne w danych strukturach pozycje bądź wymagające emfazy z racji swej funkcji, znaczenia czy jakichkolwiek innych powodów.

Pomimo istnienia w tej kwestii wątpliwości większość badaczy opowiada się za obecnością akcentu w języku mandaryńskim. Jak stwierdza Goldsmith (1994: 8), w systemach, które nazywa akcentowymi (*accentual*), pewne sylaby podkreśla się, wydłużając je, wymawiając z większą siłą, a więc głośniej lub też zwiększając wysokość głosu (rzadziej bardziej go obniżając). Z kolei w językach tonalnych występowanie akcentu na sylabach jest skorelowane z tonem i przejawia się głównie zwiększeniem zakresu F_0 oraz wydłużeniem czasu trwania danej sylaby w stosunku do jej formy nienacechowanej, podczas gdy głośność ma znaczenie drugorzędne. W efekcie tony pierwotnie wysokie wymawiane są na jeszcze większej wysokości bądź zaczynają się w wyższym rejestrze, natomiast tony niskie zaczynają się lub opadają niżej niż zwykle (Shen 1990: 59–60). Pod względem fonemicznym Chao (1970: 35) wyróżnia trzy rodzaje akcentu: normalny (*normal*), kontrastywny⁹ (*contrasting*) oraz słaby (*weak*). Poniżej znajduje się krótki opis każdego z nich.

1) Akcent normalny.

Ogólnie rzecz biorąc, wszystkie sylaby mające jeden z czterech tonów leksykalnych uważa się za akcentowane. Sylaby te Duanmu (2000) nazywa pełnymi (*full syllables*, s. 135), w opozycji do sylab niemających tonu lub inaczej – mających ton zerowy, zwanych słabymi (*weak syllables*, s. 135). Sylaby pełne otrzymują akcent w myśl sformułowanej przez badacza Reguły Ton-Akcent (*The Tone-Stress Principle*, 1996, za: Duanmu 2000: 230). Brzmi ona:

Reguła Ton-Akcent

Sylabie akcentowanej może towarzyszyć przypisany jej przebieg tonalny.

Sylaba nieakcentowana nie ma przebiegu tonalnego.

⁹ Akcent nietoniczny, aplikowany świadomie w celu zwrócenia uwagi odbiorcy na pożądane treści.

Poza różniącą obie grupy obecnością tonu w badaniach stwierdzono, że sylaby pełne są głośniejsze, mają dłuższy przebieg czasowy oraz większy zakres F_0 niż sylaby słabe (Duanmu 2000: 135).

Spośród sylab akcentowanych te, na które nie pada akcent kontrastywny, mają akcent normalny. Chao (1970) zwraca jednak uwagę, że choć opatrzone akcentem identycznym pod względem fonemicznym, sylaby te różnią się co do akcentu fonetycznego w sposób taki, że w nieprzerwanej pauzie sekwencji tychże sylab, tworzącej frazę bądź złożenie, sylaba końcowa akcentowana jest najmocniej, nieco słabiej pierwsza, natomiast sylaby znajdujące się pomiędzy otrzymują akcent najslabszy. Jak zauważa Duanmu, koncepcja ta jest w zgodzie z założeniami teorii akcentu końcowego (*the final stress theory*, 2000: 135). Częściowo zgadza się z nią także Shen (1990), pisząc, że sylaby w pozycjach końcowych niebędące enklitykami są zazwyczaj najmocniej akcentowane, co jest wynikiem wydłużenia czasu ich trwania, natomiast w przypadku, gdy na końcu znajduje się enklityka, akcent przesuwa się wówczas na sylabę przedostatnią. Inaczej rzecz się ma w teorii Firbasa (1985, za: Shen 1990: 63), zgodnie z którą cechy prozodyczne „przybierają na wadze” w miarę zbliżania się do końca wypowiedzi, Shen przyjmuje, że najslabszy akcent przypada na sylabę pierwszą. Pomimo różnicy zdań lingwistów co do kwestii dystrybucji akcentu istnieje jednak powszechna zgoda, że stopień jego natężenia jest uzależniony od pozycji sylaby, w związku z tym różne stopnie akcentowania traktowane są jako allofony jednego akcentu fonemicznego.

2) Akcent kontrastywny.

Akcent ten pełni głównie funkcję pragmatyczną, gdyż wiąże się ze świadomym, mającym konkretny cel uwypukleniem pewnych treści. Podobnie jak w innych, nietonalnych językach, sylaby akcentowane kontrastywnie w języku mandaryńskim znacznie odróżniają się od pozostałych, tak więc ich wskazanie zwykle nie przysparza rodzimym użytkownikom większych trudności (Duanmu 2000: 144). Z badań Shen (1990) wynika, że sylaby te charakteryzuje szerszy zakres F_0 oraz dłuższy czas trwania niż sylaby z akcentem normalnym, także te w pozycjach końcowych. Chao (1970) wskazuje również na towarzyszącą im zwykle większą intensywność.

3) Akcent słaby.

Przejawia się zredukowaniem tonu do jednej wartości wraz ze znacznym skróceniem czasu trwania danej sylaby, nawet do połowy długości trwania rymu sylaby pełnej (Duanmu 2000: 135); można więc stwierdzić, że akcent ten oznacza w zasadzie brak akcentu. Zgodnie z Regułą Ton-Akcent (por. s. 37) sylaby z akcentem słabym nie mają przebiegu tonalnego, zatem charakteryzuje on sylaby z tonem zerowym. Większość sylab nieakcentowanych stanowią enklityki, których wysokość uzależniona jest od poprzedzającego je tonu (por. § 1.2.1). W znacznie rzadszych przypadkach sylaba atoniczna występuje bezpośrednio przed sylabą akcentowaną, wówczas jest ona proklityką, której wysokość równa jest w przybliżeniu 3 na pięciostopniowej skali tonalnej Chao (por. § 1.1). Enklityki i proklityki należą do stosunkowo niewielkiej grupy sylab atonicznych, składającej się z morfemów gramatycznych oraz sylab zleksykalizowanych (Chao 1970). Według Shen (1990) tak jak każda sylaba toniczna traci swój ton tylko wówczas, gdy zostaje pozbawiona akcentu, tak analogicznie niemal każda sylaba z tonem zerowym może odzyskać przypisany jej ton, jeżeli zostanie zaakcentowana. Wyjątek stanowią sylaby, którym z racji ich kategorii gramatycznej przypisany jest ton zerowy, jak np. partykuła pytajna 嗎 *ma*. W swojej analizie badaczka poświęca również sporo uwagi zjawisku tonalnego *sandhi* i jego fonetycznym konsekwencjom, podsumowując ogólnym stwierdzeniem, że proces ten nie zachodzi na sylabach akcentowanych.

4) Akcent średni (*medium stress*)

Chao (1970: 37–38) wspomina o jeszcze jednym rodzaju akcentu, uwzględnianym przez niektórych badaczy, m.in. Hocketta (1947). Jest to akcent o średnim stopniu napięcia, stanowiący stadium pośrednie między akcentem normalnym a słabym. Chao stwierdza natomiast, że z uwagi na relatywną naturę akcentu prawie każdy przypadek można rozpatrywać w alternatywny sposób, np. traktując sylabę silniej akcentowaną jako mającą akcent kontrastywny zamiast normalnego, i dla uzyskania odpowiedniego kontrastu zestawiając ją z sylabą o akcencie o poziom niższym, w tym wypadku normalnym w miejsce średniego. Przeciw istnieniu akcentu średniego świadczą jednak wyniki badań percepcyjnych, w których rodzimi użytkownicy wykazują w tej kwestii znaczną niezgodność.

Na podstawie przytoczonych badań można pokusić się o konkluzję, że ton i akcent istotnie łączy bliska wzajemna zależność. Wpływ akcentu na realizację danego tonu jest w dużej mierze determinowany przez jego siłę, co Shen (1990: 60) podsumowuje stwierdzeniem, że stopień natężenia akcentu określa, do jakiego stopnia zrealizowany zostanie pewien przebieg tonalny w stosunku do jego formy nieakcentowanej. W analogiczny sposób akustyczne parametry, jakie ma dany ton, zrealizowany w sposób zamierzony lub nie, odpowiedzialne są za końcowy rezultat w postaci sylaby akcentowanej w mniejszym lub większym stopniu. Do parametrów tych należą: częstotliwość podstawowa z jej ewentualnymi zmianami, czas trwania, a także poziom głośności danej sylaby. Ogólnie rzecz ujmując, im większa wartość każdego z tych parametrów, tym akcent jest mocniejszy. Niemniej w świetle wyników badań Shen (1993, za: Jongman et al. 2006: 9–10) nad percepcją akcentu wśród rodzimych użytkowników języka mandaryńskiego, którzy wykazywali skłonność do przypisywania akcentu sylabom dłuższym, nasuwa się wniosek, że to właśnie czas trwania jest głównym korelatem akcentu w standardowej odmianie języka chińskiego.

1.6.2. Ton a intonacja

Choć w kwestiach lingwistycznych, z uwagi na ogromną różnorodność systemów językowych rzadko można pozwolić sobie na dokonanie generalizacji, w przypadku wysokości głosu jest rzeczą oczywistą i niepodważalną, że wszystkie języki wykorzystują jego modulacje do pewnych celów, bowiem w żadnym z dotychczas znanych systemów komunikacja nie odbywa się w sposób całkowicie monotony. Wprawdzie poszczególne języki różnią się w tej materii dość znacznie, można wyszczególnić dwa główne cele, którym te zmiany służą, i które są związane z wielkością struktury, będącej ich domeną – mianowicie: gdy modulacje odbywają się w granicach wyrazów, stanowią podstawę leksykalnych, a czasem gramatycznych różnic między nimi, mówimy o tonie, natomiast jeżeli obejmują one swym zasięgiem całe zdanie bądź frazę, niosąc ze sobą treści pełniące funkcję pragmatyczną, wówczas mamy do czynienia z intonacją. W języku mandaryńskim, pomimo że jest on językiem tonalnym, zmiany F_0 służą obu tym celom, gdyż

– jak zauważa Francis Katamba (1996: 239) – nie są to zjawiska wzajemnie się wykluczające. Ponieważ w systemach tonalnych swobodna manipulacja F_0 z oczywistych względów nie jest możliwa, pojawia się w związku z tym pytanie, przy użyciu jakich mechanizmów wyrazić w języku takim jak mandaryński pewne treści pozaleksykalne oraz czy i w jaki sposób intonacja może wpływać na realizację tonów.

Rozważania nad tą kwestią należałoby zacząć od próby wyjaśnienia, co pozwala intonacji funkcjonować w języku mandaryńskim. Otóż kluczem do współlistnienia obu tych zjawisk prozodycznych jest stabilność, jaką zachowuje ton pomimo oddziaływania intonacji. Jongman et al. (2006) przytaczają szereg badań, które pokazują, że podstawowe właściwości tonu takie jak jego kontur, względna długość trwania oraz amplituda nie ulegają diametralnym zmianom w ramach różnych wzorców intonacyjnych. Obszerne dane potwierdzają, że zmiana intonacji nie wpływa negatywnie również na percepcję.

Skoro wiadomo, że w języku mandaryńskim zmiany F_0 mogą nieść ze sobą zarówno informacje leksykalne, jak i pozaleksykalne, przy czym ani jedno, ani drugie nie ulegają zniekształceniom w stopniu uniemożliwiającym odbiór komunikatów, pozostaje kwestia sposobu, w jaki przekazuje się za jej pomocą treści o zabarwieniu emocjonalnym czy realizuje funkcję impresywną. Jak pisze Shen (1990: 9), każdy system wypracował sobie do tego celu pewne mechanizmy, które podlegają specyficznym dla niego zasadom. Ladd (1996) wskazuje na trzy rodzaje zjawisk towarzyszących intonacji w językach tonalnych: rozszerzenie lub kompresja amplitudy F_0 przy wyrażaniu emocji, modyfikacje poszczególnych tonów, głównie końcowych, służące sygnalizowaniu różnic pomiędzy typami wypowiedzi (np. zdaniem oznajmującym a pytaniem) oraz zmiany całościowych przebiegów tonalnych w celu przekazania specyficznych informacji. W większości odmian języka chińskiego wiele funkcji, które w innych językach spełnia intonacja, realizowanych jest przy użyciu odpowiednich partykuł (Chao 1933, za: Duanmu 2000: 235). Chao (1970) proponuje także szczegółowy opis przypadków, w których partykuły nie biorą udziału, ilustrując zjawisko interferencji za pomocą metafory morskich fal, w której tony jako drobne „zmarszczki” na powierzchni wody pokrywają grzbiety fal, reprezentujących intonację w zdaniu, przy czym „zmarszczki” te są niekiedy „większe” niż fale. Wynik interakcji obu tych czynników badacz przedstawia jako prosty

rachunek algebraiczny, w którym spotkanie dwóch grzbietów potęguje efekt końcowy, dając w rezultacie ich sumę, natomiast gdy grzbiet zbiega się z doliną, zachodzi osłabienie, będące wynikiem odejmowania. Zgodnie z tym ton 2 osiągnie wysokość wyższą niż zwykle w zdaniu o intonacji rosnącej, a ton 4 opadnie niżej w przypadku intonacji opadającej. Z kolei w sytuacji odwrotnej ton 2 zostanie zrealizowany w rejestrze nieco niższym, wciąż jednak zachowując kontur rosnący, natomiast ton 4 rozpocznie się wyżej niż na ogół, jednocześnie nie opadając do końca. Badacz wyróżnia także 13 typów intonacji, w zależności od intencji bądź nastroju nadawcy. Są to:

- 1) Intonacja normalna (*Normal Intonation*), używana do wyrażania zwyczajnych komunikatów. Stanowi sekwencję następujących po sobie tonów bez szczególnych modyfikacji.
- 2) Intonacja suspens-rozwiązanie (*Suspense-Conclusion Intonation*), w której pierwsza część zdania realizowana jest w nieco wyższym rejestrze niż część druga.
- 3) Przyspieszone tempo w fazie końcowej: proste pytania, proste polecenia (*Accelerated Tempo in the Last Few Syllables: Simple Questions, Simple Commands*). W wypowiedziach tego typu główny akcent pada na ostatnią pełną sylabę, powodując niewielkie zwiększenie tempa oraz amplitudy F_0 w końcowej części wypowiedzi. Wrażenie przyspieszenia – wyjaśnia Chao – powstaje w wyniku braku oczekiwanego wydłużenia końcowych sylab, co ma zwykle miejsce w języku mandaryńskim.
- 4) Zwiększona amplituda: wyrażenia pełne serdeczności; narzekanie (*Extra Wide Range: Airy, Breezy Expression; Complaining Expression*). Wypowiedzi pierwszego typu charakteryzuje nieco podwyższony rejestr oraz łagodniejszy ton, podczas gdy typ drugi realizowany jest w rejestrze normalnym i często towarzyszy mu „chropawy” tembr głosu.
- 5) Wysoki rejestr, spokojny ton: dystans do sytuacji, ponawianie pytania (*High Pitch, Quiet: Taking Things Lightly; Asking Questions a Second Time*). Obu sytuacjom komunikacyjnym towarzyszy intonacja opadająco-wznosząca.
- 6) Falset: zniecierpliwienie; silna dezaprobata (*Falsetto: Impatience; Strong Disapproval*); częsta obecność zajknięć w wypowiedziach.

- 7) Niski rejestr: powaga, uznanie, podniosły nastrój (*Low Pitch: Seriousness, Praise, Great Feeling*);
- 8) Niski, często „napowietrzony” ton: prośba o potwierdzenie (*Low, Often Breathy: Questions of Confirmation*). Pytaniom tym nierzadko towarzyszy partykuła emfaticzna ㄤ a.
- 9) Niski rejestr, szybkie tempo: wtrącenia (*Low and Fast: Parenthetical Insertions*).
- 10) Brak pauz: słowa nie nadążają za myślami (*Negative Pause: Words Trying to Catch Up with Thoughts*); liczne skróty myślowe, brak przerw, ucinanie zdań.
- 11) – 12) Końcowy wzrost lub spadek intonacji (*Rising and Falling Endings*). Dotyczą wyłącznie ostatniej sylaby w danej wypowiedzi, w związku z czym są przez Chao traktowane jako partykuły.
- 13) Nieoczekiwany wzrost siły głosu: gniew lub zniecierpliwienie (*Random Loudness: Anger or Impatience*); rodzaj bardzo ekspresywnej intonacji, polegającej na wymawianiu każdej lub większości sylab w wypowiedzi z większą głośnością, ale bez zwiększenia ich wysokości i długości, co odróżnia ją od akcentu kontrastycznego. (Chao 1970: 40–44).

He i Jin (1992, za: Duanmu 2000: 236) wymieniają sześć typów intonacji, stosowanych w różnych kontekstach komunikacyjnych: stwierdzenie, oczekiwanie potwierdzenia, pytanie, prośbę, polecenie oraz wykrzyknienie, przy czym, podobnie jak Chao, wskazują jedynie dwa mechanizmy biorące udział w ich realizacji, tj. rozszerzanie i zawężanie rozpiętości głosu. Shen (1990) uważa natomiast, że wszelkie emocjonalne, ekspresywne bądź emfaticzne niuanse wypowiedzi są zbyt złożone oraz charakteryzują się wieloma niestabilnymi zwrotami intonacyjnymi, co uniemożliwia ujęcie ich w kategoriach stałych, gotowych wzorców. Dlatego też proponuje wyodrębnienie „minimum intonacyjnego”¹⁰ (Shen 1990: 9), czyli neutralnej pod względem emocjonalnym, rzeczowej, dającej się sformalizować intonacji, mogącej następnie posłużyć za bazę, na którą nakładane są rozmaite wzorce o różnej modalności czy nastawieniu pragmatycznym. Baza ta, będąca pewnego rodzaju „wspólnym mianownikiem”, pozwala na porównywanie wypowiedzi o zróżnicowanym charakterze, stanowiąc narzędzie

¹⁰ Termin zaczerpnięty od Pike’a (1945).

pomocne przy identyfikacji wszelkich „dodatkowych” elementów oraz ich wnikliwej analizie.

Swoje dociekania Shen (1990) rozpoczyna od bliższego przyjrzenia się intonacji wypowiedzi interrogatywnych, a w szczególności pytaniom typu *yes-no*. Ponieważ w języku mandaryńskim jako języku tonalnym pytania zadawane są głównie z użyciem gramatycznych wykładników takich jak partykuły lub za pomocą odpowiedniego szyku zdania (A-przeczenie-A), funkcjonował niegdyś pogląd, że wypowiedzi te zachowują intonację zdania twierdzącego. Jak się jednak okazuje, nie jest to do końca prawdą, choć należy zwrócić uwagę, że za to mylne pojęcie odpowiedzialne jest pewne uproszczenie, dokonane przez innych badaczy na podstawie analizy języka kantońskiego. Otóż Shen odkryła, że w odmianie standardowej intonację pytającą charakteryzuje wzrost wysokości głosu już na samym początku, który dalej utrzymuje się przez cały przebieg tonalny wypowiedzi, a nie, jak w języku kantońskim, występuje wyłącznie na ostatniej sylabie; innymi słowy, intonacja pytająca realizowana jest w rejestrze wyższym niż twierdząca.

Kolejna kwestia podejmowana przez Shen (1990) dotyczy faktycznej relacji pomiędzy intonacją a tonem, w której zasadnicze pytanie koncentruje się wokół ewentualnej możliwości modyfikacji formy konturów tonalnych przez wzorce intonacyjne oraz – jeśli taka istnieje – stopnia, do jakiego ingerencja ta zachodzi. Inspiracją dla badaczki były dwa obecne w literaturze sprzeczne poglądy: jeden z nich zakładał, że tony mogą podlegać zmianom w wyniku interakcji z intonacją, choć nie są jej całkowicie podporządkowane, podczas gdy drugi głosił bezwzględną niezmienną podstawowych kształtów konturów tonalnych, które są na tyle trwałe, aby oprzeć się wpływowi intonacji. Z badań Shen wynika tymczasem, że wpływ ten zachodzi w dwóch wymiarach: zarówno wartości F_0 , jak i kształtu konturów niektórych tonów. Jak wspomniano, podwyższony rejestr wypowiedzi interrogatywnych wiąże się ze zwiększeniem wartości F_0 każdego tonu z osobna, a w szczególności tonu sylaby końcowej. Badania te potwierdzają częściowo spostrzeżenia Chao (1970) (por. s. 42) dotyczące wzrostu wysokości tonów rosnących w przypadku, gdy kierunek ich konturów jest zgodny z ogólnym kierunkiem przebiegu intonacyjnego w danej wypowiedzi. Wyniki pokazują również, że w zdaniach o intonacji pytającej oraz twierdzą-

cej zniekształceniu ulega kontur tonu 1, który zmienia się z równego na lekko opadający, co jednak nie wpływa negatywnie na jego percepcję.

Na uwagę zasługują także wnioski dotyczące tonu zerowego, którego wysokość w dużym stopniu zależy od wpływu intonacji. Jak stwierdzono już wcześniej, kształt przebiegu tonu zerowego jest spłaszczony niemal do jednej wartości. Wyjątek stanowią przypadki, gdy występuje on w pozycji końcowej w zdaniach interrogratywnych; wówczas przybiera kontur opadający po tonie 1, 2 i 4, natomiast rosnący – po tonie 3 (por. § 1.2.1). Z kolei w zdaniach o intonacji normalnej – twierdzącej ton zerowy stanowi niejako „przedłużenie” konturu tonu bezpośrednio poprzedzającego. Inaczej sytuacja przedstawia się w sekwencjach tonów zerowych, z których jedynie pierwszy ma wartość uzależnioną od wysokości tonu występującego przed nim, natomiast reszta podporządkowana jest intonacji, w związku z czym przyjmuje narzucone przez nią wartości (Shen 1990: 38–48).

1.7. Akcent w języku polskim

Analizując historię badań nad akcentem języka polskiego, można zauważyć, że zagadnienie to zawsze – odkąd zaczęto się nim szerzej interesować – wzbudzało w kręgach lingwistycznych dość silne kontrowersje; poza nielicznymi okresami, w których można mówić o pewnej ogólnej tendencji do akceptacji niektórych jego aspektów, trudno bowiem doszukać się w tej kwestii powszechnego konsensusu. Już samo zdefiniowanie pojęcia akcentu okazuje się zadaniem niezwykle problematycznym, czego dowodzą niekiedy znaczne różnice w funkcjonującej w literaturze językoznawczej terminologii, stosowanej przez poszczególnych badaczy do opisu zjawisk akcentowych; Karpiński (2006) wspomina tu nawet o pewnego rodzaju „chaosie terminologicznym” (s. 61). Problem ten ma jednak znacznie szersze podłoże, gdyż za wszelkie nieścisłości oraz brak jednolitego, spójnego ujęcia tego zagadnienia odpowiedzialnych jest wiele czynników, począwszy od podejścia badawczego, poprzez rodzaj i przedmiot prowadzonych badań, aż po wykorzystywany aparat pojęciowy oraz zestaw terminów, a także – przede wszystkim – sposób ich rozumienia.

Duży wpływ na zmianę podejścia badawczego do zagadnienia akcentu polskiego z pewnością miał znaczny postęp techniczny, wykazujący szczególną dynamikę w ostatnich kilku dekadach. Spowodował on, że na znaczeniu zaczęły zyskiwać badania empiryczne w postaci szczegółowych analiz fonetyczno-akustycznych oraz testów percepcyjnych, stopniowo wypierając rozprawy czysto teoretyczne. Choć podobne zjawisko obserwuje się także w innych, dających ująć się w statystyczne ramy dziedzinach nauki, w przypadku akcentu w polszczyźnie jest ono spotęgowane powszechnie niegdyś panującym przekonaniem, że akcent jako zjawisko subiektywne i przez to, być może, w dużej mierze intuicyjne, nie kwalifikuje się do bycia obiektem pomiarów (Jassem 1962). Pogląd ten niewątpliwie sprzyjał teoretyzowaniu oraz stwarzał dogodne warunki do nieco swobodnej analizy i interpretacji, często opartej na twierdzeniach innych lingwistów podejmujących to zagadnienie. Ponieważ we współczesnym świecie nauki za podstawę wiarygodności uznaje się empiryczne potwierdzenie, badacze chętnie korzystają z możliwości technicznych dzisiejszej aparatury, formułując wnioski oparte na udokumentowanych wynikach własnych dociekań. Z tego względu unikają stawiania hipotez ogólnych na rzecz hipotez roboczych, które często podlegają weryfikacji w toku prowadzonych badań, niekiedy prowadząc do zaskakujących konkluzji. W rezultacie coraz częściej kwestionuje się poglądy, które przez lata z powodzeniem funkcjonowały w tradycji językoznawczej.

Co się tyczy wspomnianej tradycji, na pierwszy rzut oka można odnieść wrażenie panującej w ujęciu tym pewnej przejrzystości, wynikającej ze zgodności znacznej większości badaczy w kwestii charakteru akcentu polskiego. Mianowicie – z uwagi na fakt, że wśród „tradycjonalistów” istotnie dominowało przekonanie o jego dynamiczności, przyjęło się tradycję łączyć właśnie z teorią ekspiracyjną. Wydawać by się zatem mogło, że była to kwestia poniekąd oczywista i przez to przesądzona, niemniej, jak zauważa Demenko (1999), już na wczesnym etapie badań nad tym zjawiskiem akcent był pojęciem dość mętnym:

Akcent był w przeszłości w literaturze lingwistycznej często traktowany tak, jakby był niezależnym parametrem opisu fonetycznego, chociaż nie przypisywano mu jakichś niezależnych cech fonetyczno-akustycznych. Bywał określany bądź w niej-

snych terminach subiektywnych [...], bądź w terminach artykulacyjnych [...]. Jeśli wiązano akcent z jakąś cechą akustyczną, to z reguły przypisywano mu korelacje z wymiarem amplitudy, co oznaczało, że przyjmowano, iż sylaba akcentowana jest głośniejsza niż nieakcentowana [...] (s. 17–18).

Należy jednak zwrócić uwagę, iż – choć cieszący się największym poparciem – pogląd ten nie był przyjmowany bezdyskusyjnie, bowiem w latach 50. i 60. pojawiły się alternatywne koncepcje, w ramach których sugerowano, że wbrew powszechnemu przekonaniu głównym korelatem akcentu nie jest energia artykulacyjna, lecz wysokość głosu bądź intonacja i iloczyn sylab będących jego nośnikiem (Demenko 1999: 18). Na gruncie polskiego językoznawstwa kontrowersyjne stanowisko jako jedyny zaprezentował Jassem (1962), twierdząc, że akcent w polszczyźnie ma charakter melodyczny czy też inaczej – tonalny, a zatem realizowany jest głównie za pomocą intonacji, czym podważył dominującą w tamtym czasie teorię dynamiczną. Pomimo że żaden z badaczy nie podzielał jego zdania, pogląd ten niewątpliwie spowodował niemałą konfuzję, jako że nie był jedynie odważną hipotezą, lecz wnioskiem sformułowanym na podstawie wyników szczegółowych badań akustycznych, przeprowadzonych przy użyciu najnowszej wówczas aparatury.

Biorąc pod uwagę całokształt prac poświęconych tematyce akcentu polskiego, warto również wspomnieć, że wraz z rozwojem badań w tej dziedzinie zmianie uległ także przedmiot zainteresowania lingwistów. Otóż mniej więcej do początku lat 80. w literaturze językoznawczej porusza się jeszcze problem natury akcentu. U Wierzchowskiej (1980) czytamy: „Akcent języka polskiego najczęściej uważa się za dynamiczny” (s. 134). Autorka pokrótce wyjaśnia, w jaki sposób akcent ten przejawia się w danym języku oraz na jakiej podstawie odbywa się przypisywanie poszczególnym językom ich typów akcentowych. W późniejszych latach natomiast stopniowo odchodzi się od podobnych uogólnień w kontekście języka polskiego na rzecz większego zainteresowania *de facto* wyłącznie strukturą akcentu. Tendencja ta dominuje także obecnie, choć równie często mówi się o akcencie jako jednym ze zjawisk towarzyszących intonacji, z którą wykazuje on bardzo bliski związek, podczas gdy problem samej jego natury w ogóle nie jest w literaturze podejmowany.

Powodem tego jest zapewne szereg niejasności związanych z postrzeganiem i ujmowaniem zagadnienia akcentu, włączając w to brak jednolitej nomenklatury; nie dziwi zatem fakt, że kwestia ta jest przez dzisiejszych lingwistów traktowana z dużą ostrożnością.

Z kolei rozważania nad strukturą akcentu polskiego w ramach tradycji fonetycznej zasadniczo koncentrują się wokół kilku głównych pojęć, które w zależności od badacza mogą, choć nie muszą przybrać odmiennego znaczenia. Niemniej, najogólniej rzecz biorąc, wyróżnić można dwa podstawowe, powszechnie uznane rodzaje akcentu: akcent wyrazowy (*word stress*), zlokalizowany w obrębie pojedynczego wyrazu, oraz akcent zdaniowy (akcent frazowy, *sentence stress*, Karpiński 2006: 61), którego dziedziną jest zazwyczaj jednostka wyższego rzędu, większa i bardziej złożona niż wyraz. Akcent ten może pełnić funkcję emfatyczną, nieść ładunek emocjonalny wypowiedzi bądź wynikać z logiki wyvodu, w związku z czym jego użycie jest zawsze intencjonalne; nie spełnia on natomiast funkcji kontrastywnej, tj. nie rozróżnia wyrazów pod względem leksykalnym, aczkolwiek w pewnym stopniu wpływa na znaczenie wypowiedzi. Natomiast akcent wyrazowy jest pewnego rodzaju własnością z góry przypisaną pojedynczym, konkretnym jednostkom leksykalnym. Ponieważ w literaturze lingwistycznej istnieje powszechne przekonanie, że jednostki te – wyrazy – są strukturami kluczowymi w procesie przetwarzania oraz odbioru mowy (Steffen-Batogowa 2000), pomocny w ich wyodrębnianiu akcent wyrazowy, ze względu na regularność, jaką wykazuje w sposobie rozmieszczenia określany jako stały¹¹ pod względem fonetycznym, spełnia dwie bardzo istotne funkcje: kulminatywną oraz delimitatywną. W teorii funkcja kulminatywna umożliwia wyznaczenie liczby wyrazów danej wypowiedzi na podstawie liczby występujących w jej ramach akcentów, natomiast delimitatywna – wskazanie granic poszczególnych wyrazów, w rzeczywistości jednak obie funkcje podlegają znacznym ograniczeniom pod wpływem różnorodnych czynników. Oprócz swej nadrzędnej semantyczno-składniowej roli akcent wyrazowy służy także celom rytmicznym, porządkując wypowiedź zgodnie z pewnym swoistym dla polszczyzny wzorcem rytmicznym.

¹¹ Warto zaznaczyć, że stabilizacja ta zaistniała dopiero na jednym z kolejnych etapów ewolucji języka polskiego, mniej więcej około wieku XV, bowiem wcześniej akcent był elementem ruchomym (Topolińska 1961).

Polski akcent główny określa się jako paroksytoniczny, gdyż zasadniczo pada on na sylabę drugą od końca. Zdarzają się jednak dość liczne i częste odstępstwa od paroksytonezy – mianowicie w pewnych przypadkach akcent zlokalizowany jest na trzeciej lub nawet czwartej sylabie od końca; wówczas zwany jest odpowiednio proparoksytonicznym oraz proproparoksytonicznym. Do grupy wyjątków należą m.in. zapożyczenia grecko-łacińskie zakończone na „-ika” lub „-yka”, takie jak „matematyka”, „logika”, „gramatyka”. W tych wyrazach akcent pada na trzecią sylabę od końca, o ile forma przypadkowa ma tyle samo sylab co mianownik. Odmienna reguła dotyczy wyrazów zakończonych na „-ik” lub „-yk”, w których akcent proparoksytoniczny występuje w formach przypadkowych równych co do liczby sylab dopełniaczowi (np. matematyk → matematyka). Odrębną grupę stanowią formy gramatyczne wyrazów rodzimych, w których do przesunięcia akcentu dochodzi w wyniku dodania określonych sufiksów i końcówek. Jak pisze Topolińska (1961), zjawisko to występuje m.in. w formach czasu przeszłego i trybu warunkowego, np. „robiliśmy”, „widzielibyśmy”. Do form podlegających nieregularnej akcentuacji należą także niektóre złożenia i skrótowce (Steffen-Batogowa 2000). Choć powyższe sposoby akcentowania są wciąż obecne w normie, coraz powszechniejsze dążenie do unifikacji systemu – wynikające z ekonomii języka – prowadzi do uproszczeń, w tym do redukcji form obocznych, co umacnia pozycję paroksytonezy we współczesnej polszczyźnie. Zjawisko to jest jednak złożone i niejednoznaczne, ponieważ obserwuje się również tendencję odwrotną: poprawnie paroksytonicznie akcentowane wyrazy, takie jak „analiza”, „liceum” czy „biblioteka”, bywają dziś wymawiane z akcentem na trzecią sylabę od końca.

Należy zwrócić uwagę, że położenie akcentu nie ulega zmianom w wyniku działania procesów derywacyjnych i zmian fleksyjnych, a zatem nie jest ono podporządkowane morfologii. W konsekwencji zarówno wyrazy o prostej budowie, jak i ich derywaty oraz formy fleksyjne będą charakteryzowały się akcentem umiejscowionym na tej samej w kolejności – tj. drugiej, względnie trzeciej lub czwartej – sylabie od końca, licząc od granicy morfologicznej danej struktury.

Kolejną właściwość polskiego akcentu głównego to jego niefakultatywność, o czym pisze w swojej rozprawie Dłuska (1976). Cecha ta ma związek z wyraźną hierarchizacją polskich akcentów wyrazowych,

u podłoża której leży zróżnicowany stopień uwydatnienia sylab akcentowanych za pomocą odpowiednich bodźców akustycznych [...] (Steffen-Batogowa 2000: 268).

Dotyczy ona wyrazów co najmniej dwusylabowych, w zestawieniach których występuje swoiste stopniowanie siły akcentu w sposób taki, że pewne sylaby w wyrazach mających dla danej wypowiedzi szczególną wartość, i przez to wymagających emfazy ze względu na ich znaczenie bądź z innych określonych powodów, otrzymują akcent najsilniejszy, natomiast w pozostałych sylabach akcentowanych położonych w sąsiedztwie znacznie traci on na sile, a nierzadko zostaje zupełnie zignorowany. Niemniej jakkolwiek mocno osłabiona sylaba taka jest wciąż nieco silniejsza niż sąsiadująca z nią sylaba nieakcentowana, gdyż mimo wszystko zachowała ona swój akcent. Taka gradacja stopnia uwypuklenia poszczególnych sylab w obrębie danej wypowiedzi jest niezwykle ważna w procesie percepcji mowy, gdyż, jak ujmuje to Steffen-Batogowa:

wskazuje [ona] bowiem na względną doniosłość informacyjną poszczególnych wyrazów, przez co ułatwia słuchaczowi orientowanie się w treściowym toku wypowiedzi (Steffen-Batogowa 2000, s. 268).

Funkcję tę określa autorka jako wartościującą.

Zgodnie z teorią Dłuskiej nefakultatywność ma ponadto jeszcze jedną istotną implikację, która przejawia się w zdolności sylab do utrzymania akcentu i ma bliski związek z zachowaniem się wyrazów jednosylabowych, odznaczających się odmiennymi ze względu na swą specyfikę właściwościami. Otóż wyrazy te, poza stosunkowo niewielką grupą wyrazów atonicznych, do których należą enklityki i proklityki (por. § 1.6.1), są strukturami ortotonicznymi, czyli samodzielnie akcentowo, i traktowane są na równi z sylabami akcentowanymi struktur dwu- i wielosylabowych. O ile wyrazy dłuższe, o liczbie sylab równej co najmniej dwóm, nie mogą utracić akcentu, jak wykazano powyżej, jednosylabowce zachowują go tylko wówczas, gdy jednocześnie pada na nie akcent zdaniowy lub rytmiczny; w pozostałych przypadkach – gdy ich akcent gramatyczny nie otrzymuje żadnego „wsparcia”, stają

się atoniczne i przyłączają się do innych akcentowanych monosylab lub dłuższych wyrazów, formując z nimi struktury zwane zestrojami akcentowymi. Struktury te, stanowiące de facto centralny punkt teorii Dłuskiej, okazują się bardzo przydatne przy wyjaśnianiu kwestii związanych z rozmieszczeniem akcentu polskiego w ujęciu tradycyjnym, dlatego też wymagają nieco szerszego omówienia.

Termin „zestrój akcentowy” Dłuska (1976) definiuje w sposób następujący:

(...) grupy sylab podporządkowane prozodyjnie wspólnemu akcentowi (...). Przy czym będą nimi zarówno wyrazy samodzielne pojedynczo brane, jak i zespoły wyrazów samodzielnych z niesamodzielnymi, gdyż najistotniejsze cechy zestroju będą się w tych i tamtych realizowały jednakowo.

I dalej:

Zestrój akcentowy nie jest li tylko sumą należących do niego sylab. Jest on zorganizowaną postacią prozodyjną (s. 13).

Jak wyjaśnia w kolejnej części rozprawy, sylaby tworzące zestrój mogą należeć nawet do kilku wyrazów, niemniej każda tego typu jednostka mowy niezmiennie koncentruje się wokół jednego tylko akcentu głównego. Podstawowym elementem zestroju jest co najmniej jedna forma wyrazowa ortotoniczna, do której bądź których często dołączają się formy atoniczne – proklityki lub enklityki. Wyjątkowo zestroje mogą składać się jedynie z dwóch występujących obok siebie atonów, jak np. „dla_mnie”¹².

Zasadność wyodrębniania tego typu struktury wynika z automatycznej i często nieuświadomionej przez użytkownika danego języka segmentacji wypowiedzi na mniejsze objętościowo fragmenty, biorące udział w procesie przetwarzania mowy (Demenko 1999). Założenie to znajduje odzwierciedlenie w badaniach percepcyjnych przeprowadzonych przez Steffen-Batogową (2000). Rezultaty pokazują, że wszyscy bez wyjątku słuchacze, bez względu na poziom kompetencji językowej, intuicyjnie i bez większego problemu dzielili frazy fonetyczne

¹² Przykład zaczerpnięty z: Wierzchowska 1980, s. 135.

na odcinki, które, choć niejednokrotnie różniły się liczbą sylab oraz umiejscowieniem granic, odznaczały się pewnymi wspólnymi właściwościami. Otóż każdy z wydzielonych odcinków zawierał przynajmniej jedną sylabę opatrzoną akcentem głównym, natomiast wiele z nich zawierało więcej niż jeden wyraz.

Zestroj charakteryzuje zwarta struktura wewnętrzna, której ciągłości nie naruszają pauzy ani żadne inne rozgraniczenia, tworząca pewną całość spajaną akcentem i mającą swoiste właściwości iloczynowo-intonacyjne, dzięki którym jednostkę tę łatwo wyodrębnić. Poszczególne zestroje są od siebie oddzielane za pomocą pauz – wprawdzie nie zawsze realnie występujących, natomiast zawsze możliwych do zrealizowania przy zachowaniu wewnętrznej spójności jednostek (Dłuska 1976: 18–19). Wynika z tego, że stanowiące granice pomiędzy sąsiednimi zestrojami pauzy pełnią dwojaką funkcję, którą Steffen-Batogowa (2000) nazywa integracyjno-delimitatywną: integrują sylaby skoncentrowane wokół jednego wspólnego akcentu, jednocześnie wydzielając z nich samodzielne jednostki powiązane zależnościami o charakterze syntaktyczno-semantycznym, dając w efekcie sensowne całości. Jak zauważa badaczka:

Im słabsze związki znaczeniowo-składniowe łączą takie występujące obok siebie (graniczące ze sobą) składniki wypowiedzi, tym wyrazistsze są oddzielające je granice międzyczestrojowe; i na odwrót, im związki te są silniejsze, tym słabiej uwydatnione są granice pomiędzy zestrojami akcentowymi (s. 275).

Niekiedy zdarza się, że reguły poprawności zostają naruszone i następuje rozpad jednej formy wyrazowej na dwa zestroje akcentowe. W tym wypadku mowa o zestrojach rozpadowych (Dłuska 1976: 19). Ponieważ zjawisko to wymaga pewnych wyjątkowych okoliczności, występuje ono stosunkowo rzadko, np. może zostać wymuszone schematem rytmicznym wiersza.

Dłuska wskazuje również na przypadki tzw. zestrojów ściągniętych, znacznie częstszych, będących połączeniem dwóch samodzielnych wyrazów, z których jeden jest akcentowo podporządkowany akcentowanej sylabie wyrazu drugiego. Dobrym przykładem zestroju ściągniętego jest przytoczone przez badaczkę powszechne pytanie „która godzina?”,

gdzie wyraz „która” w następstwie utraty samodzielności akcentowej staje się niemal proklityką. Tworzenie się zestrojów ściągniętych uzależnione jest od pewnych czynników, mogących stworzyć warunki sprzyjające bądź niekorzystne dla zaistnienia procesów konsolidacyjnych. Czynnikiemami tymi są dobitność wypowiedzi oraz tempo mowy, przy czym tendencja do łączenia zestrojów rośnie wraz ze spadkiem wyrazistości i zwiększeniem tempa. Inaczej mówiąc, będą one częstym zjawiskiem w szybkiej mowie potocznej o pobieżnie realizowanej akcentuacji. Należy jednak zaznaczyć, że aby połączenie zestrojów prymarnych (s. 20) – jak nazywa Dłuska poszczególne elementy składowe zestrojów ściągniętych – mogło w ogóle zaistnieć, musi zostać spełniony pewien wymóg – mianowicie wyrazy stanowiące zestroje prymarne muszą pozostawać w bliskim związku semantycznym. Konieczna jest także obecność akcentu znaczeniowego w danej wypowiedzi, któremu to podporządkowany jest każdy zestwór ściągnięty, gdyż, w przeciwieństwie do akcentu gramatycznego, będącego spoiwem i zarazem centralnym punktem zestroju prymarnego, jest on elementem fakultatywnym.

Należy zwrócić uwagę, że teoria zestroju akcentowego Dłuskiej, podobnie jak wiele innych teorii zaliczanych do tradycji fonetycznej, bazuje na dychotomicznym podziale wyrazów na ortotona i atona, w ramach których wyróżnia się proklityki oraz enklityki. Klasyfikacja ta, uznawana przez znaczną większość lingwistów, niemal na trwałe wpisała się w dorobek językoznawstwa polskiego, tymczasem, jak zauważa Steffen-Batogowa (2000): „znacznie gorzej przedstawia się sprawa praktycznego rozgraniczenia tych trzech postulowanych kategorii wyrazów” (s. 153). Wprawdzie literatura dysponuje już obszerną terminologią z zakresu zjawisk enklizy i proklizy, jednak pomimo tego sklasyfikowanie niektórych wyrazów wciąż okazuje się dość kłopotliwe. Według badaczki problem tkwi w braku pełnej listy klityk prymarnych, definiowanych jako człony nieleksykalnych kategorii (do których należą m.in. niektóre spójniki, przyimki, krótkie formy zaimków, zaimek zwrotny „się”, przeczenie „nie”) zachowujące się jak klityki ze względu na brak rytmicznej samodzielności (Sawicka 1995: 171–172; za: Steffen-Batogowa 2000: 161). Jest również szereg niezgodności co do kwestii przynależności pewnych jednostek leksykalnych do poszczególnych kategorii, a także odpowiednich reguł tę przynależność warunkujących.

Steffen-Batogowa (2000), dokonawszy przeglądu postulowanych przez innych badaczy koncepcji i nie znalazłszy potwierdzenia żadnej z nich w wynikach własnych badań percepcyjnych, proponuje następującą, czteroelementową klasyfikację:

O r t o t o n a obejmują wszystkie wyrazy ponaddwusylabowe, wszystkie wyrazy dwusylabowe z wyjątkiem spójników *aby*, *albo*, *ale*, *jeśli*, *oraz*, wszystkie jednosylabowe rzeczowniki, przymiotniki, przysłówki, liczebniki, czasowniki i wykrzykniki oraz jednosylabowe zaimki przeczące i nieokreślone *nikt*, *nic*, *ktoś*, *coś*, *gdzieś*. [...] S e m i o r t o t o n a to jednosylabowe formy zaimków osobowych w mianowniku oraz niektóre dwusylabowe spójniki. [...]. S e m i a t o n a to wszystkie pozostałe wyrazy jednosylabowe z wyjątkiem spójników *i* oraz *a*, klitycznych form zaimków osobowych, zaimka zwrotnego *się* oraz partykuł *bym*, *byś*, *by*, *-bym*, *-byś*, *-by*. [...]. A t o n a obejmują klityczne formy zaimków osobowych, zaimek zwrotny *się*, partykuły *bym*, *byś*, *by*, *-bym*, *-byś*, *-by*, a także spójniki *i* oraz *a* (s. 179–180).

Jak wspomniano na początku tego podrozdziału, szczegółowe badania, na które współcześnie kładzie się nacisk, niejednokrotnie prowadzą do weryfikacji tradycyjnie uznawanych poglądów. Dzieje się tak m.in. w przypadku pojęcia zestroju akcentowego, które w wyniku obszernych testów percepcyjnych swojego autorstwa redefiniuje Steffen-Batogowa (2000). Według niej:

Zestrój akcentowy jest to odcinek frazy fonetycznej złożony z pewnej liczby następujących bezpośrednio po sobie sylab, wśród których jedna jest wyróżniona akcentem głównym, a pozostałe są prozodyjnie podporządkowane tej wyróżnionej (s. 207).

Zgodnie z tą definicją pojedynczy zestrój może obejmować zarówno jedną, jak i kilkanaście sylab, a zatem jeden, kilka lub też jedynie część wyrazu. Nie ma ponadto żadnych teoretycznych ograniczeń w kwestii przynależności wyrazów, które zawierają ośrodki zestrojów w postaci sylab akcentowanych, do opisanych powyżej kategorii. Mówiąc ina-

czej, sylaba stojąca pod akcentem może być częścią wyrazu należącego do każdej z nich (Steffen-Batogowa 2000: 207).

