Daniel Śledziński

WIELOWARSTWOWY MODEL TRANSKRYPCJI TEKSTU W JĘZYKU POLSKIM

Wielowarstwowy model transkrypcji tekstu w języku polskim

Daniel Śledziński

Wielowarstwowy model transkrypcji tekstu w języku polskim

wydawnictwo

Poznań 2022

Projekt okładki:
Daniel Śledziński

Recenzja:
dr hab. Paweł Nowakowski, prof. UAM

Wydanie I
Poznań 2022

ISBN 978-83-67287-28-9

DOI 10.48226/978-83-67287-28-9

Wydanie:

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

Dziękuję Panu Profesorowi Pawłowi Nowakowskiemu
za wszechstronną i wieloaspektową pomoc.

Autor

Spis treści

Wprowadzenie

Publikacja jest poświęcona koncepcji wielowarstwowego modelu transkrypcji. Omówione rozwiązanie umożliwia automatyczną konwersję tekstu ortograficznego w języku polskim na transkrypcję fonologiczną oraz fonetyczną. Najważniejszą inspiracją była monografia autorstwa Marii Steffen-Batogowej: *Automatyzacja transkrypcji fonematycznej tekstów polskich* (dalej: *Automatyzacja…*). Została ona wydana w 1975 roku i było to pierwsze kompletne opracowanie, które dotyczyło tego zagadnienia [Steffen-Batogowa 1975; Steffen-Batóg 1997; Pluciński 2002]. Informacje oraz reguły pochodzące z tej publikacji posłużyły autorowi za punkt wyjścia dla każdej analizy omówionej w rozdziale drugim.

Podstawową koncepcję opisaną w niniejszej monografii można uznać za daleko posuniętą modyfikację rozwiązań omówionych w *Automatyzacji…* . Z punktu widzenia tej koncepcji propozycje przedstawione przez M. Steffen-Batogową można uznać za system jednowarstwowy. Oznacza to, że dany wyraz ortograficzny jest przetwarzany tylko jeden raz przy użyciu określonego zestawu reguł i po tej czynności uzyskuje się ostateczną transkrypcję fonologiczną. U podstaw zaprezentowanego w niniejszej książce modelu wielowarstwowego leży natomiast wieloetapowe przetwarzanie wyrazów ortograficznych. W kolejnych etapach działania algorytmu zapis tych wyrazów coraz bardziej upodabnia się do docelowej transkrypcji. Ma to związek z różnymi problemami, takimi jak: wydzielanie digrafemów i trigrafemów, zmiana statusu dźwięczności czy też uproszczenia grup spółgłoskowych. Model wielowarstwowy umożliwia przyporządkowywanie poszczególnych problemów do różnych warstw. Dzięki temu istnieje możliwość zwiększenia przejrzystości zbioru reguł. Poza tym reguły związane z określonym czynnikiem mogą być modyfikowane niezależnie od reguł dotyczących innych czynników, które zostały przyporządkowane do innych warstw. Z tych właściwości wynikają podstawowe zalety wielowarstwowego modelu transkrypcji – można mówić o elastyczności oraz o przejrzystości rozwiązań opartych na tym modelu.

Istotny problem związany z wielowarstwowym modelem transkrypcji dotyczy nazewnictwa lingwistycznego. Wynika on z niekonwencjonalnej istoty tego rozwiązania. Problem nie występuje w wypadku reguł omówionych w *Automatyzacji…*, ponieważ dotyczą one liter oraz fonemów. Zdefiniowane lingwistycznie ciągi jednostek są przekształcane na ciągi innych jednostek, które również są zdefiniowane w teoriach lingwistycznych. Natomiast w wypadku omówionego modelu wielowarstwowego sytuacja wygląda inaczej. Do pierwszej warstwy rozwiązania opartego na tym modelu są przekazywane wyrazy ortograficzne. Pełnowartościowy zapis transkrypcyjny jest zwracany dopiero po przetworzeniu w ostatniej warstwie. W kolejnych warstwach wyrazy stopniowo stają się docelową transkrypcją (mechanizm został dokładnie omówiony w opracowaniu). W czasie przetwarzania zapis każdego wyrazu ma zatem charakter pośredni, a jednostki, z których składa się ten zapis, nie są ani znakami ortograficznymi, ani fonemami zgodnymi z definicjami lingwistycznymi. Rozwiązanie polega na przyjęciu nazewnictwa umownego, przy czym żadnego konkretnego nazewnictwa nie można włączyć do definicji modelu, ponieważ nie zawiera on założeń dotyczących funkcji poszczególnych warstw ani nie precyzuje dopuszczalnych jednostek, z których składają się wyrazy w czasie ich przetwarzania. W przykładowym rozwiązaniu opartym na wielowarstwowym modelu transkrypcji (zob. rozdz. 4) przyjęta została następująca konwencja nazewnictwa: wyrazy przekazywane do przetwarzania w n-tej warstwie są złożone z segmentów n-tego stopnia. Jednocześnie te segmenty należą do inwentarza n-tego stopnia. Zatem wyrazy mające być przetworzone w drugiej warstwie są złożone z segmentów drugiego stopnia, które należą do inwentarza drugiego stopnia.

Informacje przedstawione w tej książce dotyczą głównie transkrypcji fonologicznej. Dotyczy to zarówno analiz omówionych w rozdziale drugim, jak i przykładowego projektu systemu transkrypcji omówionego w rozdziale czwartym. Istnieje kilka powodów związanych z takim podejściem. Pierwszy wynika z faktu, o którym już wspomniano – że inspiracją do stworzenia modelu wielowarstwowego była książka autorstwa Marii Steffen-Batogowej, a reguły w jej publikacji dotyczą transkrypcji fonologicznej. Poza tym łatwiej jest przedstawić koncepcje związane z wielowarstwowym modelem transkrypcji na przykładzie transkrypcji fonologicznej (można to zrobić w sposób bardziej przejrzysty). Kolejny istotny powód jest wyrazem przekonania, że utworzenie systemu transkrypcji fonologicznej nie wyklucza możliwości jego rozbudowy do systemu transkrypcji fonetycznej (alofonicznej). Do danego projektu wystarczy dodać jedną warstwę, która będzie używana do przekształcania transkrypcji fonologicznej na transkrypcję fonetyczną. Jest to możliwe, ponieważ w zapisie fonologicznym konkretnych wyrazów każdemu fonemowi niemal zawsze odpowiada jedna głoska. Trzeba podkreślić, że reguły umożliwiające konwersję zapisu fonologicznego na zapis fonetyczny zostały już opracowane przez Marię Steffen-Batogową – można to uznać za tę część prac autorki, która jest najbliższa koncepcji wielowarstwowego modelu transkrypcji [Steffen-Batóg 1989−1990].

Rozdział pierwszy rozpoczyna się od przytoczenia analizy porównawczej polskich inwentarzy fonologicznych. Autorką tego zestawienia jest Ewa Wolańska. Oprócz danych zgromadzonych w tabeli w skróconej formie zostały przytoczone najważniejsze wnioski autorki dotyczące różnic występujących pomiędzy poszczególnymi inwentarzami. Następnie został przedstawiony inwentarz fonologiczny oparty na przyjętym przez Marię Steffen-Batogową w *Automatyzacji…* . To rozwiązanie oraz alfabet transkrypcyjny SAMPA stanowią podstawę zapisywania wyrazów w niniejszej publikacji. W dalszym ciągu rozdziału pierwszego przedstawiono istotne definicje: grafem, diada ortograficzna i triada ortograficzna. Następnie omówiono pojęcie transkrypcji podstawowej – jest to termin związany z wielowarstwowym modelem transkrypcji, który jednak nie jest częścią tego modelu. Nie ma obowiązku używania transkrypcji podstawowej w konkretnych rozwiązaniach opartych na omówionym modelu. Jednak wykazano, że właściwa transkrypcja wyrazów może opierać się przede wszystkim na regułach transkrypcji podstawowej, natomiast wszystkie inne reguły można traktować jako wyjątki. Kolejna część rozdziału pierwszego została poświęcona zagadnieniu, o którym już wspomniano – wieloczynnikowej naturze zadań transkrypcyjnych. Następnie omówiono sposób zapisywania reguł zawartych w *Automatyzacji…* – w niniejszym opracowaniu został on zmieniony w stosunku do oryginału. Na końcu rozdziału pierwszego znajduje się opis systemów konwersji tekstu na transkrypcję, które były tworzone dla języka polskiego.

Rozdział drugi jest poświęcony lingwistycznym podstawom transkrypcji fonologicznej tekstu ortograficznego. Obejmuje on rozległe analizy słownika i korpusu tekstowego. Analizy wykonano i opisano bardzo szczegółowo, ponieważ nie są one powiązane z konkretnym projektem systemu transkrypcji. Zostały one zaprojektowane w ten sposób, żeby na ich podstawie można było stworzyć dowolny system transkrypcji tekstu w języku polskim. Całokształt badań związanych z modyfikacją statusu dźwięczności (ze zjawiskami: sonoryzacji oraz desonoryzacji) został omówiony w oddzielnej publikacji autora: *Asymilacja dźwięczności w polskich grupach spółgłoskowych* (dalej: *Asymilacja…*) [Śledziński 2019]. Tutaj powoływano się na ustalenia zawarte w tej publikacji w opisach analiz dotyczących liter oznaczających spółgłoski właściwe oraz w opisie dotyczącym przykładowego projektu systemu transkrypcji. Szczegółowe omówienie wyników analiz (w rozdziale drugim) może być również przydatne dla projektowania badań naukowych związanych z wymową, na przykład badań akustycznofonetycznych [por.: Jassem 1960, 1998; Kleśta 1998; Nowakowski 2002; Owsianny 1998; Patryn 1988].

W rozdziale trzecim przedstawiono podstawowe założenia oraz mechanizm, na którym oparty jest wielowarstwowy model transkrypcji. Zdefiniowano takie pojęcia, jak: warstwa, segment, indeks, a także dokładnie scharakteryzowano operacje, które mogą być wykonywane na segmentach. Przedstawiono również informacje dotyczące składni reguł.

Omówione analizy oraz opis dotyczący formalnych założeń modelu wielowarstwowego stanowią podstawę dla przykładowego projektu systemu transkrypcji, który został przedstawiony w rozdziale czwartym. Ten system zaprojektowano przy założeniu maksymalnej redukcji liczby reguł. Omówiono także etapy przetwarzania konkretnych wyrazów ortograficznych. Jest to dokładny opis, który obrazuje zmiany zachodzące w strukturze wyrazów w miarę przetwarzania w kolejnych warstwach.

Rozdział piąty poświęcono zaawansowanym zagadnieniom związanym z wielowarstwowym modelem transkrypcji. Opisano w nim mechanizm list odwzorowań, który umożliwia znaczną redukcję liczby reguł. Przedstawiono również koncepcję systemów probabilistycznych – zakłada ona użycie elementów rachunku prawdopodobieństwa w regułach. Inna koncepcja obejmuje także systemy generatywne. Podstawowe założenie tej koncepcji dotyczy możliwości tworzenia alternatywnych względem siebie zbiorów reguł. W rozdziale piątym zaprezentowano również mechanizm oraz składnię reguł związanych z procesami fonetyki międzywyrazowej, a także zagadnienie, które formalnie nie należy do wielowarstwowego modelu transkrypcji, dotyczące sposobu zapisywania projektów w plikach.

Wprowadzenie zakończono refleksją, która dotyczy przyczyn powstania wielowarstwowego modelu transkrypcji. Z problemem transkrypcji powiązane są zmienne czynniki: różne inwentarze i systemy fonologiczne, różne alfabety transkrypcyjne czy różne możliwości wymowy tych samych wyrazów. Wspomniano już, że model wielowarstwowy cechuje się znaczną elastycznością przejawiającą się w możliwości łatwej modyfikacji reguł, które najczęściej są tworzone pod kątem pojedynczych problemów. Dzięki temu w oparciu o model wielowarstwowy można zbudować dowolny system transkrypcyjny – można to osiągnąć poprzez zastosowanie odpowiedniej konstrukcji reguł. Ma to związek z podstawowym założeniem, które przyświecało autorowi przy projektowaniu modelu wielowarstwowego – dotyczy ono utworzenia rozwiązania uniwersalnego, które jest odpowiednie dla każdego systemu fonologicznego, inwentarza fonologicznego, alfabetu transkrypcyjnego oraz dla dowolnej wymowy. Kolejny etap rozwoju tej koncepcji obejmie systemy generatywne. Pozwolą one generować zestawy reguł dla różnych wariantów systemów transkrypcji. Nie jest wykluczone użycie omó-

wionego modelu wielowarstwowego dla transkrypcji tekstu w innych językach, w szczególności w językach słowiańskich [Sawicka 1988, 2007].

1. Podstawowe pojęcia

1.1. Inwentarze fonologiczne języka polskiego

Inwentarze fonologiczne są rozumiane jako zbiory fonemów wynikające z różnych koncepcji fonologicznych. „Polskie inwentarze fonologiczne są zróżnicowane pod względem wyróżnionych fonemów ze względu na różnice dotyczące przyjętych kryteriów, metodologii oraz sposobu interpretacji statusu fonemicznego niektórych głosek" [Lorenc, Ptaszkowska 2015: 230]. Systemy fonologiczne, które stanowią podstawę wyznaczania inwentarzy fonologicznych, w niniejszej pracy nie są szczegółowo omawiane [por.: Dukiewicz, Sawicka 1995; Dunaj 1991; Jassem 1996; Łobacz, Ciarkowski 1988; Rocławski 1986; SzpyraKozłowska 2007].

Monografia autorstwa Ewy Wolańskiej: *System grafematyczny współczesnej polszczyzny na tle innych systemów pisma* zawiera rozdział zatytułowany: *Różnice ilościowe i jakościowe w opisie inwentarza fonemów języka polskiego*. Został w nim przedstawiony wynik analizy porównawczej obejmującej 15 różnych inwentarzy fonologicznych. Są to opracowania następujących autorów: Zdzisława Stiebera [1966], Piotry Łobacz [1973], Bożeny Wierzchowskiej [1980], Jerzego Bańczerowskiego [1982], Jerzego J. Rubacha [1984], Barbary Klebanowskiej [1990], Romana Laskowskiego [1991], Christiny Y. Bethin [1992], Ireny Sawickiej [1995], Jolanty Szpyry [1995], Marka Wiśniewskiego [1998], Danuty Ostaszewskiej i Jolanty Tambor [2000], Bronisława Rocławskiego [2001], Wiktora Jassema [2003] oraz Edmunda Gussmanna [2007]. Analiza ma istotne znaczenie dla koncepcji przedstawionej w niniejszej książce, dlatego tabela zawierająca jej wynik została przytoczona w postaci oryginalnej (tabela 1.1). Również wnioski zamieszczone w dalszym ciągu tej części są oparte na informacjach pochodzących ze wspomnianej publikacji.

Tabela 1.1 – Porównanie polskich inwentarzy fonologicznych (autorka analizy: Ewa Wolańska)

Fonem		Z. Stieber (1966)	P. Łobacz (1973)	B. Wierzchowska (1980)	J. Bańczerowski (1982)	J.J. Rubach (1984)	B. Klebanowska (1990)	R. Laskowski (1991)	Ch.Y. Bethin (1992)	I. Sawicka (1995)	J. Szpyra (1995)	M. Wiśniewski (1998)	D. Ostaszewska i J. Tambor (2000)	B. Rocławski (2001)	W. Jassem (2003)	E. Gussmann (2007)	Liczba wystąpień fonemu w inwentarzach
IPA	SAF	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	*N* = 15
/a/	/a/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/b/	/b/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/bʲ/	/ḿ/	+	–	+	+	–	–	–	+	–	+	–	–	–	–	+	6
/ʦ/	/c/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/ʨ/	/ć/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/ʦʲ/	/c′/	–	–	–	–	–	–	–	–	–	+	–	–	–	–	+	2
/ʧ/	/č/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/ʧʲ/	/č′/	–	–	–	–	–	–	–	–	–	+	–	–	–	–	+	2
/d/	/d/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/dʲ/	/d′/	–	–	–	–	–	–	–	–	–	+	–	–	–	–	+	2
/ʣ/	/ʒ/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/ʥ/	/ʑ/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/ʤ/	/ǯ/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/ʤʲ/	/ǯ′/	–	–	–	–	–	–	–	–	–	+	–	–	–	–	+	2
/ɛ/	/e/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/ɛ̃/	/ę/	+	–	+	+	+	–	–	+	–	+	–	–	+	–	–	7
/f/	/f/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/fʲ/	/ḟ/	+	–	+	+	–	–	–	+	–	+	–	–	–	–	+	6

Fonem		Z. Stieber (1966)	P. Łobacz (1973)	B. Wierzchowska (1980)	J. Bańczerowski (1982)	J.J. Rubach (1984)	B. Klebanowska (1990)	R. Laskowski (1991)	Ch.Y. Bethin (1992)	I. Sawicka (1995)	J. Szpyra (1995)	M. Wiśniewski (1998)	D. Ostaszewska i J. Tambor (2000)	B. Rocławski (2001)	W. Jassem (2003)	E. Gussmann (2007)	Liczba wystąpień fonemu w inwentarzach
IPA	SAF	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	*N* = 15
/g/	/g/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/ɟ/	/ǵ/	+	+	+	+	–	+	+	+	+	+	+	+	–	+	+	13
/x/	/χ/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/ç/	/χ́/	–	–	–	–	–	–	–	+	+	+	–	–	–	–	+	4
/ɣ/	/γ/	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	1
/ɣʲ/	/γ́/	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	1
/i/	/i/	+	+	–	+	+	+	+	+	+	+	+	+	+	+	+	14
/ĩ/	/i̯/	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	1
/j/	/i̯/	+	–	+	+	+	+	+	+	+	+	+	+	+	+	+	14
/j̃/	/ĩ̯/	–	–	–	–	–	–	–	–	–	–	+	–	–	+	–	2
/k/	/k/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/c/	/ḱ/	+	+	+	+	–	+	+	+	+	+	+	+	–	+	+	13
/l/	/l/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/lʲ/	/l′/	–	–	–	–	–	–	–	–	–	+	–	–	–	–	+	2
/ɬ/	/ɬ/	+	–	–	–	–	–	–	–	+	+	+	–	–	–	–	4
/w/	/u̯/	–	–	+	+	+	+	+	+	+	+	+	+	+	+	+	13
/w̃/	/ũ̯/	–	+	–	–	–	+	–	–	–	–	+	+	–	+	+	6
/m/	/m/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/mʲ/	/ḿ/	+	–	–	+	–	–	–	+	–	+	–	–	–	–	+	5
/n/	/n/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/ŋ/	/ŋ/	–	+	–	–	–	–	+	–	+	–	+	–	–	+	+	6
/ɲ/	/ń/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/ɔ/	/o/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/ɔ̃/	/o̯/	+	–	+	+	+	–	–	+	–	+	–	–	+	–	–	7
/p/	/p/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/pʲ/	/ṕ/	+	–	+	+	–	–	–	+	–	+	–	–	–	–	+	6
/r/	/r/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/rʲ/	/r′/	–	–	–	–	–	–	–	–	–	+	–	–	–	–	+	2
/s/	/s/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/ɕ/	/ś/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/sʲ/	/s′/	–	–	–	–	–	–	–	–	–	+	–	–	–	–	+	2
/ʃ/	/š/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/ʃʲ/	/š′/	–	–	–	–	–	–	–	–	–	+	–	–	–	–	+	2
/t/	/t/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/tʲ/	/t′/	–	–	–	–	–	–	–	–	–	+	–	–	–	–	+	2
/u/	/u/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/ũ/	/u̯/	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	1
/v/	/v/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/vʲ/	/v́/	+	–	+	+	–	–	–	+	–	+	–	–	–	–	+	6
/ɨ/	/y/	–	+	–	–	+	+	+	+	+	+	+	+	+	+	+	12
/z/	/z/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15

Fonem		Z. Stieber (1966)	P. Łobacz (1973)	B. Wierzchowska (1980)	J. Bańczerowski (1982)	J.J. Rubach (1984)	B. Klebanowska (1990)	R. Laskowski (1991)	Ch.Y. Bethin (1992)	I. Sawicka (1995)	J. Szpyra (1995)	M. Wiśniewski (1998)	D. Ostaszewska i J. Tambor (2000)	B. Rocławski (2001)	W. Jassem (2003)	E. Gussmann (2007)	Liczba wystąpień fonemu w inwentarzach
IPA	SAF	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	*N* = 15
/ʑ/	/ź/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/zʲ/	/z′/	–	–	–	–	–	–	–	–	–	+	–	–	–	–	+	2
/ʒ/	/ž/	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	15
/ʒʲ/	/ž′/	–	–	–	–	–	–	–	–	–	+	–	–	–	–	+	2
–	/I/[1]	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	1
/V/	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	1
Suma		42	36	41	44	36	37	37	44	41	57	40	37	36	39	55	

[1] /I/ – wspólne oznaczenie alofonów /i/ ([i]) oraz /ɨ/ ([y]).

W przytoczonej publikacji autorka zaznaczyła, że najmniej liczne zestawienia fonemów różnią się od zestawienia najbardziej licznego (obejmuje ono 57 elementów) brakiem 17 następujących jednostek: /tsʲ ʧʲ bʲ dʲ ʤʲ fʲ ç lʲ mʲ pʲ rʲ sʲ ʃʲ tʲ vʲ zʲ ʒʲ/ (/cʹ čʹ bʹ dʹ ǯʹ fʹ χʹ lʹ mʹ pʹ rʹ sʹ šʹ tʹ vʹ zʹ žʹ/). Zbiór całkowicie niekontrowersyjnych fonemów (które pojawiły się w co najmniej czternastu z piętnastu przedstawionych zestawień) obejmuje 32 jednostki: /a b ts tɕ ʧ d dz dʑ ʤ ɛ f g x i j k l m n ɲ ɔ p r s ɕ ʃ t u v z ʑ ʒ/ (/a b c ć č d ʒ ʑ ǯ e f g χ i i̯ k l m n ń o p r s ś š t u v z ź ž/). Poza tym autorka poruszyła kilka problemów, które mają związek z różnicami występującymi pomiędzy poszczególnymi inwentarzami fonologicznymi:

– ujmowanie fonemów [i] ([i]) oraz [ɨ] ([y]) jako wariantów tego samego fonemu (w starszych opracowaniach);
– bifonematyczność tzw. polskich samogłosek nosowych (w nowszych opracowaniach) [Biedrzycki 1963, 1978; Dukiewicz 1967; Dziczek-Karlikowska 2012; Kamińska 2004; Nowakowski, Wiatrowski 2020; Zagórska-Brooks 1968]. Konsekwencją tego podejścia było wyodrębnienie niesylabicznych fonemów nosowych: /w̃/ (/ũ̯/) oraz /j̃/ (/ĩ̯/). W podejściu bifonematycznym litery *ę* oraz *ą* oznaczają ciąg zawierający fonem /ɛ/ (/e/) lub fonem /ɔ/ (/o/), po którym następuje fonem ze zbioru: /w̃ j̃ m n ɲ ŋ/ (/u̯ ũ̯ m n ń ŋ/). Jest to alternatywa dla podejścia monofonematycznego, w którym wyodrębnia się fonem /ɛ̃/ (/ę/) oraz fonem /ɔ̃/ (/ǫ/);
– asynchroniczna wymowa spółgłosek wargowych palatalnych [pʲ bʲ fʲ vʲ mʲ] ([ṕ b́ f́ v́ ḿ]) z wydzieleniem się głoski [j] ([i̯]). W większości najnowszych opracowań spółgłoski miękkie wargowe są wariantami pozycyjnymi odpowiednich fonemów twardych [Osowicka-Kondratowicz 2005; Retz 1989; Rocławski 1984; Sawicka 1999];
– wyodrębnienie przez niektórych autorów spółgłoskowego, półotwartego, nosowego, tylnojęzykowego fonemu /ŋ/ (/ŋ/). W innych ujęciach głoska [ŋ] ([ŋ]) uznawana jest za wariant pozycyjny nosowego, niesylabicznego fonemu /w̃/ (/ũ̯/) [Ostaszewska, Tambor 2000]. W jeszcze innym ujęciu rezonans nosowy [w̃] ([ũ̯]) jest alofonem fonemu /ŋ/ (/ŋ/) [Biedrzycki 1978];
– różnice dotyczą również palatalnych obstruentów /c ɟ/ (/ḱ ǵ/). Większość autorów uważa, że różnica /g/ (/g/) : /ɟ/ (/ǵ/) oraz /k/ (/k/) : /c/ (/ḱ/) jest fonologiczna;
– wyodrębnienie przez Irenę Sawicką fonemu /ɣ/ (/γ/) – według większości badaczy jest to alofon fonemu /x/ (/χ/). Autorka taką interpretację motywuje zróżnicowaniem regionalnym. Również miękkie /ç/ (/χ́/), które występuje tylko przed /i j/ (/i i̯/), zostało uznane za fonem tylko w opracowaniu tej autorki;
– większość badaczy systemu fonologicznego współczesnej polszczyzny uznaje głoskę [ɫ] ([ɫ]) za wariant fonemu /w/ (/u̯/).

W niniejszej publikacji dla omówienia koncepcji wielowarstwowego modelu transkrypcji oraz w celu przedstawienia wyników badań użyto inwentarza fonologicznego zaproponowanego przez Marię Steffen-Batogową. W wypadku odniesienia do wyników trzeba uwzględnić ewentualne różnice dotyczące inwentarzy.

Istotną kwestią jest użyty alfabet transkrypcyjny. Jest to zmodyfikowana wersja standardu SAMPA (w wersji dla języka polskiego). Nazwa SAMPA (ang. *Speech Assessment Methods Phonetic Alphabet*) została po raz

pierwszy użyta przez Johna C. Wellsa. Najważniejsze założenie tego rozwiązania dotyczy zapisywania fonemów lub głosek przy użyciu zestawu znaków ASCII [Wells 1997]. SAMPA jest pierwowzorem dla standardu XSAMPA (ang. *Extended Speech Assessment Methods Phonetic Alphabet*). Różnica pomiędzy tymi standardami polega na tym, że symbole SAMPA były ustalane indywidualnie dla poszczególnych języków, natomiast standard X–SAMPA jest uniwersalny. Należy zaznaczyć, że w dalszym ciągu publikacji w nielicznych miejscach została użyta transkrypcja alofoniczna. Głoski były zapisywane przy użyciu alfabetu slawistycznego (dalej: AS).

W tabeli 1.2 przedstawiono przyjęty inwentarz fonemów, który został zapisany przy użyciu zmodyfikowanego alfabetu SAMPA [Demenko 2003]. Zawiera on fonemy /c/ i /J/ (jak w wyrazach: *kino, gitara*). Wyodrębnienie tych fonemów ma związek z założeniem dotyczącym możliwej synchronicznej realizacji miękkości w strukturach ortograficznych typu: *kie, kia, kio, gie, gia*. Ewentualne przyjęcie konsekwentnej (bezwyjątkowej) wymowy asynchronicznej wiąże się z uznaniem głosek: [ḱ] i [ǵ] (AS) za alofony fonemów: /k/ i /g/. Trzeba zaznaczyć, że w tej książce uwzględniono obie interpretacje fonologiczne, dlatego fonemy /c/ i /J/ w tabeli 1.2 zostały specjalnie oznaczone. Warto zauważyć, że ten inwentarz zawiera również oddzielną notację /rz/ (obok /Z/) – w rzeczywistości jest to jeden fonem. To niekonwencjonalne podejście wynika z faktu, że głoska odpowiadająca diadzie ortograficznej *rz* podlega perseweracji artykulacyjnej, co ma istotne znaczenie przy przekształceniach związanych z modyfikacją statusu dźwięczności. Powód wyodrębnienia notacji /rz/ jest techniczny i ma związek z funkcjonowaniem systemu transkrypcji opartego na modelu wielowarstwowym.

Tabela 1.2 – Inwentarz fonemów przyjęty dla tej publikacji

Lp.	Fonem SAMPA	Przykład SAMPA	Przykładowy wyraz	Lp.	Fonem SAMPA	Przykład SAMPA	Przykładowy wyraz
1	/a/	/m.**a**.k/	*mak*	20	/rz/	/**rz**.e.k.a/	*rzeka*
2	/e/	/z.**e**.r.o/	*zero*	21	/Z/	/**Z**.a.b.a/	*żaba*
3	/i/	/n'.**i**.g.dz'.e/	*nigdzie*	22	/s'/	/**s'**.m.j.e.x/	*śmiech*
4	/o/	/s.**o**.v.a/	*sowa*	23	/z'/	/**z'**.a.r.n.o/	*ziarno*
5	/y/	/b.**y**.k/	*byk*	24	/x/	/k.u.**x**.n'.a/	*kuchnia*
6	/u/	/j.**u**.t.r.o/	*jutro*	25	/p/	/**p**.a.l.e.ts/	*palec*
7	/w/	/p.u.**w**.k.a/	*półka*	26	/b/	/**b**.u.d.a/	*buda*
8	/j/	/**j**.e.d.e.n/	*jeden*	27	/t/	/**t**.a.m.a/	*tama*
9	/w~/	/j.e.**w~**.z.y.k/	*język*	28	/d/	/**d**.o.m/	*dom*
10	/l/	/v.j.e.**l**.e/	*wiele*	29	/k/	/p.o.**k**.u.j/	*pokój*
11	/r/	/**r**.y.b.a/	*ryba*	30	/g/	/**g**.o.s'.ts'/	*gość*
12	/m/	/**m**.a.k/	*mak*	31	/c/*	/**c**.i.n.o/	*kino*
13	/n/	/m.o.**n**.e.t.a/	*moneta*	32	/J/*	/**J**.i.t.a.r.a/	*gitara*
14	/n'/	/k.o.**n'**/	*koński*	33	/ts/	/**ts**.y.r.k/	*cyrk*
15	/f/	/**f**.u.t.r.o/	*futro*	34	/dz/	/**dz**.v.o.n.e.k/	*dzwonek*
16	/v/	/**v**.j.a.t.r/	*wiatr*	35	/tS/	/**tS**.a.s/	*czas*
17	/s/	/s.t.**r**.o.m.y/	*stromy*	36	/dZ/	/**dZ**.u.m.a/	*dżuma*
18	/z/	/k.o.**z**.a/	*koza*	37	/ts'/	/k.o.**ts'**.o.w/	*kocioł*
19	/S/	/m.a.**S**.t/	*maszt*	38	/dz'/	/**dz'**.a.w.k.a/	*działka*

1.2. Grafem, diada ortograficzna i triada ortograficzna

W tej części omówiono kilka istotnych terminów, które są używane w monografii. Pierwsza definicja dotyczy grafemu. Jest to znak literowy lub połączenie literowe odnoszące się do jednego fonemu [Kaczmarek 2003; Polański 2003]. Podział grafemów, który jest istotny dla tej publikacji, opiera się na liczbie liter tworzących dany grafem [Wolańska 2019]. A więc można wyróżnić:

– monografemy, czyli grafemy złożone z jednej litery (na przykład *t* w wyrazie *tama*). W określonych interpretacjach fonologicznych monografem może odnosić się do dwóch fonemów (na przykład w interpretacjach dotyczących liter: *ą* oraz *ę*);
– digrafemy, które są złożone z dwóch liter (na przykład *sz* w wyrazie *szafa*);
– trigrafemy to grafemy złożone z trzech liter (na przykład *dzi* w wyrazie *dziesięć*).

W tekście używane jest również pojęcie wieloznaku. Ten termin obejmuje zarówno digrafemy, jak i trigrafemy. A więc dotyczy struktur ortograficznych złożonych z co najmniej dwóch liter, które oznaczają jeden fonem.

Uściślenia wymagają również często używane w tej pracy pojęcia: litera, diada ortograficzna oraz triada ortograficzna. Litera jest podstawową jednostką zapisu ortograficznego (w pracy używany jest też termin: znak ortograficzny). Alfabet polski liczy 32 litery: *a, ą, b, c, ć, d, e, ę, f, g, h, i, j, k, l, ł, m, n, ń, o, ó, p, r, s, ś, t, u, w, y, z, ź, ż*. Taka definicja pokrywa się z pojęciem litery pojedynczej, które było obecne w jednej z wcześniejszych prac Rocławskiego [Rocławski 1986]. W niniejszej monografii nie jest natomiast używany termin litera w odniesieniu do wszelkich połączeń znaków ortograficznych (np.: *sz, rz, dzi, dż* itp.). Dla takich konstrukcji są używane terminy: diada ortograficzna oraz triada ortograficzna. Trzeba podkreślić, że przyjęte definicje nie przyporządkowują poszczególnych liter, diad oraz triad ortograficznych do konkretnych fonemów czy głosek. Takie powiązanie może być zawarte w umownej transkrypcji podstawowej (zob. podrozdz. 1.3) oraz w regułach włączonych do konkretnego projektu systemu transkrypcji opartego na modelu wielowarstwowym. A więc diada ortograficzna jest parą dowolnych znaków ortograficznych (liter), chociaż opis najczęściej dotyczy diad, które oznaczają lub mogą oznaczać jeden fonem, a więc są lub mogą być digrafememi.

1.3. Transkrypcja podstawowa

Transkrypcja podstawowa (dalej w tym rozdziale: TP) to pojęcie, które zostało wprowadzone na potrzeby omawianych w tej publikacji rozwiązań. Obejmuje ona elementarne relacje między znakami ortograficznymi a fonemami. Można powiedzieć, że jest to pierwotny zbiór reguł transkrypcyjnych.

TP ma charakter umowny i może być modyfikowana między innymi ze względu na różnice występujące pomiędzy poszczególnymi inwentarzami fonologicznymi. Podstawowe założenie związane z TP polega na przypisywaniu określonych fonemów do znaków, diad i triad ortograficznych niezależnie od ich kontekstu. Istnieje zatem możliwość, że dana struktura ortograficzna w określonym kontekście powinna być transkrybowana inaczej, niż zostało to przyjęte w TP. Jednak często właściwa transkrypcja pokrywa się z TP.

Drugie istotne założenie dotyczy relacji między literami a określonymi diadami i triadami ortograficznymi. Zgodnie z nim reguły TP ustalone dla segmentów dłuższych (w sensie liczby znaków) mają wyższy priorytet. Na przykład reguła dotycząca diady *sz* ma wyższy priorytet, niż reguła ustalona dla litery *s* czy dla litery *z*.

W tabeli 1.3 oraz w tabeli 1.4 przedstawiono umowną TP przyjętą na potrzeby niniejszej publikacji (oddzielnie ujęto informacje dotyczące pojedynczych liter oraz diad ortograficznych). Zrezygnowano z ustalenia TP dla litery *ę* oraz dla litery *ą*.

Tabela 1.3 – TP przyjęta dla pojedynczych liter

Lp.	Ort.	TP	Lp.	Ort.	TP
1	*a*	/a/	17	*m*	/m/
2	*ą*	–	18	*n*	/n/
3	*b*	/b/	19	*ń*	/n'/
4	*c*	/ts/	20	*o*	/o/
5	*ć*	/ts'/	21	*ó*	/u/
6	*d*	/d/	22	*p*	/p/
7	*e*	/e/	23	*r*	/r/
8	*ę*	–	24	*s*	/s/
9	*f*	/f/	25	*ś*	/s'/
10	*g*	/g/	26	*t*	/t/
11	*h*	/x/	27	*u*	/u/
12	*i*	/i/	28	*w*	/v/
13	*j*	/j/	29	*y*	/y/
14	*k*	/k/	30	*z*	/z/
15	*l*	/l/	31	*ż*	/Z/
16	*ł*	/w/	32	*ź*	/z'/

Tabela 1.4 – TP przyjęta dla określonych diad ortograficznych

Lp.	Ort.	TP
1	*ch*	/x/
2	*cz*	/tS/
3	*dz*	/dz/
4	*dź*	/dz'/
5	*dż*	/dZ/
6	*rz*	/rz/
7	*sz*	/S/

Oprócz wymienionych w tabeli 1.4 struktur, które włączono do TP, system ortograficzny języka polskiego obejmuje kilka diad i jedną triadę, które są wieloznakami tylko w określonym kontekście ortograficznym – przed literą oznaczającą samogłoskę (inną niż litera *i*). Ze względu na tak silne uwarunkowanie kontekstowe te struktury nie zostały włączone do TP. Wyszczególniono je w tabeli 1.5.