W języku polskim (podobnie jak np. w języku angielskim) obok akcentu głównego występuje także akcent poboczny, właściwy dłuższym formom wyrazowym. Akcent ten sygnalizuje głównie treści o charakterze ekspresywnym, a także – z uwagi na swoje regularne rozmieszczenie – wyznacza charakterystyczny dla polszczyzny rytm, dzięki czemu bywa niekiedy nazywany akcentem rytmicznym (Dłuska 1976). Ma on drugorzędne znaczenie dla procesu komunikacji, gdyż wszelkie zakłócenia w jego odbiorze prowadzą jedynie do zubożenia warstwy estetycznej danej wypowiedzi (Steffen-Batogowa 2000). Jak sugeruje jego nazwa, akcent poboczny jest de facto elementem dodatkowym, a zatem nieobligatoryjnym, który – w przeciwieństwie do akcentu głównego – może, ale nie musi w danym zestroju wystąpić. Cecha ta, określana jako fakultatywność, jest silnie skorelowana z tempem mowy, gdyż im ono większe, tym mniejszą liczbę akcentów pobocznych zawiera wypowiedź. Akcent ten znacznie różni się w swej naturze od akcentu głównego, co jest widoczne przede wszystkim w trzech aspektach: sile, wpływie i relacji z otoczeniem oraz umiejscowieniu w wyrazie lub zestroju. Jako że akcent poboczny odgrywa niejako rolę pomocniczą, jest on słabszy i mniej wyrazisty od akcentu głównego, a co za tym idzie – zarazem mniej oczywisty. Dzięki swoim odmiennym właściwościom oraz obszarowi oddziaływania nie narusza on wewnętrznej spójności zestroju, do którego należy nawet wówczas, gdy przypada kilkakrotnie. Jak wyjaśnia Dłuska (Steffen-Batogowa 2000), dzieje się tak, ponieważ:

(...) akcent poboczny wzmacnia spółgłoskę, mniej samogłoskę, na którą pada, prawie nie wpływa na iloczas i wysokość tej ostatniej, a przede wszystkim w niczym nie osłabia charakterystycznych cech samogłoski stojącej pod akcentem głównym i nie zmienia przodowego stosunku do niej samogłosek w innych sylabach zestroju (s. 13–14).

Lokalizacja akcentu pobocznego jest kwestią nieco mniej jednoznaczną, głównie za sprawą różnic w nomenklaturze, jaką przyjmują badacze zajmujący się tym zagadnieniem. Dłuska określa akcent poboczny jako zestrojowy, w opozycji do głównego, który jest notabene

akcentem wyrazowym. Pomimo odmiennej dziedziny akcent poboczny również charakteryzuje ustabilizowana pozycja – jak twierdzi badaczka, pada on na pierwszą sylabę zestroju, jeżeli sylaba ta nie stoi bezpośrednio przed sylabą akcentowaną. Natomiast w zestrojach dłuższych, w których występuje więcej akcentów pobocznych, kolejno przypadają one regularnie na co drugą sylabę.

Mała wyrazistość polskiego akcentu pobocznego znajduje potwierdzenie w rezultatach testów percepcyjnych, a ściślej – w bardzo małej zgodności słuchaczy co do jego występowania oraz lokalizacji. Ponadto, wielu słuchaczy wykazuje trudności w odróżnianiu akcentu pobocznego od głównego. Podobnie rzecz ma się z percepcją akcentu głównego, co może być nieco zaskakujące z uwagi na fakt, że kwestia ta w teorii wydaje się dość oczywista. Okazuje się jednak, że w znacznej liczbie przypadków identyfikacja sylab naznaczonych akcentem głównym również naraża słuchaczy na pewne trudności. Jak podaje Steffen-Batogowa (2000), przyczyny tego stanu rzeczy mogą tkwić zarówno w czynnikach zależnych od odbiorcy takich jak niejednakowa wrażliwość na owe bodźce, a także wpływ własnej kompetencji językowej na ocenę zjawisk akcentowych, jak i zróżnicowanym stopniu natężenia towarzyszących akcentowi bodźców akustycznych, mogącym wynikać z akcentowego wzmocnienia sylab nieakcentogennych, które ma na celu emfazę wyrazu o kluczowym dla wypowiedzi znaczeniu (s. 181) bądź dodatkowych czynnikach zewnętrznych kształtujących struktury akcentowe – wśród nich pod uwagę należy wziąć: tempo mowy, specyfikę głosu mówcy, obecność w wypowiedzi silnego akcentu logicznego, odmianę stylistyczną języka polskiego oraz kompetencje odbiorcy związane z głosową interpretacją wypowiedzi realizowanych w różnych odmianach językowych (s. 323).

Warto zwrócić uwagę na pewne interesujące spostrzeżenie, zanotowane przez Wierzchowską (ok. 1970) w kwestii zjawisk akcentowych będących wynikiem zmian, jakie zaszły w polszczyźnie na przestrzeni kilku dekad. Zauważa ona, iż w niemalże wszystkich polisylabowych formach wyrazowych występuje tendencja do silniejszego akcentowania sylaby pierwszej, natomiast nie – jak mówi teoria – sylaby przedostatniej, na której zlokalizowany jest akcent główny. I choć spostrzeżenie to zostało poczynione na podstawie obserwacji języka z lat 60. i 70., należy zaznaczyć, że nawet wówczas nie było to zupełnie novum, gdyż

znacznie wcześniej, bo już w 1962 r., o zjawisku tym wspomina Jassem (s. 13). W ciągu kolejnych lat tendencja ta ugruntowała się w języku polskim, bowiem w drugiej połowie lat 70. była już na tyle silna i wyrazista, że Wierzchowska mogła podsumować następująco:

Można twierdzić, że dziś akcent tzw. główny spoczywa na początku wyrazu, akcent poboczny – na sylabie przedostatniej (zwykle nieco wzdluzanej). O takim sposobie akcentowania zdecydowały względy rytmiczne i melodyczne (1980 s. 134).

Trudno wyrokować, czy istotnie zaszło tu odwrócenie kolejności umiejscowienia akcentów, niemniej faktem jest, że zjawisko to obserwuje się także współcześnie, gdzie znane jest jako przesunięcie akcentu. Steffen-Batogowa (2000) wspomina o trzech sposobach wzmocnienia sylab nieakcentowanych:

Pierwszy z nich polega na akcentowym wzmocnieniu sylaby nacechowanej akcentem fonologicznym i silniejszym podporządkowaniu jej sylab ją otaczających, drugi – na przeniesieniu głównego akcentu wyrazowego na zgłoskę inicjalną, wreszcie trzeci na zachowaniu uwydatnienia akcentowego sylaby nacechowanej akcentem fonologicznym przy równoczesnym wzmocnieniu akcentowym innych sylab w wyrazie. Zgłoską najbardziej podatną na takie dodatkowe wzmocnienie jest w tym wypadku sylaba inicjalna (s. 181).

Zjawisko to, po części odpowiedzialne za rozbieżności pomiędzy teorią i praktyką – w szczególności percepcją – przyczynia się do tworzenia obrazu akcentu polskiego jako niezwykle trudnego do opisania sztywnymi regułami, gdzie kolejne wyjątki powodują jedynie kolejne wątpliwości.

Połowa lat 90. przyniosła świeże spojrzenie na kwestię akcentu, co w sposób oczywisty wywarło wpływ także na językoznawstwo polskie; wówczas to pojawiła się koncepcja Ladda (1996), a wraz z nią pojęcie akcentu potencjalnego i realnego. Odrzucenie „klasycznego” podziału na akcent zdaniowy i wyrazowy na rzecz nowych kategorii, co wiąże się z ostatecznym odstąpieniem od tradycji fonetycznej, można zaobserwować w podejmujących tematykę akcentu publikacjach po-

chodzących z okresu przełomu wieków i pierwszej dekady XXI wieku. W ujęciu tym akcent potencjalny, rozumiany jako zjawisko natury leksykalno-morfologicznej, zlokalizowany jest na leksemach, a zatem stanowi on kategorię abstrakcyjną w odróżnieniu od akcentu realnego, który spoczywa na odpowiednich sylabach fraz intonacyjnych, będących jego dziedziną. Z uwagi na fakt, że nie reprezentują one kategorii równorzędnych, istotne jest właściwe uchwycenie zależności łączącej oba te rodzaje akcentu. Demenko (1999) wyjaśnia:

Akcent potencjalny przypada mianowicie na tę sylabę w obrębie leksemu i jego alternantów fleksyjnych, która w realizacji frazy intonacyjnej niesie akcent realny (s. 54).

Można zatem przyrównać akcent potencjalny do swego rodzaju „szablonu”, którego faktycznym odwzorowaniem jest akcent realny. Nie zawsze jednak realizacja przebiega dokładnie według wzorca, bowiem zdarza się, że

[...] akcent frazowy (realny) bierze górę nad akcentem potencjalnym, pojawiając się w miejscach nietypowych, różnych od miejsc akcentu potencjalnego, z którymi w standardowych realizacjach tych fraz powinien był się pokryć (Karpiński 2006: 61).

Sytuacje takie mogą być efektem celowego wprowadzenia przez mówcę emfazy na sylabach z definicji niebędących nośnikami akcentu bądź też odzwierciedlać ładunek emocjonalny danej wypowiedzi (Karpiński 2006). Niemniej bez względu na przyczynę tego typu przypadki stanowią podstawę dla uznawanego w literaturze prymatu akcentu realnego nad potencjalnym.

Choć obecny stan badań nad zjawiskiem akcentu w języku polskim jest – m.in. dzięki wsparciu nowoczesnej technologii – niewątpliwie mocno zaawansowany, mimo wszystko trudno ocenić, na ile płynąca z nich wiedza jest w sposób satysfakcjonujący i przede wszystkim adekwatny wyjaśnić kwestie nurtujące współczesnych językoznawców i prozodystów. Analizując wypracowane dotychczas teorie, można odnieść wrażenie, że im bardziej rozbudowany aparat pojęciowy, jakim dysponuje językoznawstwo, tym więcej pojawia się

terminów do zdefiniowania. Kiedy przedmiotem badań jest język – podlegający nieustannym zmianom żywy twór – wrażenie takie wydaje się jednak uzasadnione.

Rozdział 2.

Zjawisko tonalnego *sandhi*

Termin *sandhi* (z sanskr. łącznie), pierwotnie oznaczający „dostosowanie” (*accommodation*) lub „upodobnienie” (*assimilation*), z czasem zaczęto stosować w kontekście stricte lingwistycznym, na określenie grupy zjawisk natury fonologicznej, zachodzących na granicach wyrazów bądź najmniejszych jednostek znaczeniowych – morfemów. Jego rodowód sięga czasów starożytnych, a pierwsze wzmianki na ten temat pojawiają się najprawdopodobniej w ówczesnej gramatyce sanskrytu autorstwa Paniniego. Niektórzy badacze wyróżniają natomiast dwa rodzaje *sandhi* na podstawie położenia zmian, będących wynikiem tego zjawiska: *sandhi* wewnętrzne, zachodzące w obrębie poszczególnych wyrazów, a zatem niewykraczających poza ich granice oraz *sandhi* zewnętrzne, czyli modyfikacje obejmujące końcowy segment lub segmenty jednego wyrazu i segment(y) początkowe wyrazu po nim następującego. Współcześnie terminu *sandhi* używa się w odniesieniu do zjawisk, których domeną są połączenia wyrazowe; przejął on więc znaczenie *sandhi* zewnętrznego (MacKay 1987: 144–145).

Z kolei w językach tonalnych zaobserwować można zjawisko *sandhi* tonalnego, znane pod potocznym określeniem „następstwa tonów”. Dotyczy ono wyłącznie wypowiedzi składających się z co najmniej dwóch sylab (*connected speech*), których wyjściowe tony, będące wynikiem działających na nie procesów, charakteryzują kontury intonacyjne zniekształcone w stosunku do modelowych konturów tonów produkowanych w izolacji (Yuan i Chen 2014). Mówiąc prościej, zjawisko to polega na zmianie tonu danej sylaby pod wpływem odpowiedniego tonu sylaby po niej następującej. Za dobry przykład powszechnego występowania procesów *sandhi* mogą posłużyć różne dialekty języka chińskiego, którym w okresie ostatnich dwudziestu lat poświęcono pod tym kątem sporo uwagi. Rezultatem licznych badań, jakie w tym czasie powstały, jest mnóstwo interesujących prac, zarówno o charakterze deskryptywnym, jak i teoretycznym, które znacznie przyczyniły się do zrozumienia zjawiska *sandhi* w ogóle (Yuan i Chen 2014).

2.1. *Sandhi* tonalne w języku chińskim (mandaryńskim)

Choć zjawisko *sandhi* w standardowej odmianie języka chińskiego dotyczyć może także poszczególnych wyrazów, jak ma to miejsce w przypadku 一 *yī* (*jeden*) i przeczenia 不 *bù*, które rządzą się osobnymi zasadami, to jednak znaczna większość reguł odnosi się do sylab jako takich, których tony modyfikowane są w odpowiedni sposób, w zależności od tonu sylaby występującej w ich bezpośrednim sąsiedztwie, abstrahując od ich znaczenia leksykalnego.

Obok pomniejszych i znacznie rzadziej występujących zmian tonów 2 i 4, najciekawszym przypadkiem, i zarazem niosącym ze sobą najbardziej rozbudowane implikacje, jest *sandhi* tonu 3, którym niniejsza praca zajmować się będzie w sposób nieco bardziej szczegółowy. Nie dziwi więc fakt, że termin ten, używany w odniesieniu do chińskiego, w potocznym rozumieniu oznacza głównie procesy, jakim podlega ton 3.

Zasady tonalnego *sandhi* w chińskim mandaryńskim jako pierwszy formułuje Chao (1933, za: Duanmu 2000). Przedstawiają się one w sposób następujący:

- 1) Ton 3 zmienia się na 2, gdy występuje po nim inny ton 3.
- 2) Ton 3 przed tonem 1, 2 lub 4 zostaje zredukowany o część rosnącą, dając w rezultacie ton $\frac{1}{2}$ 3.
- 3) Opadająca część tonu 4, po którym następuje kolejny ton 4, nie osiąga maksymalnie niskiego rejestru.
- 4) Pewne jednostki leksykalne, takie jak 一 *yī* i 不 *bù*, podlegają własnym, odrębnym zasadom (Duanmu 2000: 226–228):
 - a) 一 *yī*, wymawiane w izolacji na tonie 1, zmienia ton na:
 - 2, gdy poprzedza sylabę z tonem 4,
 - 4, jeżeli następuje po niej każdy inny ton niż 4, a zatem 1, 2 lub 3,
 - ton zerowy (*neutral tone*), gdy występuje pomiędzy dwoma identycznymi czasownikami, będącymi efektem reduplikacji i tworzącymi tzw. aspekt tymczasowy (*tentative aspect*), np. 想一想 *xiǎngyixiǎng* (*przemyśleć*),
 - ton 4, gdy występuje przed 千 *qiān* (*tysiąc*) lub 百 *bǎi* (*sto*) i jest częścią liczebnika głównego; natomiast jeżeli tworzy liczebnik porządkowy, zachowuje ton 1.
 - b) 不 *bù*, wymawiane w izolacji na tonie 4, zmienia ton na:

- 2, jeśli występuje przed tonem 4,
- zerowy, gdy występuje pomiędzy dwoma wyrazami, z którymi tworzy pytanie ogólne (*yes-no question*).

2.1.1. *Sandhi* tonu 3

Ton 3 w chińskim mandaryńskim, ze względu na pewne niewątpliwe właściwości nieco odróżniające go od pozostałych, wzbudza największe zainteresowanie wśród badaczy. Chao (1970) zauważa, że ton 3, określany przy użyciu jego notacji za pomocą trzech poziomów tonicznych – (214) – jako ton niski lub średni opadająco-wznoszący, jest tonem bardziej złożonym niż pozostałe, dla których określenia wystarczają maksymalnie dwa poziomy toniczne (s. 25–26).

Specyfika tonu 3 oraz jego zachowanie pod wpływem procesów *sandhi* generowało liczne kontrowersje w środowisku językoznawczym na przestrzeni ostatnich czterech dekad, co wyraźnie zaznacza się w dotychczasowych badaniach nad tą tematyką, przynosząc w efekcie wiele różnych teorii. Kontrowersje te dotyczą głównie dwóch aspektów, w ramach których zjawisko to jest rozpatrywane.

Pierwszy z nich koncentruje się na kwestii podstawowej domeny *sandhi* oraz, co za tym idzie, procesach biorących udział w jej formowaniu i zasadach, na podstawie których można tę domenę wyodrębnić. Problem zdefiniowania charakteru domeny *sandhi* zasadzał się na próbie znalezienia odpowiedzi na pytanie czy *sandhi* tonu 3 ma związek ze strukturą syntaktyczną, strukturą prozodyczną czy też kombinacją ich obu (Yuan i Chen 2014). Każda z koncepcji ma zarówno swoich zwolenników, jak i przeciwników, spośród których wymienić należy nazwiska takie, jak: Cheng, Kaisse, Shih, Zhang, Jin czy też Hsiao.

Teoriom wspomnianych badaczy przyjrzymy się bliżej w kolejnym podrozdziale.

Intensywne badania prowadzone w tym kierunku od początku lat 90. wypracowały konsensus do dziś obecny w literaturze, który za niewątpliwą domenę *sandhi* uznaje wyrazy dwusylabowe złożone z następujących po sobie tonów 3, gdzie pierwszy z nich w rezultacie modyfikacji daje pewien ton rosnący; kwestią sporną natomiast wciąż są struktury o trzech i większej liczbie sylab oraz złożenia wyrazów jednosylabowych. Badacze zgadzają się obecnie także co do faktu,

że na przebieg procesów *sandhi* tonalnego wpływ mają rozmaite czynniki takie jak: struktura syntaktyczna, rozkład informacji semantycznej, prozodia, a także tempo mowy (Yuan i Chen 2014).

Aspekt drugi zaś dotyczy fonetycznej natury tonu powstałego w efekcie zjawiska *sandhi* w relacji do leksykalnego tonu 2 – rosnącego. Tu opinie lingwistów przedstawiają się głównie dwojako. Stanowisko pierwsze zakłada zamianę pierwszego z dwóch występujących obok siebie tonów 3 na ton 2 w jego „czystej” formie, a zatem leksykalny ton 2 (*lexical rising tone*). Pogląd taki zaprezentował pionier badań nad *sandhi* tonu 3, Chao (1948, za: Yuan i Chen 2014), w pierwszej znanej pracy o tej tematyce, opublikowanej pod koniec lat 40. W opozycji do Chao stoi szereg innych badaczy, z których jako pierwsi pogląd ten zakwestionowali Hockett i Martin (odpowiednio 1947 i 1957, za: Yuan i Chen 2014) w pracach wydanych w tamtym okresie. Obaj oni twierdzili, że wyjściowa forma tonu poddanego zjawisku *sandhi* tworzy nową, osobną kategorię; w efekcie otrzymujemy ton podobny, który jednak nie może być całkowicie utożsamiany z „kanonicznym” tonem 2. Stanowisko to, będące przeciw wagą dla teorii Chao, przybierało następnie różne formy w twierdzeniach innych lingwistów zajmujących się badaniem natury tonu powstałego na skutek *sandhi*. Dociekania ostatnich lat koncentrowały się na próbie znalezienia odpowiedzi na pytanie, czy można postawić znak równości pomiędzy dwoma wspomnianymi tonami, a jeśli nie są one identyczne, wówczas jakie są akustyczne parametry, które te tony odróżniają. Spośród grona lingwistów zajmujących się tym zagadnieniem na uwagę zasługują: Zee (1980, za: Yuan i Chen 2014), autor pierwszych tego typu badań laboratoryjnych, Peng (2000, za: Yuan i Chen 2014) oraz badania Speer i Xu (2008), Yuan i Chen (2014), Lee-Schoenfeld i Kandybowicz (2008), a także nowatorskie studium Zhang i Lai (2010), badające odporność logatomów¹³ („*wug*” words) na działanie procesów *sandhi* (Yuan i Chen 2014). Badania te oraz wynikające z nich spostrzeżenia zostaną szerzej omówione w podrozdziale 2.3.

¹³ Poprawnie zbudowane sylaby, połączone z sobą zgodnie z systemem fonologicznym danego języka w nieistniejący w nim i niemający żadnego znaczenia wyraz.

2.2. Domena *sandhi* tonu 3 w badaniach lingwistycznych

Wśród wymienionych wcześniej badaczy domeny *sandhi* niektórzy opowiadali się za jej silnym związkiem z syntaktyczną strukturą jednostek leksykalnych. Zarówno Cheng (1973, za: L. Cheng 1987), jak i Kaisse (1985, za: L.Cheng 1987) proponują hipotezę bezpośredniego wpływu składni na strukturę fonologiczną oraz wskazują na cykliczny charakter występowania zjawiska *sandhi* w złożeniach sylab z tonem 3.

2.2.1. Podejście cykliczne (*The Cyclic Approach*)

Podejście to zakłada działanie reguł *sandhi* w obrębie frazy w kilku cyklach, których liczba zależy od stopnia jej złożoności. Zgodnie z tym cykl pierwszy ma miejsce w obrębie najmniejszych pod względem syntaktycznym fraz, kolejne cykle natomiast obejmują swoim zasięgiem coraz to większe struktury, aż do momentu przekształcenia wszystkich odpowiednich elementów frazy wyjściowej. Innymi słowy, zjawisko *sandhi* występuje tu w granicach struktur wyodrębnionych w oparciu o zasady składni, począwszy od tych najgłębiej zakorzenionych, konsekwentnie do coraz bardziej rozbudowanych całości (Cheng 1973, za: L.Cheng 1987).

Choć podejście cykliczne zdaje się mieć zastosowanie w przypadku znacznej liczby zróżnicowanej budowy struktur syntaktycznych, Cheng wskazuje na przypadki problematyczne, w których wygenerowane w ostatnim cyklu tony wykazują niezgodność z danymi empirycznymi. Dla ilustracji posługuje się dwoma przykładami:

- 1) *wo xiang yang xiao ma*
ja chceć hodować mały koń
3 3 3 3 3 – tony początkowe
2 2 3 2 3 – tony wyjściowe
„(Ja) chcę (wy)hodować kucyka”.

Zgodnie z podejściem cyklicznym tony w kolejnych cyklach przedstawiają się następująco:

cykl 1: 3 3 3 [2 3]¹⁴
 cykl 2: 3 3 [3 ---]
 cykl 3: 3 [2 ----]
 cykl 4: [3 -----]
 tony wyjściowe : * 3 2 3 2 3

Podobnie w kolejnym, nieco bardziej skomplikowanym przykładzie rezultat wyjściowy zawiera błędne tony.

2) *wo zhun ni yang xiao ma*
 ja pozwalać ty hodować mały koń
 3 3 3 3 3 3 – tony początkowe
 3 2 2 3 2 3 – tony wyjściowe
 „(Ja) pozwalam ci (wy)hodować kuczka”.

Zasada Chenga zakłada poniższy rezultat:

cykl 1: 3 3 3 3 [2 3]
 cykl 2: 3 3 3 [3 ---]
 cykl 3: 3 [3 2 ----]
 cykl 4: [2 -----]
 tony wyjściowe : * 2 3 2 3 2 3

Przykłady te pokazują, że podejście cykliczne nie jest w stanie w sposób adekwatny i satysfakcjonujący wyjaśnić działania zasady *sandhi* tonu 3 w chińskim mandaryńskim.

2.2.2. Podejście Kaisse – Zasada gałęzi (*The Branch Condition*)

Jak twierdzi Kaisse, *sandhi* tonalne może zajść pomiędzy wyrazami A i B wówczas, gdy A jest lewą gałęzią¹⁵ struktury zawierającej B lub B jest prawą gałęzią struktury zawierającej A – mówiąc inaczej, zjawisko to ma miejsce, jeśli para wyrazów, będących jego domeną znajduje się

¹⁴ Nawias [] oznacza frazę stanowiącą domenę *sandhi* w danym cyklu.

¹⁵ W swojej analizie zjawiska *sandhi* Kaisse posługuje się schematami drzew syntaktycznych.

na początku lub na końcu większej struktury, której jest częścią (Kaisse 1985: 175, za: L.Cheng 1987). Zgodnie z tą zasadą *sandhi* nie jest możliwe pomiędzy występującymi obok siebie wyrazami wtedy, gdy każdy z nich jest bezpośrednim składnikiem innej struktury.

Podobnie jak poprzednia koncepcja, podejście Kaisse zakłada cykliczność *sandhi* tonu 3 w obrębie struktur składniowych, zatem ma ono początek w ich najgłębiej osadzonych elementach mogących stanowić domenę, obejmując kolejno struktury wraz ze wzrostem poziomu ich skomplikowania. Cheng przytacza następujący przykład frazy, gdzie hipoteza Kaisse daje poprawny wynik:

- 1) [*shui* – *guo*] *jiu*
 ‘[woda – owoc] wino = wino owocowe’
- | | | | |
|---|---|---|-------------------|
| 3 | 3 | 3 | – tony początkowe |
| 2 | 3 | 3 | – cykl 1 |
| 2 | 2 | 3 | – cykl 2 |
| 2 | 2 | 3 | – tony wyjściowe |

Jednakże hipoteza ta ma pewne braki, co jako pierwsza zauważyła Shih, podając przykłady ilustrujące nieprawidłowe rezultaty jej zastosowania. Poniżej przykład, gdzie zgodnie z zasadą gałęzi *sandhi* nie powinno wystąpić (Shih 1985, za: L.Cheng 1987):

- 2)
- | | /\ | | /\ | |
|------------|--------------|------------|------------|--|
| NUM | CL | ADJ | N | |
| <i>san</i> | <i>zhong</i> | <i>lan</i> | <i>ren</i> | |
| trzy | rodzaj | leniwy | ludzie | |
| 1 | 3 | 3 | 2 | – tony początkowe |
| 1 | 2 | 3 | 2 | – tony wyjściowe |
| *1 | 3 | 3 | 2 | – tony wygenerowane przez model Kaisse |

‘Trzy rodzaje leniwych ludzi’

Shih eksponuje ponadto słabości koncepcji Kaisse w odniesieniu do fraz zawierających w swojej strukturze dopełnienie bliższe i dalsze. Kolejny przykład, jakim posługuje się badaczka, pokazuje natomiast, że w przypadku, gdy możliwa jest więcej niż jedna kombinacja tonów

wyjściowych, model Kaisse jest w stanie wygenerować wyłącznie jedną z nich.

3)	<i>ni</i>	<i>bi</i>	<i>gou</i>	<i>xiao</i>	
	ty	porównywać	pies	mały	
	3	3	3	3	– tony początkowe
a.	3	2	2	3	– tony wyjściowe
b.	2	2	2	3	– tony wyjściowe
c.	2	3	2	3	– tony wyjściowe

„(Ty) jesteś mniejszy niż pies”.

Zgodnie z modelem Kaisse:

<i>ni</i>	<i>bi</i>	<i>gou</i>	<i>xiao</i>	
ty	porównywać	pies	mały	
3	3	3	3	– tony początkowe
3	[2	3]	3	– cykl 1
3	[2	2	3]	– cykl 2
-----				– cykl 3
3	2	2	3	– tony wyjściowe

Z przykładów wynika, że czysto syntaktyczne podejścia do zjawiska *sandhi* tonalnego nie umożliwiają jakiegokolwiek wariacyjności ostatecznej kombinacji tonów. Odkrycie to okazało się inspiracją dla Shih (1985, za: L.Cheng 1987), która obok składni uwzględniła także wpływ struktury prozodycznej na procesy *sandhi*.

2.2.3. Podejście prozodyczne (*The Prosodic Approach*) – model Shih

Shih (1985, za: L.Cheng 1987), czerpiąc z krytyki systemów syntaktycznych, konstruuje model, który, choć nie wyklucza zupełnie wpływu składni, zdecydowany prymat w procesie formowania się domeny *sandhi* tonu 3 w standardowej odmianie chińskiego daje strukturze prozodycznej. Według badaczki *sandhi* tonalne zachodzi w obrębie struktur prozodycznych, które są wrażliwe na struktury syntaktyczne, lecz bynajmniej nie w stosunku do nich izomorficzne

(1985: 107, za L. Cheng 1987). W celu poprawnego zlokalizowania domeny *sandhi* posługuje się Regułą Tworzenia Stopy (FFR) (*Foot Formation Rule*).

Reguła Tworzenia Stopy (FFR) (Shih 1986: 110, za Zhang 1997):

I. Tworzenie Stopy (f) (*Foot Construction*)

- a. Zasada Bezpośredniego Składnika¹⁶ (IC) (*Immediate Constituency*): Połącz bezpośrednie składniki w dwusylabowe stopy.
- b. Zasada Miary Dwudzielnej (DM) (*Duple Meter*): Poruszając się od lewej do prawej, połącz pojedyncze sylaby w stopy binarne, o ile nie łączą się one z innymi elementami w kierunku przeciwnym.

II. Tworzenie Superstopy (f^{*}) (*Super-foot Construction*)

Złącz pozostałe pojedyncze sylaby z sąsiednimi stopami binarnymi zgodnie z kierunkiem, w jakim elementy te łączą się w struktury wyższego rzędu.

Przy stosowaniu reguły FFR ważne jest zachowanie odpowiedniej kolejności kroków, tj. tworząc stopę, najpierw należy połączyć składniki bezpośrednie, a następnie postępować zgodnie z zasadą miary dwudzielnej. Jak wyjaśnia autorka teorii, jeśli sąsiedni składnik jest już elementem dwusylabowym, dołączenie do niego dodatkowej sylaby spowoduje powstanie trzy- lub czterosylabowego tworu zwanego superstopą, który jest niczym innym, jak stopą prozodyczną o większych rozmiarach (Zhang 1997: 298). Można zatem wywnioskować, że, podobnie jak zakładały obie wspomniane wcześniej koncepcje, *sandhi* tonalne zachodzi cyklicznie, począwszy od stopy, następnie na poziomie superstopy, by ostatecznie objąć swym zasięgiem strukturę o większym stopniu skomplikowania – całą frazę (p, *phrase*). Działanie reguły FFR ilustruje następujący przykład, częściowo zaczerpnięty z L. Cheng (1987):

¹⁶ Bezpośrednie Składniki rozpatrujemy względem struktur bardziej złożonych; o elemencie można powiedzieć, że jest bezpośrednim składnikiem danej struktury wyższego rzędu (co jest zilustrowane za pomocą Struktury Bezpośrednich Składników (*Immediate Constituent Structure*)), jeśli jest on częścią tej struktury oraz zajmuje pozycję o jeden poziom niższą (Katamba 1996).

- 1) *wo qing Lao-li mai hao jiu*
 ja zapraszać Lao-li kupować dobry wino
 3 3 3 3 3 3 – tony początkowe
 2 3 2 2 3 2 3 – tony wyjściowe
 f [---] [----] [----] [2 3] [2 3] [3] [2 3]
 f' [-----] [3 2 3]
 p [-----] [2 3 2 2 3 2 3]
 „(Ja) proszę Lao-li, aby kupić(a) dobre wino”.

Choć, jak stwierdza L. Cheng, model Shih jest w stanie poprawnie wskazać miejsce występowania *sandhi* tonu 3 w wielu strukturach w mandaryńskim dialekcie języka chińskiego, jego słabym punktem są frazy przyimkowe. Lin (2005) ilustruje problem, zestawiając dwie frazy o identycznym rozkładzie składników bezpośrednich oraz tej samej kombinacji tonów początkowych, z których tylko jedna (3) jest frazą przyimkową. Jak przewiduje Reguła Tworzenia Stopy, domeny *sandhi* obu fraz także powinny być identyczne; jak się jednak okazuje, kombinacje tonów wyjściowych w obu przypadkach są różne.

- 2) *mao da san zou*
 kot uderzać parasol chodzić
 1 3 3 3 – tony początkowe
 1 2 2 3 – tony wyjściowe
 [-----] IC
 [-----] f'
 „Kot spaceruje z parasolem”.

- 3) *mao bi gou xiao*
 kot porównywać pies mały
 1 3 3 3 – tony początkowe
 1 3 2 3 – tony wyjściowe
 [-----] IC
 [-----] f'
 „Kot jest mniejszy od psa”.

Ponieważ pomimo identycznej struktury syntaktycznej powyższych przykładów dystrybucja *sandhi* przedstawia się odmiennie, oczywiste jest, że różnice te nie mogą wynikać ze składni.

L. Cheng (1987) dowodzi również, że w sytuacjach, gdzie możliwa jest więcej niż jedna kombinacja tonów wyjściowych, zastosowanie reguły FFR pozwala wygenerować co najwyżej jedną poprawną wersję. Autorka koncepcji wskazuje tu na dodatkowe czynniki, mogące modyfikować końcowy rezultat, takie jak choćby tempo mowy. Przy tempie normalnym – wyjaśnia – zjawisko *sandhi* ma miejsce na etapie stopy; w tempie szybszym natomiast startuje z poziomu superstopy; z kolei w przypadku tempa najszybszego strukturą wyjściową jest fraza. Analiza L. Cheng pokazuje jednak, że nawet z uwzględnieniem zróżnicowanego tempa mowy model Shih daje jedynie dwie możliwe kombinacje. W odpowiedzi Shih sugeruje kolejny czynnik, którego wpływ może być istotny dla liczby poprawnie generowanych wersji, mianowicie część mowy, do jakiej należy podmiot lub dopełnienie danego zdania; mówiąc ściśle, znaczenie ma to, czy dany składnik jest rzeczownikiem, czy też zaimkiem.

Jako rozwiązanie w przypadku struktur zawierających pełną frazę rzeczownikową badaczka proponuje pominięcie zastosowania zasady Bezpośredniego Składnika i rozbicie teź czterosylabowej struktury na dwie stopy prozodyczne zgodnie z zasadą Miary Dwudzielnej, co nie tylko nie sprawdza się w kolejnych analizowanych przez L. Cheng zdaniach, lecz także dowodzi pewnej niekonsekwencji w stosowaniu własnych reguł. Próby wyjaśnienia działania procesów *sandhi* w strukturach o liczbie sylab większej niż cztery Shih w ogóle nie podejmuje.

Należy zaznaczyć również, że różnice pomiędzy zaimkami i pełnymi frazami rzeczownikowymi nie występują jedynie we frazach przyimkowych. Zestawiając dwa zdania:

- 4) (Ja) poprosiłem Xiao-ming, aby uderzył(a) szczura (*Wǒ zhǎo Xiǎo-ming dǎ lǎo-shǔ*)

oraz

5) Xiao-li poprosił(a) Xiao-ming, aby uderzył(a) szczura (*Xiǎo-li zhǎo Xiǎo-ming dǎ lǎo-shǔ*)

o identycznej budowie składniowej, z których żadne nie jest frazą przyimkową, L. Cheng (1987) zwraca uwagę na odmienny sposób utworzenia się domen *sandhi*. W zdaniu 4) przyimek „wo” (*ja*), pełniący rolę podmiotu (w języku polskim podmiotu domyślnego) połączył się z czasownikiem „zhao” (*znajdować*), natomiast w zdaniu 5), gdzie podmiotem jest rzeczownik – nazwa własna „Xiaoli”, czasownik został złączony z dopełnieniem „Xiaoming”. Wynika stąd, że wpływ różnic związanych z rzeczownikiem i zaimkiem wykracza swym zasięgiem poza wąską kategorię fraz przyimkowych, dlatego też rozróżnienie to powinno przybrać formę nieco bardziej ogólnego czynnika warunkującego tworzenie się struktur prozodycznych.

W swojej analizie L. Cheng (1987) wskazuje na jeszcze jeden typ struktur, w przypadku których wyjaśnienie działania procesów *sandhi* stanowi problem dla modelu Shih w jego pierwotnym kształcie. Są to frazy o identycznej strukturze składniowej, zbudowane z tych samych segmentów leksykalnych, które, w zależności od kontekstu, mają różne znaczenie semantyczne. Efektem tego są odmienne konfiguracje tonalne. Dla ilustracji L. Cheng posługuje się poniższym przykładem:

- 6) *wo xiang* [*qi* – *lai* *le*]
 ja myśleć w-górę – przychodzić
 3 3 3 2 – tony początkowe
 2 2 3 2 – tony wyjściowe
 „(Ja) chcę wstać”.

wo [*xiang* – [*qi* – *lai*]] *le*
 ja myśleć w-górę – przychodzić
 3 3 3 3 – tony początkowe
 3 2 3 2 – tony wyjściowe
 „(Ja) przypomniałem/-am sobie”.

Ponieważ przyczyna różnic leży w wewnętrznej strukturze segmentów leksykalnych, w celu wyjaśnienia problemu konieczne jest odwołanie się do zasady spójności leksykalnej (*lexical integrity prin-*

cipe). Głosi ona, że wszelkie procesy zachodzące na poziomie wewnętrznych struktur leksykalnych nie podlegają żadnym regułom, przez co informacja leksykalna nie może zostać zniszczona ani zmieniona w wyniku jakichkolwiek procesów postleksykalnych (Shih 1985: 136, za L.Cheng 1987). Jako że Reguła Tworzenia Stopy może ingerować jedynie w strukturę syntaktyczną, w przypadku, gdy w zdaniu pojawia się kategoria leksykalna taka jak np. złożenie (np. *xiang qi lai*), jej zastosowanie nie gwarantuje prawidłowych rezultatów.

Pomimo znacznych niedociągnięć koncepcja Shih jest niezwykle istotnym wkładem w rozwój badań nad *sandhi* tonalnym w języku mandaryńskim, dlatego też była chętnie wykorzystywana w późniejszych pracach innych badaczy. Kilka teorii przedstawionych w kolejnych podrozdziałach, bazując na regule FFR, podejmuje próbę wyjaśnienia problematycznych kwestii, z którymi podejście prozodyczne nie było w stanie sobie poradzić.

2.2.4. Koncepcja Przypadku Specjalnego (*The Special Circumstance Approach*)

Z.S. Zhang (1988: 116–120, za N. Zhang 1997) koncentruje się na czterosylabowych frazach przyimkowych o strukturze [d [[a b] c]], gdzie [a b] jest frazą złożoną z przyimka *a* i rzeczownika *b*, które to nazywa przypadkami specjalnymi, jako że następuje wówczas pominięcie fazy grupowania bezpośrednich składników i przejście do zasady Miary Dwudzielnej. Jego zdaniem tego typu przykłady są dowodem na to, że w procesie tworzenia się struktur prozodycznych udział biorą nie jeden, lecz dwa osobne mechanizmy. Są to:

- 1) wstępne tworzenie struktur prozodycznych (jak np. przy użyciu Reguły Tworzenia Stopy (FFR) – por. § 2.2.3), stosowane w normalnych przypadkach,
- 2) modyfikacja struktur, stosowana w przypadkach specjalnych.

Jak twierdzi autor koncepcji, w przypadkach takich struktury prozodyczne podlegają zmianom mającym na celu podporządkowanie ich rytmicznemu wzorcom, charakterystycznym dla danego języka. Dobrym przykładem jest poezja, gdzie wzorce te mają szczególną wagę, a zatem główna rola przypada tu zasadzie DM, podczas gdy zasada IC ma

jedynie pomniejsze znaczenie. Pomimo niewątpliwej trafności tego spostrzeżenia teoria Przypadku Specjalnego wciąż jednak nie proponuje satysfakcjonującego wyjaśnienia, dlaczego frazy przyimkowe o powyższej strukturze powinny być traktowane inaczej niż frazy, które przyimka nie zawierają i rozpatrywane na równi z poezją. Prowadzi to do konkluzji, że wkład Z.S. Zhang w postęp badań w dziedzinie *sandhi* tonalnego jest w rzeczywistości nieistotny (N. Zhang 1997: 299).

2.2.5. Koncepcja Jednostki Sensownej (*The Sense Unit Approach*)

Hung i Chen (odpowiednio: 1989 i 1990: 34, za N. Zhang 1997: 299–301), podobnie jak Z.S. Zhang, stawiają sobie za cel wyjaśnienie problematycznej kwestii fraz przyimkowych. Twierdzą oni, że w strukturach o budowie [d [[a b] c]] łączenie się elementów w binarne jednostki w sposób dający w rezultacie strukturę typu ((d a) (b c)) można umotywić dążeniem elementów *d* i *a* do utworzenia dwusylabowej jednostki znaczącej. Ponieważ chińskie przyimki mogą także pełnić funkcję czasowników, jednostka taka stanowiłaby wówczas strukturę opartą na relacji argument – czasownik.

Jak zauważa N. Zhang (1997), przyimki nie zawsze jednak zachowują się tak jak czasowniki. Na potwierdzenie swoich słów przytacza kilka zdań, na przykładzie których obala słuszność koncepcji Jednostki Sensownej za pomocą trzech kontrargumentów.

Po pierwsze, wyjaśnia, że przyimki w odróżnieniu od czasowników nie łączą się z wykładnikami aspektu. Za ilustrację niech posłuży przykład zaczerpnięty z pracy *The Avoidance of the Third Tone Sandhi*:

- 1) *xiaqi women bi – guo liang ci*
 szachy my porównywać – ASP dwa raz(y)
 „(My) rywalizowaliśmy w szachach dwa razy”.

gou bi ma xiao
 pies niż koń mały
 „Pies jest mniejszy niż koń”.

* *gou bi – guo ma xiao*

Powyższy przykład dowodzi, że na podstawie funkcji, jaką pełni *bi* w zdaniu pierwszym, nie można zakładać, że funkcję taką spełnia także w innych zdaniach.

Po drugie, analizując frazę przyimkową [d [[a b] c]], z elementem *d* w roli podmiotu i elementem *b* w roli komplementu¹⁷ (*complement*) przyimka *a*, Zhang dochodzi do wniosku, że pozycja, jaką zajmuje *a*, nie może być pozycją czasownika. Przy opisanym wyżej rozmieszczeniu danych elementów, tłumaczy, obecność *c*, który może przybrać formę predykatu czasownikowego bądź przymiotnikowego, jest konieczna ze względu na pełnioną przezeń funkcję głównego składnika frazy; z kolei w przypadku czasowników, obecność *c* jest jedynie opcjonalna.

Po trzecie, gdyby element *a* był czasownikiem, należałoby przyjąć, że *b*, jako jego komplement, może zostać poddany topikalizacji, zjawisku obecnemu w dialekcie mandaryńskim. Umieszczenie na początku zdania komplementu przyimka skutkuje natomiast rozerwaniem frazy przyimkowej, dając w efekcie pojedynczy przyimek w dalszej części zdania, co nie jest w języku chińskim dozwolone.

Podsumowując, koncepcja Jednostki Sensownej nie wyjaśnia, na jakiej zasadzie element *a*, który w przedstawionej wyżej strukturze może pełnić jedynie funkcję przyimka, łączy się z *d*, tworząc potencjalną dwusylabową domenę, i nie podlega dzięki temu procesowi *sandhi* tonalnego. Co więcej, jednostką określaną mianem sensownej nie może być kombinacja pojedynczego argumentu i następującego po nim przyimka, który nie jest ani jego predykatem, ani też składnikiem tego argumentu wymagającym (N. Zhang 1997: 301).

2.2.6. Teoria Funktorów (*The Functor Approach*)

Jak czytamy u N. Zhang (1997: 301), w swojej teorii Hsiao (1991) posługuje się pojęciem „beatu” (z ang. *beat* – uderzenie), zapożyczonym od Selkirk (1984). Wyróżniamy dwa typy „beatów”: funktry (*functor beats*), do których należą zaimki, przyimki i wykładniki aspektu oraz „beaty” leksykalne (*lexical beats*). Procedura Tworzenia Funktoru (*Functor Beat Assignment*) odbywa się dwuetapowo, na wzór pierwszej fazy reguły FFR,

¹⁷ Według *A Dictionary of Linguistics & Phonetics*: „complement [is an] argument of a verb or predicate” (Crystal 2003).

z jej zasadami IC oraz DM (por. § 2.2.3) z tą jednak różnicą, że zachodzi ona po uprzednim uformowaniu się stóp z połączenia bezpośrednich składników (*Immediate Constituent Feet*) lub występujących obok siebie „beatów” (*Adjacent Beat Feet*). Frazy przyimkowe o budowie [d [[a b] c]] będzie w rezultacie charakteryzowała odmienna struktura prozodyczna od struktury tego typu fraz, które przyimka nie mają. N. Zhang ilustruje działanie tego procesu, posługując się niniejszym przykładem, gdzie *gei* pełni, odpowiednio, funkcję przyimka, a następnie czasownika:

1) *mao gei gou yao*
 kot przez pies gryźć
 x x x beat leksykalny
 [-----] f *Adjacent Beat Foot*
 x
 [-----] f *Adjacent Beat Foot*
 1 3 3 3 – tony początkowe
 1 3 2 3 – tony wyjściowe
 „Kot jest ugryziony przez psa”.

2) *qian gei shao hao*
 pieniądze dawać mało dobrze
 x x x x beat leksykalny
 [-----] f *Immediate Constituent Foot*
 [-----] f hiperstopa (*Jumbo Foot*)
 [-----] f hiperstopa
 2 3 3 3 – tony początkowe
 2 2 2 3 – tony wyjściowe
 „Lepiej da(wa)ć mało pieniędzy”.

Dokładniejsza analiza pozwala jednak dostrzec pewne słabości teorii funktorów: zgodnie z jej założeniami pierwsze dwa „beaty” leksykalne powinny połączyć się, dając w efekcie stopę typu *Adjacent Beat Foot* mającą postać *mao-gei-gou*. Powyższy przykład pokazuje natomiast, że jest inaczej. Co więcej, kategorie funktorów okazują się zbyt ogólne; nie uwzględniają bowiem różnic, jakie są udziałem przyimków i zaimków w procesie przebiegu *sandhi*, wpływając tym samym na ostateczny rezultat tonalny wypowiedzi (N. Zhang 1997: 303).