Tabela 1.5 – Struktury ortograficzne wykluczone z przyjętej TP

Lp.	Zapis ortograficzny	Transkrypcja w kontekście samogłoskowym	Transkrypcja w kontekście spółgłoskowym
1	*ci*	/ts'/	/ts'.i/
2	*dzi*	/dz'/	/dz'.i/
3	*ni*	/n'/	/n'.i/
4	*si*	/s'/	/s'.i/
5	*zi*	/z'/	/z'.i/

Koncepcja TP w pewnym zakresie pokrywa się z istotnymi pojęciami i koncepcjami lingwistycznymi. Podobieństwo dotyczy przede wszystkim pojęcia grafemu rozumianego jako znak literowy lub połączenie literowe odnoszące się do jednego fonemu (zob. podrozdz. 1.2). Różnice mogą natomiast dotyczyć faktu, że z TP został wyłączony czynnik kontekstu ortograficznego. Można tutaj przytoczyć zdefiniowane przez Tomasza Lisowskiego pojęcie grafemu oraz pojęcie alografu [Lisowski 2001]. Autor definiuje grafem jako dyspozycję do graficznego substytuowania fonemu, która jest realizowana w tekście przez alografy [Wolańska 2019]. Jednym z kryteriów wskazanych przez autora jest pozycja alografu w kontekście ortograficznym. W rozwiązaniach opartych na omówionym w niniejszej publikacji modelu wielowarstwowym kontekst poszczególnych segmentów może być uwzględniany w czasie przetwarzania wyrazów przy użyciu reguł, które są zdefiniowane w ramach kolejnych warstw.

1.4. Wieloczynnikowa natura transkrypcji tekstu

W języku polskim występują złożone relacje pomiędzy płaszczyzną ortograficzną i płaszczyzną fonologiczną (oraz fonetyczną). Jedna litera może oznaczać dwa fonemy. Natomiast dwie lub trzy litery mogą oznaczać pojedynczy fonem (są to digrafemy i trigrafemy). Poza tym poszczególne znaki oraz diady ortograficzne mogą oznaczać różne fonemy w zależności od kontekstu w jakim są one umiejscowione. W procesie przekształcania zapisu ortograficznego na zapis transkrypcyjny można wyodrębnić kilka zasadniczych czynników:
– wielofunkcyjność litery *i*;
– obecność digrafemów oraz trigrafemów;
– modyfikacje statusu dźwięczności;
– uproszczenia grup spółgłoskowych;
– inne modyfikacje w grupach głoskowych.

Litera *i* może oznaczać fonem (/i/ lub /j/) oraz może być jednocześnie znakiem miękkości. Może też być ona tylko znakiem miękkości. Jej obecność często ma wpływ na interpretację fonologiczną litery spółgłoskowej umiejscowionej w kontekście lewostronnym.

Do problemu digrafemów i trigrafemów odniesiono się w częściach 1.2 i 1.3 oraz w rozdziale drugim. W pewnym zakresie pokrywa się on z wielofunkcyjnością litery *i*, która również może być komponentem tych struktur. Wyznaczanie fragmentów wyrazów stanowiących digrafemy i trigrafemy jest istotną czynnością w procesie transkrypcji.

Dźwięczność jest cechą fonetyczną (fizyczną) dźwięków mowy. Ma ona związek z drganiem wiązadeł głosowych w czasie procesu artykulacji. Dźwięczność zatem jest jedną z podstawowych cech jednostek lingwistycznych związanych z dźwiękami mowy – z fonemami oraz z głoskami. W niniejszym opracowaniu w sposób umowny używane jest pojęcie dźwięczności ortograficznej liter oraz diad ortograficznych. Jednostki te funkcjonują na płaszczyźnie ortograficznej (graficznej), dlatego podane nazewnictwo wymaga dodatkowego wyjaśnienia. Pod pojęciem dźwięczności ortograficznej liter oraz diad ortograficznych rozumiany jest status dźwięczności fonemów, które odpowiadają tym strukturom w przyjętej TP (zob. podrozdz. 1.3). A więc jest to domyślny status dźwięczności segmentów ortograficznych. Informacje te zawarto w tabeli 1.6.

Tabela 1.6 – Dźwięczność ortograficzna liter i diad w przyjętej TP

Lp.	Litera lub diada ort.	TP	Status dźwięczności
1	*f*	/f/	bezdźwięczny
2	*w*	/v/	dźwięczny
3	*s*	/s/	bezdźwięczny
4	*z*	/z/	dźwięczny
5	*sz*	/S/	bezdźwięczny
6	*ż*	/Z/	dźwięczny
7	*rz*	/rz/	dźwięczny
8	*ś*	/s'/	bezdźwięczny
9	*ź*	/z'/	dźwięczny
10	*h*	/x/	bezdźwięczny
11	*ch*	/x/	bezdźwięczny
12	*p*	/p/	bezdźwięczny
13	*b*	/b/	dźwięczny
14	*t*	/t/	bezdźwięczny
15	*d*	/d/	dźwięczny
16	*k*	/k/	bezdźwięczny
17	*g*	/g/	dźwięczny
18	*c*	/ts/	bezdźwięczny
19	*dz*	/dz/	dźwięczny
20	*cz*	/tS/	bezdźwięczny
21	*dż*	/dZ/	dźwięczny
22	*ć*	/ts'/	bezdźwięczny
23	*dź*	/dz'/	dźwięczny

Tabela 1.7 dotyczy z kolei domyślnej dźwięczności diad ortograficznych (oraz jednej triady) nieuwzględnionych w przyjętej TP (podane w kolumnie drugiej struktury ortograficzne są wieloznakami tylko w kontekście prawostronnym litery oznaczającej samogłoskę inną niż *i*).

Tabela 1.7 – Dźwięczność ortograficzna struktur nieuwzględnionych w przyjętej TP

Lp.	Diada lub triada ort.	Transkrypcja	Status dźwięczności
1	*ci*	/ts'/	bezdźwięczny
2	*dzi*	/dz'/	dźwięczny
3	*si*	/s'/	bezdźwięczny
4	*zi*	/z'/	dźwięczny

Kolejny wymieniony czynnik obejmuje uproszczenia grup spółgłoskowych [por.: Bargiełówna 1950; Dobrogowska 1984, 1990, 1992; Dukiewicz 1985; Dunaj 1985, 1986; Kuryłowicz 1952; Sawicka 1974; Skulina 1964].

Wieloczynnikowa natura transkrypcji tekstu ma kluczowe znaczenie dla modelu wielowarstwowego. Model ten nie narzuca przypisywania określonych funkcji do poszczególnych warstw, jednak jest to podejście najbardziej uzasadnione. W niniejszej publikacji wykazano, że poszczególne czynniki można traktować niezależnie. Dla każdego czynnika (lub dla grupy czynników) można utworzyć oddzielny zbiór reguł i każdy z tych zbiorów może zostać powiązany z oddzielną warstwą. W przykładowym projekcie (zob. rozdz. 4) do pierwszej warstwy zostały przypisane reguły dotyczące wydzielania digrafemów i trigrafemów w wyrazach (ma to związek z wielofunkcyjnością litery *i*). Warstwa druga zawiera reguły dotyczące asymilacji dźwięczności, natomiast do warstwy trzeciej włączono reguły związane z uproszczeniami w ramach grup spółgłoskowych oraz z innymi modyfikacjami. Najistotniejszy jest fakt, że istnieje możliwość konstruowania reguł, które są związane tylko z jednym problemem. Dzięki temu zbiory reguł można łatwo interpretować i porządkować. Poza tym konstrukcja tych reguł jest najczęściej stosunkowo prosta, zatem ich ewentualna modyfikacja nie jest kłopotliwa.

1.5. Zapis reguł przytoczonych z *Automatyzacji…*

We wprowadzeniu podkreślono, że analizy wstępne omówione w rozdziale drugim są oparte na regułach pochodzących z *Automatyzacji…* . Sposób zapisywania tych reguł został jednak zmodyfikowany. W *Automatyzacji…* zapis wszystkich reguł jest ujęty w tabelach, w których pierwsza kolumna oraz pierwszy wiersz mają szczególne znaczenie. Komórki w pierwszej kolumnie zawierają definicje kontekstu lewostronnego danej litery, natomiast w komórkach w pierwszym wierszu znajdują się definicje kontekstu prawostronnego tej samej litery. Pozostałe komórki (w wierszach od drugiego wzwyż oraz w kolumnach od drugiej wzwyż) zawierają transkrypcję fonologiczną tejże litery (w określonych kontekstach). W definicjach kontekstów używane są symbole oznaczające zbiory liter:

A={*a, ą, e, ę, i, o, ó, u, y*}
D={*b, d, g, z, ź, ż*}
T={*c, ć, f, h, k, p, s, ś, t*}
R={*l, ł, r, w*}
M={*m, n, ń, j*}
V={*a, ą, e, ę, i, o, ó, u, y, l, ł, r, w, m, n, ń, j*}

Oprócz przytoczonych definicji zbiorów, w *Automatyzacji…* zdefiniowano zbiór X, który obejmuje wszystkie litery, a także dwa zbiory: O i O', które zawierają znaki interpunkcyjne i symbole specjalne (zbiór O' nie zawiera symbolu # oznaczającego granicę wyrazu ortograficznego).

W niniejszej publikacji oryginalny tabelaryczny sposób prezentacji reguł został zastąpiony przez zapis jednolity. Dzięki temu każdą regułę można pomieścić w pojedynczej komórce tabeli. Poniżej przedstawiono przykładowy, używany w tej publikacji, zapis reguły pochodzącej z opracowania Marii Steffen-Batogowej:

#o[*d*]*z*(D+M+R–*w*)=/d/

W konstrukcji tej każdy typ nawiasów ma określoną funkcję i jest to konwencja przyjęta tylko na potrzeby niniejszej publikacji (i *Asymilacji…*) i tylko dla zapisywania reguł przytaczanych z *Automatyzacji…* . W nawiasach kwadratowych umiejscowiona jest litera, której dotyczy dana reguła (przed przekształceniem). Przed otwierającym nawiasem kwadratowym definiowany jest ortograficzny kontekst lewostronny dla tej litery, natomiast po nawiasie kwadratowym zamykającym (przed znakiem równości) definiowany jest ortograficzny kontekst prawostronny. W definicjach kontekstów mogą być używane symbole oznaczające zbiory liter, które zostały przytoczone w tej części. W nawiasach okrągłych może być zamieszczony zapis definiujący operacje na zbiorach i elementach zbiorów – wyznaczanie sumy oraz różnicy. Zapis: (D+M+R–*w*) oznacza jeden dowolny element z sumy zbiorów: D, M i R oprócz litery *w*. Fragment podanej reguły: *#o* oznacza kontekst lewostronny złożony z litery *o* umiejscowionej na początku wyrazu ortograficznego. Z kolei fragment: *z*(D+M+R–*w*) oznacza ciąg złożony z dwóch liter (*z* jest w nim na pierwszej pozycji). Umieszczony za znakiem równości zapis: /d/ stanowi transkrypcję fonologiczną litery *d*.

Zbiory elementów mogą być definiowane przy użyciu nawiasów klamrowych, które umieszcza się w nawiasach okrągłych, na przykład:

(X–{*d,r,c,s*})[*z*]TD=/s/

Powyższy zapis jest równoznaczny z następującym zapisem:

(X–*d*–*r*–*c*–*s*)[*z*]TD=/s/

Obie konstrukcje (przed nawiasem kwadratowym) oznaczają jednoelementowy kontekst lewostronny (tutaj kontekst litery *z*) złożony z dowolnego znaku z wyjątkiem liter: *d, r ,c, s*. Fragment: TD oznacza natomiast ciąg dwóch liter, przy czym pierwsza z tych liter należy do zbioru T, natomiast litera druga należy do zbioru D. Po znaku równości jest umieszczony zapis transkrypcji litery *z* w podanym kontekście. W wypadku każdej reguły wynikowa transkrypcja obowiązuje tylko dla kontekstu lewostronnego i kontekstu prawostronnego, które są określone przed znakiem równości. Opisany tutaj sposób zapisywania reguł jest równoważny z zapisem odpowiednich reguł w *Automatyzacji…* .

1.6. Wcześniejsze prace dla języka polskiego

Prace dotyczące automatycznej konwersji tekstu ortograficznego w języku polskim na zapis transkrypcyjny powstawały już w pierwszej połowie lat siedemdziesiątych XX w. W 1973 roku ukazał się artykuł M.S.-B.: *The problem of automatic phonemic transcription of written Polish* [Steffen-Batogowa 1973]. Natomiast *Automatyzacja…* była pierwszym kompletnym opracowaniem dotyczącym omawianego zagadnienia dla języka polskiego. Odniesiono się w niej do całokształtu dostępnych wtedy wyników badań fonetycznych. Przedstawione metody, sposób formułowania reguł, jak i same reguły były wielokrotnie używane w pracach późniejszych. W artykule: *Implementacja algorytmu transkrypcji fonematycznej* [Wypych 1999] autor opisał własną implementację reguł transkrypcji M.S.-B. oraz zawarł przegląd prac innych autorów, które pochodzą z lat siedemdziesiątych, osiemdziesiątych oraz dziewięćdziesiątych XX w. Zestawienie to otwiera publikacja: *Program na maszynę Odra-1204 dla automatycznej transkrypcji fonematycznej tekstów języka polskiego* [Warmus 1972]. Dotyczy ona pierwszego działającego programu komputerowego opartego na wówczas jeszcze niekompletnych regułach M.S.-B. Sam system był na tyle powolny, że nie było możliwe jego używanie w czasie rzeczywistym. Miało to związek głównie z ograniczeniami sprzętowymi. Kolejne publikacje w tym zestawieniu wnoszą ulepszenia głównie techniczne [Maksymienko, Bolc 1992; Chomyszyn 1986; Nowak 1995; Jassem 1996]. W rozwiązaniu opisanym przez Krzysztofa Jassema w artykule: *A phonemic transcription – syllable division rule engine* zostały wyodrębnione moduły: reguł zapisywanych w pliku tekstowym i interpretatora reguł. Osiągnięcia z *Automatyzacji…* były także rozwijane w późniejszym czasie przez autorkę monografii. W artykule: *Rules for the mutual conversion of the phonemic and phonetic transcriptions of the Polish texts* [Steffen-Batóg 1989-1990] zostały omówione reguły umożliwiające konwersje transkrypcji fonologicznej na transkrypcję fonetyczną. Z kolei w artykule: *An algorithm for phonetic transcription of orthographic texts in Polish* [Steffen-Batóg, Nowakowski 1992] omówiono reguły konwersji tekstu na transkrypcję fonetyczną. Te prace (łącznie z monografią z 1975 roku) stanowiły podstawę kolejnych implementacji praktycznych. W 2003 roku ukazał się artykuł: *Implementation of grapheme-to-phoneme rules and extended SAMPA alphabet in Polish text-to-speech synthesis* [Demenko i in. 2003]. Zawarto w nim opis wybranych problemów lingwistycznych związanych z transkrypcją automatyczną na potrzeby syntezy mowy. Poruszono istotną kwestię alfabetu transkrypcyjnego (użyto modyfikacji standardu SAMPA). Innym przykładem pracy opartej na tabelach Batogowej jest artykuł: *Ortfon2 – tool for orthographic to phonetic transcription* [Skurzok i in. 2015]. Autorzy przedstawili własne rozwiązania, między innymi metodę binarnej reprezentacji ciągów segmentów. W artykule: *Algorithm and implementation of automatic phonemic transcription for polish* [Kłosowski 2016] autor omówił działanie programu *TransFon*. Program ten został napisany w języku Python i jego działanie polega na sekwencyjnym przetwarzaniu pojedynczych liter wyrazu przy użyciu zestawu reguł, w których jest uwzględniony kontekst. Koleje rozwiązanie zostało zaimplementowane w ramach platformy CLARIN (Common Language Resources & Technology Infrastructure) [Koržinek i in. 2016a, 2016b]. Omówione pozycje dotyczą systemów konwersji zapisu ortograficznego na transkrypcję, które są oparte na zaprojektowanych regułach. Jednak dla języka polskiego powstały również algorytmy, których celem jest automatyczne tworzenie reguł transkrypcji. W monografii: *Rekonstrukcja reguł transkrypcji fonematycznej na podstawie próby uczącej* [Pluciński 2002] autor opisał oryginalną koncepcję systemu uczącego się, który na podstawie danych empirycznych potrafi budować określone statystyki. Na ich podstawie możliwy jest adaptacyjny dobór i rekonstrukcja reguł transkrypcji. Inna publikacja, w której omówiono proces automa-

tycznego generowania reguł transkrypcji dla języka polskiego to: *The generation of letter-to-sound rules for grapheme-to-phoneme conversion* [Przybysz, Kasprzak 2013].

2. Lingwistyczne podstawy fonologicznej transkrypcji tekstu

2.1. Wprowadzenie

W tym rozdziale omówiono zasady przekształcania tekstu ortograficznego na transkrypcję fonologiczną. Zaprezentowano wyniki szczegółowych analiz słownikowych oraz korpusowych. Znaczna część tych analiz została oparta na regułach pochodzących *Automatyzacji…* . Sposób zapisywania tych reguł został omówiony w podrozdziale 1.5 (używano również przytoczonych symboli zbiorów). Najpierw zostały omówione wyniki dotyczące liter oznaczających samogłoski, półsamogłoski oraz spółgłoski sonorne. Po nich przedstawiono wyniki odnoszące się do liter, diad i triad ortograficznych oznaczających spółgłoski właściwe dźwięczne. Następnie rezultaty związane z przekształcaniem struktur ortograficznych, które w przyjętej transkrypcji podstawowej oznaczają spółgłoski właściwe bezdźwięczne.

Niemal każda część w tym rozdziale dotyczy jednej litery lub jednej diady ortograficznej. W treści odniesiono się do transkrypcji podstawowej – wykonane analizy dotyczą przede wszystkim kontekstów, w których występują odstępstwa od tej transkrypcji [por.: Karaś, Madejowa 1977; Lubaś, Urbańczyk 1990; Nowakowski, Wiatrowski 2013]. Poza tym uwzględniono szczegóły, które są istotne z punktu widzenia potrzeb związanych z projektowaniem automatycznego systemu konwersji tekstu na transkrypcję. W wielu wypadkach analizy wykonane na podstawie reguł pochodzących z *Automatyzacji…* zostały uznane za wystarczające (dotyczy to przede wszystkim liter oznaczających samogłoski, półsamogłoski oraz spółgłoski sonorne).

W omówionych badaniach użyto słownika *SJP.pl*, który jest dostępny nieodpłatnie. Zawiera on niemal wszystkie formy fleksyjne języka polskiego (słownik jest stale aktualizowany). Niektóre formy użyte w analizach można uznać za hipotetyczne. To oznacza, że prawdopodobieństwo ich użycia jest niezwykle niskie lub że nigdy nie są używane. Jednak system transkrypcji powinien obejmować również takie wyrazy. Innym istotnym zasobem użytym w tych badaniach był korpus tekstowy. Pierwotnie powstał on na potrzeby projektu *Graphogame-Fluent*, który był poświęcony ocenie skuteczności komputerowych gier edukacyjnych w terapii trudności w czytaniu [Szczerbiński i in. 2012]. Korpus obejmuje ponad cztery miliony wyrazów. Zawiera wyłącznie teksty w języku polskim pochodzące z lektur dziecięcych i młodzieżowych.

2.2. Transkrypcja liter oznaczających samogłoski

2.2.1. Litery: *a, e, o, ó*

Litery *a, e, o, ó* powinny być transkrybowane jako fonemy samogłoskowe: /a/, /e/, /o/, /u/. Autorka *Automatyzacji…* wspomina jednak o regionalnym wariancie wymowy litery *a* w wyrazach typu: *tutaj, dzisiaj, daj*. W wypadku takiej wymowy litera *a* jest transkrybowana jako fonem /e/. Została również poruszona kwestia transkrypcji wygłosowej struktury *ej* jako /i.j/ lub /u.j/.

2.2.2. Litera *y*

Litera *y* najczęściej oznacza fonem samogłoskowy /y/. W *Automatyzacji…* można znaleźć wzmiankę dotyczącą wariantu wymowy wyrazu *cyjanek*, w którym fonem /y/ jest pomijany. Do takiej wymowy dostosowana jest tabela 9a w tej publikacji, a w szczególności reguła: O*c*[*y*]*j*=1. Analiza słownika *SJP.pl* wykazała, że reguła dotyczy ponad czterystu form fleksyjnych. Liczba ta obejmuje następujące formy podstawowe:

cyja, cyjan, cyjanamid, cyjanamidek, cyjanek, cyjanian, cyjaniany, cyjanid, cyjanina, cyjaninowy, cyjanit, cyjanizacja, cyjanizować, cyjankali, cyjankowy, cyjanobakteria, cyjanohydryna, cyjanokobalamina, cyjanokwas, cyjanooctowy, cyjanotypia, cyjanować, cyjanowodorowy, cyjanowodór, cyjanowy, cyjanoza, cyjanożelazian, cyjanożelazin, cyjon

Przy wymowie niektórych wyrazów z tej listy głoska [y] (AS) nie powinna być pomijana. Dotyczy to między innymi wyrazu *cyja* (instrument muzyczny) czy *cyjon* (zwierzę). Poza tym nie jest jasne, czy alternatywna wymowa obejmuje również wyrazy spokrewnione z wyrazem *cyjanek*.

Tabela 2.1 zawiera wynik analizy wykonanej na podstawie reguł pochodzących z tabeli 9 alfa w *Automatyzacji…* . Nie uwzględniono w niej alternatywnej wymowy wyrazu *cyjanek*. Zawiera ona trzy reguły, które nakazują przekształcanie litery *y* na fonem inny niż /y/, czyli w sposób inny, niż wynika to z przyjętej transkrypcji podstawowej.

Tabela 2.1 – Analiza słownika i korpusu dotycząca litery *y*

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korpusie	Przykładowe wyrazy z korpusu
1	O[*y*]A=/j/	445	*yardem, yachting, yakuza*	220	*York, Yorkshire, yard*
2	A[*y*]A=/j/	2185	*spraye, Toyota, layoutu*	549	*Boyem, Molaya, royal*
3	A[*y*](X–A–*j*)=/j/	1506	*Zamoyskiemu, Rayman, Maybachu*	914	*Bentley, Zamoyskim, Taylor*

Reguły w tej tabeli dotyczą wyrazów, których pisownia nie jest w pełni zgodna z regułami ortofonicznymi współczesnego języka polskiego. Uwaga dotyczy transkrybowania litery *y* jako fonem /j/. Taką wymowę można zaobserwować w wypadku:

– wyrazów obcych zachowujących obcą pisownię. Spotykane jest połączenie obcego rdzenia z polską końcówką fleksyjną (np.: *longplayowy, Playem, Maybachem, keyboardem, Disneylandem*). Od obcego rdzenia może pochodzić kilka wyrazów pokrewnych zachowujących pisownię oryginalną. Na przykład na podstawie nazwiska *Keynes* (*John Maynard Keynes*) powstały następujące formy, które są częściowo niezgodne z zasadami ortofonicznymi języka polskiego: *keynesista, keynesistowski, keynesizm, keynesowski, neokeynesizm*;
– wyrazów języka polskiego, które zachowały pisownię historyczną: *Domeyko, boyista* (zwolennik Boya-Żeleńskiego), *Reymont, reymontowski, boyowski, boyizm*. Kategoria obejmuje też niektóre nazwiska zgodne z pisownią historyczną: *Mayowski, Reykowski, Szweykowski* [por.: Doroszewski, Wieczorkiewicz 1947; Klemensiewicz 1930].

Wyrazy zachowujące pisownię obcą mogą zostać uznane za formy niepoprawne [por.: Dunaj 2001a, 2001b, 2006; Madejowa 1989, 1990; Wierzchowska 1971]. Projektując konkretny system transkrypcji, trzeba jednak mieć na uwadze przede wszystkim względy praktyczne oraz fakt, że takie wyrazy mogą wystąpić w tekstach. Użycie reguł z tabeli 2.1 (lub reguł do nich zbliżonych) można zatem uznać za czynność uzasadnioną.

2.2.3. Litera *ą*

W tej części zostały przytoczone z *Automatyzacji…* reguły związane z transkrypcją litery *ą*. Przedstawiono też wynik analizy fonostatystycznej dotyczącej różnych wariantów wymowy. Podstawowy wariant został zawarty w tabeli 2.2. Wymowa jest tu uzależniona tylko i wyłącznie od kontekstu następującego. Zgodnie z przytoczonymi regułami litera *ą* może być transkrybowana jako fonem /o/ (na przykład przed literą *ł*, jak w wyrazie *zepchnął*) lub jako ciąg fonemów: /o.m/ (na przykład przed literą *p*, jak w wyrazie *kąpałem*), /o.n/ (na przykład przed literą *t*, jak w wyrazie *początek*), /o.n'/ (na przykład przed literą *ć*, jak w wyrazie *zacząć*) oraz /o.w~/ (na przykład przed literą *ż*, jak w wyrazie *zdążyłem*).

Tabela 2.2 – Analiza słownika i korpusu dotycząca litery *ą* (część 1)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	X[*ą*]{*l,ł,m*}=/o/	6725	*zepchnął, prześliznął, spocząłem*	18715	*zajął, warknął, musnął*
2	X[*ą*]{*p,b*}=/o.m/	8830	*kąpałem, głąby, dostąpiłby*	2976	*stąpałem, ząb, postąpił*
3	X[*ą*]{*t,k,g*}=/o.n/	27130	*wątli, sprzątnęła, pałąkiem*	14040	*początek, zwierzątka, okrągłą*
4	X[*ą*]*c*(X–{*i,h*})=/o.n/	237818	*witającej, panująca, zwożących*	34962	*śpiąca, wracając, widniejącego*
5	X[*ą*]*d*(X–{*z,ź*})=/o.n/	9906	*użądlić, oglądam, żądamy*	11557	*lądować, żąda, zażądać*
6	X[*ą*]*dz*(X–*i*)=/o.n/	4137	*zasądzono, wrzeciądzu, osądzenia*	1515	*przyrządzać, sądzę, przyrządzał*
7	X[*ą*]{*ć,ci,dź,dzi*}=/o.n'/	3884	*kopnąć, przesiądzie, czmychnąć*	7017	*dotknąć, zacząć, rządzić*
8	X[*ą*]{*ś,ź,f,w,s,z,ż,ch*} =/o.w~/	26273	*chrabąszczem, brązowy, związkami*	11545	*zdążyłem, wąwozów, sąsiednich*
9	X[*ą*]O=/o.w~/	203743	*wyłupią, wyrównoważą, żebrowatą*	81653	*nocą, używaną, jaką*

W tabeli 2.3 przedstawiono reguły związane z alternatywną wymową litery *ą* umiejscowionej przed literami: *ś*, *ź* lub przed diadami: *si, zi*. Zgodnie z nimi litera *ą* przed spółgłoskami szczelinowymi miękkimi powinna być odczytywana jako dyftong [oĩ] (AS). Głoska [ĩ] (AS) należy do fonemu /n'/, zatem jego transkrypcja jest następująca: /o.n'/. Aby uzyskać taką transkrypcję, należy pominąć wiersz ósmy w tabeli 2.2 i jednocześnie uwzględnić wszystkie wiersze z tabeli 2.3.

Tabela 2.3 – Analiza słownika i korpusu dotycząca litery *ą* (część 2)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	X[*ą*]{*ś,ź,si,zi*}=/o.n'/	1947	*wąsiskami, rozwiązle, sąsiadami*	1528	*zasiąść, wąziutkich, kąśliwy*
2	X[*ą*]{*f,w,ż,ch*}=/o.w~/	9751	*dążyć, obwąchującą, fąfrom*	5945	*zdążyłem, wąwozów, cofnąwszy*
3	X[*ą*]{*s,z*}(X–*i*)=/o.w~/	14575	*grząskiej, uwiązań, zawiązałaś*	4080	*wąsy, brązowa, dziąsła*

Kolejny wariant transkrypcji litery *ą* został ujęty w tabeli 2.4. Możliwość ta również dotyczy litery *ą* umiejscowionej przed literami oznaczającymi spółgłoski szczelinowe miękkie, jednak uwzględniony został również kontekst poprzedzający. Zgodnie z tymi regułami rezonans nosowy miękki występuje tylko po literach oznaczających głoski miękkie oraz przed literami oznaczającymi spółgłoski szczelinowe miękkie (wiersz pierwszy w tabeli 2.4). Aby uzyskać odpowiedni zbiór reguł, należy pominąć wiersz ósmy w tabeli 2.2 i uwzględnić wszystkie wiersze z tabeli 2.4.

Tabela 2.4 – Analiza słownika i korpusu dotycząca litery *ą* (część 3)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	*i*[*ą*]{*ś,ź,si,zi*}=/o.n'/	68	*dosiąść, rozwiązli, Miąsika, przewiązie*	190	*usiąść, przysiąść, wysiąść*
2	*i*[*ą*]{*f,w,ż,ch*}=/o.w~/	2152	*wiążmy, spiąwszy, książce*	3046	*książę, wspiąwszy, obciążyli*
3	*i*[*ą*]{*s,z*}(X–*i*)=/o.w~/	5361	*wiązanym, ksiąski, piąstką*	1692	*przywiązane, obowiązki, dziąsła*
4	(X–*i*)[*ą*]{*ś,ź,si,zi*} =/o.w~/	1879	*gąsienicą, gałąź, kąśliwi*	1338	*sąsiednich, ukąsi, gałąź*
5	(X–*i*)[*ą*]{*f,w,ż,ch*} =/o.w~/	7599	*zdążymy, grążąc, zwąchana*	2899	*zdążyłem, drążą, ująwszy*
6	(X–*i*)[*ą*]{*s,z*}(X–*i*) =/o.w~/	9214	*śląskim, brązującą, wstrząsy*	2388	*wąsy, brązowa, gąszcze*

W *Automatyzacji…* przedstawiono również drugi wariant interpretacji fonologicznej odnoszacej się do litery *ą* (w tabeli 10 alfa). Był on rozważany w latach siedemdziesiątych jako jedna z możliwych transkrypcji tej litery. Obecnie nie jest brany pod uwagę, dlatego w niniejszej publikacji został pominięty. Wyniki najnowszych badań fonetycznych zostały omówione w monografii Anity Lorenc *Wymowa normatywna polskich samogłosek nosowych i spółgłoski bocznej* [Lorenc 2016]. Autorka wykazała, że na płaszczyźnie fonetycznej (głównie artykulacyjnej i akustycznej) struktura omawianych segmentów jest znacznie bardziej złożona i niejednolita niż w rozwiązaniach, które najczęściej przyjmuje się w fonologii.

2.2.4. Litera *ę*

W *Automatyzacji…* zbiór reguł dotyczących transkrypcji litery *ę* ma podobną strukturę jak zbiór reguł utworzonych dla litery *ą*. W tabeli 2.5 zostały przytoczone reguły podstawowe (na podstawie tabeli 11 w *Automatyzacji…*). Zgodnie z nimi transkrypcja litery *ę* jest niezależna od kontekstu lewostronnego – istotny jest tylko kontekst prawostronny. Z tabeli 2.5 wynika, że litera *ę* może być transkrybowana jako fonem /e/ (na przykład przed literą *ł*, jak w wyrazie *drgnęła*) lub jako ciąg fonemów: /e.m/ (na przykład przed literą *p*, jak w wyrazie *krępująca*), /e.n/ (na przykład przed literą *t*, jak w wyrazie *zdjęta*), /e.n'/ (na przykład przed diadą *ci*, jak w wyrazie *pięciobój*) oraz /e.w~/ (na przykład przed literą *ż*, jak w wyrazie *stężały*).

Tabela 2.5 – Analiza słownika i korpusu dotycząca litery *ę* (część 1)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	X[*ę*]{*l,ł*}=/e/	18266	*drgnęła, drasnęła, wyślizgnęlibyście*	12792	*uśmiechnęła, zerknęła, minęło*
2	X[*ę*]O=/e/	52888	*nagrodę, lawinę, cynkografię*	161104	*literę, imię, podłogę*
3	X[*ę*]{*p,b*}=/e.m/	24170	*odstępach, krępującą, zębatkę*	7980	*głęboki, wstępnego, wstępowali*
4	X[*ę*]{*t,k,g*}=/e.n/	61711	*zdjęto, błękitnemu, Węgry*	25679	*pękały, wewnętrznemu, pamiętasz*
5	X[*ę*]*c*(X–{*i,h*})=/e.n/	20157	*tysięcznej, skręcałby, książęcymi*	14167	*ręce, niezręczne, zajęczych*
6	X[*ę*]d(X–{*z,ź*})=/e.n/	11234	*powiędły, niezbędne, względów*	5871	*mędrzec, względnie, oszczędnością*
7	X[*ę*]*dz*(X–*i*)=/e.n/	11143	*odpędzonej, pędzlom, prędzej*	3805	*najprędzej, pieniędzy, wstędze*
8	X[*ę*]{*ć,ci,dź,dzi*}=/e.n'/	41069	*objęciach, pięciobój, podkręcił*	15330	*będzie, zajęcia, rozpoczęcia*
9	X[*ę*]{*ś,ź,f,w,s,z,ż,ch*} =/e.w~/	27711	*stężały, przezwyciężyli, księżycowa*	14804	*często, ciężar, zawęszył*

Istnieje możliwość alternatywnej interpretacji realizacji litery *ę* przed spółgłoskami szczelinowymi miękkimi. Do takiej wymowy dostosowane są reguły zamieszczone w tabeli 2.6. Zgodnie z nimi litera *ę* powinna być transkrybowana jako sekwencja fonemów /e.n'/. Aby uzyskać taką transkrypcję należy pominąć wiersz 9 w tabeli 2.5 i jednocześnie uwzględnić wszystkie wiersze z tabeli 2.6.

Tabela 2.6 – Analiza słownika i korpusu dotycząca litery *ę* (część 2)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	X[*ę*]{*ś,ź,si,zi*}=/e.n'/	6013	*potrzęsiono, zarzęźmy, rzęzić*	4916	*częściach, gałęzie, gęsi*
2	X[*ę*]{*f,w,ż,ch*}=/e.w~/	12072	*ciężki, pęcherzycę, tęchnąc*	5635	*potężnego, księżna, pęcherz*
3	X[*ę*]{*s,z*}(X–*i*)=/e.w~/	9626	*ugrzęzną, zwęszyłoby, zwycięskie*	4253	*często, klęski, jęzor*

Kolejny wariant wymowy litery *ę* został przedstawiony w tabeli 2.7 – uwzględniono w nim również kontekst poprzedzający. W tym wariancie rezonans miękki występuje tylko po literach oznaczających głoski miękkie oraz przed literami oznaczającymi spółgłoski szczelinowe miękkie (wiersz pierwszy tabeli 2.7). Aby uzyskać odpowiedni zbiór reguł transkrypcji, należy pominąć wiersz 9 w tabeli 2.5 i jednocześnie uwzględnić wszystkie wiersze z tabeli 2.7.

Tabela 2.7 – Analiza słownika i korpusu dotycząca litery *ę* (część 3)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	*i*[*ę*]{*ś,ź,si,zi*}=/e.n'/	1482	*uwięzień, mięsistych, więzi*	744	*pięści, uwięziony, mięsistego*
2	*i*[*ę*]{*f,w,ż,ch*}=/e.w~/	3611	*ciężar, półksiężyc, zwyciężył*	2987	*księżyca, ciężki, zaprzysiężeni*
3	*i*[*ę*]{*s,z*}(X–*i*)=/e.w~/	919	*mięsa, zwycięskiej, bezmięsnych*	749	*zwycięstwa, zwięzłą, mięsnego*
4	(X–*i*)[*ę*]{*ś,ź,si,zi*} =/e.w~/	4531	*trzyczęściowej, potrzęsiecie, zawęźliłyśmy*	4172	*częściach, szczęśliwie, gałęzią*
5	(X–*i*)[*ę*]{*f,w,ż,ch*} =/e.w~/	8461	*prężyły, natężania, sprężały*	2648	*potężnego, pęcherze, stęchły, wężyk*
6	(X–*i*)[*ę*]{*s,z*}(X–*i*) =/e.w~/	8707	*węzełkom, zgęstnął, tęskniłbym*	3504	*często, językowi, zwęszyć*

Podobnie jak w wypadku litery *ą* w *Automatyzacji…* przedstawiono również drugą możliwość dotyczącą interpretacji fonologicznej litery *ę* (w tabeli 11 alfa). Współcześnie nie jest ona brana pod uwagę, dlatego w niniejszej publikacji nie została uwzględniona.