2.2.7. Teoria Efektu Semantyczno-Pragmatycznego (*The Semantic-Pragmatic Effect Approach*)

Jin (1993: 10, za N. Zhang 1997: 303), podobnie jak kilkoro jego poprzedników, analizuje frazy przyimkowe o strukturze [d [[a b] c]]. Wysuwa on tezę, zgodnie z którą powodem niepodlegania przyimka *a* zjawisku *sandhi* jest akcent kontrastywny, padający na sylabę *b*, co, w celu uniknięcia niedozwolonej konfiguracji, prowadzi do przekształcenia się frazy w strukturę o postaci ((d a)(b c)).

Zdaniem N. Zhang akcent kontrastywny nie jest konieczny, a wręcz zbędny, dla zajścia takiego przekształcenia, gdyż ma ono miejsce w przypadku wszystkich przyimków w tego typu frazach, bez udziału dodatkowych czynników. Zatem teoria efektu semantyczno-pragmatycznego również nie oferuje zadowalającej odpowiedzi na pytanie, dlaczego przyimki we frazach o strukturze [d [[a b] c]], są odporne na działanie *sandhi*, niezależnie od rozmieszczenia akcentu kontrastywnego.

2.2.8. Teoria Klityk¹⁸ (*The Clitic Approach*)

Shih (1991, za N. Zhang 1997: 303), podobnie jak Hsiao (por. Teorię Funktorów), traktuje zaimki na równi z przyimkami, umieszczając obie części mowy w jednej grupie. Jako że zaimki należą do kategorii klityk, można przyjąć, że w języku chińskim są nimi także przyimki. Biorąc pod uwagę sposób zachowania się klityk w strukturach leksykalnych, które, jako segmenty pozbawione własnego akcentu, najczęściej łączą się z poprzedzającymi je sylabami, założenie takie niewątpliwie sprzyja wyjaśnieniu formowania się dwusylabowych domen prozodycznych w czterosylabowych frazach przyimkowych w sposób opisany schematem ((d a)(b c)).

Jak się jednak okazuje, są pewne różnice pomiędzy przyimkiem nacechowanym tonem 3 a typową klityką, również na tonie 3. Mianowicie klityka zawsze tworzy domenę *sandhi* z poprzedzającą ją sylabą opatrzoną tonem 3, natomiast w przypadku przyimków jest to możliwe, choć nie jest konieczne, gdyż w dużej mierze zależy od kontekstu wypowiedzi (N. Zhang 1997: 303–304).

¹⁸ Wyraz atoniczny (proklityka lub enklityka).

2.2.9. Koncepcja Prozodycznej Hierarchii (*The Prosodic Hierarchy Proposal*)

Dokonawszy dogłębnej analizy podejścia syntaktycznego (por. § 2.2.1 i 2.2.2) oraz modelu prozodycznego Shih (por. § 2.2.3), której efektem jest dowiedzenie nieadekwatności pierwszego i wykazanie słabości drugiego, L. Cheng (1987) proponuje własną koncepcję, opierając się na modelu Prozodycznej Hierarchii Hayesa (1984, za L. Cheng 1987: 36–37), przedstawionym poniżej. Jako że słuszność niektórych założeń teorii Shih nie została podważona, a wręcz funkcjonują one w literaturze jako prawdziwe i tym samym powszechnie uznane, L. Cheng bierze sobie za cel wyjaśnienie kwestii przez poprzedniczkę nierozwiązanych. Zgodnie z tym przyjmuje pogląd, że *sandhi* tonalne w języku mandaryńskim zachodzi w obrębie struktur prozodycznych, tworzonych z uwzględnieniem zasad, jakimi rządzi się składnia.

Prozodyczna Hierarchia Hayesa

Hayes (1984, za: L.Cheng 1987: 36–37) wyróżnia dwa poziomy związane z pojemnością segmentów, w obrębie których działają procesy fonologiczne, i które to rządzą się odmiennymi zasadami, co następnie znajduje odzwierciedlenie w konsekwencjach tych procesów: poziom wyrazu oraz poziom frazy. Segmenty te dzieli natomiast w sposób następujący, poczynając od najmniejszego:

- 1) wyraz (*word*),
- 2) grupa klityczna (*clitic group*),
- 3) fraza fonologiczna (*phonological phrase*),
- 4) fraza intonacyjna (*intonational phrase*),
- 5) wypowiedź (*utterance*).

Autor twierdzi, że zmiany fonologiczne, jakim podlegają frazy, mają swoje podłoże nie bezpośrednio w strukturach syntaktycznych, lecz w sposób pośredni, poprzez struktury prozodyczne, które są w stosunku do nich wtórne. Pogląd ten przyjęła także Shih (1985, za: L. Cheng 1987) w koncepcji stóp prozodycznych, uznając struktury prozodyczne za właściwe domeny *sandhi*.

W kwestii wyrazów autor przyjmuje stanowisko prezentowane w modelu fonologii leksykalnej (Kiparsky 1982, za: L.Cheng 1987: 36–37). Zgodnie z nim wewnętrzna struktura wyrazów pozostaje nienaruszalna dla procesów postleksykalnych, a zatem zmiany fonologiczne w obrębie tychże segmentów muszą zachodzić na poziomie leksykalnym. Innymi słowy, za wszelkie modyfikacje tonalne wyrazów, a co za tym idzie – także złożenia i reduplikacji, odpowiedzialna jest nie składnia czy prozodia, lecz leksykon. Grupa klityczna natomiast okazuje się pomocna przy wyjaśnianiu różnic tonalnych pomiędzy wyrazami treściowymi (*content words*) – jak pełne frazy rzeczownikowe – o znaczeniu semantycznym, a funkcjonalnymi (fw, *function words*), pełniącymi różne funkcje – w tej grupie zaimki, przyimki oraz określniki. Różnice te wynikają ze zdolności wyrazów znaczących do tworzenia własnych grup klitycznych i konieczności „doczepiania się” wyrazów gramatycznych do sąsiednich, już istniejących grup, przy czym proces klitycyzacji wykazuje tendencje charakterystyczne w zależności od danego języka.

Jak podaje L. Cheng (1987: 40), proces klitycyzacji w chińskim mandaryńskim przebiega zgodnie z następującymi zasadami:

- 1) każdy wyraz leksykalny tworzy osobną grupę klityczną;
- 2) każda grupa klityczna inkorporuje przyległe wyrazy gramatyczne na zasadach łączliwości syntaktycznej, w granicach wyznaczonych przez zdania. Inkorporacja odbywa się wyłącznie w kierunku zgodnym z kierunkiem łączliwości syntaktycznej; inkorporacja w kierunku przeciwnym jest niedozwolona.

W przypadku fraz, zarówno fonologicznych, jak i intonacyjnych, rzecz ma się podobnie jak z klitykami, bowiem wyodrębnianie obu tych struktur oraz ich zachowanie pod wpływem modyfikacji tonalnych zależy od indywidualnego zestawu czynników. Ogółem wiadomo, że frazy fonologiczne powstają z połączenia grup klitycznych i nie przekraczają granic zdaniowych. Możliwe jest także istnienie kilku poziomów frazowych. Spośród fraz fonologicznych, te, które nie są subfrazami fraz na wyższym poziomie, mogą łączyć się w większe całości – frazy intonacyjne. Jak twierdzi L. Cheng (1987), procesy fonologiczne respektują granice frazowe także i w tym wypadku, co dotyczy m.in. tonalnego *sandhi*.

Ciekawe następstwa dla formowania się fraz intonacyjnych ma tymczasem zjawisko topikalizacji, powszechnie występujące w języku mandaryńskim. Według Huang (1984, za: L.Cheng 1987: 44), elementy

poddane topikalizacji opuszczają granice danej frazy intonacyjnej, której były częścią przed zajściem procesu. L. Cheng przyjmuje, że elementy te tworzą własne, odrębne frazy intonacyjne, a tym samym także frazy fonologiczne i grupy klityczne. Ponieważ *sandhi* nie zachodzi na granicy fraz intonacyjnych, element topikalizowany zachowuje w efekcie swój ton w formie pierwotnej. Dla ilustracji przykład zaczerpnięty z pracy *On the Prosodic Hierarchy and Tone Sandhi in Mandarin*:

1) *ni* [*bi wo xiao*]
 ty niż ja mały = „Jesteś mniejszy/a niż ja”.

w	fw	fw	w	3	3	3	3	– tony początkowe
		\\	/					
c		c		3	[2		2 3]	
Ph		Ph			-----			
IPh		IPh		[3		2	2 3]	– tony wyjściowe

Bazując na modelu Hayesa oraz własnych spostrzeżeniach, L. Cheng za pomocą szeregu przykładów ilustruje działanie zjawiska *sandhi* zarówno na poziomie wyrazów, gdzie przy użyciu złożeń i reduplikacji dowodzi jego cyklicznego charakteru, jak i fraz – tu skupia się głównie na kwestiach problematycznych dla koncepcji Shih (1985), osiągając te możliwe konfiguracje tonalne, których model poprzedniczki nie był w stanie wygenerować. Wariancyjność tonalną uzyskuje na podstawie struktury powstałej na skutek wspomnianej już topikalizacji oraz dzięki zaproponowanej przez siebie regule tworzenia Nacechowanej Grupy Klitycznej (*Marked Clitic Group Formation*), którą formułuje następująco:

Reguła Tworzenia Nacechowanej Grupy Klitycznej:

Jeśli (poruszając się od lewej do prawej) w danej strukturze występują co najmniej dwa zestawione obok siebie wyrazy gramatyczne, mogą one wówczas tworzyć własne grupy klityczne. Pozostałe wyrazy gramatyczne są inkorporowane przez przyległe grupy zgodnie z zasadami łączliwości syntaktycznej.

2.2.10. Teoria Optymalności (*Optimality Theory*)

Teoria Optymalności, w odróżnieniu od wcześniej tu zaprezentowanych koncepcji, nie oferuje w zasadzie gotowego zestawu twierdzeń na dany temat, gdyż nie odnosi się do konkretnego zagadnienia. Trafniejsze byłoby zatem określenie jej jako pewnego rodzaju podejścia, którego nadrzędnym celem jest wyłonienie właściwego rozwiązania spośród szeregu możliwych opcji przy użyciu – mówiąc językiem potocznym – rachunku zysków i strat. Choć stworzona z myślą o analizie struktur fonologicznych, obecnie jest używana także w kontekście innych poziomów języka. Teoria ta nieodłącznie wiąże się z pojęciem ograniczeń (*constraints*), nakładanych przez gramatykę danego języka na wszelkie możliwe do wygenerowania przez ten system formy, tzw. kandydatów (*candidates*), które to następnie podlegają ewaluacji zwieńczonej wyborem formy optymalnej – „kandydata” najbardziej pożądanego. Na wybór ten wpływ mają zarówno czynniki ilościowe, jak i jakościowe, tj. pod uwagę brana jest liczba naruszonych ograniczeń oraz ich miejsce w hierarchii ważności, zatem „kandydatem” najbardziej pożądanym będzie ten, który spośród wszystkich rozpatrywanych narusza najmniej ograniczeń o najniższej randze.

Teoria Optymalności jest dość powszechną i chętnie stosowaną teorią w literaturze traktującej o zagadnieniu tonalnego *sandhi*, być może z uwagi na fakt, że jest ona na tyle elastyczna, iż umożliwia uzyskanie więcej niż jednej wyjściowej kombinacji tonów. Tak czy inaczej teorię tę wykorzystują w swoich pracach m.in. N. Zhang (1997), Lee-Schoenfeld i Kandybowicz (2008) oraz Lin (2005), przy czym każdy z badaczy ujmuje problem pod nieco innym kątem.

N. Zhang (1997), pomimo obecnej już wówczas w literaturze krytyki podejścia syntaktycznego, niejako powraca w swojej koncepcji do – wydawać by się mogło – dawno porzuconego nurtu, podkreślając wpływ struktury syntaktycznej na tworzenie się domen *sandhi*. Według autorki jest to złożony, dwuwymiarowy proces, którego rezultat zależy od wzajemnego oddziaływania wielu czynników. O tym, które elementy danej frazy mogą uniknąć tonalnego *sandhi*, a które będą mu podlegać, decyduje szereg ograniczeń, co odbywa się w ramach pierwszego wymiaru. Z kolei wymiar drugi koncentruje się na kwestii możliwych końcowych przebiegów tonalnych, będących wynikiem istnienia więcej

niż jednej formy optymalnej i mających związek z rozmieszczeniem akcentu frazowego. W celu wyjaśnienia relacji składnia – akcent badaczka posługuje się pojęciem siły elementów struktury syntaktycznej (*constituent strength*), którym to przypisuje wartości „silny” (*strong*) bądź „słaby” (*weak*) na podstawie pełnionych przez nie funkcji w danej strukturze. Swoją hipotezę buduje w oparciu o teorię akcentu frazowego *Null Theory of Phrase Stress* zaproponowaną przez Cinque w pracy z 1993 roku. Zgodnie z tą teorią we frazach zawierających komplement (por. § 2.2.5) element ten jako „najsilniejszy” jest głównym nośnikiem akcentu frazowego (*phrase stress bearer*), natomiast we frazach bez komplementu rola ta przypada jej głównemu elementowi (*head*), czyli temu, od którego dana fraza przyjmuje nazwę. Wszelkie inne elementy określające lub modyfikujące element główny są niezmiennie „słabe” – stąd nieakcentowane (N. Zhang 1997: 304). Jako że nośniki akcentu frazowego nie podlegają *sandhi* tonalnemu, na podstawie przyjętych przez N. Zhang założeń logiczną konkluzją byłoby stwierdzenie, że procesom tym nie mogą zostać poddane przyimki we frazach przyimkowych. Tymczasem, jak wykazano we wcześniejszych podrozdziałach, to właśnie frazy przyimkowe są zasadniczym problemem, z którym borykają się badacze zjawiska *sandhi*, co sprawia, że stanowią one słaby punkt wielu teorii. Ponieważ zachowanie przyimków, którego teorie te często nie są w stanie wyjaśnić w sposób satysfakcjonujący, zdaje się wymykać wszelkim regułom, frazy przyimkowe w koncepcji N. Zhang są traktowane jako przypadki specjalne. Zdaniem autorki to, czy przyimek podlega czy też nie podlega modyfikacji tonalnej, wynika z faktu, że wartość owej części mowy w kategorii siły jest nieokreślona (*unspecified constituent strength*). Prowadzi to do wygenerowania większej liczby „kandydatów”, które są następnie poddawane selekcji przy użyciu szeregu ograniczeń gramatycznych, dając w rezultacie możliwe kombinacje tonalne (N. Zhang 1997: 336).

N. Zhang wyróżnia sześć ograniczeń, które określa następująco:

- 1) Zachowanie tonu składnika „najsilniejszego” (PTAS) (*Parse Underlying Tone of an Absolutely Strong Node*),
- 2) Zachowanie tonu składnika „silnego” (PTRS) (*Parse Underlying Tone of a Relatively Strong Node*),
- 3) Nierozzerwalność klityki (CI) (*Clitic Dependency*),
- 4) Niedozwolona sekwencja tonów 3 (*33),

- 5) Nienaruszalność granic składnika dwusylabowego (Align-Di-L) (*Disyllabic Constituent Alignment*),
 6) Maksymalna domena (Max) (*Maximal Domain*).

Ograniczenia te badaczka dzieli na trzy grupy, w zależności od tego, jakiego aspektu dotyczą. W grupie pierwszej, którą stanowią 1), 2) oraz 3) z powyższej listy, znajdują się ograniczenia pośrednio lub bezpośrednio związane ze strukturami syntaktycznymi. Zarówno PTAS, jak i PTRS mają na celu ochronę początkowych tonów elementów „silnych” danej struktury przed jakimikolwiek zmianami, przy czym pierwszy z nich odnosi się do elementów niebędących częścią żadnego składnika wyższego rzędu (o większym stopniu skomplikowania) tejsze struktury, któremu przypisano wartość „słaby”, natomiast drugi – do elementów przynależących do co najmniej jednego składnika „słabego”. Działanie tych ograniczeń ilustruje poniższy przykład, który wykorzystuje N. Zhang w pracy *The Avoidance of the Third Tone Sandhi in Mandarin Chinese* (1997: 306):

- 1) | |
 w s
 / \
 w s
 / \
 w s
 ma [[hen shao] hou]
 koń bardzo rzadki ryczeć = „Konie rzadko ryczą”.
 3 3 3 3 – tony początkowe
 [3] [2 2 3] – tony wyjściowe 1
 [2 2 2 3] – tony wyjściowe 2

Z kolei *Clitic Dependency* głosi, że klityka nie może zostać odłączona od domeny stworzonej przez poprzedzającą ją czasownik lub przyimek (N. Zhang 1997: 307).

Kolejne dwa ograniczenia kontrolują przebieg tonalny po zajściu zjawiska *sandhi* oraz kierunek tworzenia się domen. Warunek o symbolicznym zapisie *33 przeciwdziała występowaniu obok siebie sylab opatrzonych tonem 3, podczas gdy Align-Di-L, spełnia dwojaką funkcję. Po pierwsze, strzeże granic domeny *sandhi*, gwarantując ich pokrywanie się z granicami dwusylabowych składników; innymi słowy, nie pozwala on na rozbitcie

elementów binarnych tworzących pewną semantyczną całość. Jest to niejako rozszerzenie tradycyjnego ujęcia zasady spójności leksykalnej (por. s. 72–73), głoszącej, że jednostki leksykalne takie jak złożenia nie mogą zostać rozbite i pogrupowane osobno. Tymczasem wspomniany wyżej warunek dodaje do tej grupy także dwusylabowe frazy, stanowiące pewną całość, lecz niebędące jednostkami leksykalnymi, jak np. *dian huo* (zapalić ogień) czy też *hao jiu* (dobre wino, N. Zhang 1997: 308–309).

Druga z funkcji omawianego tu ograniczenia polega na kontrolowaniu kierunku grupowania pojedynczych sylab, mającym na celu zapewnienie nienaruszalności lewej granicy domeny *sandhi*; w praktyce oznacza to, że segment jednosylabowy znajdujący się po lewej stronie domeny nie zostanie z tą domeną złączony (N. Zhang 1997: 309).

Ostatnie z powyższej listy – ograniczenie Max – dotyczy domen prozodycznych, których maksymalna liczba elementów wynosi dwa w tempie normalnym oraz więcej w tempie szybszym lub stylu potocznym. Ograniczenie to przywołuje na myśl zasadę miary dwudzielnej, będącej częścią Reguły FFR z koncepcji Shih (por. § 2.2.3), która podporządkowuje tworzenie się stopy prozodycznej czynnikom metrycznym (N. Zhang 1997: 308).

Jak wspomniano wcześniej, zgodnie z założeniami Teorii Optymalności ograniczenia te mogą być i istotnie są naruszane, a końcowy rezultat ich działania w obrębie danej struktury syntaktycznej jest uzależniony od ich pozycji w hierarchii. N. Zhang proponuje następujące uszeregowanie:

- PTAS, *33, Cl;
- PTRS, Align-Di-L;
 - Max.

Ponieważ „kandydaci” naruszający ograniczenia na najwyższym stopniu w hierarchii są w sposób oczywisty eliminowani, w konsekwencji często zdarza się więc, że w procesie wyboru optymalnego wyniku nieuniknione stają się naruszenia ograniczeń zajmujących pozycje niższe, dzięki czemu przestrzegane są te ważniejsze.

Lee-Schoenfeld i Kandybowicz we wspólnej pracy *Sandhi Sans Derivation: Third Tone Patterns in Mandarin Chinese* (2008) przedstawiają kwestię następstwa tonów zgoła inaczej. Przede wszystkim zwracają oni uwagę na dwojaki charakter zmian, jakim może zostać poddany

ton 3 – mianowicie *sandhi* „kanoniczne”, zachodzące w sekwencjach co najmniej dwóch sylab z tonem 3 i których wynikiem jest pewien ton rosnący oraz zmiany dające w rezultacie ton określany jako $\frac{1}{2}$ 3 (por. § 2.1), któremu, choć jest on w języku mandaryńskim znacznie powszechniejszy, poświęca się zdecydowanie mniej uwagi. W swojej koncepcji badacze ci wychodzą od założenia, że zjawisko *sandhi* tonu 3 nie powoduje jego całkowitej zmiany, lecz jedynie pewne uproszczenie konturu tonalnego, spowodowane redukcją o odpowiedni poziom toniczny¹⁹. I tak: w przypadku *sandhi* „kanonicznego” wyeliminowaniu miałyby ulegać poziomy (por. § 1.5.2) środkowy – niski, pozostawiając ton realizowany przez sekwencję [MH], lecz w pewien sposób różny od leksykalnego tonu 2, o czym szerzej w następnym podrozdziale. W podobny sposób tłumaczą badacze drugi rodzaj modyfikacji tonu 3 – mianowicie jako redukcję konturu o końcowy poziom wysoki, co w efekcie przynosi pewien niski ton opadający o symbolicznym zapisie [ML].

Innym czynnikiem, który według badaczy wpływa na brak *sandhi* w sylabach z tonem 3, jest ich prozodyczne uwydatnienie, pojawiające się np. pod wpływem akcentu, zarówno tonicznego, jak i kontrastownego. Prozodia mające wpływ na tworzenie się domen, w obrębie których te procesy zachodzą, są natomiast silnie skorelowane ze składnią, co badacze usiłują wykazać, analizując struktury różnej długości. Na podstawie powyższych założeń wyodrębniają sześć ograniczeń, których wzajemne oddziaływanie odpowiada za redukcję odpowiednich poziomów, dając tym samym zmieniony końcowy przebieg tonalny. Należą do nich:

- 1) MAX – T(σ),
- 2) MAX – T,
- 3) *COMPLEX CONTOUR,
- 4) *BOUNDARY RISE,
- 5) OCP(TONE) oraz
- 6) *H >> *L >> *M.

(Lee-Schoenfeld i Kandybowicz 2008: 4).

¹⁹ Lee-Schoenfeld i Kandybowicz opierają się na założeniu Chao (1948, za: Yuan i Chen 2014, 1970), zgodnie z którym do opisu sylab z tonem 3 potrzeba trzech poziomów tonicznych, reprezentowanych przez cyfry arabskie odpowiadające wysokości głosu; zapis symboliczny ma zatem postać T3: (214). Badacze przyjmują natomiast notację przy użyciu liter alfabetu łacińskiego L, M, H oznaczających odpowiednio: niski (*low*), średni (*middle*) oraz wysoki (*high*) ton, tak więc opis tonu 3 przybiera formę MLH (2008: 1–2).

Ograniczenie wymienione jako pierwsze zapobiega redukcji tonemicznej sylab mających kluczowe pod względem prozodycznym znaczenie dla danej struktury, tj. sylab akcentowanych. Zatem aby pewien segment mógł uniknąć zjawiska *sandhi*, warunek ten nie może zostać naruszony. Ograniczenie drugie, które jest z kolei bardziej ogólną wersją poprzedniego, postuluje ochronę przed redukcją konturów tonalnych bez względu na pozycję, w jakiej dany segment jest umieszczony. Warunek *COMPLEX CONTOUR zabrania natomiast sekwencji więcej niż dwóch poziomów tonicznych w obrębie jednej sylaby, dążąc tym samym do uproszczenia konturów bardziej złożonych, takich, jak kontur tonu 3. Kolejne ograniczenie, *BOUNDARY RISE, przeciwdziała występowaniu wzrostu intonacji na końcu domeny prozodycznej, co czyni je odpowiedzialnym za powstawanie tonu $\frac{1}{2}$ 3. OCP(TONE) za niedozwolone uważa zestawienia identycznych sekwencji dwóch lub więcej poziomów tonicznych w ramach jednej domeny prozodycznej. Zgodnie z tym kombinacja taka jak np. (MLMLH) jest niepożądana. Ostatni, szósty warunek głosi, że w językach, w których występują trzy różne poziomy toniczne, H jest najbardziej, a M najmniej nacechowanym. W praktyce oznacza to, że poziom H podlega eliminacji najczęściej, nieco rzadziej dotyczy to poziomów niskich, natomiast poziomy o średniej wysokości wykazują największą odporność na redukcję. Podobnie jak w poprzednim ujęciu, autorzy koncepcji szeregują ograniczenia w kolejności od najbardziej do najmniej ważnych. Hierarchia przedstawia się następująco:

- MAX – T(σ);
- *COMPLEX CONTOUR;
 - MAX – T;
- *BOUNDARY RISE, OCP(TONE);
 - *H >> *L >> *M.

(Lee-Schoenfeld i Kandybowicz 2008: 4–5).

Lin we wstępie do pracy zatytułowanej *Prosodic Correspondence in Tone Sandhi* (2005) wprowadza – jej zdaniem – jedno z centralnych dla Teorii Optymalności pojęć, mianowicie pojęcie relacji zgodności, która ma zachodzić pomiędzy parami form. Ogólnie rzecz biorąc, zgodnie z pierwotnym założeniem zgodność ta dotyczy form początkowych, mających postać sprzed zajścia procesów, w kontekście których teoria

ta jest stosowana oraz form wyjściowych, przybieranych przez elementy podlegające danym procesom po ich wystąpieniu – jest to zatem zgodność wejście-wyjście (*Input-Output Correspondence*). Powołując się na twierdzenie, szeroko poparte w literaturze, że prozodia odgrywa bardzo istotną rolę w procesie produkcji i rozumienia mowy, Lin sugeruje, że formy wyjściowe derywowane w obrębie domen prozodycznych również mogą i powinny być badane pod kątem występowania zgodności. Opiera się na teorii Hsiao (1996, 2000b, za: Lin 2005: 234), który opowiada się za traktowaniem ograniczeń kształtujących rezultat tonalny oraz tych, którym podlegają domeny prozodyczne, jako należących do dwóch osobnych kategorii. Wykorzystuje także zaczerpnięty z Teorii Zgodności (*Correspondence Theory*) oraz przyjęty przez Hsiao pogląd, że pomiędzy poziomem syntaktycznym, prozodycznym i fonologicznym danego języka powinny występować przynajmniej trzy rodzaje zgodności, które zachodzą równolegle: zgodność syntaktyczno-prozodyczna, zgodność prozodyczno-fonologiczna (bądź też prozodyczno-tonalna) oraz zgodność pomiędzy tymi dwoma. Badaczka na tej podstawie proponuje dwa zestawy ograniczeń: ograniczenia tonalne, odpowiedzialne za zmiany przebiegu tonalnego oraz ograniczenia prozodyczne, pozwalające na wyodrębnienie domeny *sandhi*.

I. Ograniczenia tonalne:

- 1) Należy unikać sekwencji tonów niskich (*Avoid adjacent low tones at the tonal level*) (OCP-T(L)),
 - 2) Odpowiadające sobie tony początkowe i wyjściowe powinny być identyczne (*Input-Output corresponding tones (at the tonal level) are identical*) (IDENT-IO-T),
 - 3) Ton znajdujący się najdalej po prawo w danej wypowiedzi powinien być identyczny z odpowiadającym mu tonem początkowym (*The rightmost tone of an utterance (at the tonal level) is identical to its input correspondent*) (IDENT-IO-T-R).
- (Lin 2005: 235–236).

Hierarchia ważności przedstawia się następująco:

- IDENT-IO-T-R, OCP-T(L);
- IDENT-IO-T.

Zgodnie z Zasadą Właściwego Konturu (*Obligatory Contour Principle*, Katamba 1996: 193) w języku mandaryńskim powszechnie jest zjawisko unikania następujących po sobie identycznych tonów, w szczególności tonów niskich, przed czym przestrzega ograniczenie nr 1). Aby zapobiec występowaniu niepożądanych sekwencji tonów w końcowym rezultacie, ograniczenie to musi zdominować warunek kolejny, który postuluje ochronę tonów przed jakimikolwiek zmianami, będąc tym samym w sprzeczności z poprzednikiem. Natomiast ograniczenie o symbolicznym zapisie IDENT-IO-T-R głosi, że odpowiadające sobie segmenty stojące w „pozycji uprzywilejowanej” (*prominent position*), takie jak np. sylaby akcentowane, powinny być identyczne przed i po zajściu zmian. W przypadku języka mandaryńskiego, w którym „pozycje uprzywilejowane” są zlokalizowane zawsze po prawej stronie, oznacza to, że segmenty znajdujące się w takich pozycjach będą mniej podatne na zmiany, a co za tym idzie – istnieje większe prawdopodobieństwo, że *sandhi* tonalne wystąpi po lewej stronie domeny (Lin 2005: 235–236).

O ile powyższy zestaw sprawdza się w przypadku struktur dwusylabowych, nie jest on wystarczający do wyjaśnienia różnych możliwych przebiegów w sekwencjach trzytonowych; w tym celu Lin dodaje kolejne ograniczenie:

- 4) Końcowy przebieg tonalny nie powinien zawierać tonów niskich (*No L tones*) (*L).

Pomimo że zastosowanie ograniczenia nr 4) pozwala na eliminację pewnych niepożądanych „kandydatów”, wciąż jednak zdarzają się przypadkowe wykluczenia rezultatów dopuszczalnych jako jedna z możliwości. Lin wykazuje to na przykładzie dwóch fraz: *yu san xiao* („parasol jest mały”) oraz *xiao yu san* („mały parasol”) o identycznych segmentach, lecz różnym położeniu domen – odpowiednio: (($\sigma\sigma$) σ) oraz (σ ($\sigma\sigma$)). W tego rodzaju przypadkach, jak wyjaśnia badaczka, należy najpierw odpowiednio zlokalizować domenę, dzięki czemu możliwe staje się wygenerowanie jedynie pożądanego przebiegu tonalnego. Ponieważ, jej zdaniem, w końcowej ewaluacji brana jest pod uwagę zgodność sekwencji tonalnych w dwusylabowej domenie przed zaj-

ściem procesów *sandhi*, jak i po tym konieczne jest wprowadzenie jeszcze jednego warunku, który tę zgodność zapewni:

- 5) Odpowiadające sobie tony w domenach bazowych i wyjściowych, izomorficznych pod względem morfosyntaktycznym oraz prozodycznym, powinny być identyczne (*Corresponding tones in the morphosyntactically and prosodically related bases and outputs must be identical*) (IDENT-BOT).

Warunek ten stanowi dowód na korelację obu grup ograniczeń, których współpraca jest niezbędna do wygenerowania prawidłowego wyniku (Lin 2005: 237–243).

Jak już wspomniano wcześniej, domeny *sandhi*, choć podatne na wpływ struktur syntaktycznych, nie są bynajmniej w stosunku do nich izomorficzne, co jest szczególnie widoczne na przykładzie fraz przyimkowych, w których znaczny udział w tworzeniu domen przypada prozodii. W swojej koncepcji Lin (Lin 2005: 252–259) przyjmuje założenie Shih (1986), uznając stopę prozodyczną za domenę *sandhi* tonalnego w języku mandaryńskim. W wyodrębnieniu tej domeny mają pomóc następujące ograniczenia, uwzględniające odrębny charakter fraz przyimkowych i niezawierających przyimka.

II. Ograniczenia prozodyczne:

- Frazy „nieprzyimkowe”:

- 1) Każda sylaba powinna być przyporządkowana strukturom prozodycznym wyższego rzędu (*Parse every syllable into higher prosodic level*; PARSE SYLL),
- 2) Struktury fonologiczne składają się z dwóch elementów, z których każdy stanowi pewną całość (*Phonological structures are binary branching*; BIN BRAN),
- 3) Stopa jest zawsze strukturą dwusylabową (*Foot must be binary under syllabic analysis*; FT BIN),
- 4) ALIGN IC/FT
 - (a) Lewa granica każdego bezpośredniego składnika (IC) powinna pokrywać się z lewą granicą jakiejś stopy (Ft) (*The left edge of every immediate constituent is aligned with the left edge of some foot*; ALIGN(IC,Ft)L),

- (b) Prawa granica każdego bezpośredniego składnika (IC) powinna pokrywać się z prawą granicą jakiegś stopy (Ft) (*The right edge of every immediate constituent is aligned with the right edge of some foot*; ALIGN(IC,Ft)R),
- 5) ALIGN FT/IC
- (a) Lewa granica każdej stopy (Ft) powinna pokrywać się z lewą granicą jakiegś bezpośredniego składnika (IC) (*The left edge of every foot is aligned with the left edge of some immediate constituent*; ALIGN(Ft,IC)L),
- (b) Prawa granica każdej stopy (Ft) powinna pokrywać się z prawą granicą jakiegś bezpośredniego składnika (IC) (*The right edge of every foot is aligned with the right edge of some immediate constituent*; ALIGN(Ft,IC)R),
- 6) ALIGN FT/WD
- (a) Lewa granica każdej stopy (Ft) powinna pokrywać się z lewą granicą jakiegś wyrazu (gramatycznego, Wd) (*The left edge of every foot is aligned with the left edge of some (grammatical) word*; ALIGN(Ft,WD)L),
- (b) Prawa granica każdej stopy (Ft) powinna pokrywać się z prawą granicą jakiegś wyrazu (gramatycznego, Wd) (*The right edge of every foot is aligned with the right edge of some (grammatical) word*; ALIGN(Ft,WD)R);
- Frazy przyimkowe:
- 7) Lewa granica każdej stopy (Ft) powinna pokrywać się z lewą granicą jakiegś wyrazu prozodycznego (Prwd) (*The left edge of every foot is aligned with the left edge of some prosodic word*) (ALIGN(FT,PRWD)L).
(Lin 2005: 250).

Warunek pierwszy – jak wskazuje jego treść – wymaga, aby każda sylaba była częścią jakiegś struktury prozodycznej o większym stopniu skomplikowania. Pomimo pozornego podobieństwa dwóch następujących Lin twierdzi, że efekty ich zastosowania mogą być zupełnie różne, jak ma to miejsce w strukturach typu ((σ) σ), w których stopa wyższego rzędu składa się z więcej niż dwóch sylab, przez co narusza ograniczenie 3), jednocześnie przestrzegając ograniczenia 2), gdyż istotnie zawiera ona dwa elementy o niesprecyzowanej liczbie składników prostych (Lin 2005: 251).

Kolejne trzy ograniczenia to ograniczenia złożone, nakładające wymagania na obustronne granice pewnych segmentów. Pierwsze z nich – ALIGN IC/FT – pozwala na wybór odpowiedniego pogrupowania stóp prozodycznych w strukturach trzysylabowych. Decyzja, czy będzie to $(\sigma (\sigma\sigma))$ czy też $((\sigma\sigma)\sigma)$, zachodzi przy uwzględnieniu struktury morfosyntaktycznej danej wypowiedzi podlegającej zmianom. Analiza Lin pokazuje jednak, że w procesie wyboru optymalnego „kandydata” w niektórych strukturach czterosylabowych ograniczenie to jest naruszane, aczkolwiek zawsze odbywa się to kosztem przestrzegania ograniczenia FT BIN, zajmującego w hierarchii pozycję równorzędną, a nierzadko nawet wyższą. Warunek kolejny, pomimo znacznego podobieństwa do poprzednika zarówno w formie, jak i treści, również ma dla końcowego rezultatu odmienne konsekwencje. Kluczowe znaczenie mają tu wyrazy „każdy” i „jakaś”; nie jest więc istotne, czy obie granice każdego bezpośredniego składnika pokrywają się z odpowiadającymi im granicami danej stopy, tak, jak nie miało to znaczenia w poprzednim przypadku, gdyż oba te ograniczenia uzupełniają się, tworząc kombinację wymagającą całkowitej zgodności pomiędzy strukturami syntaktycznymi i prozodycznymi (Lin 2005: 253–254).

Niekiedy zdarza się, że dwie struktury charakteryzuje jednakowy układ bezpośrednich składników, natomiast odmienne pogrupowanie stóp, co może być spowodowane różnicami w ich strukturze wyrazowej; wówczas w wyjaśnieniu tego przypadku pomaga warunek nr 6), który chroni przed rozbięciem wyrazów gramatycznych na mniejsze elementy (Lin 2005: 255–257).

Ograniczenie wymienione jako ostatnie zostało zaproponowane przez autorkę w celu wyjaśnienia położenia *sandhi* tonalnego w niektórych frazach przyimkowych, które to okazują się niezwykle problematyczne. Często bowiem frazy o identycznej strukturze syntaktycznej i wyrazowej wykazują różne końcowe przebiegi tonalne. Jak tłumaczy Lin, jest to wynik odmiennego układu wyrazów prozodycznych, rozumianych przez badaczkę jako wyrazy leksykalne; wyrazy funkcyjne takie, jak np. przyimki, klasyfikatory czy zaimki, nie mogą tworzyć wyrazów prozodycznych (Lin 2005: 257–258).

Powyższe ograniczenia zajmują na skali ważności następujące pozycje:

- PARSE SYLL, BIN BRAN;
 - ALIGN FT/WD;
- FT BIN, ALIGN IC/FT;
- ALIGN(FT,PRWD)L;
 - ALIGN FT/IC.

2.3. Fonetyczna natura tonu powstałego na skutek *sandhi*

Dyskusje nad faktyczną naturą „tonu *sandhi*” rozpoczęły się wraz z opublikowaniem w późnych latach 40. dwóch prac prezentujących dwa odmienne stanowiska. Chao w swojej pracy z 1933 roku – notabene pierwszej szeroko znanej rozprawie na temat *sandhi* tonalnego w języku mandaryńskim – wyraził pogląd, że proces ten polega na zastąpieniu tonu 3 leksykalnym tonem 2²⁰. Pogląd ten stoi w sprzeczności z koncepcją Hocketta (1947, za: Yuan i Chen 2014), który uważał, że ton powstały w wyniku *sandhi* jest nowym tworem, wykazującym jedynie pewne podobieństwo do tonu 2, co zostało następnie poparte przez Martina (1957, za: Yuan i Chen 2014). Od tamtego czasu usiłowano zbadać czy ton 2 i „ton *sandhi*” to istotnie dwie różne kategorie, próbując poprzeć tezę wynikami eksperymentów naukowych, których rozkwit nastąpił z początkiem lat 80. Choć obecnie znaczna większość badaczy zdaje się przychylić do koncepcji Hocketta i Martina, wciąż brakuje jednomyślności w kwestii wielkości różnicy pomiędzy wspomnianymi tonami. Są i tacy, którzy różnice te łączą jedynie z niektórymi odmianami języka mandaryńskiego, jak ta używana przez jego rodzimych użytkowników zamieszkujących Pekin (Myers i Tsai 2003, za: Yuan i Chen 2014).

Autorem pierwszych badań laboratoryjnych nad *sandhi* tonu 3 jest (według Yuan i Chen) prof. Zee (1980), który na przykładzie dwóch mówców posługujących się dialektem pekińskim wykazał, że ton poddany zmianom jest produkowany z niższą częstotliwością niż kanoniczny ton 2, gdyż zarówno początek, jak i koniec jego konturu charakteryzuje niższa częstotliwość podstawowa (F_0). Spostrzeżenia te zostały następnie potwierdzone w szeregu eksperymentów akustycznych, wciąż jednak brak konsensusu co do szczegółowych parametrów różnic między oboma tonami.

²⁰ Większość badań dotyczy *sandhi* „kanonicznego”.

Choć obecność tych różnic w produkcji została w zasadzie stwierdzona, ich wpływ na percepcję jak dotąd pozostaje kwestią do rozstrzygnięcia. Wrażliwość słuchaczy na zmiany tonalne badała m.in. Peng (2000, za: Yuan i Chen 2014), wzorując się na pierwszym tego typu eksperymencie przeprowadzonym w 1967 roku przez Wang i Li. W obu tych eksperymentach badanych poddano odsłuchowi i poproszono o ocenę, którą z dwóch sekwencji tonalnych zawierał usłyszany wyraz: ton 2-ton 3 (LR-T3) czy też „ton *sandhi*”-ton 3 (SR-T3). Należy przy tym zaznaczyć, że nagranie zawierało pary poprawnych, istniejących w leksykonie języka mandaryńskiego homofonów w obu konfiguracjach tonalnych, jak np.: *qí mǎ* „jeździć konno” i *qǐ mǎ* „nareszcie” (Yuan i Chen 2014), dzięki czemu wyeliminowano wiele zmiennych mogących wpłynąć na wynik badania. Zarówno u Wang i Li, jak i Peng wskaźnik średniej wrażliwości na bodźce wynosił ok. 50%, co sugeruje przypadkowość w udzielaniu poprawnych odpowiedzi. Należy jednak pamiętać, że jest to wynik uśredniony, zatem nie bez znaczenia pozostaje fakt, że niektóre spośród osób badanych uzyskały dużo lepsze rezultaty (Yuan i Chen 2014).

Speer i Xu (2008) mierzyli natomiast czas potrzebny badanym na przetworzenie informacji leksykalnej i prozodycznej zawartej w tonach standardowej odmiany języka chińskiego w serii eksperymentów polegających na identyfikacji wyrazów. Materiał wykorzystany w badaniach zawierał wyrazy dwuznaczne pod względem leksykalnym, co uzyskano w efekcie wstąpienia zjawiska *sandhi* tonu 3. W jednym z eksperymentów zastosowano procedurę multimodalnego torowania semantycznego (*cross-modal semantic priming*), gdyż spodziewano się zwiększenia prawdopodobieństwa poprawnego zidentyfikowania bodźców przy jednoczesnym skróceniu czasu reakcji poprzez wcześniejszą ekspozycję na bodziec pokrewny semantycznie. Z kolei w innym monitorowano ruchy gałek ocznych uczestników podczas wizualnej i słuchowej ekspozycji na bodźce w postaci wyrazów semantycznie dwuznacznych. Badanym odtwarzano dwuznaczny pod względem znaczeniowym fragment zdania, po którym następowało zdanie wyjaśniające tę dwuznaczność za pomocą kontekstu tonalnego, po czym proszono o wskazanie spośród widocznych na monitorze znaków, reprezentacji wyrazu usłyszanego w danym zdaniu. Ruchy gałek ocznych u osób badanych śledzące obie możliwe sekwencje tonalne

(tj. LR-T3 oraz SR-T3) podczas ekspozycji na fragment dwuznaczny świadczą o świadomości występowania zjawiska *sandhi*, natomiast gdy dwuznaczność została wyjaśniona przez kontekst kolejnego zdania, badani kierowali wzrok na właściwy wyraz. Zarówno wyniki obu tych eksperymentów, jak i materiał zebrany podczas ich przeprowadzania pokazuje wrażliwość słuchaczy na drobne różnice fonetyczne oraz podświadome wykorzystanie tych różnic w procesie przetwarzania informacji leksykalnej, co pozwoliło autorom na dokonanie konkluzji, że zmiany fonetyczne mogą już na wczesnym etapie wpływać na percepcję tonów (Speer i Xu 2008).

Yuan i Chen dla celów naukowych zdecydowali się porzucić laboratorium i przeanalizować pewne fonetyczne aspekty języka w warunkach naturalnych, wykorzystując obszerny materiał z nagrań telewizyjnych serwisów informacyjnych oraz spontanicznych rozmów telefonicznych, co dostarczyło sposobności przyjrzeniu się bliżej językowi, jakim posługują się zarówno wyszkoleni profesjonalści, jak i przeciętni użytkownicy z różnymi naleciałościami akcentowymi. Analizie poddano dwusylabowe wyrazy o następujących konfiguracjach tonalnych: ton 2 + ton 2, ton 3 + ton 2, ton 2 + ton 3 oraz ton 3 + ton 3, przy czym dwie pierwsze kombinacje potraktowano głównie jako punkt odniesienia. Z dwóch pozostałych sekwencji przedmiotem zainteresowania był ton zajmujący pozycję wcześniejszą, który w obu przypadkach poddawano dokładnym pomiarom; na podstawie konturu tonalnego mierzono najmniejsze i końcowe F_0 oraz procent, o jaki parametr ten wzrósł od wartości minimalnej do maksymalnej w czasie trwania danej sylaby. Brano też pod uwagę wpływ częstości wyrazów w korpusie języka mandaryńskiego na realizację tonu 2 i „tonu *sandhi*” poprzedzających sylabę z tonem 3. Wyniki pokazują, że tony te istotnie różnią się akustycznie w obu rozpatrywanych sytuacjach komunikacyjnych, potwierdzając tym samym wnioski z dotychczasowych badań z tego obszaru. Pomimo niewątpliwej wrażliwości na drobne różnice akustyczne, która objawia się u rodzimych użytkowników nieświadomie w procesie produkcji mowy, ci sami użytkownicy często jednak nie są w stanie tych różnic poprawnie wskazać; jest to jednak dość powszechne zjawisko, mogące mieć związek z odkodowywaniem informacji na poziomie fonologicznym, zachodzącym przed uzyskaniem dostępu do zasobów leksykalnych ludzkiego mózgu. W kwestii częstości natomiast badanie wykazało,

że różnica pomiędzy tonem 2 a „tonem *sandhi*” jest większa w przypadku wyrazów powszechniej występujących w języku mandaryńskim, co jest szczególnie widoczne w wielkości przyrostu F_0 . W związku z tym narzuca się wniosek, że w produkcję wyrazów o większej i mniejszej częstości mogą być zaangażowane różne procesy. Jak domniemają autorzy badania, jest to być może konsekwencja odmiennego sposobu przechowywania wyrazów w pamięci – tych o większej częstości jako pojedyncze jednostki, a wyrazów mniej powszechnych jako dwóch osobnych sylab, łączonych dopiero w momencie produkcji (Yuan i Chen 2014).

Koncepcja Lee-Schoenfeld i Kandybowicz (por. s. 84–85) jest w tym miejscu warta przypomnienia z uwagi na pierwszeństwo, jakie badacze ci przyznają drugiemu rodzajowi *sandhi* – zjawisku znacznie częstszemu, lecz zwykle pomijanemu w literaturze – mianowicie zmianom, których wynikiem jest ton $\frac{1}{2}$ 3 (HT3S). Analizują oni wypowiedzi pięciu rodzimych użytkowników języka mandaryńskiego, próbując dostarczyć dowodów na poparcie swej tezy o redukcji konturów tonalnych. Jak twierdzą, ton będący efektem redukcji o poziom toniczny H, dla którego przyjmują oznaczenie ML (*middle low*) – średni opadający, może przybierać różne formy w zależności od sposobu jego realizacji przez poszczególnych mówców, jak np. HM (*high middle*) – wysoki opadający lub też ton równy, realizowany najczęściej w rejestrze dolnym, natomiast, co jest szczególnie istotne, nigdy nie jest on ani tonem opadająco-wznoszącym z wyraźnie zaznaczoną zmianą kierunku, jak ma to miejsce w przypadku tonu 3, ani też równomiernie rosnącym, co z kolei cechuje ton 2. Mówiąc inaczej, ogólny trend tonu $\frac{1}{2}$ 3 charakteryzuje spadek lub brak wyraźnych zmian wysokości, lecz nigdy wzrost. Różnicy między leksykalnym tonem 2 a wynikiem *sandhi* „kanonicznego” dowodzi tymczasem mniejszy niż w pierwszym przypadku przyrost F_0 , co zostało już wykazane m.in. przez Yuan i Chen (2014).

Do ciekawych wniosków dochodzą Zhang i Lai (2010), autorzy badań psycholingwistycznych nad realizacją *sandhi* tonalnego w dwusylabowych logatomach. Przy założeniu, że czynniki fonetyczne często wpływają na pewne fonologiczne aspekty języka, sprzyjając w ten sposób dominacji wzorców mających silniejszą podbudowę fonetyczną, badacze ci sprawdzają słuszność swojej hipotezy na przykładzie różnic

spowodowanych zajściem zjawiska natury fonologicznej w obrębie dwóch form w różnym stopniu umotywowanych fonetycznie. Zdaniem Zhang i Lai, w przypadku *sandhi* tonu 3 w języku mandaryńskim, które charakter fonologiczny zawdzięcza swoistym dla tego języka złożonym przemianom tonalnym, formą bardziej umotywowaną jest HT3S, gdyż kontury tonów po zajściu zmian wykazują większe podobieństwo do konturów swoich początkowych odpowiedników, jako że są wynikiem ich prostej redukcji o część rosnącą. Z kolei T3S, przekształcające ton 3 – zgodnie z przyjętą przez autorów notacją – oznaczany 213 w ton 35, powoduje wzrost częstotliwości w sposób, jaki nie może zostać wytłumaczony uproszczeniem konturu o dany poziom toniczny. W odróżnieniu od *sandhi* „kanonicznego”, HT3S rządzą procesy postleksykalne, gdyż jest ono całkowicie niezależne od czynników syntaktycznych. Biorąc to pod uwagę, badacze sprawdzają, czy różnice w podbudowie fonetycznej przekładają się na wyniki zastosowania obu rodzajów przekształceń w dwusylabowych logatomach o różnych sekwencjach tonalnych przez rodzimych użytkowników języka mandaryńskiego. Zgodnie z przewidywaniami *sandhi* o słabszej motywacji fonetycznej (T3S), pomimo że zachodzi we wszystkich wyrazach bez wyjątku, jest stosowane niedokładnie, gdyż „ton *sandhi*” wykazuje większe podobieństwo do odpowiedniego tonu początkowego, niż ma to miejsce w przypadku wyrazów sensownych. Podobieństwo to dotyczy przede wszystkim umiejscowienia oraz częstotliwości punktu, w którym następuje załamanie konturu, a także długości trwania danej sylaby – mówiąc inaczej, ton ten jest dłuższy i niższy w logatomach niż wyrazach znaczących. Natomiast zachowanie tonów pod wpływem *sandhi* drugiego typu zależy od sekwencji, w jakiej tony te są zestawione; ogółem różnice pomiędzy logatomami i wyrazami poprawnymi są co najwyżej nieznaczące, co pozwala stwierdzić, że zjawisko to przebiega jak należy. Wyjątek stanowią struktury o elementach składowych ton 3 + ton 2, charakteryzujących się najmniejszą częstością w wykorzystanym w eksperymencie korpusie. Jak podsumowują badacze, choć częstość ma pewien wpływ na sposób realizacji *sandhi* tonalnego w sztucznie skonstruowanych, nieistniejących w języku mandaryńskim wyrazach, zmienna ta rozpatrywana w izolacji nie daje jednak podstaw do wyjaśnienia zaobserwowanych różnic (Zhang i Lai 2010).

Rozdział 3.

Część eksperymentalna

W ramach badań nad produkcją oraz percepcją tonów języka mandaryńskiego stanowiących obszar dociekań tej pracy przeprowadzono dwa eksperymenty. Pierwszy z nich stanowiło dwuetapowe nagranie z udziałem rodzimych i nierodzimych użytkowników języka tonalnego (mandaryński); uzyskany w ten sposób materiał poddano następnie analizie tonograficznej. Drugim eksperymentem był test percepcyjny, w którym uczestniczyli wyłącznie rodzimi użytkownicy języka nietonalnego – w tym wypadku języka polskiego.

3.1. Analiza tonograficzna wyrazów chińskich i polskich wypowiadanych przez rodzimych użytkowników języka mandaryńskiego

3.1.1. Materiał językowy

W pierwszym etapie nagrania wykorzystano materiał składający się z 40 izolowanych wyrazów języka mandaryńskiego. Ponieważ celem badania było uzyskanie tonów w ich „czystej” postaci, mających następnie posłużyć za wzorce i zarazem punkty odniesienia w zestawieniu z materiałem otrzymanym z nagrań z etapu drugiego, starano się dobrać materiał tak, aby zwiększyć łatwość ekstrakcji poszczególnych tonów, a także zminimalizować ryzyko wystąpienia zniekształceń, mogących powstać na skutek zjawiska tonalnego *sandhi* (por. rozdz. 2) czy zmian przebiegu F_0 spowodowanych realizacją akcentu kontrastywnego bądź wynikających ze zróżnicowanych pozycji poszczególnych sylab w obrębie analizowanego segmentu językowego. Z uwagi na powyższe czynniki zdecydowano się na wybór wyrazów jednosylabowych, odpowiadających pojedynczym jednostkom nacechowanym tonami, a zarazem o typowej dla języka mandaryńskiego długości. Należy przy tym zaznaczyć, że korpus stanowiło 10 sylab homograficznych w transliteracji *pinyin* (ignorując symboliczny zapis poszczególnych

tonów), z których każda występowała w czterech różnych wariantach tonalnych. W doborze odpowiednich sylab kierowano się głównie ich jak największą różnorodnością pod względem fonetycznym, zatem starano się, aby były wśród nich zarówno sylaby zamknięte, jak i otwarte, o zróżnicowanych nagłosach i wygłosach. Brano pod uwagę także ich powszechność w języku mandaryńskim, w związku z tym wyselekcjonowano wyrazy zajmujące pierwsze pod względem częstości występowania pozycje słownikowe; za źródło danych dotyczących częstości posłużył elektroniczny słownik chińsko-angielsko-chiński *Wenlin* (*Wenlin Institute, Inc.* 1997–2007). Z wybranych wyrazów sporządzono listę (por. Aneks), na której zestawiono kolejno po cztery znaki, stanowiące możliwe warianty tonalne poszczególnych homograficznych sylab, umieszczone w dziesięciu wierszach, przy czym warianty te występowały zawsze w kolejności od tonu 1 do 4. Z uwagi na brak znakowych odpowiedników niektórych spośród wybranych sylab konieczne okazało się zastosowanie transliteracji *pinyin*; choć dla rodzimych użytkowników języka mandaryńskiego jest to nienaturalna – a dla niewykształconych użytkowników często zupełnie nieznaną – forma zapisu, pełniąca na ogół rolę narzędzia pomocnego w nauce i nauczaniu mandaryńskiego jako języka obcego, na potrzeby badania użyto jej jedynie jako uzupełnienia.

W etapie drugim, który stanowił zasadniczą część nagrania, posłużono się materiałem w języku polskim. Aby stworzyć maksymalnie zbliżone warunki, stanowiące podstawę do porównania tych dwóch diametralnie różnych języków, wybrano 10 wyrazów jednosylabowych, z których następnie skomponowano listę. Mając na uwadze cele i założenia niniejszego badania (por. Wstęp), zdecydowano, że wykorzystane w nagraniu wyrazy powinny odznaczać się cechami typowymi dla języka polskiego, zarówno pod względem morfologicznym, jak i fonetycznym, ale jednocześnie niewłaściwych dla języka mandaryńskiego, zatem główne kryterium doboru stanowiła ich swoistość dla polszczyzny. Jednocześnie starano się, aby – podobnie jak w przypadku sylab chińskich – były one między sobą możliwie najbardziej fonetycznie zróżnicowane. Odpowiednie wyrazy zaczerpnięto z korpusu udostępnionego przez dra D. Śledzińskiego.

3.1.2. Uczestnicy

Uczestnikami nagrań były cztery osoby (dwie kobiety oraz dwóch mężczyzn), dla których język chiński w odmianie mandaryńskiej był językiem natywnym. Trzy z nich pochodziły z Republiki Tajwanu, a jedna z Chińskiej Republiki Ludowej²¹. Wśród mówców znajdowało się dwóch studentów Poznańskiego Uniwersytetu Medycznego, jedna studentka Uniwersytetu Adama Mickiewicza oraz jedna lektorka języka chińskiego w Instytucie Językoznawstwa UAM. Każda z tych osób wykazywała zarówno bierną, jak i czynną znajomość języka polskiego na poziomie co najmniej podstawowym, przy czym poszczególne osoby różniły się między sobą pod tym względem w znaczący sposób. Zebrano także informacje dotyczące celu nauki oraz sposobu wykorzystywania języka polskiego przez poszczególnych mówców, na podstawie których stwierdzono w tej kwestii różnice, mogące znaleźć odzwierciedlenie w wynikach badania. Mianowicie trzech uczestników posługiwało się językiem polskim w życiu zawodowym lub studenckim, a nawet wiązało z nim przyszłość, w związku z czym osoby te przywiązywały dużą wagę do poprawności gramatycznej i artykulacyjnej, a także – co jest szczególnie istotne – pobierały w tym zakresie lekcje, podczas gdy pozostała osoba używała języka polskiego jedynie w codziennej komunikacji.

Do oznaczania poszczególnych uczestników badania przyjęto notację mającą postać dwóch liter alfabetu łacińskiego, z których pierwsza jest inicjałem, natomiast druga wskazuje płeć danej osoby. Skład badanej grupy w porządku alfabetycznym przedstawia się następująco: DM, JK, WK, YM.

3.1.3. Przebieg eksperymentu

Nagrania przeprowadzono w czterech sesjach, każdą poświęcono w całości jednej osobie badanej. Wszystkie sesje odbyły się w studiu nagraniowym Zakładu Fonetyki Instytutu Językoznawstwa UAM w Po-

²¹ Pomimo istnienia pewnych drobnych różnic w wymowie pomiędzy językiem mandaryńskim używanym na Tajwanie a tym obowiązującym w Chinach upewniono się, że materiał językowy wykorzystany w niniejszym badaniu jest od nich wolny, względnie ewentualne rozbieżności nie wpłyną w sposób istotny na wyniki.

znaniu. W pierwszej części zadaniem mówców było przeczytanie listy wyrazów chińskich z zachowaniem niewielkich odstępów pomiędzy kolejnymi wariantami tonalnymi. Uczestnikom polecono, aby używali naturalnego dla nich tempa mowy oraz w razie pomyłki, powtórzyli błędnie wypowiedziany wyraz, starając się przy tym utrzymać wcześniejszy rytm. Nagranie rozpoczynało się w momencie, gdy mówcy sygnalizowali gotowość. Po skończeniu części pierwszej następowała krótka przerwa, w czasie której badani otrzymywali instrukcje dotyczące dalszego postępowania, po czym przechodzono do kluczowej części nagrania. W części tej osoby uczestniczące zostały poproszone o przeczytanie kolejno wszystkich pozycji z listy wyrazów polskich, lecz w sposób analogiczny do zastosowanego poprzednio, tj. każdy wyraz jednosylabowy musiał zostać wypowiedziany na czterech tonach mandaryńskich w kolejności od 1 do 4. W kwestii tempa czytania oraz odstępów między poszczególnymi wyrazami obowiązywały zasady z części pierwszej.

Nagrań dokonano na komputerze MacBook, umieszczonym poza kabiną dźwiękoszczelną i podłączonym do interfejsu UA-25. Dźwięk rejestrowany był przy użyciu programu do nagrywania i edycji plików dźwiękowych Audacity przy częstotliwości próbkowania równej 44,1 kHz, a następnie zapisywany w postaci plików o formacie obsługiwanym przez ten program. Po zakończeniu każdej sesji duplikaty plików konwertowano do formatu mp3 i w tej postaci zapisywano na zewnętrznym dysku.

3.1.4. Edycja materiału dźwiękowego

Materiał uzyskany ze wszystkich sesji nagraniowych poddano następnie obróbce w celu przygotowania go do późniejszej analizy. W pierwszej kolejności przeprowadzono niezbędne korekty, w ramach których wycięto wyrazy zrealizowane błędnie w wyniku pomyłek mówców, mających miejsce w trakcie przeprowadzania nagrań, a także mlaski, szumy oraz wszelkie inne dźwięki niebędące przedmiotem analizy, które mogłyby zakłócić bądź utrudnić proces uzyskiwania odpowiednich danych. W niektórych plikach zaszła konieczność korekcy pojawiających się w nich (niekiedy licznie) nagłych, gwałtownych fluktuacji w przebiegu F_0 głosu danego mówcy, niewynikających

z celowej modulacji wysokości tego parametru i zaburzających jego płynny przebieg w czasie. Korekty te nie spowodowały znaczących dla wyników eksperymentu modyfikacji, natomiast znacznie ułatwiły dokonywanie pomiarów.

Kolejnym krokiem był podział nagrań z każdej sesji na mniejsze fragmenty odpowiadające pojedynczym wyrazom. W ten sposób uzyskano w sumie 160 plików, które następnie wczytano do programu Praat w celu otrzymania obrazów poszczególnych tonów.

3.1.5. Analiza

Analizę tonograficzną oraz związane z nią wszystkie niezbędne pomiary również przeprowadzono z wykorzystaniem programu Praat. W celu uzyskania jak największej dokładności wszelkich ustawień i pomiarów dokonano manualnie. Materiał analizowano w dwóch zakresach częstotliwości: 75–300 Hz dla głosów męskich i 100–600 Hz dla głosów żeńskich²². Aby uzyskać wspólną podstawę do porównania tonów zrealizowanych w języku mandaryńskim i polskim, w obu przypadkach zastosowano jednakowe skale. Następnie z każdego pliku wyekstrahowano szereg wartości umożliwiających rozpatrywanie poszczególnych tonów pod kątem kilku parametrów: F_0 i jego zmian w czasie, iloczasu oraz poprawności realizacji danego tonu (por. § 3.1.5 C). Wszystkich obliczeń statystycznych dokonano w programie Microsoft Excel for Mac.

A. F_0

Częstotliwość podstawowa, a w szczególności jej przebiegi czasowe, jako główny fonetyczny korelat wysokości głosu, ze względu na kluczową rolę, jaką odgrywa w językach tonalnych, stanowiła niezwykle istotny element dociekań będących przedmiotem tej pracy. W związku z tym zmierzono maksymalną i minimalną wartość F_0 dla każdego wyrazu z osobna, zrealizowanego przez każdego z czworga mówców, po czym wyliczono średnie arytmetyczne dla poszczególnych wyrazów oraz zasięg skali wykorzystanej przy produkcji każdego z nich. Następnie z otrzymanych wcześniej wielkości średnich wyciągnięto ogólne średnie arytmetyczne

²² Skale dostosowano w oparciu o średnie wartości F_0 dla obu płci podane przez Yip (2002: 12).

tyczne dla poszczególnych głosów, uzyskując w ten sposób ich przeciętną wysokość. Z danych z pomiarów $F_{0\min}$ i $F_{0\max}$ wyekstrahowano wartości ekstremalne, w oparciu o które obliczono skalę głosów osób badanych zastosowaną w nagraniu. Aby uzyskać podstawę do bardziej szczegółowych porównań, wyliczono również ogólne średnie wartości F_0 dla wszystkich czterech tonów w zależności od mówcy, a także zakresy częstotliwości, w których tony te zostały zrealizowane. Wszystkie wielkości mierzono w hercach [Hz]. Na koniec obliczono odchylenia standardowe od wartości pierwotnych uzyskanych w pierwszym pomiarze, najpierw ogółem dla każdego głosu, a następnie z podziałem na poszczególne tony.

B. iloczyn

W ramach tego parametru mierzono długość reprezentacji fali dźwiękowej w postaci konturu tonalnego. Analizie poddawano jedynie wyraźnie najdłuższe segmenty dźwięczne, które stanowiły główne obszary realizacji tonów. Pomiarów dokonywano w milisekundach [ms]. Poza wartościami przypisanymi każdemu wyrazowi z osobna oszacowano również średnie długości wszystkich tonów 1, 2, 3 oraz 4 wypowiedzianych przez danego mówcę.

C. poprawność realizacji tonu

Parametr ten, jako zdecydowanie najmniej sformalizowany spośród branych tu pod uwagę, niewątpliwie trudno przedstawić w sposób wolny od wszelkiego subiektywizmu. Niemniej, ponieważ jest on znaczący dla niniejszej analizy, postanowiono przyjąć pewne kryteria, które pozwoliłyby dokonać oceny pozytywnej lub negatywnej oraz dostrzec tendencje w ewentualnych substytucjach jednych tonów innymi, co zakładano głównie w przypadku nakładania tonów na wyrazy języka polskiego. Zgodnie z tym poprawność realizacji poszczególnych tonów oceniano na podstawie kształtu ich konturów, co następnie weryfikowano w oparciu o wyniki testów percepcyjnych (por. § 3.3). W sytuacji, gdy kontur danego tonu miał przebieg zgodny z oczekiwanym oraz nie sprawiał trudności w percepcji, ton ten został sklasyfikowany jako poprawny. Natomiast niekiedy zdarzało się, że kontury tonalne wykazywały wyraźne spadki tam, gdzie nie były one pożądane bądź kształtem przypominały kontury innych tonów, co znajdowało odzwierciedlenie również w ich odbiorze słuchowym; wówczas oznaczano je

jako zrealizowane niepoprawnie oraz określano rodzaj modyfikacji, jaka zaszła w danym przypadku. Aby wyodrębnić czynniki, które mogły mieć wpływ na realizację tonów, sporządzono procentowy wykaz poprawnych i błędnych realizacji, uzależniony od kilku zmiennych. Przy sporządzaniu wykazu zwracano uwagę na następujące kwestie:

- który ton zrealizowano najlepiej, a który najgorzej;
- czy istnieją wyrazy, które zostały zrealizowane błędnie przez wszystkich bez wyjątku mówców;
- czy tony były realizowane lepiej przez kobiety czy przez mężczyzn;
- który mówca zrealizował poprawnie najwięcej tonów.

W zależności od tego, w którym przypadku stwierdzono korelację, dominującym czynnikiem oddziałującym na poprawność realizacji mógł okazać się stopień złożoności tonu, struktura sylaby będącej jego nośnikiem, płeć czy też pewne indywidualne zdolności i predyspozycje mówcy, wynikające z szeregu różnych przyczyn.

3.2. Analiza tonograficzna wyrazów chińskich i polskich wypowiedzianych przez nierodzimych użytkowników języka mandaryńskiego

3.2.1. Materiał językowy

Ponieważ głównym celem części badawczej tej pracy jest porównanie realizacji tonów przez rodzimych i nierodzimych użytkowników języka mandaryńskiego, niezbędne było uzyskanie w obu przypadkach materiału jednolitego pod względem treści. Z tego względu ponownie wykorzystano zgromadzony wcześniej materiał językowy w niezmiennym kształcie (por. § 3.1.1).

3.2.2. Uczestnicy

W nagraniu udział wzięły cztery osoby narodowości polskiej: trzy kobiety, w tym dwie studentki Uniwersytetu Adama Mickiewicza w Poznaniu i jedna wykładowczyni w Instytucie Językoznawstwa UAM oraz jeden słuchacz studiów doktoranckich, również w Instytucie

Językoznawstwa UAM, z których wszystkie wykazują czynną i bierną znajomość języka mandaryńskiego na wysokim poziomie. Każda z nich regularnie posługuje się nim w życiu zawodowym lub w ramach zajęć z praktycznej nauki języka. Ponadto, osoby te wykazują praktyczne umiejętności w zakresie zarówno ustnego, jak i pisemnego tłumaczenia polsko-chińskiego oraz chińsko-polskiego, a także przynajmniej raz przebywały przez dłuższy czas w kraju, w którym mandaryński jest językiem urzędowym: trzy z nich odbyło roczne stypendium w Chińskiej Republice Ludowej, natomiast jedna uczestniczyła w 6-tygodniowym kursie języka mandaryńskiego na Tajwanie.

Poszczególnych uczestników oznaczono przy użyciu notacji z poprzedniego eksperymentu, złożoną z inicjału oraz litery sygnalizującej płeć osoby, otrzymując grupę w składzie: AK, EK, GK oraz TM.

3.2.3. Przebieg eksperymentu, edycja oraz analiza danych

Podczas przeprowadzania eksperymentu trzymano się procedury zastosowanej uprzednio (por. § 3.1.3). Zmianie uległo jedynie miejsce nagrań – z powodu prac remontowych, które uniemożliwiły skorzystanie ze studia, wszystkie sesje odbyły się w jednym z pomieszczeń Instytutu Językoznawstwa UAM, co nie wpłynęło znacząco na jakość nagrania. W celu uzyskania materiału do analizy posłużono się mikrofonem, za pomocą którego wprowadzono dźwięk do komputera MacBook, podłączonego do interfejsu Alesis iO26 Firewire. Nagranie rejestrowano przy użyciu wykorzystanego w poprzednim badaniu programu, a następnie utrwalano w taki sam sposób.

Również każda sesja przebiegała według ustalonego wcześniej porządku, z zachowaniem kolejności nagrywanego materiału. Gotowe nagrania poddano edycji, w ramach której przeprowadzono niezbędne korekty oraz podzielono je na mniejsze części.

Podczas analizy postępowano według zastosowanego uprzednio algorytmu, który obejmował odpowiednie ustawienia w programie Praat, sposób dokonywania pomiarów oraz rodzaj pobieranych danych. Materiał rozpatrywano pod kątem identycznych parametrów, z jedną tylko drobną modyfikacją – ponieważ zmianie uległy proporcje grupy, zrezygnowano z kryterium płci; modyfikacja ta nie spowodowała

jednak żadnych istotnych dla wyników badania konsekwencji, jako że w przypadku tej zmiennej niezależnej nie stwierdzono korelacji z jakością realizacji tonów w języku.

3.3. Badanie percepcji słuchowej tonów chińskich wśród nierodzimych użytkowników języka mandaryńskiego

3.3.1. Materiał językowy

W badaniu wykorzystano materiał językowy uzyskany z dokonanych uprzednio nagrań czterech rodzimych użytkowników języka mandaryńskiego i następnie poddany edycji (por. § 3.1.4). W ten sposób otrzymano dwa zestawy po 160 wyrazów jednosylabowych, z których pierwszy zawierał wyłącznie wyrazy języka mandaryńskiego, natomiast drugi – wyrazy polskie. Całość materiału podlegała szczegółowej analizie tonograficznej w poprzednim eksperymencie, badającym produkcję tonów w zależności od wybranych parametrów (por. § 3.1.5), w związku z czym przed przystąpieniem do interpretacji wyników niniejszego testu słuchowego dysponowano pełnym wykazem danych pomiarowych, dotyczących poszczególnych wyrazów. Dane te miały za zadanie ułatwić wskazanie źródła potencjalnych błędów w identyfikacji, a także stanowić podstawę do wnioskowania o ich przyczynach.

Do przeprowadzenia odsłuchów posłużono się programem *Testy audytywne*, autorstwa dra D. Śledzińskiego, z pomocą którego zaprojektowano dwa testy. W tym celu do bazy programu wprowadzono 160 plików odpowiadających pojedynczym wyrazom mandaryńskim z zestawu 1, zapisanych w formacie .wav, które następnie poddano randomizacji. Analogicznie postąpiono z wyrazami polskimi, po czym obie kombinacje zapisano w systemie Windows, będącym oprogramowaniem paralelnym na dysku komputera MacBook.

3.3.2. Słuchacze

W badaniu ogółem wzięło udział 10 studentów języka chińskiego (w odmianie mandaryńskiej) na Wydziale Neofilologii UAM, w tym

dziewięć kobiet i mężczyzna, przy czym osoby te wykazywały znajomość języka na zróżnicowanym poziomie: dwie z podstawową wiedzą oraz umiejętnościami w zakresie posługiwania się językiem mandaryńskim, zdobytymi w czasie rocznej nauki, dwie osoby po studiach licencjackich (trzy lata), z których jedna dodatkowo przebywała 1,5 roku w Chinach, kolejne cztery osoby po 4-letnim okresie nauki (trzy z nich w międzyczasie uczestniczyły w 6-tygodniowym wakacyjnym kursie języka mandaryńskiego na Tajwanie, natomiast pozostała osoba odbyła roczne stypendium w Pekinie) i wreszcie dwie absolwentki studiów magisterskich (pięć lat) – obie po wspomnianym kursie. Spośród badanej grupy dwie osoby zdobyły ponadto pewne doświadczenie w pracy w charakterze tłumacza ustnego języka mandaryńskiego. Wiedzano także, iż dwie uczestniczki badania mają wykształcenie muzyczne oraz praktyczne umiejętności gry na instrumencie, przy czym jedna z nich realizowała program studiów na poznańskiej Akademii Muzycznej; wiedza ta okazała się użyteczna podczas interpretacji wyników eksperymentu oraz formułowania ostatecznych konkluzji.

Do oznaczenia osób badanych wykorzystano notację zastosowaną w poprzednich eksperymentach, otrzymując grupę w składzie: AK, BK, DK, EK, GK, JK, KK, MK, PK oraz SM.

3.3.3. Przebieg eksperymentu

Słuchaczy poddawano badaniu indywidualnie. Każdy odsłuch obejmował dwa testy: pierwszy w języku mandaryńskim, a następnie identyczny w formie test w języku polskim. Zadaniem uczestników była identyfikacja poszczególnych tonów dobiegających z słuchawek podłączonych do komputera MacBook jako ton 1, 2, 3 lub 4. Test rozpoczynał się za naciśnięciem przez słuchacza odpowiedniego przycisku na ekranie komputera, co wywoływało pierwszy sygnał. Każdy wyraz odtwarzany był dwukrotnie z zachowaniem 2-sekundowej pauzy, po czym na ekranie pojawiały się cztery przyciski ponumerowane od 1 do 4, odpowiadające kolejnym tonom mandaryńskim. Naciśnięcie dowolnego z nich powodowało automatyczne wygenerowanie następnego sygnału. W celu zminimalizowania liczby ewentualnych pomyłek, wynikających z obsługiwania klawiatury, odpowiedzi udzielano przy

użyciu kursora sterowanego za pomocą panelu dotykowego *touchpad*, poprzez naciśnięcie jego lewego przycisku w odpowiednim polu ekranu; z tego względu, a także z uwagi na fakt, że czas na zastanowienie nie był ograniczony, zrezygnowano z wprowadzenia funkcji umożliwiającej powtórne odtworzenie sygnału. W czasie trwania testu zaznaczane przez słuchaczy odpowiedzi były automatycznie zapisywane w docelowym pliku tekstowym oraz – na podstawie przypisanych im wcześniej wartości – oceniane jako prawidłowe (*TRUE*) lub błędne (*FALSE*). Test kończył się z chwilą udzielenia ostatniej odpowiedzi. Pomimo że słuchaczom przysługiwał praktycznie nieograniczony czas, poszczególne sesje zajęły nie dłużej niż 30 minut każda.

3.3.4. Analiza

W poszukiwaniu zależności dane analizowano pod kątem trzech parametrów: indywidualnych kompetencji poszczególnych słuchaczy w zakresie znajomości języka mandaryńskiego, jakości i sposobu realizacji tonów przez każdego z czterech mówców oraz typu tonu (dokładniejszy opis procedury związanej z badaniem wpływu poszczególnych zmiennych zamieszczono poniżej). Do tworzenia zestawień wyników uzyskanych przez słuchaczy oraz obliczania procentowych wykazów poprawnych identyfikacji tonów wykorzystano program Microsoft Excel for Mac.

A. kompetencje językowe odbiorcy

Z uwagi na niemożność wyodrębnienia grup słuchaczy, które charakteryzowałyby się podobnymi warunkami i możliwościami uzyskiwania wiedzy i umiejętności praktycznych w zakresie posługiwania się językiem mandaryńskim, wynikającą choćby z dużego zróżnicowania liczby lat nauki oraz ewentualnym kształceniem w krajach chińskojęzycznych na różnych etapach studiów, każdy przypadek rozpatrywano indywidualnie. Informacje uzyskane od uczestników badania pozwoliły na uwzględnienie zarówno długości kształcenia danej osoby, jej pobytu w Chinach lub na Tajwanie, uczestnictwa w kursie lub realizowania programu stypendium, jak i wykonywanie przez nią pracy tłumacza, a także pewnych predyspozycji słuchowych, które mogły wzmocnić

się w wyniku kształcenia muzycznego. W ramach tego parametru dla każdego słuchacza obliczono ogólną liczbę prawidłowo zidentyfikowanych tonów, co umożliwiło zbadanie ewentualnych korelacji pomiędzy branymi pod uwagę czynnikami a zdolnością do trafnej identyfikacji. W analizie uwzględniono również zgodność ocen słuchaczy z faktyczną realizacją danego tonu, gdyż w wielu przypadkach powodem błędnych identyfikacji były zniekształcenia przebiegów powstałe na skutek błędów w ich produkcji. Końcowe rezultaty, uzyskane przez słuchaczy w każdym języku, wyrażono w procentach, a następnie zestawiono w celu porównania.

B. wpływ nadawcy

Badano poprawność identyfikacji wśród poszczególnych mówców. W tym celu odpowiedzi udzielane przez słuchaczy rozpatrywano ze względu na nadawcę, po czym zsumowano liczbę wszystkich prawidłowych identyfikacji wypowiedzianych przez nich tonów.

C. rodzaj tonu

Z korpusu danych wyekstrahowano liczbę wszystkich tonów 1, 2, 3 oraz 4, prawidłowo zidentyfikowanych przez ogół słuchaczy. Wyniki uzależniono od języka, w którym realizowane były sylaby z danym tonem.

Sporządzono również wykaz poprawnych identyfikacji poszczególnych tonów zrealizowanych na każdej z 10 homograficznych sylab, co pozwoliło na wskazanie wyrazów, które uzyskały największą oraz najmniejszą zgodność ocen słuchaczy. Na koniec zsumowano wszystkie prawidłowe odpowiedzi udzielone w każdym z testów z osobna, otrzymując ogólną liczbę poprawnie zidentyfikowanych tonów w języku mandaryńskim i polskim.

Rozdział 4.

Uzyskane rezultaty

W kolejnych częściach niniejszej pracy znajduje się szczegółowa analiza wyników każdej z badanych grup z uwzględnieniem podziału na język mandaryński i polski. Wybrane dane z obu eksperymentów zamieszczono w tabelach zbiorczych (25–28), a następnie zestawiono w celu porównania (por. § 4.3).

4.1. Realizacja tonów chińskich przez rodzimych użytkowników języka mandaryńskiego

4.1.1. Część I – język mandaryński

Wyrazy w języku mandaryńskim, jako rodzimym języku osób badanych, z uwagi na fakt bycia naturalnymi nośnikami tonów poddano analizie jako pierwsze, co pozwoliło na poznanie specyfiki poszczególnych głosów, pewnych tendencji, które mogły, lecz nie musiały uwidocznić się także w języku polskim, i dzięki temu dostarczyło adekwatnej bazy do porównania.

A. F_0

Wyniki pomiarów F_0 dla każdego z głosów zaprezentowano w tabeli 1.

Tabela 1. Średnie wartości F_0 wyrazów chińskich realizowanych przez rodzimych użytkowników.

MÓWCA		JK	WK	DM	YM
PARAMETR					
F0 min. [Hz]		99,9	100,2	75	98,8
F0 max. [Hz]		328,1	423,6	201,9	220,9
skala		228,2	323,4	126,9	122,1
śr. F0	ogółem	204,1	246,6	131,8	143,9
	TON 1	255,9	307,6	164,4	168,1
	TON 2	200,7	243,8	137,5	135,4
	TON 3	148,5	182,8	92,3	123,3
	TON 4	211,6	252,2	133	148,9
zakres F0	TON 1	33,3	49,8	21,5	21,6
	TON 2	82,8	137,4	39,6	71,7
	TON 3	88,3	135,0	39,9	59,7
	TON 4	168,4	192,8	96,3	89,3

Tabela 2 przedstawia odchylenia standardowe (SD) każdego z głosów liczone ogółem i z uwzględnieniem poszczególnych tonów oraz współczynnik modulacji częstotliwości (WMC), stanowiący stosunek ogólnej wartości odchylenia standardowego do średniej arytmetycznej częstotliwości podstawowej wyrażony w procentach.

Tabela 2. Odchylenia standardowe głosów rodzimych użytkowników języka mandaryńskiego (tony realizowane na wyrazach chińskich) wraz z współczynnikiem modulacji częstotliwości.

MÓWCA σ	JK	WK	DM	YM
ogółem	40,6	46,7	26,5	17,7
TON 1	9,6	10,9	5,0	4,6
TON 2	10,3	11,3	5,0	5,5
TON 3	8,8	12,7	1,9	4,3
TON 4	19,4	18,9	6,0	8,0
WMC [%]	19,9	18,9	20,1	12,3

Ogółem, wszystkie głosy zmieściły się w przypisanych im skalach częstotliwości, choć w przypadku jednego z głosów kobiecych (JK) zanotowano jednorazowy spadek wartości F_0 do 99,9 Hz. Wartość ta nie może być wynikiem niedokładności powstałej na skutek manualnego dokonywania pomiarów, gdyż dotyczy ona punktu zwrotnego jednego z tonów 3, jest jednak na tyle nieznaczna, że można ją potraktować jako wartość minimalną na skali głosów żeńskich. Potwierdza to pogląd wielu językoznawców wzorujących się na notacji Chao (por. § 1.1), zgodnie z którym ton 3 osiąga w jednym ze swych ekstremów dolny próg skali. Jak wyraźnie widać w wynikach pomiarów, wartości minimalne tonów realizowanych przez osobę JK często oscylują wokół 100 Hz, natomiast maksima nigdy nie przekraczają 330 Hz, co pozwala stwierdzić, że osoba ta ma głos niski, w związku z czym wykorzystuje dolny rejestr skali głosów żeńskich. Wśród mężczyzn mówcę DM podobnie charakteryzuje zauważalnie niski głos, na co wskazuje częste osiąganie przez $F_{0\min}$ wypowiedzianych przez niego wyrazów wartości krańcowych bądź zbliżonych do krańcowych, jedynie sporadyczne przekraczanie progu 180 Hz w górnym rejestrze, a także średnie wartości F_0 , które w większej części plasują się w dolnej połowie skali.

W pozostałych głosach wartości te są na ogół wyższe adekwatnie do odpowiedniej skali, choć w przypadku kobiet różnice są wyraźniejsze

i bardziej regularne. Porównując średnie wysokości F_0 dla poszczególnych tonów, można zauważyć, że badana WK konsekwentnie realizowała każdy ton z częstotliwością o około 40 Hz wyższą niż badana JK. WK posługuje się także skalą o rozpiętości niemalże o 100 Hz szerszą niż JK, a co za tym idzie – często osiąga wyższe wartości, sięgające 350 Hz z jednorazowym absolutnym maksimum równym 423,6 Hz. U badanych mężczyzn stwierdzono porównywalne wysokości tonów 1 i 2, natomiast istotną różnicę zanotowano w przypadku tonu 3, który mówca YM konsekwentnie realizował ze średnią wysokością o około 30 Hz niższą niż DM. Z kolei różnica rzędu średnio 15 Hz, jaka występowała w realizacjach tonu 4 pomiędzy analizowanymi głosami męskimi, nie jest wielkością znaczącą, gdyż w odbiorze słuchowym jest ona ledwo wykrywalna. Porównywalna jest także ogólna rozpiętość obu głosów męskich, która wynosi nieco ponad 120 Hz.

W kwestii zakresu częstotliwości z uwzględnieniem podziału na poszczególne tony różnice przedstawiają się w sposób nieregularny, wciąż jednak można zauważyć pewne ogólne prawidłowości. Otóż głosy sklasyfikowane tutaj jako wyższe (WK i YM) poza jednym drobnym wyjątkiem realizują tony wykorzystując szerszą rozpiętość na skali F_0 , tj. w taki sposób, że odległość pomiędzy ekstremami konturów tonalnych jest większa. U kobiet największa różnica między głosami występuje dla tonu 2 i 3, w obu przypadkach wynosząc około 50 Hz, natomiast u mężczyzn jest ona nieznaczna. Można również stwierdzić, że zgodnie z obserwacjami Chao (1970) największym zakresem F_0 charakteryzuje się ton 4, przy czym w głosach męskich wartości te, choć nie osiągną 100-proc. zgodności, są bardziej zbliżone do wykorzystywanych przez te głosy skali, oszacowanych na podstawie wyników niniejszej analizy. Zdecydowanie najmniejszy zakres F_0 spośród czterech tonów każdy z mówców wykorzystuje przy produkcji wyrazów z tonem 1, co jest widoczne także na odpowiadających im tonogramach. Tomy te mają równy przebieg, bez zmian kierunku i wyraźnych skoków²³. Co ciekawe, wszystkie osoby badane wykazują tendencję do realizowania kolejnych tonów 1 na bardzo zbliżonej wysokości. Zjawisko to najprawdopodobniej można tłumaczyć faktem, że wyrazy te występowały w izolacji w powtarzającym się cyklu, regularnie co cztery jednostki

²³ Wszelkie gwałtowne wahania zarejestrowane na tonogramach, często będące wynikiem bliskości segmentów bezdźwięcznych, zostały skorygowane.

językowe, co w połączeniu ze stałym rytmem i tempem mowy mogło ułatwić utrzymanie niemal jednakowej częstotliwości. W ciągłej mowie potocznej, w której dodatkowym czynnikiem są emocje oraz często realizowany jest akcent kontrastywny, wysokości te najpewniej uległyby znacznemu urozmaiceniu.

Największe międzyosobnicze zróżnicowanie panuje wśród wyrazów z tonem 3. Najogólniej rzecz biorąc, można zaobserwować dużą różnicę związaną z płcią mówcy – mianowicie kobiety posługują się znacznie szerszą skalą niż mężczyźni, co wprawdzie jest zjawiskiem bardziej uniwersalnym, widocznym także w realizacjach innych tonów, a przede wszystkim faktem wielokrotnie stwierdzanym w literaturze (G.T. Chen 1974, S. Chen 2005, Keating & Kuo 2010, Yip 2002, za: Yuan i Chen 2014), jednak w przypadku tonu 3 i 2 wyniki są najbardziej rozproszone. Znaczna rozbieżność została odnotowana pomiędzy wyrazami wypowiedzianymi przez badane kobiety; powodem tego jest niepełna realizacja wszystkich bez wyjątku tonów trzecich przez badaną JK, które są pozbawione części rosnących. Podobna sytuacja ma miejsce w głosie DM – tu również 100% tonów opadająco-wznoszących zredukowanych jest o połowę. Z tego względu skala używana do ich realizacji przez DM ma średnią rozpiętość rzędu 40 Hz. Odwrotną skłonność prezentuje osoba YM, która wszystkie te wyrazy realizuje starannie, z nad wyraz dużą dokładnością. Tendencja ta uwidacznia się na tonogramach, które charakteryzuje wyraźnie opadająco-wznoszący przebieg z widocznym punktem zwrotnym, sięgającym najniższego rejestru skali. Również przeważająca większość tonów trzecich wypowiedzianych przez WK ma przebieg zgodny z oczekiwanym, co można stwierdzić już przy powierzchniowej analizie wzrokowej. Kontury tonów rosnących w jej wykonaniu mają kształt ukośnych linii wznoszących się proporcjonalnie w czasie, o maksimach zlokalizowanych w górnej części skali. Osoba ta używa także skali o zdecydowanie największej rozpiętości równej niemal 140 Hz, zarówno w przypadku tonów 3, jak i 2. Podobnie jak wcześniej, najwęższy zakres F_0 stosuje osoba DM. Zestawienie danych dotyczących rozpiętości skali poszczególnych głosów z wartościami ich odchyłeń standardowych pozwala stwierdzić, że w przypadku niniejszego eksperymentu fakt posługiwania się szerokim spektrum częstotliwości przełożył się na małą stabilność głosów, co miało miejsce w obu głosach żeńskich.

B. Iloczas

Następnym w kolejności parametrem poddanym analizie były właściwości temporalne tonów z nagranych materiału. Wyniki zamieszczone są w tabeli 3.

Tabela 3. Iloczasy tonów wyrazów chińskich wypowiedzianych przez rodzimych użytkowników.

MÓWCA λ [ms]	JK	WK	DM	YM
TON 1	515	602	341	548
TON 2	535	599	311	521
TON 3	432	563	246	710
TON 4	391	271	234	331

Pierwsze spostrzeżenie, jakie nasuwa się przy wstępnych oględzinach tonogramów, dotyczy długości sylab z tonem 4, które są wyraźnie krótsze od pozostałych. Fakt ten stanowi potwierdzenie panującego w językoznawstwie poglądu, co do którego większość badaczy jest zgodna. Względnie najkrótsze – bo o blisko 50% od średniej długości pozostałych tonów – są tony 4 wypowiedziane przez osobę WK, najdłuższe są one w wykonaniu JK, natomiast najkrótsze przebiegi wykazują tony 4 badanego DM. Wśród ogółu tonów realizowanych przez DM występuje najmniejsze zróżnicowanie pod względem długości, bowiem najdłuższe tony 1 są średnio zaledwie o 100 ms dłuższe od tonów 4. Wszystkie cztery tony DM są także krótsze od odpowiednich tonów pozostałych osób badanych. Zaobserwowano natomiast, że sylaby [jen] (*yàn*) oraz [maj] (*mài*) z tonem 4, w przypadku każdej z osób są zauważalnie dłuższe od pozostałych sylab noszących ten ton i wypowiedzianych przez daną osobę. Zjawisko to trudno wytłumaczyć, zważywszy na fakt, że sylaba [gen] (*gèn*) o identycznym jak *yàn* wygłosie takiej tendencji nie wykazuje; być może wpływ na zwiększenie iloczasu samogłoski ma

występowanie przed nią spółgłoski półotwartej. Natomiast u wszystkich mówców na ogół bardzo zbliżoną długością odznaczają się tony 1 i 2.

W przypadku niniejszej analizy nie sprawdziła się teza Woo (1969, za: Anderson 1982), zgodnie z którą izolowane sylaby z tonem 3 są zazwyczaj o 50% dłuższe od sylab noszących pozostałe tony. Rezultat ten został osiągnięty jedynie przez mówcę YM, u którego stosunek długości tonu 3 do średniej arytmetycznej długości tonów 1, 2 i 4 wynosi około 52%. U osoby WK stosunek ten równy jest zaledwie 15%, podczas gdy u pozostałych dwóch osób (JK oraz DM) wartość tej średniej przewyższa przeciętną długość tonu 3. W obu przypadkach powodem tego jest niepełna realizacja wszystkich sylab noszących ton 3, co na tonogramach uwidacznia się w postaci braku części rosnących. Można przypuszczać, że zjawisko to jest jedną z dwóch form procesu znanego pod nazwą tonalnego *sandhi* (por. rozdz. 2). W formie tej *sandhi* zakłada redukcję tonu 3 o część rosnącą w sytuacji, gdy poprzedza on sylabę z tonem 1, 2 lub – jak w tym wypadku – 4, dając w efekcie ton $\frac{1}{2}$ 3. Jest to o tyle ciekawe, że proces ten występuje na ogół w mowie ciągłej, co najmniej w obrębie segmentów dwusylabowych, natomiast nigdy w przypadku sylab izolowanych. Zatem być może jest to spowodowane potraktowaniem przez niektórych badanych sekwencji czterech tonów jako pewnej całości, zakończonej tonem 4, przed którym występuje sylaba z tonem 3, stanowiąca jedynie element tejże struktury wyższego rzędu. Wówczas następujące po sobie tony mogłyby podlegać procesom występującym w mowie ciągłej. Zastanawia również fakt, że zjawisko *sandhi* tonu 3 dotyczy wyłącznie tonów wypowiedzianych przez Tajwańczyków, natomiast nie pojawia się ani razu u osoby z Chin (YM), co może sugerować istnienie pewnych drobnych różnic pomiędzy językiem mandaryńskim w Chinach i na Tajwanie. Możliwym i bodaj najbardziej prawdopodobnym wyjaśnieniem tego faktu jest wpływ tempa mowy na zmianę segmentacji jednostek leksykalnych występujących w sekwencji, które w miarę jego zwiększania mogą wykazywać tendencję do tworzenia coraz bardziej rozbudowanych struktur o charakterze fraz intonacyjnych. Wówczas w szybkim tempie niektóre jednostki – szczególnie te najbardziej złożone – mogłyby ulec redukcji; proces ten byłby zatem analogiczny do zjawiska utraty akcentu przez niektóre sylaby ortotoniczne, typowego dla szybkiego stylu potocznego, o czym wspomina w swojej rozprawie Dłuska (1976). Wy tłumaczenie

to znajduje potwierdzenie w niniejszej analizie, gdyż mówca YM, który wszystkie sylaby z tonem 3 wypowiedział modelowo, mówił względnie wolnym tempem, w przeciwieństwie do mówców JK i DM. Efektem tego jest zwiększenie iloczynów tonów opadająco-wznoszących YM do wyjątkowo dużych wartości, przewyższających średnie długości odpowiednich tonów w wykonaniu pozostałych osób o co najmniej 20%. Z kolei osoba WK, u której wpływ *sandhi* zanotowano w 30% realizacji, charakteryzowała się umiarkowanym tempem mowy.