2.2.5. Litera *u*

Litera *u*, która znajduje się w kontekście lewostronnym litery innej niż litera *a* oraz innej niż litera *e*, zawsze oznacza fonem /u/. Można przyjąć, że litera *u* w diadzie ortograficznej *eu* oznacza fonem /w/. M.S.-B. wskazuje wyjątki od tej reguły:

a) jeżeli *eu* stanowi fragment jednej z następujących sekwencji: *ieu* (np.: *nieubłagany, nieuchronny, nieuk*), *rzeu* (np.: *przeubogi, przeuroczy*), *eusz* (np.: *chudeusz, jubileuszowy*), *eus* (np.: *kaduceus, Zeus*), *eum#* (np.: *apogeum, liceum*). W tabeli 2.8 został przedstawiony wynik analizy słownika *SJP.pl*. Obejmuje ona wszystkie formy podstawowe wyrazów zawierających wymienione struktury ortograficzne.

Tabela 2.8 – Analiza słownika dotycząca diady ortograficznej *eu*

Lp.	Struktura ort.	Formy podstawowe w słowniku *SJP.pl* zawierające strukturę z kolumny drugiej
1	*ieu*	*najnieuczciwiej, najnieuczciwszy, najnieudolniej, najnieudolniejszy, najnieuprzejmiej, najnieuprzejmiejszy, najnieużyteczniej, nieubłagany, nieuchronność, nieuchronny, nieuchwytność, nieuciążliwość, nieuctwo, nieuczciwiej, nieuczciwość, nieuczciwy, nieuczciwszy, nieuczony, nieuczynność, nieudacznica, nieudacznictwo, nieudacznik, nieudaczność, nieudaczny, nieudatność, nieudawanie, nieudolniej, nieudolniejszy, nieudolność, nieuległy, nieudolny, nieufność, nieufny, nieugiętość, nieugięty, nieuk, nieukładność, nieukontentowanie, nieuleczalność, nieuległość, nieulękły, nieułomek, nieumiarkowanie, nieumiarkowany, nieumiejętność, nieumyślność, nieuniknioność, nieunikniony, nieuprzejmiej, nieuprzejmiejszy, nieuprzejmość, nieuprzejmy, nieurodzaj, nieurodzajność, nieurodziwość, nieurzędowość, nieustający, nieustanny, nieustawność, nieustępliwość, nieustraszoność, nieustraszony, nieusuwalność, nieutrudzenie, nieutrudzony, nieuwaga, nieuważny, nieużyteczniej, nieużyteczność, nieużyteczny, nieużytek, nieużytkowość, nieużyty, nieużywalność*
2	*rzeu*	*przeuczać, przeucztować, przeuczyć, przeuroczy*
3	*eusz*	*Aggeusz, Amadeusz, Amadeuszostwo, Amadeuszowie, Amadeuszowy, Anteusz, Anteuszowy, ateusz, ateuszka, ateuszostwo, ateuszowski, Atreusz, Bartymeusz, Boromeusz, boromeuszka, Cefeusz, chudeusz, Doroteusz, Doroteuszostwo, Doroteuszowie, Doroteuszowy, dziadeusz, Egeusz, Egeuszowy, Elizeusz, Elizeuszostwo, Elizeuszowie, Elizeuszowy, Epimeteusz, Erechteusz, faryzeusz, faryzeuszostwo, faryzeuszowski, Galileusz, Hasmoneusze, Ireneusz, Ireneuszostwo, Ireneuszowie, Ireneuszowy, jubileusz, jubileuszować, jubileuszowy, kaduceusz, Kondeusz, Kondeusze, Kondeuszowie, koryfeusz, koryfeuszowy, Linneusz, Machabeusz, Machabeusze, Manicheusz, Manicheuszowy, Mateusz, Mateuszczyk, Mateuszek, Mateuszostwo, Mateuszowie, Mateuszowy, Medyceusze, Morfeusz, Odyseusz, Orfeusz, orfeuszowy, Perseusz, Prometeusz, prometeuszowy, Proteusz, proteuszowy, Ptolemeusz, ptolemeuszowski, ptolemeuszowy, saduceusz, skarabeusz, słabeusz, srebrnojubileuszowy, Tadeusz, Tadeuszak, Tadeuszek, Tadeuszostwo, Tadeuszowie, Tadeuszowo, Tadeuszowy, Tezeusz, Tezeuszowy, Tymoteusz, Tymoteuszostwo, Tymoteuszowie, Tymoteuszowy, Tyrteusz, ureusz, Zacheusz, Zacheuszek, Zacheuszka, Zacheuszkowy, Zacheuszostwo, Zacheuszowie, Zacheuszowy, Zebedeusz*
4	*eus*	*archaeus, aureus, balteus, basileus, bazyleus, cereus, echinocereus, Erechteus, ileus, koleus, Nereus, Peleus, pileus, Pireus, proteus, ureus, aureus, Zagreus, Zeus*
5	*eum#*	*apogeum, ateneum, petroleum, Colosseum, dacholeum, empireum, gineceum, gumoleum, gyneceum, hipogeum, hypogeum, karbolineum, kastoreum, koloseum, konopeum, liceum, linoleum, mauzoleum, mitreum, motomuzeum, muzeum, nimfeum, nitroleum, oleum, Ossolineum, panaceum, perigeum, periosteum, perygeum, petroleum, trofeum, winileum, winyleum*

b) we wszystkich formach wyrazów: *Pentateuch, reunion, teurgia.*

Diada *au* powinna być transkrybowana jako: /a.w/ z wyjątkiem następujących wypadków:

a) jeżeli jest ona częścią jednej ze struktur: *pozau* (np. *pozauczelniany*), *prau* (np. *praugrofiński*), *dnau* (np. *przednaukowy*), *unau* (np. *unaukowić*), *enau* (np. *nienaukowy*), *ezau* (np. *niezauważalny*), *onau* (np. *popularnonaukowy*), *anau* (np. *pozanaukowy*), *ynau* (np. *antynaukowy*). W tabeli 2.9 ujęto informacje o formach podstawowych wyrazów zawierających wymienione struktury. Z tych danych wynika, że przytoczone zasady wymagają uzupełnienia – diada *au* w niektórych wyrazach obcych powinna być transkrybowana jako sekwencja fonemów /a.w/ pomimo obecności jednego z wymienionych ciągów ortograficznych (wyrazy przekreślono). Na podstawie tych informacji można utworzyć listę umożliwiającą identyfikację wszystkich form fleksyjnych zawierających jedną z wymienionych struktur ortograficznych i jednocześnie zawierających diadę *au*, która powinna być transkrybowana jako /a.w/ (można powiedzieć, że są to wyjątki od wyjątków). Nagłosowy fragment umożliwiający identyfikację wyjątków zawierających ciąg *pozau* to: *#metopozaur*. W wypadku sekwencji *enau* identyfikacja wyjątków jest możliwa na podstawie nagłosowego ciągu: *#adenauer, #birkenau*. W wypadku sekwencji *ezau* identyfikacja form stanowiących wyjątki odbywa się na podstawie nagłosowych struktur: *#stezaury, #tezauru*. Natomiast w wypadku sekwencji *onau* dla identyfikacji wyjątków może być użyty następujący zbiór: *#aeronaut, #argonaut, #astroaeronaut, #astronaut, #bioastronaut, #biokosmonaut, #hiponaut, #infonaut, #kosmonaut, #lunonaut, #paleoastronautyczny, #paleoastronaut, #selenonaut, #taikonaut, #ufonaut*. Nagłosowa sekwencja umożliwiająca identyfikację wyjątków dla ciągu *anau* to: *#akwanaut*.

Tabela 2.9 – Analiza słownika dotycząca diady ortograficznej *au*

Lp.	Struktura ort.	Formy podstawowe w słowniku *SJP.pl* zawierające strukturę z kolumny drugiej
1	*pozau*	~~*metopozaury*~~, *pozauczelniany, pozaukładowy, pozaultrakrótki, pozaumowny, pozauniwersytecki, pozauranowy, pozaurlopowy, pozaurzędowy, pozaustawowy, pozaustny, pozaustrojowy, pozautylitarny*
2	*prau*	*praugrofiński, prausta*
3	*dnau*	*przednaukowy*
4	*unau*	*unaukowiać, unaukowić*
5	*enau*	~~*Adenauer*~~, ~~*Birkenau*~~, *nienaukowość, przenaukowienie*
6	*ezau*	*niezauważalność, niezauważenie,* ~~*stezauryzować*~~, ~~*tezaurus*~~, ~~*tezauryzacja*~~, ~~*tezauryzacyjny*~~, ~~*tezauryzować*~~
7	*onau*	~~*aeronauta*~~, ~~*aeronautka*~~, ~~*aeronautyczny*~~, ~~*aeronautyka*~~, ~~*argonauta*~~, ~~*astroaeronautyka*~~, ~~*astronauta*~~, ~~*astronautka*~~, ~~*astronautyczny*~~, ~~*astronautyka*~~, ~~*bioastronauta*~~, ~~*bioastronautyczny*~~, ~~*bioastronautyka*~~, ~~*biokosmonautyka*~~, *fantastycznonaukowy,* ~~*hiponautyka*~~, ~~*infonauta*~~, ~~*kosmonauta*~~, ~~*kosmonautka*~~, ~~*kosmonautyczny*~~, ~~*kosmonautyka*~~, *literackonaukowy,* ~~*lunonauta*~~, ~~*paleoastronautyczny*~~, ~~*paleoastronautyk*~~, ~~*paleoastronautyka*~~, *ponauczać, popularnonaukowy, pseudonauka, pseudonaukowiec, pseudonaukowość, pseudonaukowy,* ~~*selenonauta*~~, ~~*selenonautyczny*~~, ~~*selenonautyka*~~, ~~*taikonauta*~~, *teoretycznonaukowy,* ~~*ufonauta*~~
8	*anau*	~~*akwanauta*~~, ~~*akwanautka*~~, ~~*akwanautyczny*~~, ~~*akwanautyka*~~, *metanauka, paranauka, paranaukowiec, paranaukowy, pozanaukowy*
9	*ynau*	*antynaukowość, antynaukowy*

b) wszystkie formy fleksyjne wyrazów rozpoczynających się od *#nau* (np.: *naubliżać, nauka, nausznik*). Autorka *Automatyzacji…* podała następujące wyjątki: *nauplius, Nautilus, nautofon, nautologia, nautologiczny, nautyczny, nautyka* (zatem można stwierdzić, że lista również obejmuje wyjątki od wyjątków). Można ją rozszerzyć o kolejne wyrazy: *naumachia, Nauranka, Naurańczyk, naurański, Nauru, nautykwariat, Nauzykaa*. Poniższa lista została utworzona na podstawie słownika *SJP.pl* i obejmuje wszystkie formy podstawowe wyrazów zawierających triadę *#nau*, w ramach której diada *au* powinna być transkrybowana jako /a.u/:

> *naukowy, nauka, naukowiec, nauczycielski, naubijać, naubliżać, nauczać, nauczanie, nauczka, nauczyciel, nauczycielka, nauczycielowy, nauczycielski, nauczycielstwo, nauczyć, nauganiać, naujadać, naukawy, naukładać, naukometria, naukoznawca, naukoznawczy, naukoznawczyni, naukoznawstwo, naumyślny, naurągać, naurządzać, naustawiać, nausznik, nauszny, nautykać, nautyskiwać, nauwijać, naużerać, naużywać*

c) w wyrazach o strukturze: *#zau*(X–*t*) literze *u* odpowiada fonem /u/ – zasada dotyczy między innymi następujących wyrazów: *zaufać, zaułek, zauroczyć*, ale nie obejmuje form fleksyjnych wyrazów: *zautomatyzować* i *zautonomizować*.
d) fonem /u/ odpowiada literze *u* w formach fleksyjnych wyrazów: *laurka, laurkowy, asauł, auł, esauł, saksauł*.

2.2.6. Litera *i*

Opracowanie dotyczące litery *i* oparto na części siódmej w rozdziale VII w *Automatyzacji…* . Zgodnie z przyjętą transkrypcją podstawową literze *i* odpowiada fonem /i/. Litera *i* ma wyjątkowe znaczenie w systemie ortofonicznym języka polskiego ze względu na swoją wielofunkcyjność. Dlatego opracowanie w tej części obejmuje zarówno reguły niezgodne z przyjętą transkrypcja podstawową, jak i reguły, które nakazują przekształcanie tej litery w fonem /i/. W dalszym ciągu w sposób usystematyzowany przedstawiono informacje dotyczące transkrypcji litery *i*:

a) litera *i* w wyrazach typu: *stoi, naiwny, naigrywać* oznacza fonem /i/. W tabeli 2.10 został przedstawiony wynik analizy fonostatystycznej wykonanej na podstawie słownika *SJP.pl* oraz na podstawie korpusu. Wyszczególniono litery oznaczające samogłoski, które są umiejscowione w kontekście poprzedzającym literę *i*. Jest to interpretacja właściwa dla wymowy starannej. W wymowie nieco mniej starannej niektórych wyrazów (np.: *dwói, trói, Judei*) w wygłosie przed głoską [i] (AS) może pojawiać się dodatkowo [i̯] niesylabiczne.

Tabela 2.10 – Analiza słownika i korpusu dotycząca litery *i* (część 1)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	*a*[*i*]=/i/	11425	*zainkasowały, naiwniutki, wyczaiły*	2051	*rozmaite, krainie, odzwyczai*
2	*e*[*i*]=/i/	20181	*nieizolowanej, Judei, skleiłaby*	943	*nadziei, odkleił, przeinaczyć*
3	*o*[*i*]=/i/	16578	*samoistnie, spoinowaniu, spoiłyby*	6064	*swoich, uspokoiła, zagoi*
4	*ó*[*i*]=/i/	3	*dwói, samobói, trói*	9	*dwói*
5	*u*[*i*]=/i/	3245	*fluidów, dwuizbowy, uiszczałbym*	511	*ruinę, wiekuisty, genui*
6	*y*[*i*]=/i/	1829	*wyizolowali, antyimplozyjny, długoszyi*	358	*czyichkolwiek, czyichś, wyimaginowane*
7	*ę*[*i*]=/i/	0		0	
8	*ą*[*i*]=/i/	0		0	

b) w wyrazach zawierających konstrukcję: {*p, b, f, w, m, t, d, h, l, r*}*i*(A–*i*) literze *i* odpowiada fonem /j/ (np.: *bariera, miliard, tiara*). W ramach podanej struktury w kontekście lewostronnym litery *i* najczęściej występują litery: *w, p, b, m*. Najrzadziej w tym kontekście są spotykane litery: *h, t, f.* Wynik analizy został przedstawiony w tabeli 2.11.

Tabela 2.11 – Analiza słownika i korpusu dotycząca litery *i* (część 2)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	*p*[*i*](A–*i*)=/j/	88877	*zalepiającym, pierwszeństwu, powypiekam*	33746	*trzypiętrowy, śpiewu, wspiąć*
2	*b*[*i*](A–*i*)=/j/	66930	*podbiegła, biurokratycznym, wyżłabiali*	32270	*obiad, biały, biegaliśmy*
3	*f*[*i*](A–*i*)=/j/	8691	*faktografię, kalafiorowy, fiołkiem*	1610	*ofiarowaną, kaligrafia, fioletowa*
4	*w*[*i*](A–*i*)=/j/	182872	*przestawią, wywiązują, odwiedzała*	111069	*powiada, świata, uczniowie*
5	*m*[*i*](A–*i*)=/j/	91991	*wyszumiały, zmiękczana, mierząc*	58027	*akademia, kamienie, śmiechem*
6	*t*[*i*](A–*i*)=/j/	6226	*kwestionujmy, bestialskiego, Koryntian*	1660	*portier, apatię, partia*
7	*d*[*i*](A–*i*)=/j/	18715	*diagnostyczną, studiach, diamentowy*	2502	*melodię, Grenlandia, radioaktywne*
8	*h*[*i*](A–*i*)=/j/	2811	*hierarchią, psychiatrze, hienach*	97	*monarchię, hieroglify, melchior*
9	*l*[*i*](A–*i*)=/j/	13493	*taliom, Natalia, foliowaną*	1561	*emaliowany, biblioteka, waniliowe*
10	*r*[*i*](A–*i*)=/j/	24379	*wariackie, periodyzował, solaria*	3934	*historię, zorientować, seminarium*

W wyrazach zawierających w wygłosie strukturę: {*p, b, f, w, m, t, d, h, l, r*}*ii* (np.: *kopii, komedii, monarchii*) pierwszej literze *i* może odpowiadać fonem /j/ lub fonem /i/ – autorzy, na których powołuje się M.S.-B. dopuszczają obie możliwości transkrypcji. W tabeli 2.12 przedstawiono wynik analizy dotyczącej tych wersji.

Tabela 2.12 – Analiza słownika i korpusu dotycząca litery *i* (część 3)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	*p*[*i*]*i*#=/i/ lub *p*[*i*]*i*#=/j/	298	*aeroterapii, homotopii, monotypii*	10	*kopii, rupii, Etiopii*
2	*b*[*i*]*i*#=/i/ lub *b*[*i*]*i*#=/j/	91	*Zambii, cyberfobii, aerofobii*	15	*ksenofobii, Arabii, Nubii*
3	*f*[*i*]*i*#=/i/ lub *f*[*i*]*i*#=/j/	377	*chorografii, cellografii, bibliozofii*	179	*geografii, ortografii, fotografii*
4	*w*[*i*]*i*#=/i/ lub *w*[*i*]*i*#=/j/	26	*Oliwii, Oktawii, Sylwii*	9	*Boliwii, szałwii, Skandynawii*
5	*m*[*i*]*i*#=/i/ lub *m*[*i*]*i*#=/j/	384	*agronomii, autotomii, pandemii*	179	*akademii, armii, chemii*
6	*t*[*i*]*i*#=/i/ lub *t*[*i*]*i*#=/j/	126	*amnestii, kwestii, perypetii*	102	*partii, bestii, kwestii*
7	*d*[*i*]*i*#=/i/ lub *d*[*i*]*i*#=/j/	110	*gwardii, Irlandii, Arkadii*	231	*Grenlandii, tragedii, encyklopedii*
8	*h*[*i*]*i*#=/i/ lub *h*[*i*]*i*#=/j/	51	*anarchii, marchii, oligarchii*	8	*monarchii, hierarchii, Lechii*
9	*l*[*i*]*i*#=/i/ lub *l*[*i*]*i*#=/j/	259	*Marsylii, Anglii, Kornelii*	549	*Anglii, Australii, melancholii*
10	*r*[*i*]*i*#=/i/ lub *r*[*i*]*i*#=/j/	667	*bakterii, biżuterii, pirometrii*	736	*historii, teorii, Syberii*

c) litera *i* we wszystkich wyrazach pochodnych od: *ekspiacja, kwietyzm, klient, miastenia, patriarcha, patriota* oznacza sekwencję fonemów /i.j/. Warto zauważyć, że taki zapis nie jest zgodny z regułami przytoczonymi w punkcie b (tabela 2.11). Transkrypcję wymienionych wyrazów można zatem uznać za wyjątkową względem transkrypcji wynikającej z tabeli 2.11.

d) w wyrazach zawierających strukturę: {*#,a*}{*t,p*}*ri*A (np.: *trio, triumf, priorytet, patriota*) literze *i* odpowiada sekwencja fonemów /i.j/. W tabeli 2.13 ujęto wynik analizy, w której powyższą strukturę zastąpiono czterema konstrukcjami szczegółowymi. Dokładna analiza słownika wykazała, że omawiany sposób transkrypcji nie może być stosowany w wypadku wygłosu wyrazów (np.: *foniatrią, psychiatrię, pediatria*). Dlatego dwie ostatnie reguły w tabeli zostały zmodyfikowane.