C. Poprawność realizacji

Tabela 4 prezentuje wyniki analizy tonogramów sylab chińskich.

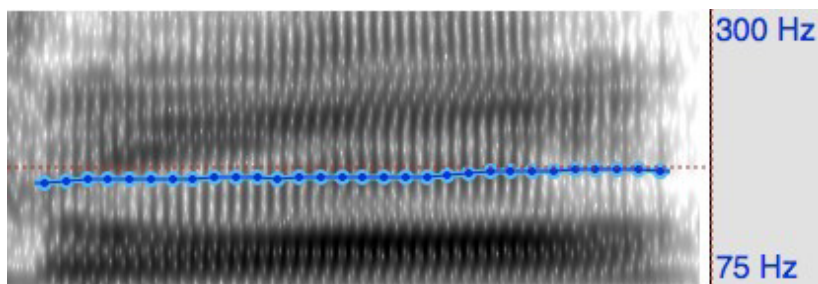
Tabela 4. Wykaz prawidłowo zrealizowanych tonów na sylabach chińskich wypowiedzianych przez rodzimych użytkowników.

TON	LICZBA POPRAWNYCH REALIZACJI			
	JK	WK	DM	YM
1	10	10	10	10
2	6	9	10	3
3	0	7	0	10
4	10	10	10	10
%	65	90	75	82,5
K vs. M [%]	77,5		78,75	

Jak pokazują dane zamieszczone w tabeli 4, ogółem najwięcej tonów prawidłowo zrealizowała osoba WK, osiągając tym samym poprawność rzędu 90%. Do rezultatu tego zbliżył się mówca YM z wynikiem 82,5%, następnie DM – 75% oraz JK – 65%. Przy uwzględnieniu kryterium płci rezultaty są niemal identyczne, z nieznaczną przewagą na korzyść mężczyzn, która przekłada się na jeden prawidłowo wypowiedziany ton więcej. Na tej podstawie można wywnioskować, że płeć nie była w niniejszym badaniu czynnikiem mającym znaczący wpływ na wyniki.

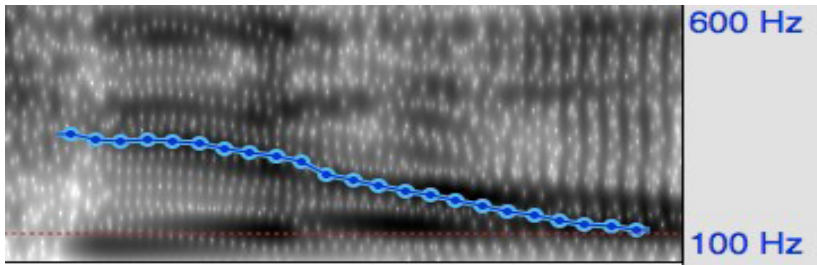
Przyjmując za kryterium oceny poprawności realizacji tonów kształt ich konturów, pierwsze ogólne spostrzeżenie, jakie można poczynić przy analizie tonogramów zarejestrowanego materiału językowego dotyczy tonów, które okazały się najmniej problematyczne dla osób uczestniczących w badaniu, mianowicie – tonów 1 i 4. Oba te tony, wypowiedziane na 10 różnych sylabach, zostały w 100% poprawnie wypowiedziane przez wszystkich czterech mówców. Tony 1 w wykonaniu badanych mają równy, niezakłócony przebieg, utrzymujący się na względnie stałej częstotliwości. Na ryc. 1 widoczny jest obraz graficzny tonu 1 zrealizowanego na sylabie [ʃwo] (*shuō*) wypowiedzianej przez DM.

Ryc. 1. Prawidłowy przebieg tonu 1 – sylaba [ʃwo] (*shuō*).



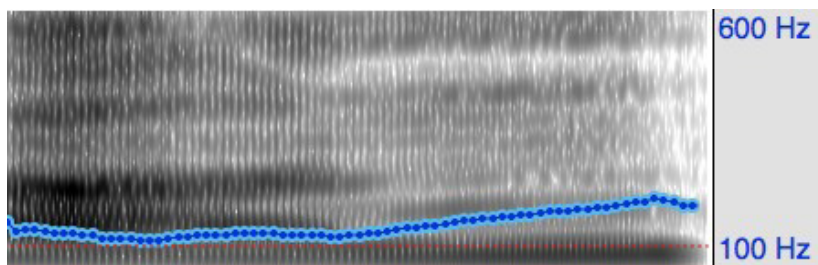
Kontury tonów 4 u niektórych osób charakteryzuje bardziej gwałtowny spadek, u innych zaś znacznie łagodniejszy, jednak zawsze mają one kierunek wyraźnie opadający. Niekiedy na linii obrazującej przebieg tonu zdarzają się drobne wahania, będące wynikiem obecności w danej sylabie głosek nosowych bądź bezdźwięcznych przydechowych. Spektrogram na ryc. 2 przedstawia realizację tonu 4 na sylabie [təow] (*qiū*) przez badaną WK.

Ryc. 2. Prawidłowy przebieg tonu 4 – sylaba [təow] (qiu).



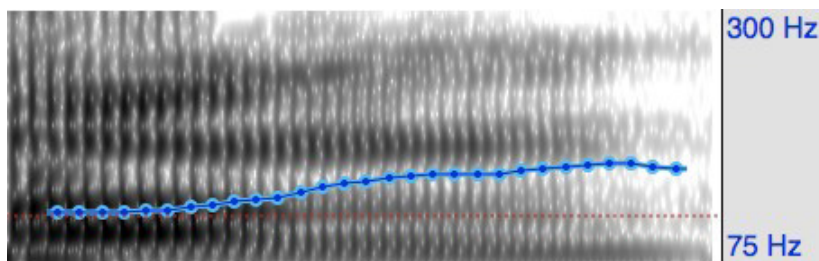
Ogółem rzecz biorąc, zdecydowanie najwięcej trudności sprawiały mówcom sylaby z tonami 2 i 3, choć byli wśród nich tacy, którzy poprawnie wypowiedzieli wszystkie wyrazy z jednej z tych grup. Ton 2, podobnie nieskomplikowany w swej formie jak 4, charakteryzujący się regularnym przyrostem F_0 w czasie, z założenia nie powinien być dla rodzimych użytkowników języka mandaryńskiego problematyczny w realizacji. Jak się jednak okazuje, obraz graficzny wyrazów z tonem 2 wypowiedzianych przez osoby badane dość często przybierał kształt nieco zaokrąglony, bardziej zbliżony do kształtu tonu 3, z zaznaczonym spadkiem i punktem zmiany kierunku. Niejednokrotnie trudno było zdecydować, czy dany ton można uznać za wypowiedziany poprawnie, czy też nie, tym bardziej, że niektóre kontrowersyjne przypadki nie budziły większych wątpliwości co do ich ewidentnie rosnącego charakteru w odbiorze słuchowym. I tak jedyną osobą, która wypowiedziała wszystkie 10 sylab z tonem 2 w sposób poprawny, był DM; kontury tonalne tych sylab nie wykazują spadków ani punktów zwrotnych. Bardzo dobrą realizacją wykazała się także osoba WK, u której zanotowano tylko jeden przypadek tonu o konturze opadająco-wznoszącym. Wyraźne trudności z prawidłową realizacją można natomiast stwierdzić w przypadku badanej JK, co jest widoczne na tonogramach odpowiednich tonów. U JK blisko połowa (4 na 10) sylab ma kontury o dyskusyjnych przebiegach. Ryc. 3 przedstawia jeden z przykładów – sylabę [t^han] (tán).

Ryc. 3. Wątpliwy przebieg tonu 2 – sylaba [tʰan] (*tán*).



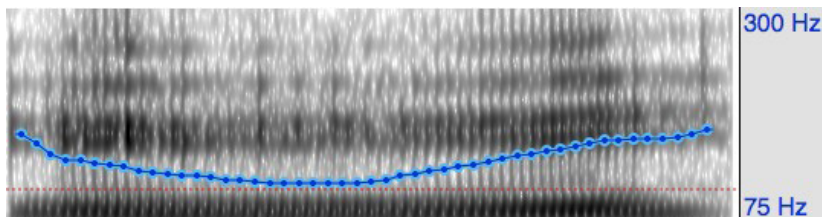
Ryc. 4 prezentuje ton 2 o prawidłowym przebiegu, wypowiedziany przez mówcę DM na sylabie [pʰeŋ] (*péng*).

Ryc. 4. Prawidłowy przebieg tonu 2 – sylaba [pʰeŋ] (*péng*).



Zdecydowanie najslabiej z realizacją tonów 2 radzi sobie badany YM, gdyż aż 7 z 10 sylab w jego wykonaniu odznacza się wątpliwą lub nieprawidłową realizacją, z czego w czterech przypadkach kontury tonalne mają wyraźny kształt o przebiegu typowym dla tonu 3. Błędnie zrealizowane tony w znacznej większości przekładają się na błędy w percepcji, bowiem przez wielu słuchaczy zostają one zidentyfikowane jako tony 3. Ryc. 5 ilustruje nieprawidłową realizację sylaby [ɣ:] (*yú*).

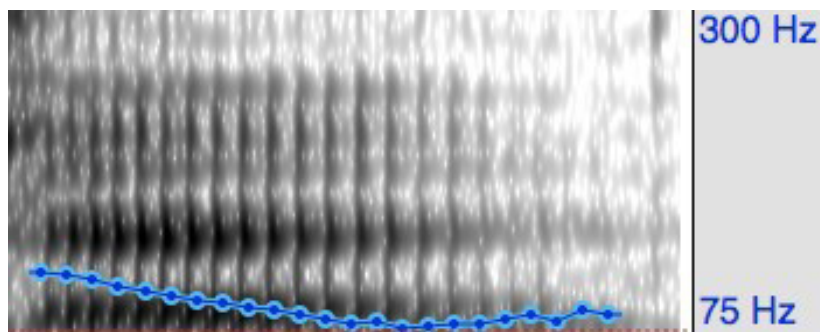
Ryc. 5. Nieprawidłowy przebieg tonu 2 – sylaba [ɣ:] (yú).



Jeszcze bardziej problematyczny okazuje się ton 3, który w sumie został błędnie zrealizowany aż w 57,5%. Podobnie jak w przypadku tonu 2, przeważa jeden typ zniekształceń – oprócz kilku konturów tonalnych o nieokreślonym kształcie, wszystkie pozostałe są pozbawione części rosnącej. W wyniku tego ich kontury przybierają formę nieco zbliżoną do tonu 4, lecz o znacznie łagodniejszym spadku, krótszym przebiegu i mniejszym zakresie F_0 . W percepcji trudno przyrównać je do jakiegokolwiek innego tonu, bowiem sprawiają wrażenie niedokończonych, jednak z uwagi na swój zdecydowanie opadający kierunek, a także z braku bardziej adekwatnego odpowiednika najczęściej zostają sklasyfikowane przez słuchaczy jako tony 4. Dwoje mówców (JK i DM) zrealizowało wszystkie tony 3 jako tony $\frac{1}{2}$ 3, przy czym tony w wykonaniu DM mają względnie regularną formę, podczas gdy kształty tonów JK w większości są pełne drobnych wahań i załamań, co jest wynikiem częstej obecności dźwięczności typu *creaky voice* w dolnym rejestrze tego głosu. U badanej WK w 30% sylab z tonem 3 zanotowano wystąpienie zjawiska *sandhi*.

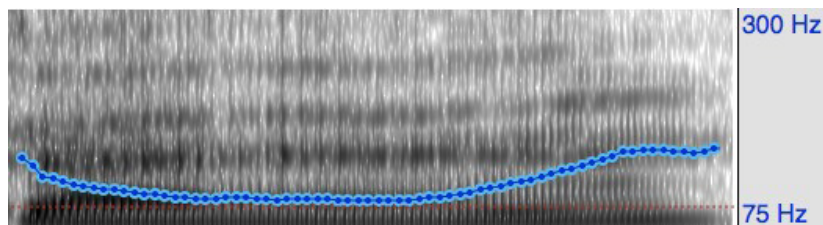
Spektrogram na ryc. 6 ilustruje ton $\frac{1}{2}$ 3 zrealizowany na sylabie [ʂɻ] (*shi*) przez osobę DM.

Ryc. 6. Przebieg tonu ½ 3 – sylaba [ʃɿ] (*shi*).



Dla porównania na ryc. 7 przedstawiono ton 3 w formie kanonicznej, wypowiedziany przez badanego YM na sylabie [jen] (*yǎn*):

Ryc. 7. Modelowy przebieg tonu 3 – sylaba [jen] (*yǎn*).



Spośród wyrazów najwięcej trudności sprawiały mówcom: *qǐu*, *gěh* oraz *tǎn* wypowiedziane błędnie przez trzy na cztery osoby; brak natomiast wyrazów, które zostałyby zrealizowane nieprawidłowo przez wszystkich badanych.

Pomimo że osoby te posługiwały się w badaniu swoim rodzimym językiem, u każdej z nich można dostrzec różne tendencje do realizowania pewnych tonów w sposób poprawny, zaś innych błędny. Z tego względu trudno orzekać o wpływie indywidualnych predyspozycji i zdolności. Niewątpliwie do powstania choćby części błędów w wymowie przy-

czyniła się sztuczność warunków laboratoryjnych, a także do pewnego stopnia nienaturalność materiału językowego, jako że na ogół posługujemy się mową w formie ciągłej, natomiast nie izolowanymi wyrazami. Powyższe wyniki nie świadczą też o niedostatecznej kompetencji językowej żadnej z osób biorących udział w nagraniu. Należy wziąć pod uwagę, że w językach tonalnych ton (lub jego brak, czyli tzw. ton zerowy – por. § 1.1) stanowi nieodłączny element każdej sylaby, tworząc z nią całość, która ma pewne konkretne znaczenie semantyczne bądź gramatyczne. W konsekwencji akwizycja tonów jako elementów umożliwiających różnicowanie poszczególnych jednostek leksykalnych odbywa się równoległe z akwizycją pojedynczych sylab – jednostek TBU. W związku z tym w przypadku dorosłych, wykształconych osób, dla których język mandaryński jest językiem natywnym, należy wykluczyć możliwość, że błędy w realizacji były wynikiem niewiedzy. Bardzo prawdopodobne, że przyczyną gorszej realizacji tonów 2 i 3 jest fakt ich częstego podlegania modyfikacjom wynikającym z procesów działających w mowie ciągłej, do których badani z pewnością są przyzwyczajeni i mogli zastosować je w sposób nieświadomy.

4.1.2. Część II – język polski

A. F_0

Średnie wyników badania, wartości skrajne oraz rozpiętość skali dla każdego głosu zestawiono w tabeli 5.

Tabela 5. Średnie wartości F_0 wyrazów polskich realizowanych przez rodzimych użytkowników języka mandaryńskiego.

MÓWCA		JK	WK	DM	YM
PARAMETR					
F0 min. [Hz]		113,1	101,5	75,1	97,6
F0 max. [Hz]		345,6	486,4	193,5	260,6
skala		232,5	384,9	118,4	163
śr. F0	ogółem	221,1	263,5	123,6	150,8
	TON 1	268	286,3	150,2	165,7
	TON 2	224,4	277,1	133,5	144,6
	TON 3	172,0	268,5	93,6	132,5
	TON 4	220,2	222,3	117,3	160,3
zakres F0	TON 1	58,8	149,8	28,1	56,3
	TON 2	68,9	175	35	70,3
	TON 3	78,6	178,7	30,9	68
	TON 4	153,8	157,1	69,8	83,7

W tabeli 6 znajdują się wartości odchyłeń standardowych oraz współczynników modulacji częstotliwości poszczególnych głosew.

Tabela 6. Odchylenia standardowe głosew rodzimych użytkowników języka mandaryńskiego (tony realizowane na wyrazach polskich) wraz z współczynnikiem modulacji częstotliwości.

MÓWCA		JK	WK	DM	YM
σ					
ogółem		37,8	30,4	23,1	16,9
TON 1		12,6	17,7	11,2	17,4
TON 2		12,4	16,7	10,7	6,1
TON 3		15,1	20,0	8,2	5,6
TON 4		22,6	17,8	7,9	10,3
WMC [%]		17,1	11,5	18,7	11,2

Podobnie jak w przypadku wyrazów chińskich, wszystkie głosy mieściły się w odpowiednich skalach częstotliwości, przy czym absolutna minimalna wartość F_0 , zanotowana w całym nagraniu dla poszczególnych osób, u każdej z nich praktycznie odpowiada dolnej granicy przypisanej danej osobie skali; wysokości maksymalnej nie osiągnęła natomiast żadna z nich. U kobiet, WK wypowiada tony w wyższych niż JK częstotliwościach, używa także skali o rozpiętości 384,9 Hz, czyli o przeszło 150 Hz szerszej niż JK, z ekstremum o wartości 486,4 Hz. Ogółem, w głosie WK dominują dźwięki o wysokich częstotliwościach, podczas gdy wartości średnie oscylują wokół 260 Hz. U badanej JK średnia wartość F_0 wynosi około 220 Hz. Osoba ta posługuje się skalą o względnie wąskim zakresie, wynoszącym 232,5 Hz, z punktami krańcowymi o wartościach F_{0min} równym 113,1 Hz i F_{0max} – 345,6 Hz. W obu głosach występował duży rozrzut wartości F_0 , co znacznie utrudnia ich porównanie pod tym względem, niemniej na podstawie danych surowych można stwierdzić, że tony osoby WK na ogół zaczynają się w wyższym rejestrze, natomiast równie często opadają do wartości niższych. Analogiczna sytuacja ma miejsce w głosach męskich; choć mówca YM na ogół wypowiada tony w wyższych niż DM częstotliwościach, różnice te ulegają zatarciu w wartościach średnich arytmetycznych, czego przyczyną jest stosowanie przez YM nieco szerszej skali. I tak, w produkcji tonów YM wykorzystuje zakres 163 Hz o wartości szczytowej wynoszącej 260,6 Hz, tymczasem DM ani razu nie przekracza progu 200 Hz, posługując się skalą o rozpiętości 118,4 Hz – tj. o blisko 50 Hz węższą niż YM – z maksimum równym 193,5 Hz. Analizując poszczególne tony z osobna, uwagę zwracają tony 3 i 4, z których 4 w realizacji DM opada do znacznie niższych wysokości, charakteryzując się przy tym porównywalnym zakresem do tonu 4 w wykonaniu YM, podczas gdy ton 3 tego mówcy z kolei osiąga wartości wyższe z uwagi na fakt, że w przeciwieństwie do tonu 3 DM, ma on część rosnącą. Usrednione wartości zakresu F_0 nie są natomiast wiarygodnym wyznacznikiem w przypadku tonów 1 i 2, bowiem nie odzwierciedlają one dużych nieregularności w głosie YM – wartości tego parametru wahają się od niespełna 20 Hz do ponad 100 Hz w sylabach z tonem 1 oraz od około 30 Hz do przeszło 120 Hz w sylabach z tonem 2. U DM wahania są niewielkie, a oba te tony realizowane są ze względnie stałą rozpiętością, wynoszącą odpowiednio około 30 Hz oraz 40 Hz.

Natomiast w głosach żeńskich bezsprzecznie szerszym – i zarazem najszerszym spośród wszystkich osób badanych – zakresem odznaczają się tony WK, z których największe wartości osiągają ton 2 i 3 – w obu przypadkach jest to niemal 180 Hz. Ciekawe i jednocześnie zaskakujące jest to, że wszystkie tony w wykonaniu WK charakteryzują się bardzo szeroką rozpiętością F_0 , co szczególnie dziwi w przypadku tonu 1. Przy szczegółowej analizie każdego wyrazu z osobna można zauważyć, że omawiany parametr jedynie sporadycznie osiąga wartość poniżej 100 Hz. Z kolei u osoby JK poszczególne tony charakteryzują się bardziej proporcjonalnymi zakresami, gdzie zdecydowanie największą rozpiętość ma ton 4, na ogół opadający do względnie stałej wartości. Spory i w miarę regularny zakres ma także ton 3, który w realizacji badanej JK jest tonem opadającym, powstałym na skutek redukcji jego kanonicznej postaci o część wznoszącą, co miało miejsce również w sylabach mandaryńskich. Trudno scharakteryzować ton 1, którego realizacja przebiega z dużą nieregularnością, co przejawia się w wartościach od ok. 25 Hz do aż 110 Hz w pierwszym wypowiedzianym wyrazie. Podobnie sytuacja przedstawia się w wyrazach z tonem 2, gdzie zdarzają się skoki do rozpiętości oscylującej wokół 100 Hz, podczas gdy średni zakres wynosi 60 Hz.

Podsumowując, w każdym głosie występują mniej lub bardziej częste wahania F_0 , które trudno ująć w postaci ogólnych tendencji. Największą stabilność pod tym względem zachowuje osoba DM, u której dobrym tego wyznacznikiem są wartości średnie. Najmniejsze zróżnicowanie pomiędzy poszczególnymi tonami występuje w głosie WK, gdzie można zaobserwować pewne zaburzenie proporcji w wyniku zbliżenia wartości zakresów wszystkich tonów. Głosy określone jako niskie na podstawie analizy materiału w języku mandaryńskim (JK oraz DM) zachowały pod tym względem swój charakter także w języku polskim.

B. Iloczasy

Średnie z pomiarów długości tonów znajdują się w tabeli 7.

Tabela 7. Iloczasy tonów wyrazów polskich wypowiedzianych przez rodzimych użytkowników języka mandaryńskiego.

MÓWCA λ [ms]	JK	WK	DM	YM
TON 1	334	363	248	335
TON 2	348	436	250	370
TON 3	262	434	239	499
TON 4	286	280	233	347

Dokonując pobieżnej analizy danych dotyczących średniej długości poszczególnych tonów, na pierwszy rzut oka daje się zauważyć niemal brak zróżnicowania pod tym względem tonów realizowanych przez mówcę DM; w rzeczywistości, wartość uśredniona spłaszcza istniejące wahania czasu trwania segmentów dźwięcznych, bowiem na szczegółowym wykazie przedstawiają się one w sposób bardzo nieregularny, a w percepcji audytywnej różnice tego rzędu są na ogół dość słabo wykrywalne. Niemniej przy bardziej dogłębnej analizie wyłania się pewna interesująca zależność, mianowicie istnieją sylaby, które u każdego mówcy powodują proporcjonalne wydłużenie bądź skrócenie wszystkich tonów wypowiedzianych na tej sylabie. Do wyrazów z tonami o iloczasy dłuższym niż przeciętny należą: *trzymiel*, *hymn*, *lgnąć*, *schnąć* oraz *tchnąć*, natomiast krótszym czasem trwania na ogół charakteryzowały się tony realizowane na wyrazie *czcić*.

W odróżnieniu od tonów na sylabach chińskich tony „polskie” wykazują znacznie mniejsze zróżnicowanie w kwestii właściwości temporalnych, co wynika przede wszystkim z dużych wahań w obrębie tonów jednego rodzaju, w efekcie prowadząc do zaburzenia proporcji pomiędzy nimi. W związku z tym niemożliwe staje się podanie przeciętnych stosunków długości poszczególnych tonów, a tym bardziej

wyszczególnienie dominujących wśród wszystkich badanych tendencji, gdyż parametr ten wykazuje duże międzyosobnicze zróżnicowanie. Tak więc u osoby DM tonem najkrótszym na ogół jest ton 4, choć w stopniu nieznacznym. Podobnie jest w przypadku osoby WK, jednak tu nie ma w tej kwestii żadnych wątpliwości. Spośród tonów w wykonaniu JK w 9 na 10 przypadków ewidentnie najkrótszy jest zredukowany ton 3, podczas gdy u drugiego z głosów męskich (YM) jest to ton 1. YM jest także jedynym głosem, w którym ton 3 jest wyraźnie dłuższy od pozostałych; wprawdzie owa różnica w niewielu przypadkach sięga aż 50%, natomiast z reguły wynosi 150–200 ms więcej niż długość tonu 1. U WK dominuje ton 2, choć kilkakrotnie dłuższy bywa ton 3. Pozostaje jeszcze mówca DM, u którego w zasadzie wszystkie tony są dość krótkie, z wyjątkiem tonu 3 najkrótsze w stosunku do tonów innych osób, gdyż jedynie w 25% przypadków trwają one ponad 300 ms. Brak natomiast tonu o wyraźnie największym iloczynie – zapisy dotyczące długości poszczególnych tonów stanowią szereg przypadkowych wartości, w których trudno dopatrzeć się jakiegokolwiek prawidłowości.

C. Poprawność realizacji

Wyniki poszczególnych mówców przedstawiają się następująco:

Tabela 8. Wykaz prawidłowo zrealizowanych tonów na sylabach polskich wypowiedzianych przez rodzimych użytkowników języka mandaryńskiego.

TON	LICZBA POPRAWNYCH REALIZACJI			
	JK	WK	DM	YM
1	5	2	6	1
2	3	9	7	3
3	0	2	1	10
4	6	6	8	3
SUMA [%]	35	47,5	55	42,5
K vs. M [%]	41,25		48,75	

Pod względem poprawności realizacji zwraca uwagę niski odsetek tonów wypowiedzianych prawidłowo (45%). Najwięcej tonów spośród osób badanych poprawnie zrealizował mówca DM, osiągając rezultat w wysokości 55%. Nieco mniej niż połowę – dokładnie 47,5% ogólnej liczby tonów – zostało wypowiedzianych poprawnie przez osobę WK, natomiast 42,5% – przez osobę YM. Podobnie jak w poprzedniej części badania, najniższą poprawność osiągnęła JK, której rezultat znacznie zaniża konsekwentna realizacja tonu $\frac{1}{2} 3$, w niniejszej analizie uznano za formę niepożądaną w przypadku sylab izolowanych i przez to nieprawidłową. Pod względem płci mężczyźni ponownie okazali się lepsi, choć tym razem różnica ta jest nieco bardziej znacząca i wynosi 7,5 punktu procentowego. Niemniej nierównomierny rozkład danych oraz znaczne rozbieżności pomiędzy liczbą poszczególnych tonów, które zostały prawidłowo zrealizowane przez każdego z mężczyzn, pozwalają stwierdzić, że także i w tej części badania płeć nie była czynnikiem decydującym.

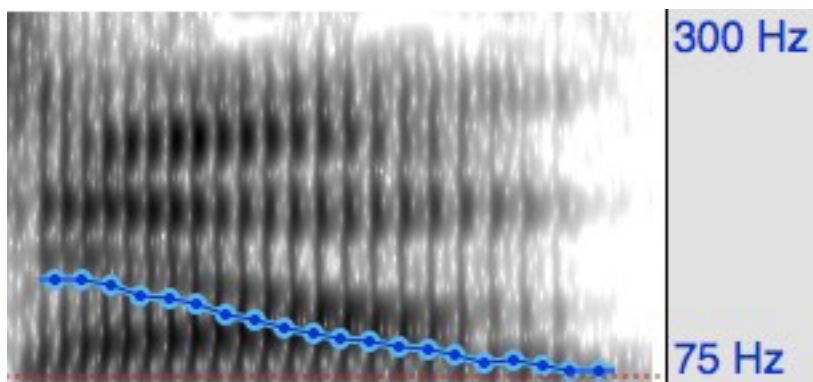
W odróżnieniu od języka mandaryńskiego w zasadzie brak takiego tonu, który nie sprawiałby mówcom trudności w realizacji. W każdym bowiem przypadku można zaobserwować liczne błędy powodujące zniekształcenia konturów tonalnych w stopniu niekiedy uniemożliwiającym identyfikację. Wiele z tych błędów wynika z pewnej tendencji, zaobserwowanej u trzech badanych, która nie tylko przełożyła się na niską poprawność w wynikach analizy, ale też w znacznym stopniu utrudniła dokonywanie pomiarów oraz kodowanie danych. Otóż mowa o zjawisku przypominającym anaptyksę, które objawia się tu wstawianiem pomiędzy dwie spółgłoski niektórych wyrazów elementu wokalicznego [ɨ] /y/, słyszalnego w wymowie wszystkich polskich spółgłosek.

Zabieg ten ma za zadanie ułatwić rodzimym użytkownikom języka mandaryńskiego wymowę zbitek spółgłoskowych, w które obfituje materiał wybrany do nagrania, a które w języku mandaryńskim nie występują, w związku z czym mogą stanowić dla tych osób niemałą trudność artykulacyjną. Konsekwencją tego zjawiska były kontury tonalne mające na tonogramach postać kilku odcinków, poprzerywane segmentami bezdźwięcznymi i pełne gwałtownych wahań F_0 . Przebiegi te okazały się niezwykle problematyczne przy próbach ekstrakcji tonów zasadniczych, zlokalizowanych na samogłoskach właściwych. Z uwagi

na zaistniałą segmentację wyrazów jednosylabowych w powyższy sposób postanowiono przy każdej realizacji danego wyrazu notować liczbę odcinków, jakie uwidoczniły się na tonogramach. Zauważono także zależność pomiędzy poziomem znajomości języka polskiego a częstością występowania niepożądanego segmentacji, co pozwoliło stwierdzić istnienie ujemnej korelacji łączącej obie te zmienne. Zgodnie z tym u osoby, która odznaczała się wysoką kompetencją w zakresie znajomości języka polskiego (JK), zdecydowanie najlepszą spośród wszystkich osób badanych, stwierdzono przypadek segmentacji 1 spośród 40 wypowiedzianych wyrazów. Z kolei mówca znający język polski najslabiej (YM) dokładnie połowę ogółu sylab zrealizował jako segmenty złożone, w tym 16 potraktował jako mające dwie jednostki TBU, natomiast pozostałe 4 – trzy jednostki TBU. Należy przy tym zaznaczyć, że sylab wypowiedzianych w ten sposób nie traktowano automatycznie jako błędnych; ponieważ w niniejszej analizie istotne jest to, jak zrealizowany został ton, brano pod uwagę jedynie zasadniczy fragment konturu tonalnego, podczas gdy resztę ignorowano, notując jedynie liczbę „sylab”, na które posegmentowany został dany wyraz, gdyż informacje te mogły okazać się przydatne podczas formułowania ostatecznych konkluzji. Co ciekawe, wyrazy, które zwykle dzielono na mniejsze jednostki, powielały się u różnych osób, co pozwala stwierdzić, że pewne właściwości związane z ich budową mogły stanowić dla rodzimych użytkowników języka tonalnego barierę trudną do pokonania w procesie nakładania cech tego języka na jednostki o innej strukturze. W grupie problematycznych wyrazów znajdowały się: *lgnąć*, *wprost*, *wszcząć*, *tchnąć*, *czcić* oraz *schnąć*.

Choć wszystkie tony okazały się dla mówców kłopotliwe, wśród nich najlepiej zrealizowane zostały tony 2 i 4, w sumie wypowiedziane prawidłowo odpowiednio 22 i 23 razy. Kontury poszczególnych realizacji tonu 4 rzadko wykazują niezakończony przebieg w postaci regularnie opadającej prostej skośnej – wprawdzie takowe zdarzają się kilkakrotnie, najczęściej jest to linia delikatnie falista, jednakże zawsze charakteryzująca się zdecydowanym spadkiem. Spektrogram na ryc. 8 ilustruje jeden z niewielu przypadków tonu 4 o właściwym konturze – tu w wykonaniu osoby DM na sylabie *strzęp*:

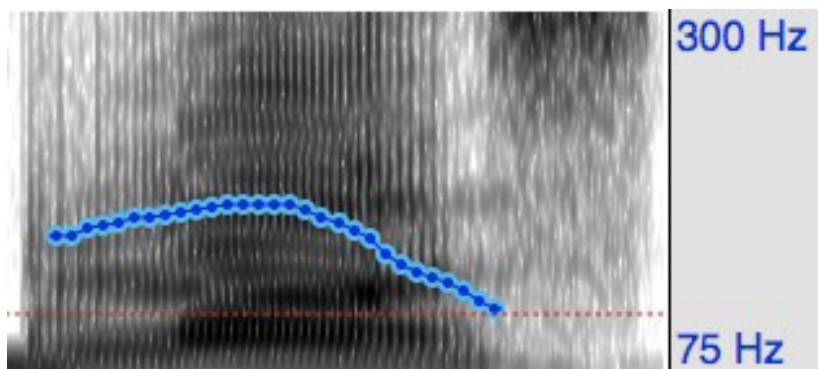
Ryc. 8. Prawidłowy przebieg tonu 4 – sylaba *strzęp*.



Należy zaznaczyć, że DM jest osobą, która realizowała sylaby z tonem 4 najlepiej ze wszystkich mówców, wypowiadając je prawidłowo aż 8 na 10 razy. Najwięcej trudności w realizacji tego tonu wykazywał natomiast mówca YM, u którego poprawną formę zyskały jedynie trzy realizacje. W pozostałych przypadkach kontury przybierały kształt przypominający symbol fonetyczny angielskiej samogłoski *wedge*, tworząc tym samym ton wznosząco-opadający, nieistniejący w języku mandaryńskim. Forma ta charakteryzowała w zasadzie wszystkie tony 4 wypowiedziane błędnie przez ogół badanych, choć pojawiały się także realizacje przypominające znak „~”, z dodatkową częścią wznoszącą na końcu.

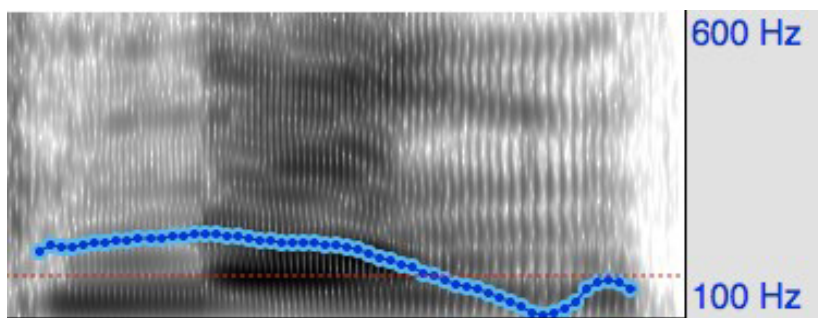
Spektrogramy prezentujące obie formy widoczne są na ryc. 9 i 10. Ton 4 na sylabie *wprost*, błędnie zrealizowany przez osobę YM:

Ryc. 9. Nieprawidłowy przebieg tonu 4 – sylaba *wprost*.



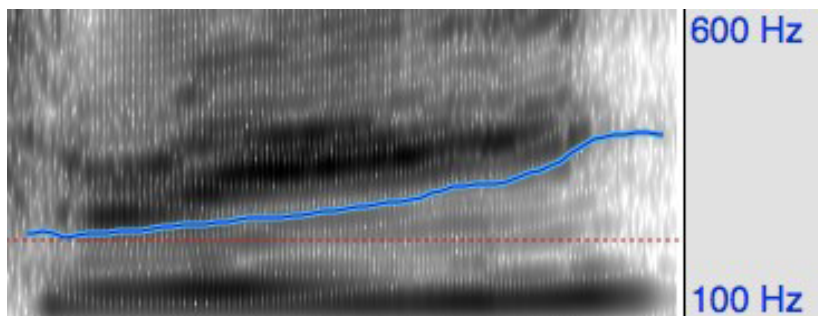
Zniekształcony kontur tonu 4, wypowiedzianego na sylabie *schnąć* przez badaną WK:

Ryc. 10. Nieprawidłowy przebieg tonu 4 – sylaba *schnąć*.



Ton 2 zdecydowanie najlepiej realizowała osoba WK, w wykonaniu której aż dziewięć sylab ma kontury tonalne o przebiegu zbliżonym do pożądanego. Na tonogramach bardzo często mają one kształt lekko wygięty, natomiast nie wykazują żadnych spadków, przez co są łatwo identyfikowalne w odbiorze słuchowym. Na ryc. 11 widoczny jest przykład jednej z najlepszych realizacji WK – ton 2 na sylabie *pstry*.

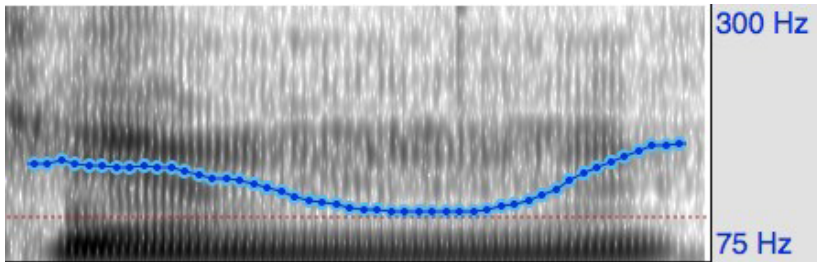
Ryc. 11. Prawidłowy przebieg tonu 2 – sylaba *pstry*.



Wyraźne problemy z prawidłową artykulacją tonu 2 na polskich sylabach zaobserwowano u dwojga badanych – JK oraz YM – u których został on zrealizowany w sposób właściwy jedynie w 3 z 10 przypadków. U JK większość tonów ma stosunkowo równy, „spłaszczony” kształt konturu, charakterystyczny dla przebiegu tonu 1; są również dwa przypadki wątpliwe, których kontury mają formę bliżej nieokreśloną. Mówca YM wykazuje tendencję do nadawania tonom drugim konturu zbliżonego do tonu 3, przy czym aż w 50% realizacji nie ulega wątpliwości, że są to tony opadająco-wznoszące. Ich obrazy graficzne odznaczają się zdecydowanym spadkiem F_0 z wyraźną zmianą kierunku i wartościami sięgającymi dolnych granic skali głosu. W kwestii tej można dostrzec analogię do sytuacji zanotowanej w języku mandaryńskim, gdzie identyczna liczba tonów drugich została nieprawidłowo zrealizowana przez YM dokładnie w ten sam sposób; zatem istnieje duże prawdopodobieństwo, że jest to pewna cecha idiolektalna, która ujawniła się u mówcy również w języku obcym.

Ryc. 12 przedstawia obraz graficzny tonu 2 o wyraźnym przebiegu opadająco-wznoszącym, zrealizowanego na sylabie *hymn* przez mówcę YM:

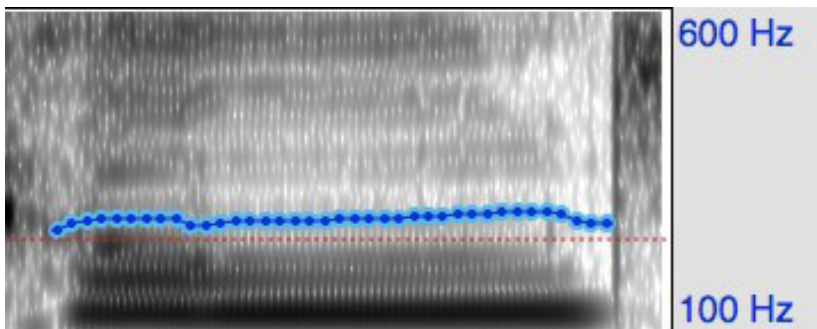
Ryc. 12. Zniekształcony przebieg tonu 2 – sylaba *hymn*.



Podobną formę prezentują niektóre tony 2 w wykonaniu DM, choć z grupy tej zdecydowaną większość (7) wypowiedział on poprawnie.

W odróżnieniu od sylab języka mandaryńskiego ton 1 na sylabach polskich okazał się dla osób badanych zadaniem niełatwym, bowiem w sumie został on prawidłowo wypowiedziany 14 razy, co wynosi nieco ponad 1/3 ogólnej liczby wszystkich realizacji. Również jego forma jest mniej stabilna niż w poprzedniej części badania – na tonogramach można zauważyć więcej nieregularności oraz drobnych wahań w przebiegu F_0 . Wśród mówców najlepszą realizacją w tym wypadku wykazali się DM oraz JK, wypowiadając w sposób właściwy odpowiednio sześć i pięć sylab z tonem 1. Spektrogram na ryc. 13 ukazuje obraz tonu 1 na sylabie *wszcząć*, o równym, regularnym przebiegu, zrealizowanego przez JK.

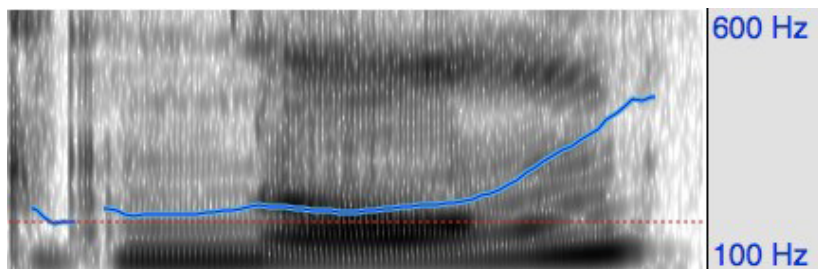
Ryc. 13. Prawidłowy przebieg tonu 1 – sylaba *wszcząć*.



Natomiast sylaby z tonem 1 zrealizowane błędnie, u obojga mówców charakteryzują się lekko bądź silnie rosnącą intonacją, w związku z czym ich kontury mają przebiegi typowe dla tonu 2. Tendencja ta jest jeszcze bardziej widoczna w drugim z głosów żeńskich (WK), gdzie aż 7 na 10 sylab z tonem 1 zostaje wypowiedzianych na wyraźnym tonie 2, o średnim zakresie wynoszącym ponad 200 Hz; w głosie tym ton 1 przybiera formę zbliżoną do pożądanej jedynie w dwóch przypadkach. Praktycznie nie istnieje on również w głosie YM – tu tylko jedna sylaba z tonem 1 wykazuje równy przebieg. Wśród pozostałych sześć tonów nawiązuje kształtem do konturu tonu 2, dwie wykazują znamiona tonu 3, natomiast jedna ma przebieg, który można określić jako wątpliwy.

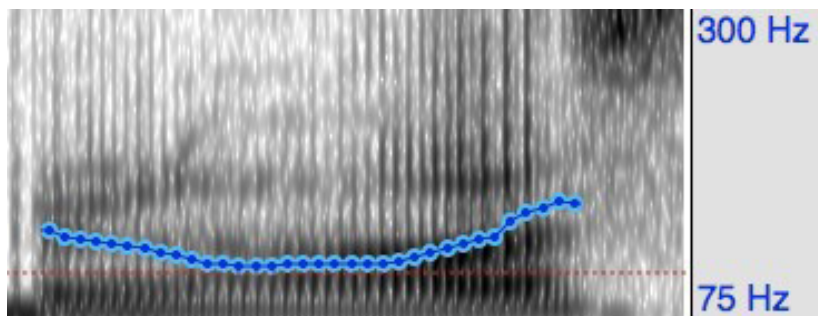
Spektrogram na ryc. 14 obrazuje zniekształcenia tonu 1 w wyniku zastosowania przez mówcę (WK) intonacji rosnącej. Ton został zrealizowany na sylabie *lgnąć*.

Ryc. 14. Nieprawidłowy przebieg tonu 1 – sylaba *lgnąć*.



Tak jak poprzednio, najwięcej trudności okazuje się przysparzać badanym realizacja tonu 3, który został uznany za prawidłowy jedynie w 32,5% przypadków. Poza niższym odsetkiem poprawnych realizacji dane dotyczące tonu 3 niemal odzwierciedlają wyniki analizy dokonanej na odpowiednich sylabach języka mandaryńskiego. Ponownie bezbłędnym rezultatem wykazał się mówca YM – wszystkie tony 3 w jego wykonaniu mają przebieg wyraźnie opadająco-wznoszący. Kontur jednego z takich tonów, zrealizowanych na sylabie *wprost*, widoczny jest na ryc. 15.

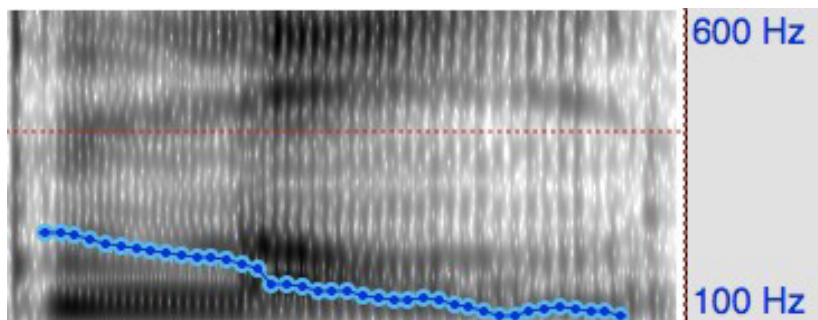
Ryc. 15. Prawidłowy kontur tonu 3 – sylaba *wprost*.



Odrotna sytuacja ma miejsce w przypadku osoby JK, u której – podobnie jak w języku mandaryńskim – wszystkie bez wyjątku tony 3 noszą wpływ procesu *sandhi* tonalnego, a więc są zredukowane o część wznoszącą. Analogicznie, również mówca DM – poza jednym tylko przypadkiem, w którym kontur tonu można było uznać za opadająco-wznoszącą – konsekwentnie realizuje tony 3 jako $\frac{1}{2}$ 3. Fakt ten może sugerować, że zjawisko *sandhi* jest dla niektórych użytkowników języka tonalnego tak naturalne, że realizowane jest w sposób mechaniczny, nieświadomie, i najwyraźniej może występować w oderwaniu od języka macierzystego.

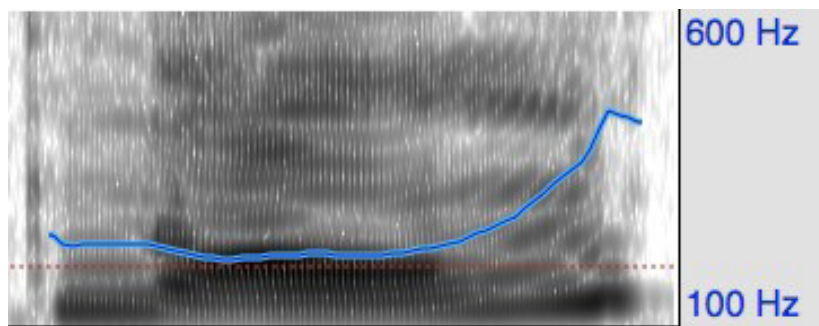
Ryc. 16 przedstawia ton $\frac{1}{2}$ 3, zrealizowany przez JK na sylabie *tchnąć*.

Ryc. 16. Ton $\frac{1}{2}$ 3 – sylaba *tchnąć*.



Osoba WK, pomimo że udało jej się prawidłowo zrealizować zaledwie 20% sylab z tonem 3, statystycznie poradziła sobie z tym najlepiej. Niemniej po dokładnej analizie tonogramów obu tonów można wywnioskować, że tony te zostały wyartykułowane z dużą dozą niepewności, jak zresztą większość ogółu tonów wypowiedzianych przez tę osobę na wyrazach polskich, w związku z czym ich poprawna forma może być w dużej mierze dziełem przypadku. Świadczą o tym łagodnie zaokrąglone kształty ich konturów, rozciągnięte w czasie punkty zwrotne zlokalizowane na częstotliwości ponad 200 Hz oraz przede wszystkim użycie intonacji pytającej. Jeden z opisanych tonów (*schnąć*) widoczny jest na ryc. 17.

Ryc. 17. Prawidłowy kontur tonu 3 – sylaba *schnąć*.



Analizując całokształt tonów zrealizowanych przez WK na sylabach polskich, można zauważyć, że wykazuje ona skłonność do nadawania charakteru wznoszącego wszystkim tonom z wyjątkiem 4, gdyż aż 25 spośród 40 ma formę typową dla tonu 2. Także w przypadku tonu 3 osiem zostało wypowiedzianych jako ton 2. Jak wspomniano, skłonność ta może być wynikiem zarejestrowanej w głosie niepewności, objawiającej się częstym stosowaniem przez badaną intonacji pytającej, która powoduje wzrost F_0 i uwidacznia się na tonogramach w postaci konturu rosnącego.

W ogólnym zestawieniu wyrazami najbardziej uciążliwymi w realizacji okazały się dla badanych: *lgnąć* (ton 1), *schnąć* (ton 4), *tchnąć* (ton 4) oraz *trzymiel* (ton 1), z których każdy został wypowiedziany

błędnie przez wszystkie osoby. Analogicznie, bezbłędnie zrealizowane zostały: *strzep* (ton 4), *pstry* (ton 2), *wszcząć* (ton 4), a także *czcić* (ton 4). Natomiast spośród sylab ogółem najwięcej błędów (12) popełniono przy produkcji wariantów tonalnych wyrazu *lgnąć*, a następnie: *schnąć*, *tchnąć* i *trzymiel* (po 11). Z kolei najmniej kłopotliwe były: *hymn* oraz *czcić* – po 6 błędów na 40 realizacji.

Dysponując wzorcem w postaci zestawu sylab chińskich, wypowiedzianych na poszczególnych tonach przez rodzimych użytkowników oraz wiedzą na temat umiejętności tychże osób w zakresie posługiwania się językiem polskim, nie stwierdzono korelacji pomiędzy poziomem językowym i zdolnością do realizacji tonów w języku obcym. Znaczenie mogą mieć natomiast pewne indywidualne zdolności imitacyjne badanych, które nie muszą wiązać się z dobrą znajomością żadnego z tych języków.

Dużą liczbę nieprawidłowości zanotowanych w eksperymencie można tłumaczyć dezorientacją mówców, spowodowaną brakiem wiedzy na temat właściwej lokalizacji jednostki TBU w wyselekcjonowanych wyrazach, do czego troje z nich przyznało się po nagraniu. Wynika to przede wszystkim z różnic w budowie sylaby pomiędzy badanymi językami – w języku mandaryńskim, w którym zbitki spółgłoskowe są kombinacją niedozwoloną, identyfikacja jednostki TBU nie sprawia większego problemu. Natomiast napotykać zbitkę w sylabach polskich, w których są one niezwykle powszechne, osoby badane mogły w sposób automatyczny próbować rozdzielić niepożądaną strukturę, wstawiając pomiędzy kolejne spółgłoski element wokaliczny /y/. Efektem tego były widoczne na tonogramach kontury tonalne złożone z kilku fragmentów, co mogło w sposób istotny zniekształcać kontur tonu właściwego. Pewien wpływ na wyniki mogły mieć także warunki laboratoryjne.

4.1.3. Porównanie wyników obu części eksperymentu

Różnice w realizacji tonów w obu językach można przedstawić w postaci kilku ogólnych sformułowań:

- 1) W obu językach osoby badane posługiwały się skalą o porównywalnej rozpiętości oraz realizowały poszczególne tony na zbliżonych częstotliwościach.

- 2) Pewne tendencje wykazywane przez poszczególnych mówców w ich języku rodzimym uwidoczniły się także w języku obcym. Sytuacja taka miała miejsce w przypadku występowania zjawiska *sandhi* tonu 3 oraz skłonności do zniekształcania niektórych tonów w jednakowy sposób.
- 3) Zachowały się także pewne naturalne właściwości poszczególnych głosem: głosy posługujące się rejestrem niskim podczas używania języka mandaryńskiego posługiwały się nim także w języku polskim, podobnie głosy wysokie pozostały wysokie bez względu na zmianę języka.
- 4) Ogółem, większość tonów w języku polskim zostało zrealizowanych niepoprawnie. W liczbie prawidłowo wypowiedzianych tonów zanotowano spadek o 42,3% w porównaniu do rezultatu uzyskanego przez mówców w języku mandaryńskim.
- 5) Tony na sylabach polskich wykazywały mniejszą stabilność zarówno w przebiegu F_0 , jak i sposobie realizacji. W przeciwieństwie do tonów na sylabach mandaryńskich, wykazują one ponadto znacznie mniejszą regularność pod względem iloczasu oraz zakresu F_0 .
- 6) Poza jednym głosem (WK) współczynnik modulacji częstotliwości u poszczególnych osób badanych nie uległ znaczącej zmianie.
- 7) W części drugiej eksperymentu (imitacja tonów na wyrazach języka polskiego), zaobserwowano zjawisko segmentacji wyrazów na mniejsze elementy, negatywnie skorelowane z poziomem znajomości języka przez osoby badane. Przyczyną tego była najprawdopodobniej próba rozbicia niedozwolonych w języku mandaryńskim zbitek spółgłoskowych poprzez realizację dodatkowej samogłoski [i] y.

4.2. Realizacja tonów chińskich przez nierodzimych użytkowników języka mandaryńskiego

4.2.1. Część I – język mandaryński

A. F_0

Tabela 9 stanowi ogólne zestawienie danych z pomiarów F_0 w grupie polskiej.

Tabela 9. Średnie wartości F_0 wyrazów chińskich realizowanych przez Polaków.

MÓWCA		AK	EK	GK	TM
PARAMETR					
F0 min. [Hz]		100,4	152	163,7	74,2
F0 max. [Hz]		481,5	325,9	593,4	157,9
skala		381,1	173,9	429,7	83,7
śr. F0	TON 1	282,7	254,3	253,2	128,4
	TON 2	206,8	206,2	234,2	111,4
	TON 3	183,4	200,7	216,8	88,2
	TON 4	244,4	236,8	221,8	120,5
zakres F0	TON 1	34,8	36,2	50	23,3
	TON 2	112,4	83,1	140,3	37,8
	TON 3	90,4	99,9	151,2	33,5
	TON 4	186,8	109,3	154,9	61

Tabela 10 przedstawia wartości odchyłeń standardowych oraz współczynników modulacji częstotliwości.

Tabela 10. Odchylenia standardowe głosów nierodzimych użytkowników języka mandaryńskiego (tony realizowane na wyrazach chińskich) wraz z współczynnikiem modulacji częstotliwości.

MÓWCZA σ	AK	EK	GK	TM
ogółem	42,1	24,1	17,1	15,7
TON 1	3,9	7,4	8,8	2,1
TON 2	15,6	11,6	14,4	4,5
TON 3	25,0	6,7	5,3	2,4
TON 4	21,3	11,5	8,8	5,5
WMC [%]	18,4	10,7	7,4	14

Spśród analizowanych mówców wszystkie głosy wykorzystują zakresy częstotliwości mieszczące się w odpowiednich skalach. Głos AK okazał się dość kłopotliwy w analizie ze względu na częstą obecność zjawiska *creaky voice* w realizacjach sylab z tonem 3, konsekwencją czego było widoczne na tonogramach przerwanie ciągłości konturów tonalnych, znaczne wahania F_0 oraz ogólny brak stabilności przebiegów, co znacznie utrudniło dokonywanie pomiarów. Z tego względu, przed przystąpieniem do analizy, wymagał on przeprowadzenia większej liczby korekt niż w przypadku innych głosów. Osoba AK posługuje się szeroką skalą o ogólnej rozpiętości 381,1 Hz, z ekstremum o wartości 481,5 Hz, choć najczęściej wykorzystywany przez nią zakres obejmuje częstotliwości średnie. Minimalne F_0 z reguły wahają się na poziomie 150–170 Hz, natomiast w przypadku tonów 4 kilkakrotnie zdarzają się spadki do wartości osadzonych w rejestrze niskim, podczas gdy maksima na ogół oscylują wokół 300 Hz.

Podobnym zakresem F_0 posługuje się osoba EK, aczkolwiek jej skala ogólna, rozpięta pomiędzy wartościami skrajnymi w punktach $F_{0\min}$ równym 152 Hz oraz $F_{0\max}$ wynoszącym 325,9 Hz, zanotowana w całym nagraniu, jest o nieco ponad 200 Hz mniejsza niż u AK. Fakt ten świadczy o większej stabilności głosu EK, w którym brak fonetycznie nieuzasadnionych, gwałtownych wahań częstotliwości,

natomiast wielkości F_0 charakteryzują się małym rozrzutem, czego dowodzi niemal dwukrotnie mniejsza wartość odchylenia standardowego niż w przypadku AK. Oba głosy odznaczają się natomiast dużą zgodnością pod względem średnich wysokości, na których wypowiedzane są poszczególne tony, przy czym wartości te stanowią adekwatne odzwierciedlenie szczegółowego zapisu pomiarów.

Trzeci z głosów żeńskich (GK) wyraźnie odbiega od pozostałych, co widoczne jest zarówno w osiągniętych przezeń wartościach F_0 , jak i przeciętnej rozpiętości skali, jaką operuje. W głosie tym zdecydowanie dominują wysokie częstotliwości – absolutne minimum, zarejestrowane w całym nagraniu wynosi 163,7 Hz, średnie wysokości poszczególnych tonów nie spadają poniżej 210 Hz, natomiast maksima w przeważającej większości przekraczają próg 300 Hz, z jednorazowym skokiem F_0 do rekordowej wartości 593,4 Hz, będącej jednym z ekstremów tonu 4. Wartość ta, zlokalizowana tuż przy górnej granicy skali głosów żeńskich, poszerza skalę badanej do imponujących rozmiarów.

Ostatni z mówców (TM) – jedyny mężczyzna w grupie Polaków – ma typowo niski głos o wyjątkowo wąskiej skali, obejmującej zaledwie 83,7 Hz. Wartości minimalne wypowiedzanych przez niego poszczególnych tonów często sięgają poziomu granicznego, z jednokrotnym spadkiem do 74,2 Hz, przy czym jedynie w 10 na 40 przypadków stanowią one liczby powyżej 100 Hz. Maksima tylko czterokrotnie przekraczają poziom 150 Hz, ustanawiając w jednym z nich najwyższą wartość równą 157,9 Hz. Zadziwia natomiast precyzja, z jaką udaje się mówcy utrzymywać stałą przeciętną wysokość każdego z tonów, gdyż wahania pomiędzy poszczególnymi realizacjami z reguły są nie większe niż 15 Hz. Dokładność ta uwidacznia się także w ich niezwykle regularnych zakresach, rozpiętych pomiędzy zbliżonymi wartościami.

Analizując zestawienie średnich zakresów F_0 z uwzględnieniem podziału na tony, wypowiedzianych przez wszystkie osoby badane, można zauważyć, że u każdej z nich bezwzględnie najmniejszym zakresem charakteryzuje się ton 1, przy czym ton ten jest najstabilniejszy w wykonaniu głosu męskiego (TM), natomiast największe wahania w jego przebiegu zaobserwowano w głosie GK. Największy zakres z kolei na ogół ma ton 4, o odpowiednio szerszej lub węższej rozpiętości w zależności od skali danego mówcy. Największą średnią wartość osiąga ton 4 osoby AK, która do jego produkcji wykorzystuje

blisko 50% swojej ogólnej skali, natomiast proporcjonalnie największą częścią skali operuje w tym wypadku mówca TM, u którego wartość ta sięga 73%. U wszystkich osób w następnej kolejności znajdują się tony 2 i 3, które charakteryzują się porównywalnymi zakresami. Wśród mówców najmniejszy rozrzut wykazują rezultaty uzyskane przez głos TM, natomiast głosem najmniej stabilnym okazuje się AK.

B. Iloczas

Tabela 11 przedstawia średnie długości poszczególnych tonów w wykonaniu mówców polskich.

Tabela 11. Iloczasy tonów wyrazów chińskich wypowiedzianych przez nierodzimykh użytkowników języka mandaryńskiego.

MÓWCA λ [ms]	AK	EK	GK	TM
TON 1	633	480	646	414
TON 2	665	497	469	371
TON 3	678	521	546	359
TON 4	314	272	317	251

Już przy wstępnej ewaluacji powyższych danych daje się zauważyć pewną analogię pomiędzy grupą rodzimych i nierodzimykh użytkowników. Dotyczy ona tonu 4, który w obu przypadkach jest zdecydowanie najkrótszym spośród realizowanych przez osoby badane czterech tonów mandaryńskich. Dodatkowo, u każdego mówcy z grupy polskiej jest on średnio o blisko połowę krótszy od odpowiedniego tonu najdłuższego, którym u AK i EK jest ton 3, natomiast u GK i TM – ton 1. Spośród wszystkich Polaków bezwzględnie najkrótszym iloczasem odznaczają się tony wypowiedziane przez TM. Poszczególne tony tego mówcy są również najbardziej ustabilizowane pod względem długości, gdyż zanotowane pomiędzy nimi różnice czasu trwania zazwyczaj nie przekraczają 150 ms; jedynie w przypadku sylab z tonem 2 zaobserwowano

sporadycznie występujący wzrost długości ponad wartość średnią. Tony o wyraźnie najdłuższych iloczasach realizowane są przez dwa głosy kobiece – AK oraz GK, przy czym w każdym z nich trudno o stałe wartości, z uwagi na duży rozrzut danych surowych. Do obu głosów należą także najdłuższe rezultaty zarejestrowane w całym nagraniu, z rekordowym iloczasem jednego z tonów 1 równym 847 ms, uzyskanym przez GK na sylabie *māi*.

Tony osoby EK również wykazują w tej kwestii znaczną nieregularność, niemniej ich długości są zdecydowanie mniej zróżnicowane niż w pozostałych głosach żeńskich.

W badaniu stwierdzono ponadto pewną interesującą zależność, właściwą wszystkim czterem osobom badanym, która – co jest szczególnie istotne – wystąpiła także wśród ogółu rodzimych użytkowników języka mandaryńskiego. Mowa tu o wyraźnym zwiększeniu długości wariantów tonalnych określonych sylab względem przeciętnych długości odpowiadających im tonów sylab pozostałych. Mianowicie, w obu badaniach efekt taki zanotowano w przypadku dokładnie tych samych sylab – [maj] *mai* oraz [jen] *yan*. Zbieżność ta niewątpliwie wyklucza zupełną przypadkowość zjawiska, jednakże zbyt skąpe dane uniemożliwiają zaferowanie satysfakcjonującego wyjaśnienia.

C. Poprawność realizacji

Rezultaty poszczególnych mówców zestawiono w tabeli 12.

Tabela 12. Wykaz prawidłowo zrealizowanych tonów na sylabach chińskich wypowiedzianych przez nierodzimych użytkowników języka mandaryńskiego.

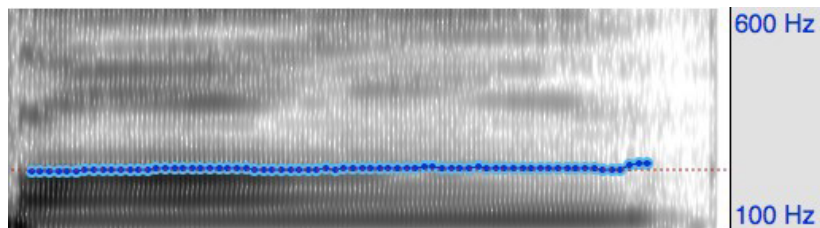
TON	LICZBA POPRAWNYCH REALIZACJI			
	AK	EK	GK	TM
1	10	10	10	10
2	1	3	6	6
3	8	10	10	9
4	10	10	10	8
[%]	72,5	82,5	90	82,5

Ogółem Polacy wykazali się bardzo dobrą realizacją tonów, co potwierdza wysoki odsetek tonów wypowiedzianych poprawnie, wynoszący 81,9%. Najlepszy rezultat w grupie osiągnęła osoba GK – aż 90% spośród wszystkich zrealizowanych przez nią tonów zostało uznanych za poprawne. Nieco słabszy wynik należy do dwóch badanych (EK oraz TM), którzy wypowiedzieli prawidłowo po 33 sylaby, co daje 82,5%, z kolei statystycznie najgorzej wypadła osoba AK, uzyskując 72,5% poprawność.

Podobnie jak w przypadku użytkowników rodzimych, w grupie polskiej nie zaobserwowano trudności w realizacji sylab z tonem 1, które w 100% uznano za prawidłowe. Przebiegi tonalne widoczne na tonogramach są ciągłe, regularne oraz wolne od zauważalnych zmian F_0 .

Na ryc. 18 przedstawiono obraz tonu 1 wypowiedzianego na sylabie [paŋ] (*bāng*) przez osobę GK. Do produkcji tonu wykorzystano zakres o jednej z najmniejszych wartości zanotowanych w całym nagraniu (17,4 Hz), co na spektrogramie widoczne jest w postaci niemal idealnie prostej linii poziomej.

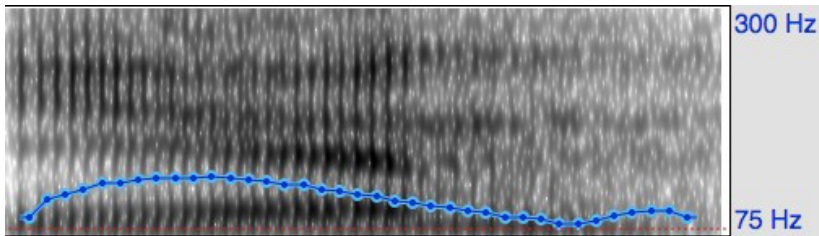
Ryc. 18. Prawidłowy przebieg tonu 1 – sylaba [paŋ] (*bāng*).



Również z realizacją tonu 4 badani poradzi sobie dobrze, chociaż zdarzyły się dwa przypadki zakwalifikowane jako błędne. Na tonogramach obu tonów, wypowiedzianych przez mówcę TM, stwierdzono nieprawidłowe przebiegi F_0 , o wyraźnie wznosząco-opadającym konturze, a więc niepożądanym w języku mandaryńskim.

Spektrogram na ryc. 19 przedstawia ton 4 o konturze uznanym za niewłaściwy, zrealizowany na sylabie [jen] (*yān*):

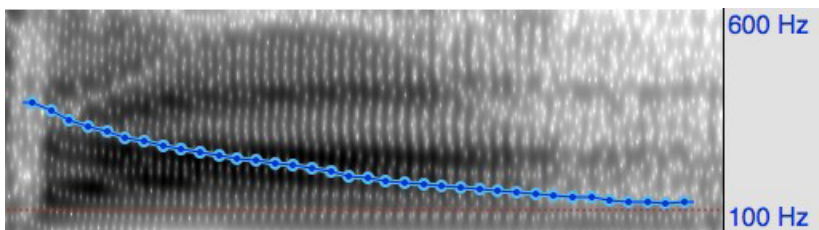
Ryc. 19. Nieprawidłowy kontur tonu 1 – sylaba [jen] (*yān*).



Wszystkie pozostałe tony 4 mają przebiegi charakteryzujące się regularnym spadkiem F_0 do wartości zlokalizowanych w rejestrze niskim. Wprawdzie ich kontury często wykazują pewne nieznaczne fluktuacje, w szczególności w punktach krańcowych, które w większości przypadków wynikają z bliskości spółgłosek dźwięcznych bądź skłonności mówców do wydłużania spółgłosek nosowych w wygłosach niektórych sylab, niemniej wahania te znajdują się w granicach akceptowalnych dla przebiegów tonu opadającego, co potwierdza również weryfikacja słuchowa.

Spektrogram na ryc. 20 jest obrazem prawidłowego, niezakłóconego przebiegu tonu 4, zrealizowanego przez badaną AK na wyrazie [t^han] (*tàn*).

Ryc. 20. Prawidłowy kontur tonu 4 – sylaba [t^han] (*tàn*).

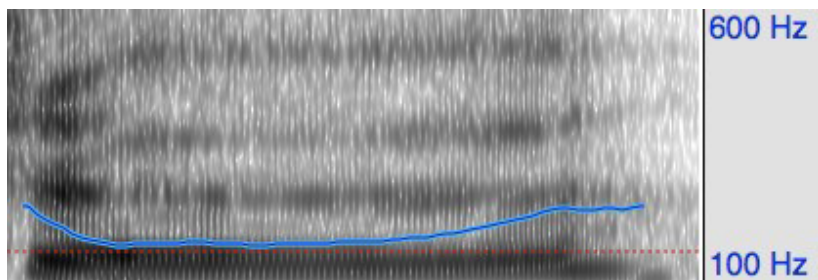


Spśród tonów 2 i 3, które z reguły sprawiają mówcom najwięcej trudności, jedynie ton 2 okazał się dla Polaków problematyczny. Ogółem, na 40 sylab z tonem wznoszącym wypowiedzianych przez czworga

badanych, prawidłowy kontur miało jedynie 16, co daje względnie niski odsetek w wysokości 40% poprawnych realizacji. Najlepsze wyniki w grupie osiągnęły dwie osoby – GK oraz TM, wypowiadając poprawnie po sześć sylab. W następnej kolejności znalazła się badana EK – trzy sylaby, natomiast najslabszy rezultat jest udziałem AK, u której ton 2 praktycznie nie występuje, bowiem prawidłowy przebieg stwierdzono w 1 na 10 zrealizowanych sylab. Podobnie jak w poprzednim badaniu, wyraźnie dominował jeden rodzaj zniekształceń – mianowicie przebiegi uznane za błędne wykazywały widoczne cechy typowe dla tonu 3. Zanotowano także dwa przypadki wątpliwe, o konturach silnie spłaszczonych, przypominających kształtem kontur tonu 1, lecz o percepcyjnie uchwytnej intonacji rosnącej; jednakże z uwagi na ich niewielkie zakresy, o rozpiętości charakterystycznej dla tonu 1, realizacje te ostatecznie zakwalifikowano jako błędne. Oba rodzaje nieprawidłowości przedstawiono na zamieszczonych niżej spektrogramach.

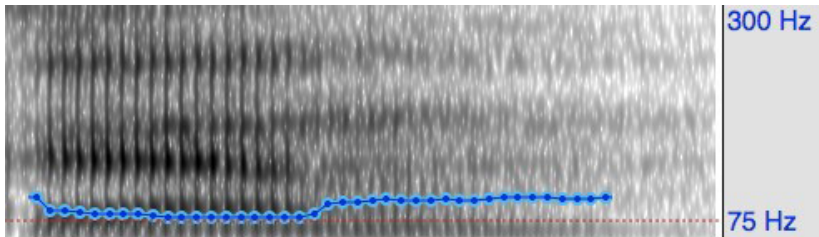
Ryc. 21 przedstawia ton 2 o wyraźnym konturze opadająco-wznośnym, wypowiedziany przez mówcę AK na sylabie [ʃɪ] (*shi*):

Ryc. 21. Nieprawidłowy przebieg tonu 2 – sylaba [ʃɪ] (*shi*).



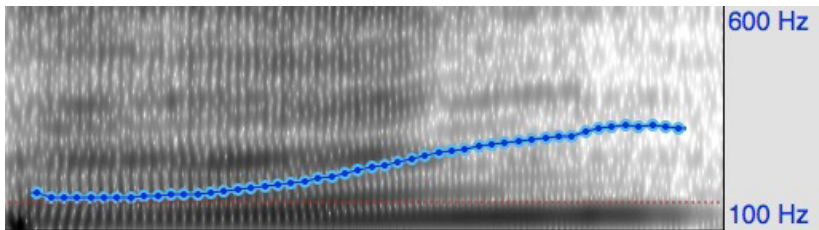
Na ryc. 22 widoczna jest nieprawidłowa realizacja tonu 2 na sylabie [gen] (*gén*) w wykonaniu mówcy TM:

Ryc. 22. Nieprawidłowy kontur tonu 2 – sylaba [gen] (*gén*).



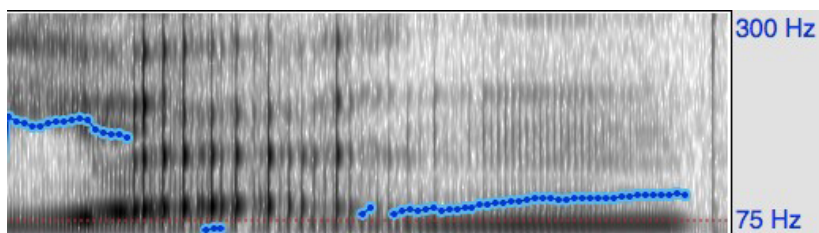
Dla porównania obraz właściwego przebiegu tonu wznoszącego – sylaba [p^heŋ] (*péng*), zrealizowana przez badaną GK (ryc. 23):

Ryc. 23. Prawidłowy kontur tonu 2 – sylaba [p^heŋ] (*péng*).



Wbrew przewidywaniom badani wykazali się natomiast bardzo dobrą realizacją sylab z tonem 3 – w sumie na 40 realizacji aż 36 nie budziło wątpliwości co do poprawności kształtów konturów zarejestrowanych na tonogramach. Ponadto dwie osoby (EK i GK) wypowiedziały wszystkie sylaby bezbłędnie. Do rezultatu tego zbliżył się TM, osiągając wynik 90% poprawnych realizacji, natomiast statystycznie najgorzej wypadł trzeci z głosów żeńskich – AK. Powodem tego jest wspomniana skłonność osoby AK do realizowania *creaky voice*, w wyniku czego powstaje kontur o nieciągłym przebiegu, naznaczony licznymi wahaniami F_0 . Dla ilustracji na ryc. 24 zamieszczono obraz graficzny tonu 3 na sylabie [jen] (*yěŋ*).

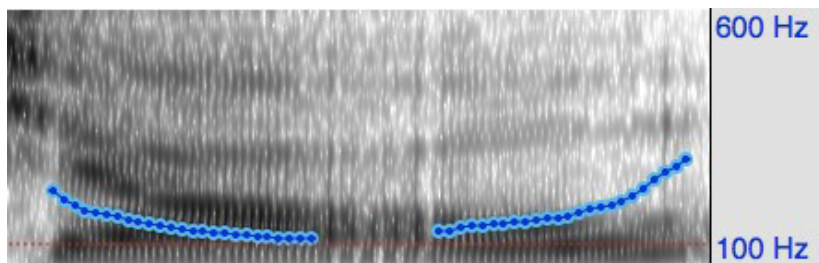
Ryc. 24. Efekt zjawiska *creaky voice* – zniekształcony kontur tonu 3 na sylabie [jen] (*yě̃n*).



Co ciekawe, wśród mówców z grupy polskiej nie zanotowano ani jednego przypadku realizacji *sandhi* tonalnego. Fakt ten może stanowić podstawę do wnioskowania, iż zjawisko to występuje u rodzimych użytkowników języka tonalnego nawykowo, a zatem jest procesem w dużej mierze zautomatyzowanym, nieświadomym, podczas gdy wyuczone reguły dotyczące następstwa tonów nie pozwalają użytkownikom nierodzimym na zastosowanie ich tam, gdzie zjawisko *sandhi* z zasady nie powinno mieć miejsca – tj. w sylabach izolowanych.

Spektrogram na ryc. 25 jest obrazem graficznym tonu 3, zrealizowanego na sylabie [təow] (*qǐu*) przez osobę EK – pomimo obecności w głosie *creaky voice*, którego efektem jest przerwanie ciągłości konturu, ton ten wciąż zachowuje wyraźnie opadająco-wznoszący przebieg.

Ryc. 25. Prawidłowy kontur tonu 3 – sylaba [təow] (*qǐu*).



Wyrazem najbardziej kłopotliwym okazał się *gén* (ton 2), gdyż jako jedyny spośród 40 został zrealizowany błędnie przez wszystkich mówców. Dalej znalazły się: *shí*, *báng*, *mái* oraz *yú* – wszystkie z tonem 2, z realizacją których nie poradziły sobie trzy na cztery osoby badane. Po zsumowaniu nieprawidłowych bądź wątpliwych realizacji ogółu wariantów tonalnych poszczególnych sylab otrzymano cztery sylaby szczególnie problematyczne: *gen*, *mai*, *yan* oraz *yu*, z których każdą wypowiedziano niepoprawnie cztery razy. Brak natomiast sylab całkowicie wolnych od błędów, choć każda miała przynajmniej jeden wariant tonalny (najczęściej ton 1 lub 4), który nie sprawił trudności w realizacji żadnemu z mówców. Do sylab zrealizowanych najlepiej należą: *qiu*, *peng* oraz *shuo* – w przypadku każdej z nich ogółem zanotowano po dwa błędy.

Pomimo bardzo dobrej produkcji tonów przez Polaków, do czego zapewne przyczyniła się ich wysoka kompetencja w zakresie języka mandaryńskiego, wciąż można było zaobserwować u nich znaczne trudności, które w przeważającej części związane były z realizacją sylab z tonem 2. Fakt ten trudno wyjaśnić, szczególnie z uwagi na dominującą wśród badanych skłonność do zniekształcania konturów tonalnych w podobny sposób. Ponieważ tendencja ta ulegała nasileniu w przypadku niektórych sylab, być może w pewnym stopniu było to wynikiem wpływu znajdujących się w ich nagłosach spółgłosek, które mogły przyczyniać się do obniżenia F_0 podczas artykulacji, dając w rezultacie spadek charakterystyczny dla przebiegu tonu 3²⁴. Część błędów można zapewne wytłumaczyć przesadną artykulacją sylab, spowodowaną nienaturalnym spowolnieniem tempa mowy, wydłużaniem wypowiedzianych jednostek leksykalnych oraz realizacją pauz, które nie występują z taką częstotliwością w mowie ciągłej. Istotną rolę mogły także odegrać pewne indywidualne możliwości głosowe mówców.

4.2.2. Część II – język polski

A. F_0

Średnie wartości F_0 , zakresy tonów oraz ekstrema każdego z analizowanych głosów znajdują się w tabeli 13.

²⁴ Według Duanmu (2000) efekt taki powodują dźwięczne obstruenty (s. 232).

Tabela 13. Średnie wartości F_0 wyrazów polskich realizowanych przez Polaków.

MÓWCA		AK	EK	GK	TM
PARAMETR					
F0 min. [Hz]		100,1	144,7	162,3	75,8
F0 max. [Hz]		396,2	345,5	412,3	168
skala		296,1	200,8	250	92,2
śr. F0	ogółem	228	223,5	239,7	111,7
	TON 1	280,6	247,5	250,3	130,1
	TON 2	196,8	203,6	249,8	110
	TON 3	189,5	198,6	225,8	92,9
	TON 4	246	244,2	232,7	113,9
zakres F0	TON 1	47,7	47,2	55,6	14,1
	TON 2	108,5	92,6	166,5	39,7
	TON 3	98,5	88,8	196,6	30,4
	TON 4	186,7	126,9	157,1	56,7

W tabeli 14 zestawiono wartości odchyłeń standardowych oraz współczynników modulacji częstotliwości obliczonych dla poszczególnych głosów.

Tabela 14. Odchylenia standardowe głosów Polaków (tony na wyrazach polskich) wraz z współczynnikiem modulacji częstotliwości.

MÓWCA		AK	EK	GK	TM
σ					
ogółem		40,7	24	15,9	14,3
	TON 1	5,5	5,5	2,9	2,8
	TON 2	16,4	4,8	13,2	5,5
	TON 3	21,5	3,5	8,7	2,2
	TON 4	13,6	13,7	18,4	7,5
WMC [%]		17,9	10,7	6,7	12,8

W produkcji tonów w języku polskim dwie badane osoby (EK oraz TM) posłużyły się skalą o bardzo zbliżonym zakresie do używanej w poprzedniej części nagrania; istotną zmianę zanotowano natomiast w przypadku pozostałych osób (AK i GK), u których nastąpiło silne zawężenie skali – u AK o blisko 100 Hz, natomiast u GK aż o 180 Hz, co stanowi 40% początkowej rozpiętości. W dodatku w głosie AK, w porównaniu z tonami wyrazów mandaryńskich, następuje ogólne obniżenie wysokości tonów „polskich”, o czym świadczą niższe wartości F_0 przeważającej większości ekstremów, natomiast jedynym tonem, który utrzymał porównywalną wysokość, jest ton 4. W dolnym rejestrze minima kilkakrotnie oscylują wokół wartości granicznych, podczas gdy maksima przekraczają próg 300 Hz w mniej niż 50% realizacji, osiągając wartość najwyższą równą 396,2 Hz. Poszczególne tony wykazują umiarkowaną stabilność pod względem wysokości, spośród których największe wahania zaobserwować można pomiędzy wyrazami z tonem 3, co znajduje odzwierciedlenie w dużej wartości odchylenia standardowego.

Znacznie większą regularnością odznaczają się tony osoby EK, rozpięte pomiędzy ekstremami o względnie stałych wysokościach, czego adekwatny obraz dają bardzo zbliżone wartości średnich arytmetycznych F_0 poszczególnych tonów. Osoba ta najczęściej operuje zakresem osadzonym w rejestrze średnim, o czym świadczą wartości minimalne na ogół utrzymujące się powyżej 150 Hz. Górną granicę skali EK stanowi wartość 345,5 Hz, wyznaczając zakres o rozpiętości 200,8 Hz. Zestawiając wyniki z obu części eksperymentu, można zaobserwować początkowe obniżenie wysokości tonów „polskich” względem tonów mandaryńskich, produkowanych z częstotliwością nawet do kilkudziesięciu herców większą, aczkolwiek różnica ta szybko ulega zatarciu pomiędzy zbliżonymi wartościami F_0 odpowiadających sobie tonów.

Wprawdzie trzeci z głosów żeńskich (GK) ogółem należy do stabilniejszych pod względem poziomu częstotliwości, de facto wykazuje on dość znaczną niekonsekwencję pomiędzy realizacją poszczególnych tonów. Dużą regularnością charakteryzują się tony 1 i 3, co na tonogramach widoczne jest w stałych wartościach ich ekstremów, podczas gdy w pozostałych tonach można zaobserwować w tej kwestii spore różnice. Jak pokazuje wartość odchylenia standardowego, zdecydowanie najmniej konsekwentnie realizowane są wyrazy z tonem 4.

W porównaniu z materiałem w języku mandaryńskim przeważająca liczba tonów „polskich” w rejestrze dolnym – w zależności od kierunku przebiegu – rozpoczyna się bądź opada do wartości wyższych. W przypadku wartości maksymalnych nie odnotowano prawidłowości, co jest spowodowane dużymi różnicami wysokości w obrębie poszczególnych tonów sylab polskich, sięgającymi niekiedy ponad 100 Hz. W wyniku zmniejszenia skali głosu badanej nastąpiło znaczne obniżenie wysokości absolutnego maksimum do 412,3 Hz, podczas gdy minimum pozostało stałe (162,3 Hz).

Ostatni z mówców – TM – podobnie jak w części poprzedniej, wykazuje największą regularność w produkcji tonów, o czym świadczą najniższe spośród wszystkich osób badanych wartości odchyień standardowych, zarówno ogółem, jak i z uwzględnieniem poszczególnych tonów (jedynym wyjątkiem jest ton 2). Z największą konsekwencją realizowane są sylaby z tonem 3, które odznaczają się bardzo zbliżonymi częstotliwościami. W obu częściach badania TM posługuje się skalą o porównywalnej rozpiętości, która dla tonów „polskich” wynosi 92,2 Hz, z wartościami krańcowymi $F_{0 \min}$ oraz $F_{0 \max}$ równymi odpowiednio 75,8 Hz i 168 Hz, przy czym poszczególne minima z reguły zlokalizowane są na nieco wyższych częstotliwościach niż w języku mandaryńskim.

Rozpatrując poszczególne tony pod względem średnich zakresów częstotliwości, można zauważyć, że u wszystkich mówców zdecydowanie najmniejszy zakres mają sylaby z tonem 1, przy czym u badanych kobiet wartość ta jest bardzo zbliżona. U trzech z czterech osób wyraźnie najszerszym średnim zakresem charakteryzuje się ton 4, podczas gdy w przypadku GK jest to ton 3. Większość badanych wykorzystuje do produkcji sylab z tonem 4 ok. 60% rozpiętości ogólnej skali głosu. Analogicznie do sytuacji w języku mandaryńskim, najmniejszy rozrzut wykazują zakresy tonów w wykonaniu TM, które są jednocześnie najwęższe w badanej grupie, natomiast głosem najmniej stabilnym pod tym względem ponownie okazuje się AK, w którym zanotowano wielokrotne znaczne wahania F_0 w obrębie poszczególnych tonów. W ogólnym zestawieniu bezwzględnie największą rozpiętością odznacza się ton 3 osoby GK, którego zakres ponaddwukrotnie przewyższa zakres skali głosu TM.

B. Iloczas

Tabela 15 zawiera uśrednione wartości iloczasów poszczególnych tonów zrealizowanych na sylabach języka polskiego.

Tabela 15. Iloczasy tonów wyrazów polskich wypowiedzianych przez nierodzimych użytkowników języka mandaryńskiego.

MÓWCA λ [ms]	AK	EK	GK	TM
TON 1	494	372	538	311
TON 2	510	415	355	256
TON 3	510	473	482	315
TON 4	269	278	269	206

Analiza danych z pomiarów długości tonów „polskich” okazała się – podobnie jak w przypadku rodzimych użytkowników języka mandaryńskiego – zadaniem dość kłopotliwym. Główną tego przyczyną były znaczne wahania tego parametru w obrębie poszczególnych tonów, zanotowane u każdej z nagranych osób. Duże rozproszenie wartości oraz wynikający z tego brak powtarzalnych wzorców w znacznym stopniu utrudniły uporządkowanie danych w sposób umożliwiający dostrzeżenie pewnych ogólnych prawidłowości. W związku z tym przedstawione tu średnie arytmetyczne danych szczegółowych należy traktować jako wartości orientacyjne.

Niemniej porównanie przeciętnych długości każdego z czterech tonów względem siebie pozwala stwierdzić pewne zależności. I tak, pomimo dużych wahań wartości iloczasu, tonem najkrótszym w stosunku do pozostałych u wszystkich mówców jest ton 4, który wykazuje największą regularność w wykonaniu głosu męskiego (TM). Ton najdłuższy jest natomiast w znacznym stopniu międzyosobniczo różnicowany: u EK oraz TM na ogół jest to ton 3, u AK – ton 2, z kolei w przypadku GK – najczęściej ton 1. Biorąc pod uwagę całokształt realizacji, tony o najdłuższym iloczacie produkują dwa głosy kobiece: AK i GK – u AK blisko połowa realizacji przekracza 500 ms, podczas

gdy u GK zanotowano aż pięć tonów o czasie trwania dłuższym niż 700 ms. W grupie Polaków zaobserwowano również zjawisko, które wystąpiło także wśród rodzimych użytkowników języka mandaryńskiego, dotyczące wpływu niektórych sylab na zwiększenie lub zmniejszenie iloczasu realizowanych na nich tonów, choć w tym wypadku miało ono nieco bardziej okrojona formę. Wyraźny wzrost długości tonów u wszystkich mówców miał miejsce na sylabie *trzymiel* oraz w trochę mniejszym stopniu na sylabach *hymn* i *schnąć*. U trzech na czterech mówców efekt ten spowodowały także wyrazy *lgnąć* oraz *tchnąć*, przy czym w drugim przypadku różnica była mniej znaczna. Podobnie jak w poprzednim badaniu, odwrotny efekt dał jedynie wyraz *czcić*, czego rezultatem były tony o iloczasach najkrótszych spośród ogółu wypowiedzianych przez daną osobę.

Powyższe rezultaty dostarczają solidnych podstaw, by przypuszczać, że iloczas danej sylaby jest ściśle związany z jej strukturą w sposób taki, że w dużej mierze zależy od rodzaju oraz konfiguracji głosek wchodzących w jej skład. Jak zaobserwowano, pewne spółgłoski – w szczególności półotwarte – umieszczone w bezpośrednim sąsiedztwie samogłoski, sprzyjają zwiększeniu jej iloczasu poprzez wydłużenie segmentu dźwięcznego, rozszerzając tym samym zasięg jednostki TBU. Do spółgłosek tych należą m.in. spółgłoski nosowe /m/ i /n/, a także samogłoska niesylabiczna /j/, w przypadku których zanotowano powyższy efekt. Z kolei spółgłoski bezdźwięczne, w szczególności, gdy otaczają samogłoskę krótką z natury, powodują skrócenie czasu trwania segmentu dźwięcznego, jak miało to miejsce w sylabie *czcić*.

C. Poprawność realizacji

Wyniki ogólne oraz liczbę poprawnych realizacji z uwzględnieniem podziału na poszczególne tony przedstawia tabela 16.

Tabela 16. Wykaz prawidłowo zrealizowanych tonów na sylabach polskich wypowiedzianych przez nierodzimych użytkowników języka mandaryńskiego.

TON	LICZBA POPRAWNYCH REALIZACJI			
	AK	EK	GK	TM
1	10	10	10	10
2	1	4	9	8
3	6	10	9	10
4	9	9	9	9
Σ [%]	65	82,5	92,5	92,5

Zanim omówione zostaną wyniki z tabeli 16, należy wspomnieć o pewnej analogii, jaką zaobserwowano pomiędzy obydwooma badanymi grupami. Dotyczy ona zjawiska segmentacji niektórych wyrazów jednosylabowych na mniejsze jednostki, przy czym podział ten jest zgodny z liczbą realizowanych w danej sylabie samogłosek, identyfikowalnych słuchowo, lecz niekoniecznie występujących w zapisie. W efekcie w obrębie jednej sylaby tworzy się kilka jednostek TBU, co na tonogramach widoczne jest w postaci konturu złożonego z więcej niż jednego odcinka. W przypadku Polaków zjawisko to należy jednak interpretować inaczej z uwagi na fakt, że wystąpiło ono w bardzo ograniczonym stopniu oraz było spowodowane innymi czynnikami. W odróżnieniu od rodzimych użytkowników języka mandaryńskiego, dla których polski jest językiem obcym, źródłem segmentacji w przypadku mówców polskich bez wyraźnych zaburzeń w wymowie nie mogły być trudności artykulacyjne, związane z produkcją zbitek spółgłoskowych. Jak wynika z analizy, w grupie polskiej w sumie sześć sylab zostało zrealizowanych jako zawierające dwie jednostki TBU, z czego wszystkie stanowiły warianty tonalne wyrazu *lgnąć*. Najwięcej błędnych wystąpień zanotowano w głosie AK, gdzie wszystkie cztery tony charakteryzowały się nieciągłym przebiegiem, natomiast dwa pozostałe dotyczyły realizacji tonów 2 przez EK oraz GK; w głosie męskim (TM) zjawisko to nie wystąpiło w ogóle. Biorąc pod uwagę budowę sylaby, która uległa segmentacji, a także zmiany parametrów towarzyszących jej realizacjom, można stwierdzić, że przyczyną zjawiska była zbitka spółgłosek dźwięcznych

oraz zwolnione tempo artykulacji. W sylabie *Ignąć* właściwą jednostkę TBU, którą stanowi samogłoska /o/²⁵, poprzedza ciąg dwóch spółgłosek półotwartych // i /n/ oraz spółgłoski zwartej dźwięcznej /g/. Ponieważ z każdej z tych głosek można wyekstrahować F_0 , podczas artykulacji tego typu kombinacji, kontury tonalne będące obrazami graficznymi zmian częstotliwości w czasie łączą się, tworząc przebieg ciągły. Proces ten dodatkowo ułatwiają elementy wokaliczne /y/ – pewnego rodzaju „efekty uboczne” pojawiające się przy produkcji spółgłosek, a zatem realizacji podlega de facto sekwencja [l-y-g-y-n-o-ŋ], której rezultatem jest ton o znacznie wydłużonym iloczasiu, co zaobserwowano w niniejszym badaniu. Natomiast w przypadku, gdy realizacja dotyczy sylaby izolowanej i odbywa się w nieco spowolnionym tempie, zwanie narządów mowy, które następuje przy artykulacji głoski /g/, objawia się na tonogramie w postaci wyraźnej przerwy w przebiegu konturu tonalnego. Jak pokazują wyniki analizy iloczasu poszczególnych sylab, tony jedyne go mówcy, u którego nie wystąpiło powyższe zjawisko (TM), odznaczały się wyraźnie krótszym czasem trwania niż tony pozostałych mówców.