Tabela 2.13 – Analiza słownika i korpusu dotycząca litery *i* (część 4)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	#*tr*[*i*]A=/i.j/	673	*trialistyczny, triada, triangiel*	190	*triumf, triest, triumfująco*
2	#*pr*[*i*]A=/i.j/	175	*priorytet, Priam, priorytetowy*	0	
3	*atr*[*i*]A(X–#)=/i.j/	1224	*patriarchalnym, repatriant, patriotko*	27	*patriotyczny, patriarcha, patriotyzmu*
4	*apr*[*i*]A(X–#)=/i.j/	84	*aprioryczność, apriorystycznym, aprioryzmy*	0	

e) w formacjach: *biandria, miliamper, poliester* litera *i* oznacza fonem /i/. Identyfikacja tego typu wyrazów może być problematyczna. Nie można zastosować reguł: O*b*[*i*]A=/i/, O*mil*[*i*]A=/i/, O*pol*[*i*]A=/i/, ponieważ odnoszą się one również do wyrazów, w których litera *i* powinna być transkrybowana jako fonem /j/. W dalszym ciągu przedstawiono informacje, które dotyczą transkrypcji wyrazów rozpoczynających się od prefiksów: *bi, mili, poli* oraz *mini*.

Wyrazy ortograficzne, które zawierają w nagłosie ciąg pokrywający się ze strukturą: #*bi*A, występują w 906 formach podstawowych. Lista nie została tutaj przytoczona, ponieważ jej analiza wykazała, że tylko w wypadku wyrazu *biandria* litera *i* powinna być transkrybowana jako fonem /i/. W pozostałych wyrazach z tego zbioru litera *i* oznacza [i̯] (AS) niesylabiczne, a w konsekwencji fonem /j/ (np.: *biały, biografia, biuro*). Wszystkie formy fleksyjne wyrazu *biandria* można zidentyfikować na podstawie nagłosowej sekwencji #*biandr*.

W zbiorze wyrazów mających nagłos o strukturze: #*mili*A, występują tylko dwie formy podstawowe (*miliamper* oraz *miliamperomierz*), w których druga litera *i* powinna być transkrybowana jako fonem /i/.

W pozostałych wypadkach oznacza ona fonem /j/ (np.: *milion, milioner, miliard*). Wszystkie formy fleksyjne wyrazów: *miliamper* oraz *miliamperomierz* można zidentyfikować na podstawie sekwencji: *#miliamp*.
Stosunkowo liczny jest zbiór wyrazów zawierających w nagłosie strukturę *#poli*A, w której litera *i* powinna być transkrybowana jako fonem /i/. Poniższa lista obejmuje formy podstawowe takich wyrazów:

poliacetal, poliaddukt, poliaddycja, poliakrylan, poliakrylonitryl, poliakrylonitrylowy, poliakrylowy, polialkohol, poliamid, poliamidowy, poliandria, poliandryczny, polianit, poliarchia, poliekran, polielektrolit, poliembrionia, polien, polienergida, polienowy, poliera, poliester, polioctan, poliestezja, poliestralny, poliestrowy, polieter, polietylen, polietylenowy, poliimid, poliiterofilia, poliizobutylen, poliizopren, poliuretan, poliuretanowy, poliuria, poliuronidy

W słowniku *SJP.pl* występują tylko dwie formy podstawowe z omawianą strukturą nagłosu, w których litera *i* oznacza fonem /j/: *poliowirus, polioza*. Na podstawie tych informacji można wyznaczyć nagłosowe sekwencje umożliwiające identyfikację wyrazów, w których litera *i* zawsze oznacza fonem /i/: *#polia, #polie, #polii, #poliu, #polioc*. W nagłosowych strukturach: *#poliow, #polioz* litera *i* zawsze powinna być transkrybowana jako fonem /j/.
Dalszy ciąg opisu dotyczy zbioru wyrazów zawierających w nagłosie strukturę: *#mini*A. Przedrostek *mini* cechuje się znaczną produktywnością słowotwórczą. Występuje on przed samogłoską w następujących formach podstawowych:

minialoha, miniankieta, miniaparat, miniaudycja, miniautobus, minielektrownia, miniesej, miniopowiadanie, miniosiedle

Listę łatwo można powiększyć o nowe pozycje (np.: *miniauto, miniantena, miniukład* itp.). Druga litera *i* w ramach przedrostka *mini* powinna być transkrybowana jako fonem /i/. Słownik *SJP.pl* zawiera również stosunkowo dużą liczbę form podstawowych ze strukturą *#mini*A, w których druga litera *i* jest tylko znakiem miękkości:

minia, miniator, miniatorski, miniatorstwo, miniatura, miniaturka, miniaturować, miniaturowość, miniaturowy, miniaturyzacja, miniaturyzacyjny, miniaturyzować, miniaturzysta, miniaturzystka, minierski, minięcie, minione, miniony, miniować, miniowy, miniówa, miniówka

Na podstawie podanych informacji można wyznaczyć nagłosowe struktury istotne dla transkrypcji. Można przyjąć, że w wyrazach zawierających w nagłosie strukturę *#mini*A druga litera *i* powinna być transkrybowana jako fonem /i/. Wyjątki stanowią wszystkie formy fleksyjne zawierające w nagłosie jedną z sekwencji: *#miniator, #miniatur, #minier, #minięc, #minion, #miniow, #miniów.* Lista wyjątków obejmuje też całe formy: *#minięć#, #minier#, #minia#, #minią#.*

f) według *Automatyzacji…* w wyrazach zawierających strukturę {*s,c,z*}*i*A literze *i* odpowiada najczęściej słowo puste (np.: *siano, ziemia*). Wyjątek od tej reguły stanowią wyrazy: *sial, scientyfizm* (współcześnie używana jest forma ortograficzna: *scjentyfizm*). W tabeli 2.14 zawarto reguły dotyczące litery *i* w podanym kontekście (z kontekstu prawostronnego wykluczono drugą literę *i*, pomimo że takie konstrukcje nie występują).

Tabela 2.14 – Analiza słownika i korpusu dotycząca litery *i* (część 5)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	*s*[*i*](A–*i*)=1	64152	*łosiem, księżyc, Sieradzkiego*	127810	*posiada, miesiąc, sąsiednie*
2	*c*[*i*](A–*i*)=1	299412	*odległościami, napięciowej, niecierpiącego*	84288	*dwanaście, wreszcie, pociąg*
3	(X–*d*)*z*[*i*](A–*i*)=1	21863	*ziarnowemu, zapyziałą, uziemionym*	12570	*wziął, gałęziami, jeziorko*

g) litera *i* występująca w strukturze *ni*(A–*i*) w wyrazach rodzimych języka polskiego jest tylko znakiem miękkości (np.: *korzenia, warzenia, zrozumienia*). Zasada nie obejmuje wyrazów obcych, które w mianowniku

liczby pojedynczej są zakończone jedną z następujących struktur: *nia#* (np.: *mania, linia*), *nior#* (np. *monsinior*), *nion#* (np.: *kanion, kompanion*), *nium#* (np.: *milenium, aluminium*), *nian#* (np.: *cynian, cytrynian*), *nianin#* (np. *fenianin*). Zgodnie z *Automatyzacją…* takie wyrazy mają wymowę asynchroniczną, zatem litera *i* oznacza fonem /j/.

Biorąc pod uwagę zmiany zachodzące we współczesnej wymowie, zagadnienie to nie jest jednoznaczne. Część wyrazów wykazuje silne przystosowanie do reguł ortofonicznych polszczyzny, co sprawia że diada *ni* jest wymawiana synchronicznie. Tendencja ta jest zauważalna między innymi w wypadku nazw miejsc (np.: *sortownia, elektrownia, serwerownia, emaliernia*). Trudno rozstrzygnąć, jakie czynniki decydują o silnej albo słabej adaptacji fonetycznej tych wyrazów do języka polskiego (w młodym pokoleniu daje się zauważyć tendencja do wymowy synchronicznej). Poza tym wielu wyrazów obcych nigdy lub prawie nigdy nie używa się. Jest to kolejny czynnik, który sprawia, że rzeczywista ich wymowa jest trudna do oszacowania.

Zagadnienia omówione w poprzednim akapicie wymagają dodatkowych badań lingwistycznych, natomiast dla celów praktycznych można przyjmować rozwiązania o charakterze umownym i tymczasowym. Przy tworzeniu reguł transkrypcji uwidacznia się kolejny problem, który dotyczy właściwej identyfikacji niektórych wyrazów. Dotyczy on w szczególności wyrazu *mania*, który ma wymowę asynchroniczną i cechuje się znaczną produktywnością słowotwórczą. Stanowi on drugi człon w złożeniach: *anglomania, arytmomania, bibliomania, dekalkomania, dromomania, dypsomania, dipsomania, erotomania, farmakomania, gigantomania, grafomania, kalkomania, kinomania, kleptomania, megalomania, melomania, mikromania, mitomania, monomania, narkomania, opiomania, piromania, poriomania, sparmania, teatromania, telemania, toksykomania, tytułomania, wideomania*. Poza tym występują formy pochodne wyrazu *mania*: *maniactwo*, *maniakalny*. Sam wyraz jest odmienny: *manią, mianiami, manię* itp. Jednocześnie występują liczne wyrazy rodzime w języku polskim, które zawierają ciąg *mania* i które mają wymowę synchroniczną, np.: *rozłamania, podtrzymania, utrzymania, przetrzymania*.

Mając na uwadze omówione problemy, można wskazać rozwiązanie polegające na uwzględnianiu w regułach odpowiednio długich nagłosowych sekwencji pozwalających na identyfikację odpowiednich wyrazów i ich form, np.: *#grafomani, #kalkomani, #kinomani, #maniactw, #maniakaln*. Należy wykluczyć formy zakończone na *manij#*. Natomiast trzeba uwzględnić w całości formy fleksyjne wyrazu *mania*.

W wyniku analizy wykonanej przy użyciu *Słownika wyrazów obcych* PWN [Bańko 2003] uzyskano kilka dłuższych struktur zawierających triadę *nia.* Można przyjąć, że w strukturach tych litera *i* powinna być transkrybowana jako fonem /j/. W każdym wypadku identyfikacja wszystkich form fleksyjnych może odbywać się na podstawie kilku sekwencji umiejscowionych w wygłosie:

- *fonia#: fonie#, foniach#, foniami#, foniom#, fonią#, fonię#, fonii#, fonio#* (występuje w wyrazach: *afonia, ambiofonia, apofonia, diafonia, dodekafonia, dysfonia, elektrofonia, eufonia, frankofonia, heterofonia, homofonia, iluminofonia, kakofonia, kalafonia, kwadrofonia, monofonia, polifonia, radiofonia, radiotelefonia, stereofonia, symfonia, telefonia, wideofonia*);
- *frenia#: frenie#, freniach#, freniami#, freniom#, frenią#, frenię#, frenii#, frenio#* (występuje w wyrazach: *cyklofrenia, hebefrenia, parafrenia, tachyfrenia, oligofrenia, schizofrenia*);
- *genia#: genie#, geniach#, geniami#, geniom#, genią#, genię#, genii#, genio#* (występuje w wyrazach: *bergenia, embriogenia, filogenia, glottogenia, heterogenia, homogenia, malagenia, ontogenia, teratogenia*); wyjątek: *Genia* (zdrobnienie imienia);
- *gonia#: gonie#, goniach#, goniami#, goniom#, gonią#, gonię#, gonii#, gonio#* (występuje w wyrazach: *dragonia, glottogonia, kosmogonia, monogonia, schizogonia, sporogonia, teogonia, wigonia*); wyjątki: *Pigonia, Świrgonia, Świgonia* (nazwiska w dopełniaczu);
- *tonia#: tonie#, toniach#, toniami#, toniom#, tonią#, tonię#, tonii#, tonio#* (występuje w wyrazach: *ceratonia, chirotonia, darlingtonia, dystonia, fitonia, hipertonia, hipotonia, kajtonia, katatonia, liwistonia, monotonia, sympatykotonia, wagotonia*); wyjątek: *Świtonia* (nazwisko w dopełniaczu);
- *linia#: linie#, liniach#, liniami#, liniom#, linią#, linię#, linii#, linio#* (występuje w wyrazach: *infolinia, interlinia, paulinia, waterlinia*).

Na ich podstawie nie da się jednak zidentyfikować wszystkich wyrazów obcych zawierających triadę *nia* w wygłosie. Wyrazy takie (inne niż już omówione) zostały wymienione w pierwszym wierszu w tabeli 2.15. Wykaz obejmuje również wyrazy, które z dużą dozą prawdopodobieństwa wymawiane są w sposób synchroniczny (przekreślono je). Zatem identyfikacja form fleksyjnych, w których ortograficzna diada *ni* oznacza sekwencję fonemów /n.j/ możliwa jest na podstawie odpowiednio długich fragmentów, np.: *#alpini, #harmoni, #diachroni, #dysharmoni* (z wyłączeniem form zakończonych na *nij#*).

Analiza wykazała, że struktury: *nium#, nion#* występują wyłącznie w wyrazach obcych i mają wymowę asynchroniczną. Sekwencja *nior#* również ma wymowę asynchroniczną, jednak istnieje niewielka liczba wyrazów z tym sufiksem. Z dużym prawdopodobieństwem większość z nich współcześnie jest wymawiana w sposób synchroniczny, np.: *junior, senior, Pinior* i *Konior* (nazwiska). Wygłosowa struktura *nian#* występuje zarówno w wyrazach rodzimych, jak i w wyrazach obcych. Drugi wiersz w tabeli 2.15 zawiera wykaz wyrazów obcych, w których litera *i* powinna być transkrybowana jako fonem /j/. W wierszu trzecim w tej samej tabeli zostały wyszczególnione wyrazy obce zawierające sekwencję *nianin#*. Dla wyrazów z drugiego oraz z trzeciego wiersza tabeli 2.15 można przyjąć wymowę asynchroniczną. Ich identyfikacja jest możliwa na podstawie odpowiednich fragmentów występujących we wszystkich formach fleksyjnych.

Tabela 2.15 – Analiza słownika i korpusu dotycząca litery *i* (część 6)

Lp.	Struktura wygłosu	Wykaz wyrazów obcych
1	*nia#*	*akrania, alpinia, ~~amplifikatornia~~, androginia, androgynia, anhedonia, ~~antykwarnia~~, astenia, awicenia, baronia, bonia, brekinia, ~~brojlernia~~, chrystofania, cynia, cyrkonia, diachronia, diakonia, dysharmonia, ~~dyspozytornia~~, eichhornia, ~~elektrownia~~, ~~emaliernia~~, epifania, eudajmonia, eudemonia, felonia, filharmonia, fisharmonia, georginia, gisernia, glicynia, gloksynia, ~~gubernia~~, hamernia, harmonia, heterochronia, hierofania, hipersomnia, holendernia, ~~hydroelektrownia~~, ~~hydrofornia~~, ~~inhalatornia~~, ~~inkubatornia~~, ~~introligatornia~~, ironia, kalumnia, kampania, kanonia, kapitania, ~~kasztelania~~, kierznia, klimenia, kolonia, ~~kolokolnia~~, kompania, konfraternia, ~~kostiumernia~~, omnia, latania, leukopenia, lewkonia, litania, litofania, ludwisarnia, luśnia, mahonia, ~~masarnia~~, miastenia, minia, mizoginia, monoginia, nenia, neotenia, neurastenia, omnia, onania, petunia, pinia, pirania, pneumonia, Polihymnia, Polonia, ~~probiernia~~, proteroginia, protoginia, ~~radiolatarnia~~, ~~rekwizytornia~~, rodymenia, ~~rusznikarnia~~, salwinia, saracenia, ~~serwerownia~~, ~~sortownia~~, sufragania, surfinia, symonia, synchronia, szambelania, ~~szulernia~~, teofania, traktiernia, trinia, trombocytopenia, turnia, tuwalnia, tyrania, unia, urania, ~~waltornia~~, wikunia, Wirginia, ~~zecernia~~*
2	*nian#*	*adypinian, alginian, aminofosfonian, amonian, antymonian, arsenian, banian, cyjanian, cyklaminian, cynamonian, cynian, cytrynian, glutaminian, hamiltonian, izocyjanian, karbaminian, kazeinian, ksantogenian, manganian, molibdenian, nadmanganian, naftenian, pikrynian, selenian, stearynian, sulfonian, tytanian, uranian, uranian, urokanian, uronian, wersenian, winian*
3	*nianin#*	*fenianin, socynianin, pilźnianin, prisztinianin, talinianin, kownianin, wilnianin*

h) w strukturze *gi*A litera *i* najczęściej oznacza fonem /j/. Ta zasada nie dotyczy triady *gie*. Litera *i* jest w niej tylko znakiem miękkości za wyjątkiem wyrazu *higiena*. Poza tym z opisu zamieszczonego w *Automatyzacji…* wynika, że w mianowniku, bierniku oraz wołaczu liczby mnogiej wyrazów obcych zakończonych w mianowniku liczby pojedynczej na *gia#* litera *i* również oznacza fonem /j/. Analiza słownika *SJP.pl* wykazała, że większość wyrazów zakończonych na *gie#* należy do tej grupy. Zawiera ona kilkaset form fleksyjnych, między innymi: *magie, alergie, filologie, mikrologie, socjologie, anergie*. Natomiast zbiór wyrazów rodzimych mających wymowę synchroniczną i zakończonych na *gie#* obejmuje kilkadziesiąt wyrazów. Są to:
   – przymiotniki w mianowniku, bierniku i wołaczu liczby mnogiej zawierające w wygłosie sekwencję *nogie#* (np.: *bystronogie, czworonogie, płetwonogie, dwunogie* itp.). Te wyrazy zostały utworzone na podstawie rzeczownika *nogi*. Istnieją dwie formy rodzime z taką samą strukturą wygłosu, które nie są derywatami wyrazu *nogi*: *mnogie, niemnogie*;
   – przymiotniki w mianowniku, bierniku i wołaczu liczby mnogiej zawierające w wygłosie sekwencję *rogie#* (np.: *długorogie, blaszkorogie, dwurogie, wielorogie*). Są to derywaty rzeczownika *rogi*. Występuje kilka innych rodzimych form, które nie są derywatami: *drogie, niedrogie, srogie, niesrogie, wrogie, niewrogie, złowrogie, niezłowrogie*;
   – formy fleksyjne: *bezmózgie, niebezmózgie, całobrzegie, niecałobrzegie, chędogie, niechędogie, cienkowargie, niecienkowargie, długie, niedługie, drugie, niedrugie, nagie, nienagie, połogie, niepołogie, półdługie, niepółdługie, półnagie niepółnagie, przydługie, nieprzydługie, tęgie, nietęgie, ubogie, nieubogie, Wielgie, wpółnagie, niewpółnagie, złotopręgie, niezłotopręgie.*