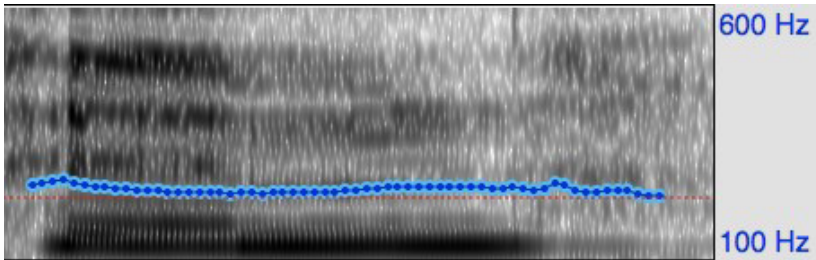
Podobnie jak w poprzedniej części badania, Polacy osiągnęli wspólnie wysoki stopień poprawności realizacji, równy 83,1%. Co zaskakujące, odsetek tonów wypowiedzianych prawidłowo na sylabach polskich nieco przewyższa rezultat osiągnięty przez grupę na materiale w języku mandaryńskim (por. s. 143–144). Najlepszą produkcją wykazało się dwoje mówców: GK oraz TM, którzy zrealizowali prawidłowo po 92,5% tonów. W następnej kolejności znalazła się osoba EK z wynikiem 82,5%, natomiast najniższy odsetek poprawności ponownie uzyskała AK, osiągając rezultat w wysokości 65%.

Spośród poszczególnych tonów najmniej problematyczny okazał się dla mówców ton 1, gdyż nie zanotowano ani jednego przypadku błędnej realizacji. Wprawdzie nie zawsze kontury sylab z tonem 1 miały formę równych linii prostych – niekiedy zdarzało się, że wykazywały one pewne drobne fluktuacje, niemniej były to nieznaczne wahania, nie powodujące wyraźnej zmiany kierunku, w związku z czym realizacje te uznano za poprawne.

Spektrogram na ryc. 26 przedstawia obraz graficzny tonu 1 wypowiedzianego na sylabie *strzęp* przez osobę EK.

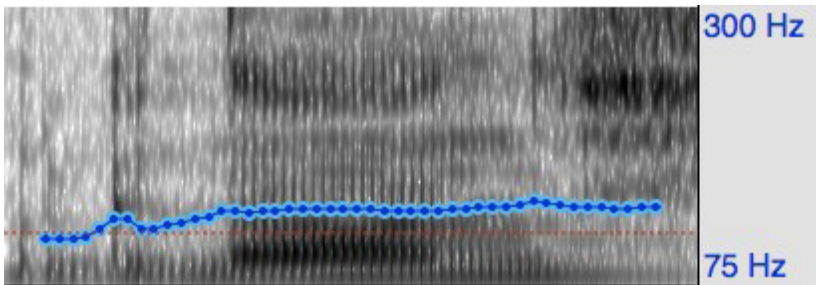
²⁵ Z fonetycznego punktu widzenia wyraz ten realizowany jest jako [lgnontɛ], z ośrodkiem zlokalizowanym na samogłosce /o/.

Ryc. 26. Prawidłowy przebieg tonu 1 – sylaba *strzęp*.



Na kolejnym spektrogramie (ryc. 27) widoczny jest rezultat jednej z prawidłowych realizacji tonu 1 w wykonaniu mówcy TM na problematycznej sylabie *lgnąć*. Należy zwrócić uwagę, że pomimo zwarcia narządów mowy na głosce /g/ kontur zachował ciągłość. Fakt, że w produkcji sylaby wykorzystano zakres typowy dla tonu 1 (30 Hz), pozwala na zakwalifikowanie poniższego przebiegu jako zgodnego z oczekiwanym.

Ryc. 27. Prawidłowy kontur tonu 1 – sylaba *lgnąć*.

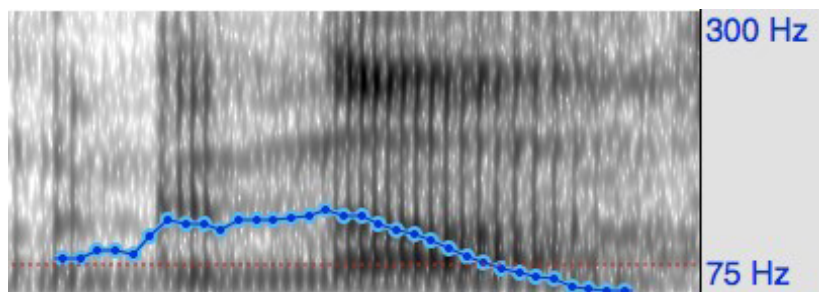


Drugim w kolejności tonem, który sprawił badanym najmniej trudności, jest ton opadający. W tym wypadku u wszystkich osób zanotowano po jednym tonie o niewłaściwym konturze. Co ciekawe, każdy z tych tonów był realizacją wspomnianej sylaby *lgnąć*. Można zatem przypuszczać, że podczas artykulacji danej sekwencji głosek dźwięcz-

nych znaczne wahania w przebiegu F_0 , objawiające się na tonogramach w postaci niepożądanych zmian kierunku, są następstwem trudnym lub wręcz niemożliwym do uniknięcia.

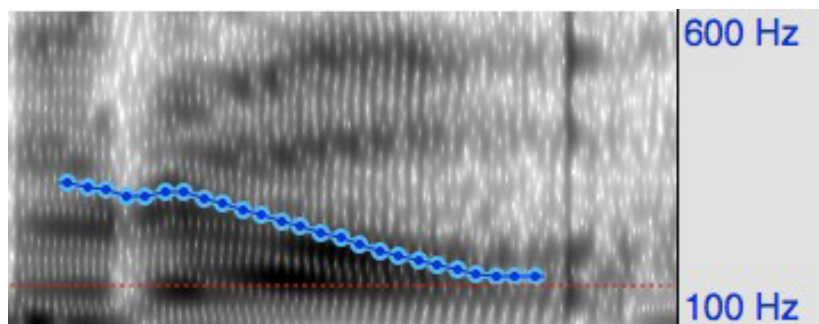
Ryc. 28 obrazuje zniekształcony kontur tonu 4, wypowiedzianego przez mówcę TM na sylabie *lgnąć*.

Ryc. 28. Nieprawidłowy przebieg tonu 4 – sylaba *lgnąć*.



Pozostałe sylaby z tonem 4 mają przebiegi charakteryzujące się wyraźnym i regularnym spadkiem F_0 , pozbawionym nieoczekiwanych skoków i zmian kierunku. Dla przykładu – ton 4 na sylabie *wprost*, zrealizowany przez osobę AK (ryc. 29):

Ryc. 29. Prawidłowy kontur tonu 4 – sylaba *wprost*.

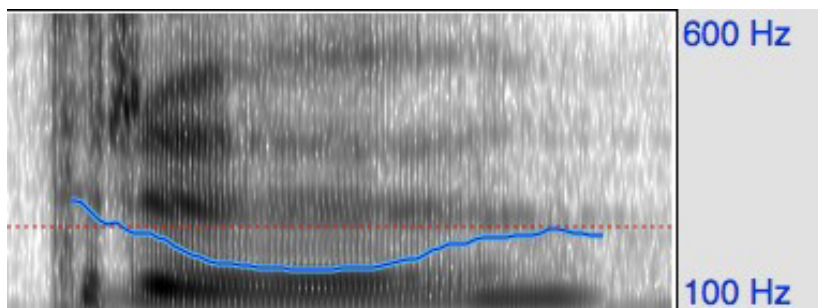


Również w kwestii tonów problematycznych sytuacja przedstawia się bardzo podobnie jak w przypadku języka mandaryńskiego: tonem o najmniejszej liczbie poprawnych realizacji po raz drugi okazał się ton rosnący, natomiast ton 3, choć jednej z badanych osób sprawiał pewne trudności, wciąż odznacza się względnie dobrą produkcją. Ujawniły się także wykazane wcześniej tendencje poszczególnych osób do zniekształcania konturów tonalnych w pewien określony sposób, co sugeruje, że niektóre cechy idiolektalne mogą być odporne na zmianę struktury i właściwości języka, jakim w danym momencie posługuje się mówca.

Ogółem osoby badane zrealizowały prawidłowo 22 na 40 sylab z tonem 2, co daje rezultat w wysokości 55%. Wśród nich najlepszą realizacją wykazała się osoba GK, w wykonaniu której tylko jeden ton miał kontur o wątpliwym kształcie. Zbliżony wynik osiągnął głos męski (TM), wypowiadając poprawnie 8 z 10 sylab. U pozostałych głosów żeńskich (AK i EK) stwierdzono natomiast wyraźne trudności z produkcją tonów 2 – u badanej EK tylko cztery charakteryzował przebieg zgodny z oczekiwanym, natomiast statystycznie najslabszą realizacją wykazała się AK, która powieliła swój rezultat uzyskany w części poprzedniej, wypowiadając prawidłowo jeden ton. Osoba ta konsekwentnie nadawała sylabom z tonem 2 przebiegi typowe dla tonu opadająco-wznoszącego, z niekiedy wyraźnie zaznaczonym spadkiem F_0 , po którym następowała zmiana kierunku. W analogiczny sposób zniekształcono wszystkie pozostałe tony uznane za błędne, co w zestawieniu z danymi dotyczącymi produkcji tonów na sylabach języka mandaryńskiego pozwala zasugerować, że poprawna realizacja tonu rosnącego może być dla Polaków zadaniem niełatwym. W celu weryfikacji przedstawionej tu hipotezy niezbędne jest jednak przeprowadzenie szczegółowych badań z udziałem większej liczby mówców.

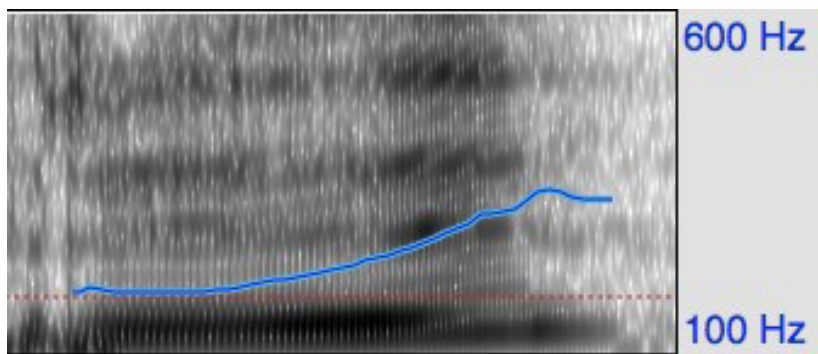
Spektrogram na ryc. 30 obrazuje zniekształcenie tonu 2, którego efektem jest kontur charakterystyczny dla tonu opadająco-wznoszącego. Ton został wypowiedziany przez osobę AK na sylabie *pstry*.

Ryc. 30. Zniekształcony kontur tonu 2 – sylaba *pstry*.



Dla porównania prawidłowa realizacja tej samej sylaby w wykonaniu osoby GK (ryc. 31):

Ryc. 31. Prawidłowy przebieg tonu 2 – sylaba *pstry*.

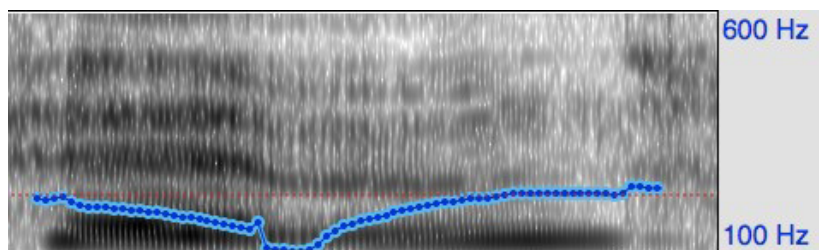


Tak jak w języku mandaryńskim produkcja tonu 3 w wykonaniu grupy mówców polskich okazała się lepsza niż przewidywano. Ponownie dwie spośród osób badanych (EK oraz TM) wykazały się bezbłędną realizacją; do wyniku tego zbliżyła się także osoba GK, wypowiadając prawidłowo 9 z 10 sylab z tonem 3. Wszystkie tony zakwalifikowane jako właściwe charakteryzują się konturem o kierunku wyraźnie opa-

dająco-wznoszącym, choć w zależności od mówcy nabierają pewnych odmiennych cech, typowych dla każdego z nich. I tak tony w wykonaniu obu głosów żeńskich (EK i GK) na ogół są bardziej wyraziste niż te wypowiedziane przez głos męski (TM), co na tonogramach objawia się szerszym zakresem F_0 , mocniej zaznaczonym spadkiem oraz częścią rosnącą sięgającą stosunkowo wyższego rejestru. Pomimo częstej obecności *creaky voice* w głosie badanej EK zrealizowane przez nią tony mają kontury o kształcie najbardziej zbliżonym do pożądanego, podczas gdy te wypowiedziane przez GK mają kształt znacznie łagodniejszy i bardziej zaokrąglony. Z kolei wszystkie tony w wykonaniu mówcy TM charakteryzują się mocno spłaszczonymi konturami, co zwraca uwagę mimo dużo mniejszej rozpiętości skali głosów męskich.

Spektrogram przedstawiony na ryc. 32 obrazuje przebieg tonu 3, wypowiedzianego przez osobę EK na sylabie *hymn*.

Ryc. 32. Modelowy kontur tonu 3 – sylaba *hymn*.



Analiza sylab z tonem 3 w wykonaniu badanej AK ponownie okazała się zadaniem problematycznym z uwagi na zniekształcenia ich konturów, będące następstwem fonacji *creaky voice*. Na tonogramach objawia się to w postaci urywanych, fragmentarycznych przebiegów, o kierunkach niekiedy trudnych do określenia. Z uwagi na fakt, że przebiegów tych nie poddano późniejszej weryfikacji audytywnej słuchaczy, decyzję o jakości realizacji podjęto wyłącznie na podstawie stopnia podobieństwa kształtu konturu danego tonu do formy pożądanego, w wyniku czego cztery kontury zakwalifikowano jako nieprawidłowe. W dwóch przypadkach po dokonaniu niezbędnych korekt analizie poddano je-

dynie części rosnące konturów, gdyż fragmenty objęte *creaky voice* nie podlegały pomiarom. Jeden z tych konturów ma silnie spłaszczony kształt, typowy dla tonu 1. Podobne zniekształcenia zaobserwowano w przebiegu kolejnego z błędnych tonów, który natomiast pozbawiony jest części rosnącej.

Łącznie w grupie polskiej zrealizowano nieprawidłowo 27 tonów, co daje odsetek błędów w wysokości 16,9%. Spośród wybranych sylab tylko *lgnąć* z tonem 4 sprawiła trudności w realizacji każdemu z czterech mówców, natomiast aż 26 na 40 wariantów tonalnych zostało przez nich wypowiedzianych bezbłędnie. Biorąc pod uwagę ogólną liczbę błędnych wystąpień każdego z 10 różnych wyrazów, do najbardziej problematycznych ponownie należał wyraz *lgnąć*, który uznano za niewłaściwy w sześciu przypadkach, a następnie *wprost* – w czterech przypadkach. Zdecydowanie najmniej kłopotliwe okazały się wyrazy *strzęp* oraz *pstry*, dla których zanotowano po jednej nieprawidłowej realizacji, z czego obie były udziałem badanej AK.

Opierając się na wynikach części II badania z udziałem rodzimych użytkowników języka mandaryńskiego oraz rezultatach uzyskanych przez mówców polskich w obu częściach eksperymentu, można stwierdzić, że budowa sylaby ma istotny związek z jej funkcją jako nośnika tonu, a także w sposób znaczący wpływa na jego formę. W związku z tym za sylabę wyjątkowo niesprzyjającą stabilności realizowanego na niej tonu należy uznać sylabę *lgnąć*, która z uwagi na swą strukturę (por. s. 155-156) stała się źródłem licznych trudności obu badanych grup. Niewątpliwie czynnikiem, który odegrał pewną rolę w niniejszym badaniu, były indywidualne predyspozycje mówców, związane z ich zdolnością do imitacji tonów na sylabach dobrze znanego im języka nietonalnego, przy czym umiejętności te nie zawsze idą w parze z kompetencją tychże osób w zakresie posługiwania się językiem tonalnym – w tym wypadku językiem mandaryńskim. Jest natomiast wysoce prawdopodobne, że są one powiązane z częstością i poziomem zaawansowania używanego języka, charakterem pracy z nim, celami, jakim służy, ewentualnymi kontaktami mówcy z rodzimymi użytkownikami itd., co – zgodnie z informacjami uzyskanymi przed przystąpieniem do nagrania – mogło w pewnym stopniu zadecydować o lepszych wynikach osób GK i TM.

4.2.3. Porównanie wyników obu części eksperymentu

- 1) Spośród mówców polskich dwoje (EK i TM) posługiwało się w obu językach skalą o względnie stałej rozpiętości; u pozostałych osób zanotowano natomiast zauważalne zmniejszenie wykorzystywanego zakresu F_0 .
- 2) Tendencje do zniekształcania niektórych konturów, a także swoiste trudności poszczególnych mówców, zaobserwowane podczas realizacji tonów na sylabach mandaryńskich, uległy przeniesieniu także na sylaby polskie. U każdego mówcy wszelkie nieprawidłowości występowały w obu tych językach z podobnym nasileniem, przez co odsetek tonów wypowiedzianych przez nich poprawnie pozostał względnie stały; wyjątek stanowi mówca TM, u którego zanotowano różnicę w wysokości 10% na korzyść tonów realizowanych na sylabach języka polskiego.
- 3) Ogółem Polacy wykazali się bardzo dobrą produkcją tonów, zarówno w języku mandaryńskim, jak i polskim, osiągając w obu przypadkach ponad 80-proc. poprawność, przy czym rezultat uzyskany na sylabach polskich nieznacznie przewyższa odsetek prawidłowych realizacji sylab mandaryńskich.
- 4) W odróżnieniu od rodzimych użytkowników języka mandaryńskiego, u mówców polskich stabilność tonów nie wykazuje związku z językiem, w którym są one realizowane, lecz ulega zmianom, co jest widoczne w wartościach parametrów branych pod uwagę w niniejszej analizie. Wbrew przewidywaniom większą konsekwencją pod względem wysokości F_0 odznaczały się tony wypowiedane na sylabach polskich, co znajduje odzwierciedlenie w na ogół niższych niż w przypadku sylab mandaryńskich wartościach odchyień standardowych. Jednocześnie, tony te wykazują większe wahania pod względem iloczasu i zakresu F_0 , natomiast ich kontury charakteryzują się nieco mniej regularnymi przebiegami.
- 5) U wszystkich osób badanych współczynnik modulacji częstotliwości zachował zbliżoną wartość w obu językach.
- 6) Podobnie jak w grupie rodzimych użytkowników języka mandaryńskiego, wśród Polaków także zanotowano przypadki segmentacji polskich wyrazów jednosylabowych na mniejsze elementy, niemniej

zjawisko to wystąpiło w znacznie ograniczonym zakresie oraz spowodowane było innymi czynnikami.

4.3. Porównanie wyników rodzimych i nierodzimych użytkowników języka mandaryńskiego

Szczegółowa analiza tonograficzna, przeprowadzona na materiale pochodzącym z nagrań obu badanych grup (por. § 4.1 i 4.2), pozwoliła uzyskać ogólny obraz sposobu i jakości realizacji tonów w każdym z języków poprzez poznanie pewnych idiolektalnych cech poszczególnych mówców, zidentyfikowanie trudności, które napotkali oraz błędów, które popełnili, i dzięki temu umożliwiła dostrzeżenie istotnych różnic pomiędzy rodzimymi i nierodzimymi użytkownikami języka mandaryńskiego. Analiza ta wykazała ponadto, że tonalne *sandhi*, licznie występujące w realizacjach *native speakerów*, jest zjawiskiem niezwykle powszechnym, którego nie należy ignorować, sygnalizując tym samym potrzebę dokładniejszego zbadania natury tonu $\frac{1}{2}$ 3. Tabele 17–22 prezentują końcowe rezultaty każdej z grup, zestawione w celu porównania.

Tabela 17. Porównanie wyników realizacji tonów na sylabach chińskich przez obie grupy mówców.

JĘZYK MANDARYŃSKI

MÓWCA	śr. F0 [Hz]	SKALA	σ	L. PRAWDŁOWO ZREALI- ZOWANYCH TONÓW				
				1	2	3 ²⁶		4
						3 (V) ²⁷	½ 3 (\) ²⁸	
DM	131,8	126,9	26,5	10	10	0	10	10
JK	204,1	228,2	40,6	10	6	0	10	10
WK	246,6	323,4	46,7	10	9	7	3	10
YM	143,9	122,1	17,7	10	3	10	0	10
śr.	181,6	$\Sigma \rightarrow$		40	28	17	23	40
AK	229,3	381,1	42,1	10	1	8		10
EK	224,5	173,9	24,1	10	3	10		10
GK	231,5	429,7	17,1	10	6	10		10
TM	112,1	83,7	15,7	10	6	9		8
śr.	199,4	$\Sigma \rightarrow$		40	16	37		38

²⁶ Pomimo uznania tonu ½ 3 za błędny, na potrzeby niniejszej analizy oraz w celach porównawczych został on uwzględniony w końcowym zestawieniu (dot. tabel 17–22).

²⁷ Symbol V oznacza ton zrealizowany modelowo (dot. tabel 17–22).

²⁸ Symbol \ oznacza część opadającą tonu 3, uwzględnianą w pomiarach tonów, w przypadku których nie zanotowano wystąpienia zjawiska *sandhi* (dot. tabel 17–22).

Tabela 18. Porównanie wyników realizacji tonów na sylabach polskich przez obie grupy mówców.

JĘZYK POLSKI

MÓWCA	śr. F0 [Hz]	SKALA	σ	L. PRAWIDŁOWO ZREALIZOWANYCH TONÓW				
				1	2	3		4
						3 (V)	$\frac{1}{2}$ 3 (l)	
DM	123,6	118,4	23,1	6	7	1	9	8
JK	221,1	232,5	37,8	5	3	0	10	6
WK	263,5	384,9	30,4	2	9	2	0	6
YM	150,8	163	16,9	1	3	10	0	3
śr.	189,8	$\Sigma \rightarrow$		14	22	13	19	23
AK	228	296,1	40,7	10	1	6		9
EK	223,5	200,8	24	10	4	10		9
GK	239,7	250	15,9	10	9	9		9
TM	111,7	92,2	14,3	10	8	10		9
śr.	200,7	$\Sigma \rightarrow$		40	22	35		36

Tabela 19. Porównanie średnich zakresów poszczególnych tonów zrealizowanych na sylabach chińskich przez mówców z obu grup.

JĘZYK MANDARYŃSKI

MÓWCA	śr. zakres F0 [Hz]				
	TON 1	TON 2	TON 3		TON 4
			3 (V)	½ 3 (V)	
DM	21,5	39,6		39,9 (s)	96,3
JK	33,3	82,8		88,3 (s)	168,4
WK	49,8	137,4	133,5	126,1	192,8
YM	21,6	71,7	59,7	36,1	89,3
AK	34,8	112,4	90,4	54,7	186,8
EK	36,2	83,1	99,9	74,4	109,3
GK	50	140,3	151,2	55,9	154,9
TM	23,3	37,8	33,5	19,2	61

Tabela 20. Porównanie średnich zakresów poszczególnych tonów zrealizowanych na sylabach polskich przez mówców z obu grup.

JĘZYK POLSKI

MÓWCA	śr. zakres F0 [Hz]				
	TON 1	TON 2	TON 3		TON 4
			3 (V)	½ 3 (V)	
DM	28,1	35	25,9 ²⁹	30,9 (s)	69,8
JK	58,8	68,9		78,6 (s)	153,8
WK	149,8	175	178,7	74,9	157,1
YM	56,3	70,3	68	26,2	83,7
AK	47,7	108,5	98,5	72,7	186,7
EK	47,2	92,6	88,8	72,9	126,9
GK	55,6	166,5	196,6	58,6	157,1
TM	14,1	39,7	30,4	23,5	56,7

²⁹ Wartość dotyczy jednego tonu zrealizowanego modelowo.

Tabela 21. Porównanie średnich iloczynów sylab chińskich w realizacji obu grup.

JĘZYK MANDARYŃSKI

MÓWCA	śr. λ [m/s]				
	TON 1	TON 2	TON 3		TON 4
			3 (V)	½ 3 (V)	
DM	341	311		246 (s)	234
JK	515	535		432(s)	391
WK	602	599	659	334 ³⁰	271
YM	548	521	710	335	331
AK	633	665	678	203	314
EK	480	497	521	210	272
GK	646	469	546	216	317
TM	414	371	359	154	251

³⁰ Wartość jest średnią długości trzech tonów ½ 3 oraz części opadających 7 tonów zrealizowanych modelowo.

Tabela 22. Porównanie średnich iloczynów sylab polskich w realizacji obu grup.

JĘZYK POLSKI

MÓWCA	śr. λ [m/s]				
	TON 1	TON 2	TON 3		TON 4
			3 (V)	½ 3 (V)	
DM	248	250	189	239 (s)	233
JK	334	348		262 (s)	286
WK	363	436	434 ³¹	184	280
YM	335	370	499	248	347
AK	494	510	510	211	269
EK	372	415	473	209	278
GK	538	355	482	170	269
TM	311	256	315	156	206

Pierwszą wyraźną różnicą pomiędzy rodzimymi użytkownikami języka mandaryńskiego a Polakami, jaką ujawniły podane dane, jest brak występowania zjawiska *sandhi* u mówców polskich, co stanowi główną przyczynę lepszych wyników Polaków w realizacji sylab chińskich. Uznając ton ½ 3 za dopuszczalną formę tonu 3, mówcy „chińscy” uzyskaliby rezultat w wysokości 148 poprawnie wypowiedzianych tonów na 160 wszystkich sylab, co przewyższa rezultat grupy polskiej, wynoszący 131 poprawnych realizacji.

Znacznie gorzej przedstawiają się wyniki grupy „chińskiej” w imitacji tonów na sylabach polskich – liczba tonów zrealizowanych prawidłowo wynosi 72, natomiast po uwzględnieniu poprawności tonu ½ 3–91, podczas gdy mówcy polscy utrzymali rezultat z poprzedniej części badania, realizując bezbłędnie 132 tony. Fakt ten pozwala przy-

³¹ Wartość jest średnią długości dwóch tonów zrealizowanych modelowo oraz ośmiu tonów zrealizowanych jako drugie.

puszczać, że przyczyną zniekształceń tonów w wykonaniu mówców „chińskich” były trudności w artykulacji zbitek spółgłoskowych, licznie występujących w materiale w języku polskim.

Jak wynika z tabel 17 i 18, mówcy polscy ogółem realizowali tony w wyższych częstotliwościach niż mówcy „chińscy”, co jest zapewne wynikiem nierównych proporcji grup pod względem płci. Nie stwierdzono natomiast korelacji pomiędzy rozpiętością skali a poprawnością realizacji tonów, choć ogółem głosy męskie odznaczały się większą stabilnością. Brak regularności danych pod względem zakresu F_0 uniemożliwia wyszczególnienie różnic w realizacji tonów pomiędzy obiema grupami. Porównawszy wartości z obu tabel, można zauważyć, że w języku polskim nastąpiło ogólne (poza mówcą TM) rozszerzenie zakresu sylab z tonem 1 względem zakresów sylab chińskich, co na tonogramach uwidacznia się w postaci nieregularnych przebiegów. Odwrotną tendencję zaobserwowano w przypadku sylab z tonem 4, których zakresy w przeważającej większości uległy zmniejszeniu w drugiej części badania.

Ciekawych spostrzeżeń dostarcza dokładna analiza tonów 3. Mianowicie, porównanie przeciętnych zakresów tonów $\frac{1}{2}$ 3 z zakresami części opadających modelowo zrealizowanych tonów 3 pokazuje, że u obu płci tony $\frac{1}{2}$ 3 wypowiedane były z użyciem szerszej skali F_0 niż w przypadku odpowiednich części tonów zrealizowanych jako dwukonturowe. Podobna prawidłowość uwidacznia się także pod względem iloczasu – tony będące efektem zjawiska *sandhi* są z reguły dłuższe niż kontury opadające modelowych tonów 3, przy czym w języku polskim różnica ta jest wyraźniejsza. Dodatkowo, tony $\frac{1}{2}$ 3 wypowiedane przez badaną JK na sylabach chińskich są średnio dłuższe niż tony 4 w wykonaniu wszystkich mówców w tym języku, natomiast w języku polskim długości obu tych tonów są porównywalne. Powyższe obserwacje mogą stanowić podstawę do stwierdzenia, że ton $\frac{1}{2}$ 3 nie jest de facto tonem „urwanym”, będącym wynikiem prostej redukcji tonu opadająco-wznoszącego o jeden z konturów, ani też tonem typowo opadającym, lecz pewną odrębną formą, wykazującą cechy tonu „pełnowymiarowego” o charakterze odmiennym od tonów pozostałych.

4.4. Test słuchowy

Pomimo że – według przyjętych w niniejszym badaniu kryteriów oceny tonów wypowiedzianych przez mówców z obu badanych grup – wyższą poprawność realizacji osiągnęli łącznie Polacy, do odsłuchów postanowiono wykorzystać materiał z nagrań rodzimych użytkowników języka mandaryńskiego, dla których tony są zjawiskiem bardziej naturalnym i oczywistym niż dla osób uczących się języka tonalnego jako obcego. Należy również zauważyć, że słabszy rezultat *native speakerów* jest przede wszystkim wynikiem częstej realizacji przez nich tonów $\frac{1}{2}$ 3, będących następstwem *sandhi* tonalnego, a ponieważ, zgodnie z teorią, nie stwierdzono wystąpienia warunków koniecznych do zaistnienia tego zjawiska, tony te uznano za błędne. Niemniej *sandhi* tonalne jest zjawiskiem powszechnie występującym w języku mandaryńskim, co potwierdza wysoki odsetek wystąpień tonów $\frac{1}{2}$ 3, wynoszący 52,5% ogólnej liczby sylab z tonem 3. W związku z tym, uznając ten sposób realizacji za dopuszczalny, mówcy natywni uzyskaliby poprawność w wysokości 92,5%. Oprócz głównego celu niniejszego badania, jakim jest sprawdzenie percypowalności tonów realizowanych na sylabach języka nietonalnego, wykorzystany materiał dodatkowo stworzył możliwość przetestowania wrażliwości słuchaczy na zjawisko *sandhi* i związanej z tym zdolności identyfikacji jego efektów w ogóle.

Tabele 21 i 22 stanowią zestawienia ogólnej liczby tonów prawidłowo zidentyfikowanych przez słuchaczy w zależności od nadawców. W środkowej części obu tabel przedstawiono liczbę identyfikacji na 40 tonów wypowiedzianych przez danego mówcę w każdym z testów. Końcowe rezultaty, uzyskane przez poszczególnych słuchaczy w całym teście (tj. na 160 realizacji), uszeregowano malejąco w przedostatniej kolumnie. W przedostatnim rzędzie znajduje się liczba trafnych identyfikacji tonów zrealizowanych przez każdego z mówców na 400 wystąpień w 10 odtworzeniach testu ogółem.

Wszystkie końcowe wyniki wyrażono także procentowo.

Tabela 23. Liczba i odsetek poprawnie zidentyfikowanych tonów zrealizowanych na sylabach chińskich przez rodzimych użytkowników języka mandaryńskiego.

TEST 1 – WYRAZY MANDARYŃSKIE

MÓWCA SŁUCHACZ	JK	WK	DM	YM	SUMA	%
	LICZBA POPRAWNYCH ODPOWIEDZI					
EK	36	38	30	36	140	87,5
JK	34	35	30	37	136	85
AK	34	36	26	35	131	81,9
PK	29	32	30	33	124	77,5
SM	30	29	26	38	123	76,9
KK	32	31	25	33	121	75,6
DK	29	33	17	26	105	65,6
BK	27	23	14	29	93	58,1
MK	19	29	19	26	93	58,1
GK	16	21	14	19	70	43,8
SUMA	286	307	231	312	1136	
%	71,5	76,8	57,8	78		71
SD	6,5	5,4	6,5	6,1		

Tabela 24. Liczba i odsetek poprawnie zidentyfikowanych tonów zrealizowanych na sylabach polskich przez rodzimych użytkowników języka mandaryńskiego.

TEST 2 – WYRAZY POLSKIE

MÓWCZA SŁUCHACZ	JK	WK	DM	YM	SUMA	%
	LICZBA POPRAWNYCH ODPOWIEDZI					
AK	26	24	26	28	104	65
JK	27	21	27	28	103	64,4
EK	27	20	22	26	95	59,4
PK	20	19	24	21	84	52,5
SM	20	18	22	22	82	51,3
MK	16	24	16	20	76	47,5
KK	19	19	15	18	71	44,4
DK	18	15	20	9	62	38,8
BK	19	15	11	12	57	35,6
GK	9	17	13	14	53	33,2
SUMA	201	192	196	198	787	
%	50,3	48	49	49,5		49,2
SD	5,5	3,2	5,6	6,6		

Jak pokazują wartości umieszczone w prawych dolnych rogach tabel, znacznie lepiej zidentyfikowano tony realizowane na sylabach języka mandaryńskiego, będącego rodzimym językiem mówców – w teście 1 dziesięć słuchaczy w sumie zidentyfikowało poprawnie 1136 tonów spośród wszystkich 1600 odtworzeń, czyli o 349 tonów więcej niż w teście 2, co daje wynik wyższy o ponad 20%. Odsetek poprawnie zidentyfikowanych tonów na sylabach polskich, sięgający niespełna 50%, sugeruje, że trafne identyfikacje mogą być w dużej mierze dziełem przypadku. Biorąc pod uwagę dane z analizy produkcji tonów

przez mówców (por. § 4.1), można stwierdzić, że na wyniki testów niewątpliwie wpływ miała poprawność realizacji tonów, która znalazła odzwierciedlenie w odbiorze słuchowym. I tak, w języku mandaryńskim zrealizowano prawidłowo 125 z 160 tonów, podczas gdy w języku polskim stanowiły one mniej niż połowę ogólnej liczby (72). Uznawszy ton $\frac{1}{2}$ 3 za dopuszczalną formę tonu 3, liczby te wzrastają odpowiednio do 148 oraz 91 tonów, przez co różnica pomiędzy oboma językami ulega pogłębieniu. W związku z tym nawet bezbłędna identyfikacja rzeczywistych przebiegów tonalnych nie pozwoliłaby uzyskać 100% poprawności. Należy również rozważyć wpływ błędów w fonetycznej realizacji wyrazów, wynikających z niedokładnej artykulacji niektórych głosek języka polskiego przez jego nierodzimych użytkowników, co mogło w istotny sposób utrudniać identyfikację tonów. Dokładniejsza analiza z uwzględnieniem czynników kształtujących percepcję słuchową znajduje się poniżej.

4.4.1. Kompetencja językowa oraz indywidualne predyspozycje słuchowe odbiorcy

Pobieżne oględziny tabel wyników obu testów pozwalają zauważyć, że kolejność słuchaczy, odpowiadająca uzyskanym przez nich rezultatom od statystycznie najlepszego do najslabszego, nie uległa radykalnej zmianie po drugiej części odsłuchu (test w języku polskim). W obu przypadkach trzy najwyższe wyniki należą do tych samych osób: AK, EK oraz JK, przy czym JK utrzymała drugą pozycję, EK, która w pierwszym teście zidentyfikowała o dziewięć tonów więcej niż AK, o tyle samo tonów mniej od współuczestniczki zidentyfikowała za drugim razem. Stałe miejsca w tabeli zachowały również osoby PK i SM, które w każdym z testów osiągnęły odpowiednio 4. i 5. z kolei rezultat. Podobnie sytuacja przedstawia się na końcu rankingu, gdzie ostatnia pozycja niezmiennie należy do GK, natomiast tuż przed nią znalazła się BK, która notabene uzyskała wynik *ex aequo* z badaną MK, lecz z uwagi na pierwszeństwo w kolejności alfabetycznej umieszczona została jako pierwsza. Niewielką rotację w rankingu zaobserwować można w środkowej części obu zestawień pomiędzy badanymi DK, KK oraz MK, przy czym różnica pomiędzy skrajnymi wynikami jest

w tym wypadku znaczna. Rozrzut wartości w obu testach można uznać za dość duży, przy czym w każdym z nich wynik najniższy stanowi niemal dokładnie połowę wartości najlepszego. I tak, w teście 1 rozstęp wynosi 43,7% przy najniższym uzyskanym wyniku w wysokości 43,8%, natomiast w teście 2 jest to 31,8% przy odsetku poprawnych identyfikacji równym 33,2% .

Zarówno AK, jak i EK (do których należą czołowe rezultaty) odznaczają się wysoką kompetencją w zakresie znajomości języka mandaryńskiego, co stwierdzono na podstawie informacji uzyskanych przed przystąpieniem do odsłuchów. Zgodnie z nimi w momencie uczestnictwa w odsłuchu obie studentki uczyły się języka od pięciu lat, przy czym jedna z nich odbyła roczne stypendium w Chinach, a druga dodatkowo uczestniczyła w kursie na Tajwanie. Każda z nich ma również pewne doświadczenie w pracy tłumacza języka chińskiego, w związku z czym wyniki uzyskane przez obie badane nie stanowią zaskoczenia. Interesujący jest natomiast przypadek JK, która w każdym z testów osiągnęła rezultat niewiele odbiegający od najlepszego, pomimo że miała za sobą jedynie roczną naukę języka mandaryńskiego. Fakt ten świadczy o tym, że u osoby JK większą rolę w słuchowym odbiorze przebiegów tonalnych odegrały indywidualne predyspozycje i zdolności, nabyte lub wrodzone, które mogą być związane z posiadanym wykształceniem muzycznym oraz aktywnie rozwijanymi umiejętnościami gry na instrumencie. Należy również wspomnieć, że tego typu wiedzą i umiejętnościami dysponuje także AK, choć od paru lat nie były one doskonalone.

Nieco wyższy wynik niż uzyskany prognozowano dla osób BK i DK, z których pierwsza ukończyła 5-letnie studia, w międzyczasie uczestnicząc w kursie językowym na Tajwanie, natomiast druga przebywała w Chinach 1,5 roku po odbyciu 3-letniej nauki. Rezultat GK, która wykazała się statystycznie najsłabszą zdolnością identyfikacji, potwierdził przewidywania z uwagi na fakt, że osoba ta miała styczność z językiem mandaryńskim dopiero od roku. Pozostali słuchacze również uzyskali spodziewane wyniki.

Jak wspomniano, wiele błędów w identyfikacji wynika z niedokładnej realizacji niektórych tonów przez mówców, co jest zjawiskiem szczególnie częstym w przypadku imitacji tonów na wyrazach języka polskiego. Wniosek ten sformułowano na podstawie konfrontacji błędnych odpowiedzi słuchaczy z obrazami graficznymi odpowied-

nich tonów zarejestrowanymi na tonogramach, które potwierdziły, że w wielu przypadkach ich przebiegi wykazywały cechy typowe dla przebiegów innych tonów. Zniekształcenia dotyczyły głównie tonu 3, który mówcy realizowali jako ton $\frac{1}{2}$ 3, kształtem przypominający kontur tonu opadającego bądź jako ton 2, co częściej miało miejsce na sylabach polskich. W drugiej kolejności modyfikacjom podlegał ton 2, w dużej części identyfikowany przez słuchaczy jako ton 3, ze względu na wyraźnie opadająco-wznoszący przebieg. Sporadycznie zniekształcano także ton 1, który często przybierał postać tonu rosnącego. Uwzględniając odpowiedzi zgodne z faktyczną realizacją, wyniki słuchaczy przedstawiają się następująco:

Tabela 25. Liczba i odsetek tonów na sylabach chińskich, zidentyfikowanych zgodnie z faktyczną realizacją.

TEST 1		
SŁU-CHACZ	LICZBA POPRAWNYCH ODPOWIEDZI	%
EK	157	98
JK	154	96,3
AK	147	91,9
PK	142	88,8
KK	136	85
SM	136	85
DK	117	73,1
MK	115	71,9
BK	109	68,1
GK	89	55,6

Tabela 26. Liczba i odsetek tonów na sylabach polskich, zidentyfikowanych zgodnie z faktyczną realizacją.

TEST 2		
SŁU-CHACZ	LICZBA POPRAWNYCH ODPOWIEDZI	%
JK	145	90,6
AK	142	88,8
EK	139	86,9
PK	121	75,6
SM	116	72,5
MK	112	70
KK	102	63,8
DK	97	60,6
BK	83	51,9
GK	80	50

Jak widać, wyniki każdego z słuchaczy wyraźnie wzrosły w stosunku do rezultatów sprzed weryfikacji z użyciem wykresów tonograficznych, przy czym w przypadku sylab polskich różnica ta wynosi średnio 20–25%, czyli nieco ponad dwukrotnie więcej niż dla sylab chińskich (10%). Częściowo tłumaczy to większa liczba niewłaściwie zrealizowanych tonów w języku polskim, co przy dobrym odbiorze słuchowym przełożyło się na liczne błędy w udzielanych odpowiedziach, będące efektem niezgodności pomiędzy oczekiwaną formą tonów a ich rzeczywistym przebiegiem. Otrzymane dane pozwalają stwierdzić, że ogółem grupa wykazała się dobrą percepcją, identyfikując łącznie 81,4% tonów w języku mandaryńskim oraz 71,1% tonów w języku polskim. Najwyższy uzyskany wynik (EK) sięga 98%, natomiast żaden z rezultatów nie spada poniżej 50%. W przeważającej części zachowała się kolejność słuchaczy względem liczby poprawnych odpowiedzi, co w dużej mierze wyklucza przypadkowość osiągniętych przez poszczególne osoby rezultatów i stanowi niejako potwierdzenie jakości ich kompetencji językowych i/lub pewnych indywidualnych predyspozycji; jedyną istotną zmianą, jaką zanotowano, było uzyskanie przez osobę JK najwyższego wyniku w teście 2, która zidentyfikowała dodatkowo 42 tony, wyprzedzając tym samym badaną AK. Dodatkowo, osoba ta zidentyfikowała prawidłowo dziewięć tonów trzecich, zrealizowanych zgodnie z zasadami *sandhi* tonalnego jako tony $\frac{1}{2}$ 3. Należy przypomnieć, że JK jest uzdolnioną muzycznie studentką po zaledwie rocznej nauce języka chińskiego.

Na podstawie wyników powyższej analizy nie stwierdzono korelacji pomiędzy liczbą lat nauki a zdolnością do identyfikacji tonów; choć w obu testach osoby zajmujące skrajne pozycje w rankingu osiągnęły oczekiwane rezultaty, zdarzały się przypadki stanowiące zaprzeczenie związku tych dwóch zmiennych. Należy jednak zaznaczyć, że istnienie takiej zależności nie jest tym samym negowane, jedynie ze względu na zbyt małą liczbę przebadanych osób uzyskane dane nie są wystarczająco obszerne do ekstrapolacji. Analiza wykazała natomiast możliwość dodatniego wpływu czynników pozajęzykowych na jakość odbioru słuchowego takich jak np. rozwinięty słuch muzyczny bądź inne, nabyte lub wrodzone, indywidualne uwarunkowania odbiorcy, do stwierdzenia których konieczne jest przeprowadzenie dalszych szczegółowych badań.

4.4.2. Wpływ nadawcy

Istnienie zależności pomiędzy poszczególnymi mówcami a liczbą trafnych identyfikacji wypowiedzianych przez nich tonów badano przy użyciu metod statystycznych. W tym celu dla wyników obu testów przeprowadzono test niezależności chi-kwadrat, umożliwiającą sprawdzenie czy istniejące pomiędzy nimi różnice są znaczące. W obu przypadkach przyjęto tę samą hipotezę zerową (H_0) oraz hipotezę alternatywną (H_1), które sformułowano w sposób następujący:

H_0 : Zmienne są niezależne.

H_1 : Zmienne nie są niezależne.

Korpus danych stanowiły tabele 21 i 22 (por. s. 169–170), które skonfrontowano z wynikami badania nad realizacją tonów z wykorzystanego w testach materiału (por. § 4.1). Rezultaty pokazują, że w teście 1 najlepszy rezultat uzyskał głos męski YM, którego tony zostały poprawnie zidentyfikowane w 78%, przy czym z analizy danych z poprzedniego badania wynika, że mówca ten wykazał się dobrą produkcją tonów i nie realizował *sandhi* tonalnego. Porównywalny wynik (76,8%) zanotowano dla jednego z głosów kobiecych (WK), który odznaczał się najwyższą poprawnością realizacji. W następnej kolejności znalazła się osoba JK z wynikiem 71,5%, a na końcu DM (57,8%). Natomiast w teście 2 wszyscy mówcy uzyskali podobny odsetek poprawnych identyfikacji, z niewielką przewagą głosu JK (50,1%), za którym znalazł się YM (49,5%), następnie DM (49%) oraz WK (48%), przy czym kolejność według kryterium poprawności realizacji przedstawia się niemal zupełnie odwrotnie, z najniższym rezultatem głosu JK. Przy założonym poziomie istotności 5%, otrzymane prawdopodobieństwa $p = 0,79$ oraz $p = 0,77$ odpowiednio dla testów 1 i 2 nie dają podstaw do odrzucenia hipotezy zerowej w żadnym z przypadków.

Podczas analizy danych z tabeli 21 i 22 zwraca uwagę różnica w rozrzucie wartości sumarycznych dla każdego z mówców pomiędzy oboma testami, który w teście z wykorzystaniem sylab języka mandaryńskiego jest znaczący, natomiast w teście „polskim” – praktycznie nieistniejący, co potwierdzają wielkości rozstępów równe odpowiednio 81 oraz 9 poprawnych identyfikacji. Wyniki uzyskane dla mówców wykazują także różnice pod względem liczby poprawnych ocen dokonanych przez poszczególnych słuchaczy. Jak pokazują wartości

odchylen standardowych, największe zróżnicowanie ocen w teście 1 występuje w głosach DM oraz JK, podczas gdy najbardziej zbliżone wyniki uzyskał głos WK. Z kolei w teście 2, głosem najstabilniejszym pod tym względem ponownie okazuje się WK, natomiast największym rozrzutem charakteryzują się wartości otrzymane dla mówcy YM.

Wpływ nadawcy widać natomiast wyraźnie na przykładzie identyfikacji sylab z tonem 3, które zostały zrealizowane z zastosowaniem reguł *sandhi* tonalnego. Na sylabach chińskich zjawisko to wystąpiło u trzech mówców w 23 z 40 realizacji tonu 3, przy czym powstały w efekcie ton $\frac{1}{2}$ 3 zdecydowanie najmniej problemów w odbiorze sprawiał w wykonaniu JK. Na wszystkie 10 sylab, których obrazy graficzne przypominały przebiegi tonu opadającego, tony te zostały zidentyfikowane jako 3 w sumie 52 razy w 10 odtworzeniach. Znacznie gorszy był odbiór tonu 3 w wykonaniu mówcy DM, u którego *sandhi* tonalne występowało z taką samą częstotliwością – spośród 10 sylab zdarzył się jeden przypadek poprawnej identyfikacji. Osoba WK, która ton $\frac{1}{2}$ 3 zrealizowała trzykrotnie, uzyskała ogółem sześć poprawnych ocen. Z kolei w języku polskim *sandhi* było zjawiskiem nieco rzadszym, natomiast okazało się dużo bardziej kłopotliwe w percepcji – z 19 przypadków ton 3 został prawidłowo zidentyfikowany 17 razy we wszystkich 10 odtworzeniach, w tym 6 razy w 3 sylabach wypowiedzianych przez mówcę DM oraz 11 razy w 7 sylabach wypowiedzianych przez osobę JK.

4.4.3. Zależność od rodzaju tonu

W celu sprawdzenia, jak polscy słuchacze radzili sobie z identyfikacją poszczególnych tonów mandaryńskich, zsumowano liczbę wszystkich tonów 1, 2, 3 i 4 prawidłowo zidentyfikowanych przez każdą z osób badanych. Otrzymane wartości stanowią liczby prawidłowych odpowiedzi na 400 odtworzeń i nie uwzględniają zgodności z rzeczywistymi przebiegami tonów zarejestrowanymi na tonogramach. Wyniki analizy przedstawiono w tabeli 27 i 28.

Tabela 27. Liczba i odsetek prawidłowo zidentyfikowanych tonów na sylabach chińskich.

TEST 1		
TON	LICZBA POPRAW- NYCH ODP. /400	%
1	327	81,8
2	249	62,3
3	202	50,5
4	358	89,5

Tabela 28. Liczba i odsetek prawidłowo zidentyfikowanych tonów na sylabach polskich.

TEST 2		
TON	LICZBA POPRAW- NYCH ODP. /400	%
1	167	41,8
2	194	48,5
3	122	30,5
4	304	76

Tabele ukazują wyraźne trudności słuchaczy z identyfikacją tonu 3, który z racji częstego podlegania zjawisku *sandhi* realizowany był jako $\frac{1}{2}$ 3, i przez to z reguły mylony z tonem 4 – opadającym. W teście 1 ton 3 zidentyfikowano prawidłowo w 50,5%, podczas gdy w teście 2 odsetek ten spadł do 30,5%. W obu przypadkach zdecydowanie najmniej problemów w odbiorze sprawiał ton 4, przy czym w teście „polskim” odsetek identyfikacji (76%) znacznie przewyższa wartości uzyskane dla pozostałych tonów. Duży odsetek poprawnych identyfikacji miał także ton 1 realizowany na sylabach chińskich, gdzie uzyskał 81,8-proc. zgodność ocen słuchaczy. Ton 2, trafnie zidentyfikowany w 62,3% wystąpień, również okazał się dość problematyczny. Poza przypadkami, w których błędy w identyfikacji wynikały z niedoskonałej realizacji mówców, dość często zdarzało się, że ton ten interpretowano jako 3. Wśród słuchaczy zaobserwowano również tendencję odwrotną, która dotyczyła prawidłowo wypowiedzianych tonów 3, niemniej z uwagi na liczne występowanie zjawiska tonalnego *sandhi* w zarejestrowanym materiale dźwiękowym nie była ona szczególnie częsta. Fakt ten jest zgodny z obserwacjami Chuang et al. oraz Kiriloffa (odpowiednio 1972 i 1969, za: Gandour 1978), którzy zwracają uwagę na powszechność substytucji tonów 2 trzecimi oraz tonów 3 drugimi, z czym trudności niekiedy miewają także sami rodzimi użytkownicy języka mandaryń-

skiego. Jako wyjaśnienie Gandour wskazuje na pewne wspólne właściwości obu tonów, które znajdują potwierdzenie w niniejszej analizie³²: każdy z nich wykazuje przynajmniej częściowo rosnący przebieg, oba zaczynają się na porównywalnej wysokości oraz charakteryzują się zbliżonym iloczasem. Jak wynika z tabeli 28, ton 2 był percypowany relatywnie lepiej na sylabach polskich (48,5%), przewyższając pod tym względem odsetek poprawnych identyfikacji tonu 1 (41,8%).

Poniżej zamieszczono szczegółowy wykaz liczby trafnych identyfikacji z podziałem na pojedyncze jednostki leksykalne. Wyniki rozpisano z uwzględnieniem poszczególnych mówców, a następnie zsumowano i przedstawiono procentowo. Dodatkowo zacieniowano pola o wartościach skrajnych – w teście 1 kolorem ciemniejszym oznaczono wyniki powyżej 30 identyfikacji, natomiast jaśniejszym – te uzyskane dla sylab zidentyfikowanych w mniej niż 50% przypadków, z kolei w teście 2, ze względu na ogólnie niższe wyniki, za wartości graniczne przyjęto odpowiednio ≥ 30 oraz ≤ 10 .

Tabela 29. Liczba prawidłowych identyfikacji tonów na poszczególnych sylabach chińskich.

WYRAZ	MÓWCA TON	DM	JK	WK	YM	SUMA	%
		L. POPRAWNYCH IDENTYFIKACJI					
shi	1	6	10	7	8	31	77,5
	2	9	8	8	5	30	75
	3	0	6	10	9	25	62,5
	4	10	3	7	10	30	75
qiu	1	6	9	7	8	30	75
	2	9	8	9	4	30	75
	3	0	6	2	7	15	37,5
	4	10	8	10	10	38	95

³² Właściwości te są zgodne jedynie dla tonów zrealizowanych poprawnie, które stanowiły względnie mały odsetek w wykorzystanym materiale badawczym.

WYRAZ	MÓWCA	DM	JK	WK	YM	SUMA	%
	TON	L. POPRAWNYCH IDENTYFIKACJI					
gen	1	7	9	9	10	35	87,5
	2	4	7	7	6	24	60
	3	0	7	1	5	13	32,5
	4	9	9	9	10	37	92,5
bang	1	7	9	9	9	34	85
	2	4	7	8	7	26	65
	3	0	4	10	10	24	60
	4	10	9	9	10	38	95
mai	1	7	9	8	7	31	77,5
	2	7	6	3	4	20	50
	3	0	5	10	9	24	60
	4	10	9	10	9	38	95
peng	1	8	8	9	10	35	87,5
	2	5	3	5	7	20	50
	3	0	7	6	6	19	47,5
	4	10	8	10	9	37	92,5
yan	1	6	8	7	9	30	75
	2	10	8	7	8	33	82,5
	3	0	1	10	8	19	47,5
	4	10	8	6	7	31	77,5
shuo	1	9	10	10	9	38	95
	2	5	5	6	6	22	55
	3	0	5	9	8	22	55
	4	10	9	10	8	37	92,5
tan	1	6	9	7	10	32	80
	2	4	7	8	6	25	62,5
	3	0	6	3	10	19	47,5
	4	10	9	9	9	37	92,5

WYRAZ	MÓWCA TON	DM	JK	WK	YM	SUMA	%
		L. POPRAWNYCH IDENTYFIKACJI					
yu	1	6	9	8	8	31	77,5
	2	6	4	6	3	19	47,5
	3	1	5	9	7	22	55
	4	10	9	9	7	35	87,5

Tabela 30. Liczba prawidłowych identyfikacji tonów na poszczególnych sylabach polskich.

WYRAZ	MÓWCA TON	DM	JK	WK	YM	SUMA	%
		L. POPRAWNYCH IDENTYFIKACJI					
Igańć	1	1	5	2	2	10	25
	2	4	6	6	4	20	50
	3	0	0	4	8	12	30
	4	10	7	9	2	28	70
schnać	1	5	2	0	2	9	22,5
	2	5	3	4	2	14	35
	3	0	1	3	10	14	35
	4	9	9	9	8	35	87,5
wprost	1	6	7	1	2	16	40
	2	4	7	1	6	18	45
	3	0	1	4	7	12	30
	4	9	7	9	5	30	75
strzęp	1	6	7	6	6	25	62,5
	2	6	4	7	6	23	57,5
	3	0	0	0	8	8	20
	4	9	4	10	7	30	75

WYRAZ	MÓWCA	DM	JK	WK	YM	SUMA	%
	TON	L. POPRAWNYCH IDENTYFIKACJI					
pstry	1	7	9	3	6	25	62,5
	2	8	7	4	7	26	65
	3	3	2	1	6	12	30
	4	8	8	10	7	33	82,5
tchnąc	1	4	10	2	1	17	42,5
	2	3	5	3	6	17	42,5
	3	0	0	8	6	14	35
	4	8	6	9	5	28	70
trzymiel	1	6	4	1	1	12	30
	2	7	4	5	0	16	40
	3	2	2	4	5	13	32,5
	4	9	6	8	4	27	67,5
wszczać	1	2	9	1	1	13	32,5
	2	5	4	7	4	20	50
	3	0	1	4	10	15	37,5
	4	10	5	7	7	29	72,5
hymn	1	5	6	1	6	18	45
	2	4	7	8	3	22	55
	3	0	2	1	3	6	15
	4	9	8	9	3	29	72,5
czcić	1	7	9	5	1	22	55
	2	4	6	2	6	18	45
	3	1	2	4	9	16	40
	4	10	9	10	6	35	87,5

Na podstawie danych zawartych w dwóch ostatnich kolumnach tabel można oszacować, że średnio najlepiej identyfikowane spośród wyrazów chińskich były: *qiu* z wynikiem 81,7% oraz *tan* – 80%

poprawnych odpowiedzi, z kolei najmniejszą zgodnością ocen słuchaczy charakteryzowały się: *peng* – zidentyfikowany prawidłowo średnio w 50% wszystkich wystąpień – oraz *yu* (55%). Rozmieszczenie zacieniowanych pól pozwala zauważyć, że najwyższe wyniki regularnie przypadają na sylaby z tonami 1 i 4, z których najlepiej zidentyfikowano:

- *qiù* (ton 4), *bàng* (ton 4), *mài* (ton 4), *shuō* (ton 1): po 38 poprawnych odpowiedzi z 40, trafnie zidentyfikowane przez wszystkich słuchaczy w wykonaniu co najmniej dwóch mówców,
- *gèn*, *pèng*, *shuò*, *tàn* (każda z tonem 4): 37 poprawnych odpowiedzi z 40, prawidłowo zidentyfikowane przez wszystkich słuchaczy w wykonaniu co najmniej jednego mówcy.

Najmniejszą zgodność ocen zanotowano dla sylab:

- *gěn* (ton 3): 13 poprawnych odpowiedzi, *qiǔ* (ton 3): 15 poprawnych odpowiedzi, *pěng* (ton 3), *yǎn* (ton 3), *tǎn* (ton 3): po 19 poprawnych odpowiedzi, w wykonaniu mówcy DM niezidentyfikowanych przez żadnego z słuchaczy,
- *yú* (ton 2): 19 poprawnych odpowiedzi, zidentyfikowanych przez co najwyżej sześciu słuchaczy.

Spośród wyrazów polskich przeciętnie najmniej problemów w identyfikacji sprawiły: *tchnąć* (52,5%) oraz *wszcząć* (50%); średni odsetek poprawnych odpowiedzi dla pozostałych wyrazów nie przekroczył 50%. Zdecydowanie najgorzej identyfikowane były: *schnąć* i *trzymiel* (średnio po 35% pozytywnych ocen) oraz w następnej kolejności *hymn* (38,3%). Podobnie jak w przypadku wyrazów chińskich, najwyższą zgodność ocen zanotowano dla wyrazów z tonem 4 opadającym. Wśród nich należy wyróżnić:

- *schnąć* (ton 4): 35 poprawnych odpowiedzi z 40, zidentyfikowany przez 9 na 10 słuchaczy w wykonaniu 3 mówców (DM, JK, WK),
- *czcić* (ton 4): 35 poprawnych odpowiedzi z 40, zidentyfikowany przez wszystkich słuchaczy w wykonaniu 2 mówców (DM i WK),
- *pstry* (ton 4): 33 poprawne odpowiedzi z 40, zidentyfikowany przez wszystkich słuchaczy w wykonaniu mówcy WK.

Najbardziej problematyczne okazały się:

- *hymn* (ton 3): 6 poprawnych odpowiedzi z 40, w wykonaniu mówcy DM niezidentyfikowany przez żadnego z słuchaczy; najwyższa liczba prawidłowych identyfikacji: 3,

- *strzep* (ton 3): 8 poprawnych odpowiedzi z 40, zidentyfikowany jedynie w wykonaniu mówcy YM,
- *schnąć* (ton 1): 9 poprawnych odpowiedzi z 40, w wykonaniu WK niezidentyfikowany przez żadnego z słuchaczy; najwyższa liczba prawidłowych identyfikacji: 5.

Należy zwrócić uwagę, że jedną z przyczyn niskiej poprawności identyfikacji każdego z wymienionych wyrazów jest ich błędna realizacja. Po uwzględnieniu identyfikacji zgodnej z rzeczywistymi przebiegami poszczególnych tonów zarejestrowanymi na tonogramach otrzymujemy:

- *gěn* (ton 3): 13 + 17, *qiū* (ton 3): 15 + 13, *pěng* (ton 3): 19 + 10, *yǎn* (ton 3): 19 + 15, *tǎn* (ton 3): 19 + 17, *yú* (ton 2): 19 + 7 poprawnych odpowiedzi;
- *hymn* (ton 3): 6 + 16, *strzep* (ton 3): 8 + 22, *schnąć* (ton 1): 9 + 11 poprawnych odpowiedzi.

Podsumowując, jak wynika z analizy, w ocenie percepcji audytywnej tonów mandaryńskich kluczowe jest uwzględnienie kryterium poprawności realizacji. Wyniki otrzymane po zsumowaniu prawidłowych odpowiedzi słuchaczy znacząco różnią się od tych uzyskanych po dokonaniu weryfikacji odpowiednich tonogramów – w przypadku tonów na sylabach mandaryńskich, ogółem realizowanych lepiej, różnica ta wynosi nieco ponad 10%, natomiast wyraźną poprawę zanotowano dla tonów „polskich”, których poprawna identyfikacja wzrosła do wartości o 22% wyższej w porównaniu do stanu poprzedniego. Zatem uznając rzeczywiste przebiegi zarejestrowanych tonów można stwierdzić, że badana grupa Polaków wykazała się dobrą percepcją tonów, zarówno na materiale w języku mandaryńskim, jak i w języku polskim. Fakt ten może stanowić podstawę do wnioskowania, że polski słuchacz, który dysponuje co najmniej podstawową wiedzą i umiejętnościami w zakresie posługiwania się językiem chińskim, jest w stanie – przynajmniej w pewnym stopniu – trafnie zidentyfikować poprawnie zrealizowane tony, nałożone na jednostki języka nietonalnego – w tym wypadku języka polskiego, pomimo że oba te systemy diametralnie różnią się pod względem struktury. Jakość percepcji mogą z kolei regulować rozmaite czynniki, niemniej, jak pokazują wyniki powyższego eksperymentu, ich wpływ jest w dużej mierze międzyosobniczo zróżnicowany. Choć nie stwierdzono wyraźnej korelacji pomiędzy liczbą lat nauki

a zdolnością do poprawnej identyfikacji tonów, jej istnienie nie zostało wykluczone z uwagi na względnie małą liczebność próby. Zaobserwowano natomiast istotną rolę indywidualnych zdolności percepcyjnych słuchaczy, które z dużym prawdopodobieństwem mogą warunkować wpływ innych parametrów takich jak np. osłuchanie z językiem bądź pobyt w kraju chińskojęzycznym. Oprócz zniekształceń powstałych na skutek niedokładnej realizacji, które niewątpliwie utrudniały odbiór słuchowy tonów, przyczyną błędów w identyfikacji mogła być także niedostateczna jakość odtwarzanych sygnałów czy choćby mechaniczne pomyłki podczas operowania panelem dotykowym *touchpad*. Nie można również wykluczyć wpływu dyspozycji poszczególnych słuchaczy w chwili badania, związanej z ich samopoczuciem oraz zdolnością koncentracji uwagi.

Obszar dociekań niniejszego eksperymentu stanowiła również umiejętność nierodzimych użytkowników języka mandaryńskiego do identyfikacji tonów powstałych w wyniku zjawiska tonalnego *sandhi*, co było możliwe dzięki jego licznemu występowaniu w zgromadzonym materiale. Jak się okazało, zdolność Polaków do identyfikacji tonu ½ 3 jako jednej z możliwych form tonu 3 ogółem jest niewielka, choć zauważalnie różni się w obu językach. W teście z wykorzystaniem sylab chińskich ton ½ 3 stanowił 57,5% wszystkich tonów opadająco-wznoszących, z czego został zidentyfikowany w 59 przypadkach (25,7%), w przeważającej większości zrealizowanych przez osobę JK. Z kolei w teście z sylabami polskimi spośród 47,5% wystąpień *sandhi* poprawnie zidentyfikowano jedynie znikomą część równą 9,5%. Na tej podstawie można stwierdzić, że badane osoby wykazały niewielką wrażliwość na zjawisko *sandhi*. Jak pokazują dane, ton ½ 3 ze względu na swój opadający kierunek najczęściej interpretowano jako ton 4, pomimo że ich przebiegi różnią się znacząco. Fakt ten można tłumaczyć w dwojaki sposób: jako wynik niedoskonałości kompetencji nienatywnej lub też niezgodności posiadanej wiedzy teoretycznej z napotkanym stanem rzeczy. Pierwsze wyjaśnienie zakłada, że polscy użytkownicy języka mandaryńskiego nie są przyzwyczajeni do formy i brzmienia tonu ½ 3, co potwierdza brak *sandhi* w realizacjach tonu 3 w grupie Polaków. Swobodne stosowanie reguł następstwa tonów wymaga bowiem intensywnej praktyki językowej i – w przeciwieństwie do *native speakerów* – nie stanowi dla nich procesu naturalnego

ani automatycznego. Trudności z identyfikacją mogą również wynikać z faktu, że tonalne *sandhi* jest zjawiskiem z definicji właściwym tylko mowie ciągłej, w związku z czym pozbawiony kontekstu ton $\frac{1}{2}$ 3 może być najzwyczajniej „nieczytelny”. Wyjaśnienie drugie związane jest z nieumiejętnością wyzwolenia się słuchaczy spod wpływu własnej kompetencji językowej, która nie dopuszcza możliwości zaistnienia zjawiska *sandhi* w sylabach izolowanych. Uznawszy taką ewentualność za błąd, słuchacz poszukuje możliwej alternatywy dla oceny tonu $\frac{1}{2}$ 3, któremu zdecydowanie najbliższej do tonu 4 opadającego.

Konkluzje

Bez względu na zasadność istnienia tonów w polszczyźnie w świetle badań będących przedmiotem niniejszej pracy ich realizacja na izolowanych sylabach polskich okazuje się możliwa. Wniosek ten wymaga natomiast sprecyzowania. Pomimo że w badaniu nad produkcją tonów mówcy łącznie uzyskali rezultat w wysokości 70% poprawnych realizacji, istnieją pod tym względem znaczne rozbieżności pomiędzy grupą rodzimych użytkowników języka mandaryńskiego (45-proc. poprawność) i grupą Polaków (83,1%). Fakt ten pozwala stwierdzić, że istotne znaczenie w procesie produkcji tonu ma poprawność artykulacyjna sylaby będącej jego nośnikiem. Jak wynika z analizy tonograficznej, błędy artykulacyjne, które wielokrotnie były udziałem mówców „chińskich” i często wynikały z niedoskonałości ich kompetencji z zakresie posługiwania się językiem polskim, powodowały znaczne wahania w przebiegu F_0 , w konsekwencji prowadząc do zniekształceń konturów tonalnych oraz trudności w odbiorze słuchowym. Należy przy tym zauważyć, że w przypadku rodzimych użytkowników języka mandaryńskiego zniekształcenia dotyczyły głównie dwóch typów tonów – tonu 1 i 3, podczas gdy u Polaków koncentrowały się niemal wyłącznie na tonie 2. Poza niewielkim odsetkiem błędnych realizacji można stwierdzić, że tony w wykonaniu mówców polskich nie różniły się w znaczący sposób od tych realizowanych na sylabach języka mandaryńskiego przez jego rodzimych użytkowników, które poza tonem 3 służyły tutaj za modele. Poczynione obserwacje dają podstawy do spekulacji, że testowana w niniejszej monografii zdolność słuchaczy do imitacji tonów na sylabach języka rodzimego mogłaby okazać się pomocna w szybszym opanowaniu języka tonalnego, ulepszając zrozumienie zasad i sposobu ich realizacji poprzez poznanie fonetycznych następstw, jakie niesie ze sobą nałożenie poszczególnych tonów na materiał dobrze znanego języka. Weryfikacja tej tezy wymaga jednak dowodu w postaci wyników bardziej szczegółowych badań. Interesujące jest również, jak dużą rolę w procesie identyfikacji odgrywa aspekt semantyczny danej jednostki, na której zlokalizowany jest ton, w szczególności wyrazów jednosylabowych, tworzących pewne całości pod względem znaczeniowym. Mianowicie zastanawia, czy tony realizowane na jednosyla-

bowych wyrazach polskich byłyby percypowalne dla użytkowników nieznających języka polskiego. Zagadnienie to może stanowić ciekawy przedmiot dalszych badań psycholingwistycznych w tym zakresie.

Studium nad percepcją tonów, stanowiące drugi z eksperymentów przeprowadzonych na potrzeby niniejszej pracy, wykazało natomiast, że słuchacze polscy są w stanie poprawnie zidentyfikować tony nałożone na sylaby języka polskiego pod warunkiem, że tony te zostały zrealizowane prawidłowo. Świadczy o tym wysoki wynik (71,1%) uzyskany przez grupę po uwzględnieniu odpowiedzi zgodnych z rzeczywistymi przebiegami tonów z odtwarzanego materiału dźwiękowego. Otrzymane dane nie wykazują natomiast wyraźnej korelacji pomiędzy liczbą lat nauki a zdolnością słuchaczy do trafnej identyfikacji tonów, choć wysoki poziom wiedzy i umiejętności w zakresie posługiwania się językiem mandaryńskim niewątpliwie zwiększa prawdopodobieństwo uzyskania dobrego rezultatu. Stwierdzono w tej kwestii natomiast istotny wpływ pewnych indywidualnych predyspozycji bądź zdolności percepcyjnych słuchaczy, które w niektórych przypadkach okazały się górować nad wiedzą teoretyczną i praktyczną znajomością języka mandaryńskiego.

Przeprowadzone badania dostarczyły również interesujących spostrzeżeń na temat zjawiska tonalnego *sandhi*, a ściślej – jego bardziej powszechnej formy, zazwyczaj pomijanej w literaturze językoznawczej. Otóż dane pochodzące z nagrań rodzimych użytkowników języka mandaryńskiego wskazują na odrębny charakter tonu $\frac{1}{2}$ 3, którego, z uwagi na pewne jego właściwości, nie należy traktować jako wyniku prostej redukcji tonu 3 o część rosnącą, ani też jako tonu 4 opadającego. Dokładna analiza tonogramów wykazała, że kontury tonów $\frac{1}{2}$ 3 są na ogół dłuższe niż części opadające tonów trzecich, co sugeruje, że są to przebiegi „pełnowymiarowe”, nie zaś fragmentaryczne, i jednocześnie obejmują węższy zakres F_0 niż w przypadku tonów 4. Wniosek ten stoi w sprzeczności z obserwacjami Zhang i Lai (2010, por. § 2.3), którzy twierdzą, że ton $\frac{1}{2}$ 3 jest w rzeczywistości efektem uproszczenia tonu złożonego, w wyniku pozbawienia go konturu rosnącego, przez co wykazuje on duże podobieństwo do odpowiedniej części modelowej formy tonu 3. W obliczu istniejących rozbieżności problem natury diskutowanego tonu wymaga zatem weryfikacji z wykorzystaniem obszerniejszego materiału. Jak zauważyli Lee-Schoenfeld i Kandybowicz (2008, por. § 2.3), ton $\frac{1}{2}$ 3 może natomiast przybierać różną

postać w zależności od mówcy, tj. może zostać zrealizowany w rejestrze niskim lub średnim, lub też wykazywać przebieg względnie równy, co znajduje potwierdzenie w niniejszym eksperymencie. Fakt ten wywiera również wpływ na percepcję, gdyż, jak pokazują uzyskane wyniki, tony w wykonaniu głosu kobiecego (JK), były znacznie częściej poprawnie identyfikowane niż te realizowane przez mężczyznę (DM). Pomimo tego ton powstały na skutek zjawiska *sandhi* wciąż sprawiał polskim słuchaczom znaczne trudności w identyfikacji, co – jak zauważa Xu (1994) – najprawdopodobniej było wynikiem pozbawienia go kontekstu.

Aneks

Tabela 31. Materiał słownikowy wykorzystany w badaniach.

ZNAK	PINYIN	ZNACZENIE
师	<i>shī</i>	<i>teacher</i>
十	<i>shí</i>	<i>ten (10)</i>
使	<i>shǐ</i>	<i>to make, cause to</i>
是	<i>shì</i>	<i>to be</i>
秋	<i>qiū</i>	<i>autumn</i>
求	<i>qiú</i>	<i>to request</i>
糗	<i>qiǔ</i>	<i>provisions (dry food)</i>
蝮*	<i>qiù</i>	
跟	<i>gēn</i>	<i>to follow</i>
眼	<i>gén</i>	<i><coll.> funny</i>
艮	<i>gěnn</i>	<i>straightforward</i>
亘	<i>gèn</i>	<i>to extend</i>
帮	<i>bāng</i>	<i>to help</i>
绑	<i>bǎng</i>	<i>to bind</i>
棒	<i>bàng</i>	<i>stick</i>
霾	<i>mái</i>	<i>haze</i>
买	<i>mǎi</i>	<i>to buy</i>
卖	<i>mài</i>	<i>to sell</i>
烹	<i>pēng</i>	<i>to boil</i>
棚	<i>péng</i>	<i>canopy</i>
捧	<i>pěng</i>	<i>to hold in both hands</i>
碰	<i>pèng</i>	<i>to bump into</i>
烟	<i>yān</i>	<i>smoke</i>
沿	<i>yán</i>	<i>edge, border</i>
眼	<i>yǎn</i>	<i>eye</i>
燕	<i>yàn</i>	<i>swallow</i>

说	<i>shuō</i>	<i>to speak</i>
硕	<i>shuò</i>	<i>large</i>
摊	<i>tān</i>	<i>to spread out</i>
谈	<i>tán</i>	<i>to chat</i>
毯	<i>tǎn</i>	<i>blanket</i>
叹	<i>tàn</i>	<i>to sigh</i>
迂	<i>yū</i>	<i>roundabout</i>
鱼	<i>yú</i>	<i>fish</i>
与	<i>yǔ</i>	<i>and</i>
遇	<i>yù</i>	<i>to encounter</i>

*Sylaba ta występuje jedynie w złożeniach.

Bibliografia

- Anderson, S.R. 1978. *Tone features*, s. 133–175. w: Fromkin, V. 1978. *Tone: A linguistic survey*. New York: Academic Press.
- Anderson, S.R. 1982. *Wprowadzenie do fonologii*. Przeł. E. Gussmann. Wrocław: Zakład Narodowy im. Ossolińskich, s. 316–370.
- Chao, Y.R. 1970. *A Grammar of Spoken Chinese*. (wyd.2) Berkeley–Los Angeles: University of California Press.
- Cheng, C.C. 1973. *A Synchronic Phonology of Mandarin Chinese* (Monographs on Linguistic Analysis, No. 4). The Hague: Mouton, s. 4–63.
- Cheng, L.L. 1987. *On the prosodic hierarchy and tone sandhi in Mandarin*. w: P. Avery (ed.) *Toronto Working Papers in Linguistics*, 7, s. 24–52.
- Crystal, D. 2003. *A Dictionary of Linguistics & Phonetics* (5th ed.). Malden, MA & Oxford: Blackwell Publishing.
- Demenko, G. 1999. *Analiza cech suprasegmentalnych języka polskiego na potrzeby technologii mowy*. Poznań: Wydawnictwo Naukowe UAM.
- Dłuska, M. 1976. *Prozodia języka polskiego*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Duanmu, S. 2000. *The Phonology of Standard Chinese*. New York: Oxford University Press.
- Gandour, J.T. 1978. *The perception of tone*, s. 41–76 w: Fromkin, V. 1978. *Tone: A linguistic survey*. New York: Academic Press.
- Goldsmith, J.A. 1979. *Autosegmental Phonology*. New York & London: Garland Publishing, s. 5–55.
- Goldsmith, J. A. 1992. *Autosegmental and metrical phonology*. Oxford & Cambridge, MA: Blackwell.
- Goldsmith, J.A. 1994. Tone languages. W: R.E. Asher & J.M.Y. Simpson (red.), *The Encyclopedia of Language and Linguistics* (s. 4626–4628). New York: Elsevier Science (Pergamon Press).
- Jassem, W. 1962. *Akcent języka polskiego*. Wrocław: Zakład Narodowy im. Ossolińskich, s. 5–14.
- Jassem W. 2003. Illustration of the IPA: Polish. *Journal of the International Phonetic Association* 33/1, s. 103–107. <https://doi.org/10.1017/S0025100303001191>
- Jongman, A., Wang, Y., Moore, C. B., & Sereno, J. A. 2006. Perception and production of Mandarin Chinese tones. In P. Li, L. H. Tan, E. Bates, & O.J.L. Tzeng (RED.), *The handbook of East Asian psycholinguistics: Vol. 1. Chinese* (s. 209–224). Cambridge, UK: Cambridge University Press. <https://doi.org/10.1017/CBO9780511550751.015>.

- Karpiński, M. 2006. *Struktura i intonacja polskiego dialogu zadaniowego*. Poznań: Wydawnictwo Naukowe UAM.
- Katamba, F. 1996. *An introduction to phonology*. London–New York: Longman, s. 53–240.
- Ladd, R. 1996. *Intonational phonology*. Cambridge: Cambridge University Press, s. 255–274.
- Lee-Schoenfeld, V., & Kandybowicz, J. (2009). Sandhi sans derivation: Third tone patterns in Mandarin Chinese. *University of Pennsylvania Working Papers in Linguistics*, 14(2), 1–16.
- Lin, H.S. 2005. *Prosodic correspondence in tone sandhi*. w: *UST Working Papers In Linguistics. 1*, s. 229–265.
- MacKay, I.R.A. 1987. *Phonetics: The science of speech production* (2nd ed.). Boston, MA: Little, Brown and Company.
- Majewicz, A.F. 1989. *Języki świata i ich klasyfikowanie*. Warszawa: Państwowe Wydawnictwo Naukowe, s. 66–193.
- McCawley, J.D. 1978. *What is a tone language?*s. 119–121. w: Fromkin, V. 1978. *Tone: A linguistic survey*. Academic Press. <https://doi.org/10.1016/B978-0-12-267350-4.50009-1>
- Michaud, A. 2012. *The complex tones of East/Southeast Asian languages: Current challenges for typology and modelling*. w: *Proceedings of the 3rd International Symposium on Tonal Aspects of Languages (TAL 2012)* (artykuł KN-6, s. 1–7). ISCA Archive.
- Pike, K.L. 1945. *The Intonation of American English*. Ann Arbor, MI: University of Michigan Press.
- Pike, K.L. 1966. *Phonetics: A critical analysis of phonetic theory and a technic for the practical description of sounds*. Ann Arbor, Mi: The University of Michigan, s. 27–31.
- Pulleyblank, E.G. 2003. *Outline of classical Chinese grammar*. Vancouver: UBC Press.
- Sagart, L. 1998. *The origin of Chinese tones*. International Symposium on “Tone Languages in the World”, Tokyo, December 10–12. Proceeding published by the Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Shen, X.N. S. 1990. *The prosody of Mandarin Chinese*. University of California Press.
- Shuh, R.G. 1978. *Tone rules*, s. 221–255. w: Fromkin, V. 1978. *Tone: A linguistic survey*. New York: Academic Press.
- Speer, S.R., Xu, L. 2008. Processing lexical tone in third-tone sandhi. *Laboratory Phonology 11*, s. 131–132.
- Steffen-Batogowa, M. 2000. *Struktura akcentowa języka polskiego*. Warszawa: Wydawnictwo Naukowe PWN.

- Szymczak, W. 2010. *Podstawy statystyki dla psychologów. Podręcznik akademicki* (wyd. 2). Warszawa: Difin.
- Śledziński, D. 2008. Fonetyczno-akustyczna struktura sylaby w języku polskim na potrzeby technologii mowy, s. 219–224. *Investigationes Linguisticae, vol. 16*.
- Topolińska, Z. 1961. *Z historii akcentu polskiego od wieku XVI do dziś*. Wrocław: Zakład Narodowy im. Ossolińskich, s. 5–28.
- Wierzchowska, B. 1980. *Fonetyka i fonologia języka polskiego*. Wrocław: Zakład Narodowy im. Ossolińskich, s. 133–135.
- Xu, Y. 1994. Production and perception of coarticulated tones. *Journal of the Acoustical Society of America* 95(4): 2240–2253. <https://doi.org/10.1121/1.408684>
- Yip, M. 2002. *Tone*. Cambridge: Cambridge University Press.
- Yip, M. 2007. *Tone*. w: P. de Lacy (red.), *The Cambridge handbook of phonology* (s. 229–252). New York: Cambridge University Press.
- Yuan, J. H., & Chen, Y. Y. 2014. 3rd tone sandhi in Standard Chinese: A corpus approach. *Journal of Chinese Linguistics*, 42(1), 218–237.
- Zhang, J. 1996. On the syllable structures of Chinese relating to speech recognition. w: *Proceedings of the 4th International Conference on Spoken Language Processing (ICSLP 1996)*, Vol. 4, s. 2450–2453. Philadelphia, PA: Institute for Spoken Language Processing / ISCA.
- Zhang, J., Lai, Y. 2010. Testing the role of phonetic knowledge in Mandarin tone sandhi. *Phonology* 27. 1, 153–201. <https://doi.org/10.1017/S0952675710000060>
- Zhang, N. 1997. The Avoidance of the Third Tone Sandhi in Mandarin Chinese. *Journal of East Asian Linguistics* 6, 293–338. <https://doi.org/10.1023/A:1008232121848>
- Zuo, P. 2002. *Tonal coarticulation. Contextual F0 realization of Mandarin Chinese tones*. Stuttgart: Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart.

Multimedialne:

Wenlin Software for learning Chinese, 1997–2007. Wenlin Institute, Inc.

Abstract

This monograph explores the interaction between tonal and non-tonal language systems through an experimental investigation of Mandarin Chinese tones as produced and perceived by native speakers of Mandarin and Polish. The central research question was whether speakers of a non-tonal, consonant-dominated language such as Polish are capable of accurately reproducing and perceiving the four canonical Mandarin tones, and conversely, how Mandarin speakers handle tonal imitation when applied to Polish monosyllabic words. A further aim was to examine the extent to which linguistic competence, language proficiency, and imitative aptitude affect both tone production and perception.

The theoretical framework provides a comprehensive overview of tonal systems, their typological classifications, and their treatment in linguistic research. Mandarin Chinese is analyzed not only as a phonological system but also in terms of its acoustic correlates, especially fundamental frequency (F0). The study also reviews transcription systems and prosodic structures, culminating in a detailed discussion of tone sandhi as one of the most complex and theoretically significant aspects of Mandarin prosody.

The empirical part consists of two complementary experiments: (1) a production study, in which Mandarin and Polish speakers were asked to pronounce Polish monosyllabic words with Mandarin tones, and (2) a perception study, in which Polish learners of Mandarin assessed tonal realizations. Tonographic and spectrographic analyses were conducted to evaluate tonal contours against the Mandarin norm.

The findings show that overall production accuracy reached about 70%, but the distribution across groups was unexpected: Mandarin speakers achieved less than half of their tonal realizations correctly, while Polish speakers exceeded 80% accuracy. Polish participants performed slightly better even on Mandarin syllables (once sandhi was excluded), and much better when tones were imposed on Polish words. The main difficulty for Polish speakers was the rising tone (Tone 2), whereas Mandarin speakers most often misproduced Tones 1 and 3, largely due to problems with Polish consonant clusters. Perceptual data confirmed these tendencies: tone identification by Polish listeners

was strongly influenced by proficiency, individual auditory aptitude, and the acoustic quality of the tokens, with tone recognition being significantly more reliable for tones produced on Mandarin syllables than on Polish ones.

The study yields several key conclusions. Cross-linguistic tonal imitation is possible but constrained by the prosodic and segmental properties of the carrier language as well as the articulatory competence of the speaker. The rising tone constitutes the greatest challenge for Polish speakers, while Mandarin speakers' lower performance on Polish material stems from articulatory distortions rather than tonal competence. Individual aptitude and speech rate play a decisive role in tonal accuracy. Finally, despite the absence of tonal contrasts in Polish, Polish speakers are able to partially internalize Mandarin tonal patterns, suggesting that cross-linguistic phonetic transfer is achievable under appropriate conditions.

Overall, this research enhances our understanding of tone acquisition and perception in cross-linguistic contexts. It not only contributes to theoretical discussions in phonetics, phonology, and language typology, but it also suggests possible directions for applying the findings to the teaching of Mandarin Chinese to speakers of non-tonal languages.

Spis tabel

Tabela 1. Średnie wartości F_0 wyrazów chińskich realizowanych przez rodzimych użytkowników.	110
Tabela 2. Odchylenia standardowe głosów rodzimych użytkowników języka mandaryńskiego (tony realizowane na wyrazach chińskich) wraz z współczynnikiem modulacji częstotliwości.	111
Tabela 3. Iloczasy tonów wyrazów chińskich wypowiedzianych przez rodzimych użytkowników.	114
Tabela 4. Wykaz prawidłowo zrealizowanych tonów na sylabach chińskich wypowiedzianych przez rodzimych użytkowników.	116
Tabela 5. Średnie wartości F_0 wyrazów polskich realizowanych przez rodzimych użytkowników języka mandaryńskiego.	123
Tabela 6. Odchylenia standardowe głosów rodzimych użytkowników języka mandaryńskiego (tony realizowane na wyrazach polskich) wraz z współczynnikiem modulacji częstotliwości.	123
Tabela 7. Iloczasy tonów wyrazów polskich wypowiedzianych przez rodzimych użytkowników języka mandaryńskiego.	126
Tabela 8. Wykaz prawidłowo zrealizowanych tonów na sylabach polskich wypowiedzianych przez rodzimych użytkowników języka mandaryńskiego.	127
Tabela 9. Średnie wartości F_0 wyrazów chińskich realizowanych przez Polaków.	139
Tabela 10. Odchylenia standardowe głosów nierodzimych użytkowników języka mandaryńskiego (tony realizowane na wyrazach chińskich) wraz z współczynnikiem modulacji częstotliwości.	140
Tabela 11. Iloczasy tonów wyrazów chińskich wypowiedzianych przez nierodzimych użytkowników języka mandaryńskiego.	142
Tabela 12. Wykaz prawidłowo zrealizowanych tonów na sylabach chińskich wypowiedzianych przez nierodzimych użytkowników języka mandaryńskiego.	143
Tabela 13. Średnie wartości F_0 wyrazów polskich realizowanych przez Polaków.	150
Tabela 14. Odchylenia standardowe głosów Polaków (tony na wyrazach polskich) wraz z współczynnikiem modulacji częstotliwości. .	150

Tabela 15. Iloczasy tonów wyrazów polskich wypowiedzianych przez nierodzimych użytkowników języka mandaryńskiego.	153
Tabela 16. Wykaz prawidłowo zrealizowanych tonów na sylabach polskich wypowiedzianych przez nierodzimych użytkowników języka mandaryńskiego.	155
Tabela 17. Porównanie wyników realizacji tonów na sylabach chińskich przez obie grupy mówców.	165
Tabela 18. Porównanie wyników realizacji tonów na sylabach polskich przez obie grupy mówców.	166
Tabela 19. Porównanie średnich zakresów poszczególnych tonów zrealizowanych na sylabach chińskich przez mówców z obu grup.	167
Tabela 20. Porównanie średnich zakresów poszczególnych tonów zrealizowanych na sylabach polskich przez mówców z obu grup.	168
Tabela 21. Porównanie średnich iloczasów sylab chińskich w realizacji obu grup.	169
Tabela 22. Porównanie średnich iloczasów sylab polskich w realizacji obu grup.	170
Tabela 23. Liczba i odsetek poprawnie zidentyfikowanych tonów zrealizowanych na sylabach chińskich przez rodzimych użytkowników języka mandaryńskiego.	173
Tabela 24. Liczba i odsetek poprawnie zidentyfikowanych tonów zrealizowanych na sylabach polskich przez rodzimych użytkowników języka mandaryńskiego.	174
Tabela 25. Liczba i odsetek tonów na sylabach chińskich, zidentyfikowanych zgodnie z faktyczną realizacją.	177
Tabela 26. Liczba i odsetek tonów na sylabach polskich, zidentyfikowanych zgodnie z faktyczną realizacją.	177
Tabela 27. Liczba i odsetek prawidłowo zidentyfikowanych tonów na sylabach chińskich.	181
Tabela 28. Liczba i odsetek prawidłowo zidentyfikowanych tonów na sylabach polskich.	181
Tabela 29. Liczba prawidłowych identyfikacji tonów na poszczególnych sylabach chińskich.	182
Tabela 30. Liczba prawidłowych identyfikacji tonów na poszczególnych sylabach polskich.	184
Tabela 31. Materiał słownikowy wykorzystany w badaniach	195

Spis rycin

Ryc. 1. Prawidłowy przebieg tonu 1 – sylaba [ʃwo] (<i>shuō</i>).	117
Ryc. 2. Prawidłowy przebieg tonu 4 – sylaba [təow] (<i>qiū</i>).	118
Ryc. 3. Wątpliwy przebieg tonu 2 – sylaba [tʰan] (<i>tán</i>).	119
Ryc. 4. Prawidłowy przebieg tonu 2 – sylaba [pʰeŋ] (<i>péng</i>).	119
Ryc. 5. Nieprawidłowy przebieg tonu 2 – sylaba [y:] (<i>yú</i>).	120
Ryc. 6. Przebieg tonu ½ 3 – sylaba [ʂʅ] (<i>shì</i>).	121
Ryc. 7. Modelowy przebieg tonu 3 – sylaba [jen] (<i>yǎn</i>).	121
Ryc. 8. Prawidłowy przebieg tonu 4 – sylaba <i>strzep</i>	130
Ryc. 9. Nieprawidłowy przebieg tonu 4 – sylaba <i>wprost</i>	131
Ryc. 10. Nieprawidłowy przebieg tonu 4 – sylaba <i>schnąc</i>	131
Ryc. 11. Prawidłowy przebieg tonu 2 – sylaba <i>pstry</i>	132
Ryc. 12. Zniekształcony przebieg tonu 2 – sylaba <i>hymn</i>	133
Ryc. 13. Prawidłowy przebieg tonu 1 – sylaba <i>wszcząc</i>	133
Ryc. 14. Nieprawidłowy przebieg tonu 1 – sylaba <i>lgnać</i>	134
Ryc. 15. Prawidłowy kontur tonu 3 – sylaba <i>wprost</i>	135
Ryc. 16. Ton ½ 3 – sylaba <i>tchnąc</i>	135
Ryc. 17. Prawidłowy kontur tonu 3 – sylaba <i>schnąc</i>	136
Ryc. 18. Prawidłowy przebieg tonu 1 – sylaba [paŋ] (<i>bāng</i>).	144
Ryc. 19. Nieprawidłowy kontur tonu 1 – sylaba [jen] (<i>yān</i>).	145
Ryc. 20. Prawidłowy kontur tonu 4 – sylaba [tʰan] (<i>tàn</i>).	145
Ryc. 21. Nieprawidłowy przebieg tonu 2 – sylaba [ʃɿ] (<i>shǐ</i>).	146
Ryc. 22. Nieprawidłowy kontur tonu 2 – sylaba [gen] (<i>gén</i>).	147
Ryc. 23. Prawidłowy kontur tonu 2 – sylaba [pʰeŋ] (<i>péng</i>).	147
Ryc. 24. Efekt zjawiska <i>creaky voice</i> – zniekształcony kontur tonu 3 na sylabie [jen] (<i>yěn</i>).	148
Ryc. 25. Prawidłowy kontur tonu 3 – sylaba [təow] (<i>qiū</i>).	158
Ryc. 26. Prawidłowy przebieg tonu 1 – sylaba <i>strzep</i>	157
Ryc. 27. Prawidłowy kontur tonu 1 – sylaba <i>lgnać</i>	157
Ryc. 28. Nieprawidłowy przebieg tonu 4 – sylaba <i>lgnać</i>	158
Ryc. 29. Prawidłowy kontur tonu 4 – sylaba <i>wprost</i>	158
Ryc. 30. Zniekształcony kontur tonu 2 – sylaba <i>pstry</i>	160
Ryc. 31. Prawidłowy przebieg tonu 2 – sylaba <i>pstry</i>	160
Ryc. 32. Modelowy kontur tonu 3 – sylaba <i>hymn</i>	161

ISBN 978-83-68668-03-2

DOI 10.48226/978-83-68668-03-2