   Można przyjąć, że triada *gie* w wygłosie ma wymowę asynchroniczną, natomiast na podstawie wymienionych wyrazów rodzimych zostały wyznaczone sekwencje wygłosowe umożliwiające identyfikację wyjątków: *nogie#, rogie#, mózgie#, brzegie#, chędogie#, wargie#, długie#, drugie#, nagie#, połogie#, tęgie#, ubogie#, wielgie#, pręgie#.*

W tabeli 2.16 reguła dotycząca transkrypcji triady *gie* została oznaczona, ponieważ taka postać nie obejmuje omówionych wypadków wymowy asynchronicznej. Poza tym obecność tej reguły wiąże się z wyodrębnieniem fonemu /J/. Natomiast uznanie głoski [ǵ] (AS) za alofon fonemu /g/ jest możliwe tylko przy założeniu konsekwentnej (każdorazowej) asynchronicznej wymowy struktur *gi*A. Do takiej wymowy jest dostosowana druga wersja reguły znajdującej się w drugim wierszu. Ostatnia reguła w tabeli występuje w dwóch wariantach. Wariant drugi dotyczy wymowy mniej starannej związanej z redukcją fonemu odpowiadającego pierwszej literze *i*.

Tabela 2.16 – Analiza słownika i korpusu dotycząca litery *i* (część 7)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	g[*i*]*a*=/j/	2956	*teologia, nostalgiach, terminologia*	48	*energia, zoologia, kolegialny*
2	g[*i*]*e*=1* lub g[*i*]*e*=/j/	7819	*giełda, gierki, metrologiem*	4744	*cegiełek, drugiego, węgierskie*
3	g[*i*]*o*=/j/	2865	*religiom, regionalnych, region*	58	*legion, regionu, legioniści*
4	g[*i*]*u*=/j/	208	*Eligiusz, plagiuje, pedagogium,*	15	*kolegium, Sergiusza, Remigiusz*
5	g[*i*]*y*=/j/	0		0	
6	g[*i*]*ó*=/j/	32	*pedagogiów, kolegiów, elogiów*	0	
7	g[*i*]*ą*=/j/	812	*alergią, Belgią, frazeologią*	90	*wygiąć, energią, chronologią*
8	g[*i*]*ę*=/j/	2026	*giętarce, strategię, urologię*	227	*strategię, nieugięta, socjologię*
9	g[*i*]*i*#=/j/ lub g[*i*]*i*#=1	688	*analogii, biologii, etymologii*	163	*religii, energii, strategii*

i) literze *i* występującej w strukturze *ki*A odpowiada fonem /j/. Wyjątkiem jest ortograficzna triada *kie*, w której litera ta jest tylko znakiem miękkości ze względu na wymowę synchroniczną. Taka wymowa jest zgodna z *Automatyzacją…* i dostosowana jest do niej pierwsza wersja reguły zamieszczonej w drugim wierszu w tabeli 2.17.
Można również odnieść się do bardziej współczesnej interpretacji, w której zakłada się pełną asynchronię wymowy. W tym wariancie fonem /c/ nie zostaje wyodrębniony i obowiązuje druga postać reguły w drugim wierszu tabeli. Oprócz tego istnieją dwa warianty ostatniej wyszczególnionej reguły. Drugi wariant dotyczy wymowy mniej starannej związanej z redukcją fonemu odpowiadającego pierwszej literze *i*.

Tabela 2.17 – Analiza słownika i korpusu dotycząca litery *i* (część 8)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	k[*i*]*a*=/j/	392	*kiat, skiagrafię, rejkiawiczanka*	75	*Pinokia, makiawelizm, spartakiady*
2	k[*i*]*e*=1 lub k[*i*]*e*=/j/	118154	*koleżeńskiego, tybetańskie, malarskie*	37712	*angielskiej, kiedy, pobliskiego*
3	k[*i*]*o*=/j/	227	*barokiosk, kioskarką, Tokio*	387	*Pinokio, kiosk, kioskiem*
4	k[*i*]*u*=/j/	190	*kiur, manikiurzystką, Monteskiusz*	14	*manikiur, Pinokiu, manikiurzystki*
5	k[*i*]*y*=/j/	0		0	
6	k[*i*]*ó*=/j/	0		1	*Pinokiów*
7	k[*i*]*ą*=/j/	31	*Nokią, merkią, arrenotokią*	0	
8	k[*i*]*ę*=/j/	26	*celiakię, akię, makię*	0	
9	k[*i*]*i*#=/j/ lub k[*i*]*i*#=1	27	*makii, merkii, celiakii*	0	

2.3. Transkrypcja liter oznaczających półsamogłoski

2.3.1. Litera *j*

Podstawowa reguła transkrypcji X[*j*]X=/j/ jest zgodna z transkrypcją podstawową litery *j*. Taka reguła jest właściwa dla starannej wymowy wyrazów typu: *sesji, kolacji, okazji*. Tabela 2.18 zawiera dwie reguły związane z alternatywną wymową wyrazów tego typu. W takiej wymowie fonem /j/ nie jest realizowany.

Tabela 2.18 – Analiza słownika i korpusu dotycząca litery *j* (część 1)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	T[*j*]*i*O=1	3053	*izolacji, kondycji, rywalizacji*	2553	*lekcji, kolekcji, stacji*
2	D[*j*]*i*O=1	268	*dysfazji, fuzji, Eurazji*	417	*rewizji, telewizji, poezji*

Kolejna tabela (2.19) zawiera wynik analizy opartej na tych samych regułach, jednak wyszczególniono w niej litery, które mogą znajdować się w kontekście lewostronnym litery *j* (litery należące do zbioru T oraz do zbioru D). Na podstawie tych danych można stwierdzić, że najczęściej są to litery: *c, s, z*. Pozostałe litery należące do tych zbiorów nie występują w podanym kontekście lub występują w nim w stopniu marginalnym. Informacja ta może być istotna dla projektowania badań dotyczących prawdopodobieństwa poszczególnych wariantów wymowy.

Tabela 2.19 – Analiza słownika i korpusu dotycząca litery *j* (część 2)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	*c*[*j*]*i*O=1	2797	*izolacji, redakcji, restauracji*	2363	*lekcji, kolekcji, sytuacji*
2	*ć*[*j*]*i*O=1	0		0	
3	*f*[*j*]*i*O=1	0		0	
4	*h*[*j*]*i*O=1	0		0	
5	*k*[*j*]*i*O=1	0		0	
6	*p*[*j*]*i*O=1	0		0	
7	*s*[*j*]*i*O=1	251	*admisji, dysleksji, dystorsji*	190	*dyskusji, opresji, wersji*
8	*ś*[*j*]*i*O=1	0		0	
9	*t*[*j*]*i*O=1	2	*czajtji, hechtji*	0	
10	*b*[*j*]*i*O=1	0		0	
11	*d*[*j*]*i*O=1	0		0	
12	*g*[*j*]*i*O=1	0		0	
13	*z*[*j*]*i*O=1	267	*atanazji, Gruzji, iluzji*	417	*okazji, fantazji, poezji*
14	*ź*[*j*]*i*O=1	0		0	
15	*ż*[*j*]*i*O=1	0		0	

2.3.2. Litera *ł*

Reguła X[*ł*]X=/w/ pokrywa się z transkrypcją podstawową. Pierwsza modyfikacja reguły dotyczy redukcji geminaty *łł* umiejscowionej przed literą oznaczającą spółgłoskę lub w wygłosie wyrazu (tabela 2.20).

Tabela 2.20 – Analiza słownika i korpusu dotycząca litery *ł* (część 1)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	*ł*[*ł*](X–A)=1	55	*mełłby, rozmełłszy, Radziwiłł*	5	*pełł, Radziwiłł*

Kolejna modyfikacja dotyczy grup spółgłoskowych, w których litera *ł* występuje w sąsiedztwie liter oznaczających spółgłoski właściwe lub w sąsiedztwie litery *r* oraz jednej spółgłoski właściwej. Półsamogłoska [u̯] (AS) w otoczeniu głosek o niższej sonorności często nie jest realizowana. Tabela 2.21 zawiera również reguły związane z redukcją geminaty umiejscowionej w wygłosie lub w prawostronnym kontekście spółgłoskowym. Tabela 2.21 nie jest więc uzupełnieniem tabeli 2.20, tylko stanowi dla niej alternatywę. Redukcja fonemu /w/ jest opcjonalna i jej prawdopodobieństwo wynika z frekwencji wyrazów [Mańczak 1988].

Tabela 2.21 – Analiza słownika i korpusu dotycząca litery *ł* (część 2)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	D[*ł*]T(X–D)=1	316	*jabłka, bedłka, mydłka*	190	*jabłka, jabłkami, jabłkach*
2	R[*ł*]TD=1	83	*namełłszy, odparłszy, zwarłszy*	90	*oparłszy, dotarłszy, starłszy*
3	R[*ł*]O=1	95	*zmełł, poparł, uparł*	2197	*pożarł, otarł, poparł*
4	R[*ł*]*b*=1	270	*mełłbym, dodarłby, nadżarłbym*	6	*pożarłbym, podarłby, wyparłby*
5	T[*ł*]TD=1	346	*pocichłszy, wyschłszy, nawisłszy*	134	*podniósłszy, odniósłszy, uląkłszy*
6	T[*ł*]O=1	428	*głuchł, dociekł, przycichł*	6310	*rzekł, wzniósł, zanikł*
7	T[*ł*]*b*=1	1188	*wzniósłbym, przyrósłbyś, przemokłby*	55	*tłukłbym, doniósłby, wsiąkłby*
8	D[*ł*]TD=1	414	*jabłczanka, przepadłszy, przybladłszy*	377	*znalazłszy, zabiegłszy, doszedłszy*
9	D[*ł*]O=1	413	*osłabł, naszedł, ostygł*	11755	*przyszedł, wszedł, postrzegł*
10	D[*ł*]*b*=1	1221	*przysiągłby, schudłbyś, uległbyś*	701	*zjadłby, pomógłby, usiadłbym*

Reguły z tabeli 2.21 zostały użyte dla zaprojektowania dwóch kolejnych analiz. Utworzono listę szczegółowych reguł, które odnoszą się do kontekstu litery *ł* zapisanego przy użyciu poszczególnych liter, a nie zbiorów liter. Dzięki takiemu podejściu ewentualna modyfikacja reguł staje się znacznie łatwiejsza. Dwa początkowe wiersze w tabeli 2.22 dotyczą redukcji w ramach geminaty znajdującej się przed literą *b* lub przed diadą *sz*. Pozostałe reguły dotyczą redukcji fonemu /w/ umiejscowionego w obustronnym kontekście spółgłosek.

Tabela 2.22 – Analiza słownika i korpusu dotycząca litery *ł* (część 3)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	*ł*[*ł*]*b*=1	33	*mełłby, opełłby, rozmełłbym*	0	
2	*ł*[*ł*]*sz*=1	9	*zmełłszy, opełłszy, popełłszy*	0	
3	*b*[*ł*]*k*=1	180	*bebłka, jabłko, jabłkowiec*	188	*jabłka, jabłkowych, jabłkową*
4	*d*[*ł*]*k*=1	99	*bedłka, bedłkach, bedłkowce*	1	*bedłkami*
5	*z*[*ł*]*k*=1	35	*kozłka, kozłkom, kozłków*	0	
6	*d*[*ł*]*c*=1	1	*bedłce*	0	
7	*r*[*ł*]*sz*=1	74	*dodarłszy, otarłszy, potarłszy*	90	*oparłszy, dotarłszy, starłszy*
8	*r*[*ł*]*b*=1	237	*dodarłby, nażarłby, wsparłbyś*	6	*pożarłbym, podarłby, wyparłby*
9	*ch*[*ł*]*sz*=1	36	*doschłszy, napuchłszy, przycichłszy*	2	*przycichłszy, zdechłszy*
10	*k*[*ł*]*sz*=1	159	*dociekłszy, obtłukłszy, przywlekłszy*	65	*ukląkłszy, rzekłszy, zwlókłszy*
11	*p*[*ł*]*sz*=1	12	*ochrypłszy, oklapłszy zachrypłszy*	0	
12	*s*[*ł*]*sz*=1	96	*dogasłszy, potrząsłszy, wzniósłszy*	58	*podniósłszy, dorósłszy, wyrósłszy*
13	*t*[*ł*]*sz*=1	43	*rozkwitłszy, wgniótłszy, zmiótłszy*	9	*oplótłszy, zmiótłszy, rozplótłszy*
14	*s*[*ł*]*b*=1	315	*dogasłby, narósłby, pasłbyś*	15	*uniósłby, zniósłbym przyniósłbyś*
15	*k*[*ł*]*b*=1	543	*ciekłby, miękłby, żółkłby*	34	*tłukłbym, rzekłbyś, wsiąkłby*
16	*ch*[*ł*]*b*=1	135	*cichłby, opuchłby, wysechłby*	3	*zdechłbyś, zdechłby, spuchłbym*
17	*t*[*ł*]*b*=1	144	*dogniótłby, kwitłbyś, splótłbyś*	3	*zmiótłbym, plótłbyś, rozgniótłby*
18	*p*[*ł*]*b*=1	51	*chrypłby, oślepłby, zakrzepłby*	0	
19	*b*[*ł*]*cz*=1	54	*jabłczanka, jabłczanką, jabłczanowe*	0	
20	*d*[*ł*]*sz*=1	133	*dojadłszy, odszedłszy, wybladłszy*	220	*opadłszy, pobladłszy, wsiadłszy*
21	*g*[*ł*]*sz*=1	135	*dobiegłszy, rozbiegłszy, zmógłszy*	86	*dobiegłszy, ległszy, ubiegłszy*
22	*z*[*ł*]*sz*=1	83	*dogryzłszy, uwiązłszy, zwiózłszy*	71	*znalazłszy, przelazłszy, zlazłszy*
23	*g*[*ł*]*b*=1	450	*biegłby, pomógłby, sprzągłbyś*	581	*mógłbym, przysiągłby, pomógłbym*
24	*d*[*ł*]*b*=1	450	*bladłby, doszedłby, podpadłby*	84	*poszedłby, zjadłby, usiadłbym*
25	*z*[*ł*]*b*=1	276	*dogryzłby, pełzłby, znalazłbyś*	36	*zamarzłby, wygryzłby, przewiózłbym*
26	*b*[*ł*]*b*=1	45	*osłabłbym, zziębłbyś, wyziąbłby*	0	

Tabela 2.23 zawiera uszczegółowiony zapis tych reguł z tabeli 2.21, które odnoszą się do litery *ł* w wygłosie wyrazu. Reguła zamieszczona w pierwszej linii dotyczy geminaty ortograficznej, natomiast pozostałe reguły odnoszą się do litery *ł* umiejscowionej za spółgłoską.

Tabela 2.23 – Analiza słownika i korpusu dotycząca litery *ł* (część 4)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	*ł*[*ł*]O=1	13	*mełł, rozmełł, Radziwiłł*	4	*pełł, Radziwiłł*
2	*r*[*ł*]O=1	79	*dodarł, zdarł, żarł*	2188	*rozdarł, oparł, starł*
3	*ch*[*ł*]O=1	45	*cichł, wybuchł, zacichł*	101	*wybuchł, wysechł, przycichł*
4	*k*[*ł*]O=1	182	*blakł, owlekł, odrzekł*	3972	*pękł, upiekł, uciekł*
5	*p*[*ł*]O=1	21	*zachrypł, oklapł, okrzepł*	13	*oślepł, okrzepł, ścierpł*
6	*s*[*ł*]O=1	130	*dogasł, rozniósł, rósł*	2172	*przyniósł, pasł, rozrósł*
7	*t*[*ł*]O=1	50	*dokwitł, wymiótł, zgniótł*	49	*zmiótł, zakwitł, zgniótł*
8	*b*[*ł*]O=1	17	*osłabł, oziębł, wyziębł*	21	*osłabł, słabł, zasłabł*
9	*d*[*ł*]O=1	151	*bladł, dosiadł, przepadł*	6360	*spadł, przyszedł, pobladł*
10	*g*[*ł*]O=1	150	*biegł, obstrzygł, pomógł*	4155	*pobiegł, ostrzygł, mógł*
11	*z*[*ł*]O=1	95	*dogryzł, nawiózł, wygryzł*	1219	*znalazł, przywiózł, odmarzł*

2.4. Transkrypcja liter oznaczających spółgłoski sonorne

2.4.1. Litera *m*

Podstawowa reguła w *Automatyzacji…*, która dotyczy transkrypcji litery *m*, to: X[*m*]X=/m/. Pokrywa się ona z przyjętą transkrypcją podstawową. Reguły alternatywne dotyczą litery *m* umiejscowionej przed literami oznaczającymi przednie spółgłoski szczelinowe: przed *f* lub przed *w*. Zgodnie z nimi litera *m* powinna być odczytywana jako głoska [ɱ] (AS). Przy założeniu, że należy ona do fonemu /w~/, transkrypcja litery *m* jest zgodna z pierwszym wariantem reguł podanych w wierszach tabeli 2.24. Przy założeniu przynależności głoski [ɱ] do fonemu /n/ trzeba odnieść się do wariantu drugiego.

Tabela 2.24 – Analiza słownika i korpusu dotycząca litery *m*

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	X[*m*]*f*=/w~/ lub X[*m*]*f*=/n/	2614	*limfatyczny, komfort, symfoniczny*	317	*triumfalnie, amfiteatr, amfibia*
2	X[*m*]*w*=/w~/ lub X[*m*]*w*=/n/	418	*tramwaj, triumwiry, decemwiry*	98	*tramwajem, tramwajowy, triumwirat*

2.4.2. Litera *n*

Możliwości transkrypcji litery *n* są zróżnicowane i w znacznym stopniu uzależnione od kontekstu prawostronnego. W tabeli 2.25 przedstawiono wynik analizy wykonanej na podstawie reguł zamieszczonych w tabeli 16 w *Automatyzacji…* . W badaniu uwzględniono tylko te reguły, które nie są zgodne z przyjętą transkrypcją podstawową. Litera *n* umiejscowiona przed literą lub diadą ortograficzną oznaczającymi jeden z fonemów: /s/, /z/, /Z/ oraz /rz/ (zob. podrozdz. 1.1) jest transkrybowana jako fonem /w~/. Przed literą, diadą lub triadą kodującymi: /i/, /ts'/, /dz'/, /s'/, /z'/, /n'/ litera *n* oznacza fonem /n'/. Reguły w takiej postaci obowiązują przy założeniu, że głoska [ɱ] (AS) jest alofonem fonemu /w~/.

Tabela 2.25 – Analiza słownika i korpusu dotycząca litery *n*

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	X[*n*]*s*(X–*i*)=/w~/	34246	*zniuansuję, konsulenty, transylwański*	2407	*kwadrans, pensji, instrumenty*
2	X[*n*]*z*(X–*i*)=/w~/	3109	*hanza, cenzurowych, tranzystorowej*	126	*benzyna, cenzurze, tranzyt*
3	X[*n*]*ż*=/w~/	2300	*aranżują, inżynier, rewanżująca*	187	*inżynier, oranżerii, rewanż*
4	X[*n*]*rz*=/w~/	7	*żanrze, mikrohenrze, henrze*	0	
5	X[*n*]*i*=/n'/	1732885	*nieskażony, zniekształcana, niestarannie*	306727	*nikt, sypialnie, zapomniał*
6	X[*n*]*ci*=/n'/	2881	*kanciarskie, kontyngencie, rencisto*	852	*gruncie, horyzoncie, kanciastej*
7	X[*n*]*ć*=/n'/	0		0	
8	X[*n*]*dzi*=/n'/	904	*almandzie, bandzie, Wandzia*	317	*sekundzie, komendzie, legendzie*
9	X[*n*]*dź*=/n'/	22	*grandź, grandźmy, mendźcie*	0	
10	X[*n*]*si*=/n'/	652	*ambulansie, bilansik, szympansich*	111	*kwadransie, sensie, szansie*
11	X[*n*]*ś*=/n'/	3	*Ernście, Nernście, dunście*	0	
12	X[*n*]*zi*=/n'/	1	*hanzie*	0	
13	X[*n*]*ź*=/n'/	0		0	
14	X[*n*]*ni*=/n'/	7390	*półpustynni, mennica, sutannie*	2444	*gościnnie, niezmiennie, zwinni*

W tabeli 2.26 przedstawiono dopuszczalne modyfikacje reguł zamieszczonych w tabeli 2.25. Kolumna druga: *Kontekst litery n* pokrywa się z fragmentami reguł zamieszczonych w tabeli 2.25 (przed znakiem równości). Dodano jeden kontekst: X[*n*]*n*O. Kolumna *Bez zmian* zawiera transkrypcję litery *n*, która pokrywa się z transkrypcją wynikającą z reguł ujętych w tabeli 2.25. Kolumny: *Modyfikacja 1*, *Modyfikacja 2* oraz *Modyfikacja 3* zawierają transkrypcję alternatywą względem transkrypcji podanej w kolumnie *Bez zmian*. Uzyskanie transkrypcji alternatywnej wiąże się z następującymi założeniami:

– *Modyfikacja 1* – w słowach typu *pannie, wannie, rynnie* pierwszej literze *n* odpowiada fonem /n/;
– *Modyfikacja 2* – geminaty złożone z dwóch liter *n* w wygłosie wyrazu (np. *fontann*) są wymawiane jako pojedyncza głoska;
– *Modyfikacja 3* – litera *n* umiejscowiona przed literą lub diadą ortograficzną oznaczającymi fonem /s'/ lub fonem /z'/ jest transkrybowana jako fonem /w~/.

Omówione trzy modyfikacje można stosować jednocześnie w dowolnej kombinacji. Są to zmiany fakultatywne i wzajemnie od siebie niezależne.

Tabela 2.26 – Warianty modyfikacji transkrypcji litery *n* (część 1)

Lp.	Kontekst litery *n*	Bez zmian	Modyfikacja 1	Modyfikacja 2	Modyfikacja 3
1	X[*n*]*s*(X–*i*)	/w~/			
2	X[*n*]*z*(X–*i*)	/w~/			
3	X[*n*]*ż*	/w~/			
4	X[*n*]*rz*	/w~/			
5	X[*n*]*i*	/n'/			
6	X[*n*]*ci*	/n'/			
7	X[*n*]*ć*	/n'/			
8	X[*n*]*dzi*	/n'/			
9	X[*n*]*dź*	/n'/			
10	X[*n*]*si*	/n'/			/w~/
11	X[*n*]*ś*	/n'/			/w~/
12	X[*n*]*zi*	/n'/			/w~/
13	X[*n*]*ź*	/n'/			/w~/
14	X[*n*]*ni*	/n'/	/n/		
15	X[*n*]*n*O	/n/		1	

Kolejna tabela (2.27) jest analogiczna względem tabeli 2.26 przy założeniu, że głoska [ɰ̃] (AS) należy do fonemu /n/. W tej tabeli wiersze dotyczące umiejscowienia litery *n* przed literą lub diadą ortograficzną oznaczającymi fonemy: /s/, /z/, /Z/ oraz /rz/ zostały pominięte, ponieważ przy powyższym założeniu transkrypcja litery *n* jest zgodna z transkrypcją podstawową (uwaga dotyczy kolumny *Bez zmian* w tabeli 2.27). Poza tym w wypadku modyfikacji trzeciej litera *n* również jest transkrybowana jako fonem /n/.

Tabela 2.27 – Warianty modyfikacji transkrypcji litery *n* (część 2)

Lp.	Kontekst litery *n*	bez zmian	Modyfikacja 1	Modyfikacja 2	Modyfikacja 3
1	X[*n*]*i*	/n'/			
2	X[*n*]*ci*	/n'/			
3	X[*n*]*ć*	/n'/			
4	X[*n*]*dzi*	/n'/			
5	X[*n*]*dź*	/n'/			
6	X[*n*]*si*	/n'/			/n/
7	X[*n*]*ś*	/n'/			/n/
8	X[*n*]*zi*	/n'/			/n/
9	X[*n*]*ź*	/n'/			/n/
10	X[*n*]*ni*	/n'/	/n/		
11	X[*n*]*n*O	/n/		1	

2.4.3. Litera *ń*

Zasadnicza reguła umożliwiająca transkrypcję litery *ń* to: X[*ń*]X=/n'/ (tabela 17 w *Automatyzacji…*). Jest ona zgodna z transkrypcją podstawową. Autorka zaznaczyła jednak, że taka reguła dostosowana jest do wymowy starannej. Zgodnie z regułami alternatywnymi w kontekstach wymienionych w tabeli 2.28 litera *ń* powinna być transkrybowana jako sekwencja fonemów /j.m/ lub /j.n/. Jest to konsekwencja rozsunięcia artykulacyjnego w pozycji przed głoską zwartą.

Tabela 2.28 – Analiza słownika i korpusu dotycząca litery *ń*

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	X[*ń*]*b*=/j.m/	289	*hańba, zhańbione, zhańbił*	52	*hańba, hańbie, hańbiący*
2	X[*ń*]*t*=/j.n/	30	*rańtuch, wańtuch, maleńtas*	2	*rańtuchu, rańtucha*
3	X[*ń*]*k*=/j.n/	3901	*świeżuteńki, maleńkimi, cieniusieńki*	499	*maleńkim, niańkę, Heńkowi*
4	X[*ń*]*d*(X–*z*) =/j.n/	1	*dzieńdoberek*	0	
5	X[*ń*]*c*(X–*i*)=/j.n/	17136	*gońcom, tańcujmy, słońcem*	8336	*końcówki, dziedzińcu, łańcuszka*

2.4.4. Litera *l*

Reguła transkrypcji dla litery *l*: X[*l*]X=/l/ (tabela 18 w *Automatyzacji…*) jest zgodna z transkrypcją podstawową. W tabeli 2.29 przedstawiono regułę alternatywną, zgodnie z którą diada ortograficzna *ll* powinna być transkrybowana jako pojedynczy fonem /l/.

Tabela 2.29 – Analiza słownika i korpusu dotycząca litery *l*

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	*l*[*l*](X–A)=1	621	*billboard, rollborderem, billboardowe*	443	*troll, moll*

Przytoczona reguła dotyczy przede wszystkim wyrazów obcych lub wyrazów, w których rdzeń niezgodny z zasadami ortofonicznymi języka polskiego jest połączony z polską końcówką fleksyjną.

2.4.5. Litera *r*

Litera *r*, jeżeli nie jest komponentem digrafemu *rz*, oznacza fonem /r/. Natomiast jeżeli jest ona częścią diady *rz*, to w określonych wypadkach może oznaczać fonem /r/. Dokładna analiza dotycząca różnych możliwości transkrypcji diady *rz* została zawarta w podrozdziale 2.5.12.

2.5. Transkrypcja liter, diad i triad ortograficznych oznaczających spółgłoski właściwe dźwięczne

2.5.1. Litera *w*

Litera *w* oznacza dźwięczny szczelinowy fonem /v/. Jednak często jest ona realizowana jako głoska bezdźwięczna w wyniku umiejscowienia w kontekście ubezdźwięczniającym prawostronnym lub lewostronnym. Dodatkowa komplikacja związana jest z faktem, że struktura morfologiczna wyrazu (obecność junktury) może mieć wpływ na status dźwięczności fonemu odpowiadającego literze *w*. Problemy te zostały uwidocznione w szczegółowej analizie zawartej w podrozdziale 2.1 w *Asymilacji…* .

W *Automatyzacji…* został poruszony problem wymowy wyrazów typu: *krakowski, krakowscy, Beniowski, Poniatowscy, warszawska, pierwszy, pierwsi*. Przy nie bardzo starannej wymowie tych wyrazów głoska odpowiadająca literze *w* może nie być realizowana. Autorka zaproponowała odrębny zestaw reguł, które dotyczą wymowy mniej starannej. Alternatywne reguły nie zostały tutaj przytoczone, ale w kolejnym akapicie przedstawiono rozwiązania umożliwiające identyfikację form fleksyjnych, które mogą być wymawiane w sposób uproszczony.

W systemie transkrypcji opartym na modelu wielowarstwowym nie można zastosować reguły w postaci: A[*w*]*sk*A=1 (reguła pochodzi z części 2a w tabeli 21 alfa w *Automatyzacji…*), ponieważ odnosi się ona również do wyrazów typu: *niewskazany, powskakiwać, przeciwskarpa, przeciwskurczowy* (w ich wymowie redukcja nie występuje). Przyjęte w *Automatyzacji…* rozwiązanie polega na wstawianiu znaczników między przedrostkiem a rdzeniem – dzięki temu reguła: A[*w*]*sk*A=1 nie mogłaby być użyta. Jednak w niniejszym opracowaniu takie podejście nie jest brane pod uwagę, ponieważ zakłada się pełną automatyzację konwersji tekstu ortograficznego na transkrypcję. Dlatego musi być użyte inne rozwiązanie. Na podstawie wszystkich form wyrazu *krakowski* można wyznaczyć zbiór sekwencji ortograficznych umożliwiających identyfikację wyrazów posiadających omawianą konstrukcję wygłosu: *wski#, wskich#, wskiego#, wskiemu#, wskim#, wskimi#, wscy#, wsku#, wska#, wską#, wskiej#, wsko#.* Trzeba jednak wykluczyć formy wyrazów: *szewski* oraz *szewstwo*, w których głoska [f] (AS) jest realizowana.

W inny sposób można rozwiązać problem alternatywnej wymowy wyrazu *pierwszy.* Wymowa ta obejmuje zarówno różne formy (np.: *pierwszemu, pierwszych, pierwszym*), jak i wyrazy pochodne (np.: *pierwszaki, pierwszomajowe*). W wypadku takich wyrazów nie występują trudności omówione w poprzednim akapicie, dlatego można zastosować reguły pochodzące z części 2a w tabeli 21 alfa w *Automatyzacji…*: *r*[*w*]*s*V=1, *r*[*w*]*sz*A=1 lub odpowiednie modyfikacje.

2.5.2. Litera *b*

Zgodnie z informacjami zawartymi w tabeli 22 w *Automatyzacji…* transkrypcja litery *b* inna niż transkrypcja podstawowa może mieć związek tylko ze zjawiskiem desonoryzacji. Możliwość transkrypcji jako fonem /p/ została omówiona w podrozdziale 2.2 w *Asymilacji…* .

2.5.3. Litera *g*

Większość reguł w *Automatyzacji…*, które dotyczą litery *g* (niezgodnych z przyjętą transkrypcją podstawową) umożliwia przekształcenie tej litery w fonem: /k/. Zostały one przytoczone w *Asymilacji…* w podrozdziale 2.3.

W *Automatyzacji…* zawarto regułę (tutaj przytoczoną w tabeli 2.30), która dotyczy litery *g* umiejscowionej przed literą *i*. Zgodnie z nią litera *g* powinna być transkrybowana jako fonem /J/ (np. /J.i.Z.y.ts.k.o/). Jednak taki sposób transkrypcji jest związany z założeniem, że /J/ jest wyodrębnionym fonemem. Natomiast przy założeniu, że głoska [ǵ] (AS) jest alofonem fonemu /g/ obowiązuje podany drugi wariant.

Tabela 2.30 – Analiza słownika i korpusu dotycząca litery *g*

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	X[*g*]*i*=/J/ lub X[*g*]*i*=/g/	54540	*filologiczne, wyczołgiwałaby, regionalnej*	16083	*cegiełek, szeregi, powyginały*

2.5.4. Litera *ź*

Transkrypcja litry *ź* inna niż transkrypcja podstawowa może mieć związek tylko ze zjawiskiem desonoryzacji. Możliwość transkrypcji jako fonem /s'/ została omówiona w *Asymilacji…* w podrozdziale 2.4.

2.5.5. Litera *d*

Litera *d* może oznaczać fonem /d/ (jest to wariant zgodny z przyjętą transkrypcją podstawową) lub fonem /t/ (jeżeli jest ona umiejscowiona w kontekście ubezdźwięczniającym). Analiza związana ze zjawiskiem desono-

ryzacji została omówiona w podrozdziale 2.5 w *Asymilacji…* . Istnieje znacznie więcej możliwości transkrypcji litery *d*, ponieważ może być ona komponentem wieloznaków lub ulegać redukcji.

W tabeli 2.31 przedstawiono wynik analizy opartej na regułach zamieszczonych w pierwszej części tabeli 25 w *Automatyzacji…* . Pominięto wszystkie reguły dotyczące diad ortograficznych: *dz, dż, dź* oraz triady *dzi.* W niniejszym opracowaniu zostały one omówione oddzielnie. Wiersze 1–4 tabeli 2.31 dotyczą transkrypcji ortograficznej triady *drz*. Jeżeli znajduje się ona przed literą oznaczającą samogłoskę, to litera *d* powinna być transkrybowana jako fonem /d/. Jeżeli triada *drz* jest umiejscowiona przed spółgłoską, to w całości powinna być transkrybowana jako fonem /dZ/ lub jako fonem /tS/ (w zależności od tego, czy kontekst następujący jest ubezdźwięczniający).

Tabela 2.31 – Analiza słownika i korpusu dotycząca litery *d*

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	X[*d*]*rz*A=/d/	7789	*nadrzędnego, turboodrzutowemu, drzewo*	4199	*odrzekł, podrzucać, zmądrzeje*
2	X[*drz*]R=/dZ/	212	*trójdrzwiowy, bezdrzwiowa, czterod-rzwiowym*	2751	*drzwi, drzwiowe, drzwiczkom*
3	X[*drz*]M=/dZ/	18	*mądrzmy, guzdrzmy, wyguzdrzmy*	0	
4	X[*drz*]T=/tS/	18	*guzdrzcie, mądrzcie, zmądrzcie*	0	

2.5.6. Litera *z*

Zgodnie z przyjętą transkrypcją podstawową litera *z* powinna być transkrybowana jako fonem /z/. Jest ona komponentem wieloznaków: *rz, dz, sz, cz oraz dzi* i *zi* (przed samogłoską) – w niniejszej publikacji zostały one omówione oddzielnie. Wynik analizy dotyczącej zjawiska desonoryzacji oraz transkrypcji litery *z* jako fonem /s/ został przedstawiony w *Asymilacji…* w podrozdziale 2.6.

Osobnego omówienia wymaga diada ortograficzna *zi*. Zgodnie z regułami ortofonicznymi języka polskiego umiejscowiona przed literą oznaczającą spółgłoskę powinna być transkrybowana jako sekwencja fonemów /z'.i/. Czasami jednak specyficzna struktura morfologiczna wyrazu sprawia, że oznacza ona sekwencję fonemów: /z.i/. Najczęściej dotyczy to połączenia przedrostka zakończonego literą *z* lub przedrostka złożonego tylko z tej litery z rdzeniem rozpoczynającym się od litery *i*. W tabeli 2.32 przedstawiono wynik analizy słownika *SJP.pl*. Zbiór użytych konstrukcji (kolumna druga) został utworzony na podstawie reguł zawartych w siódmej części tabeli 26 w *Automatyzacji…* . Uwzględnione zostały tylko formy podstawowe, ponieważ badanie dotyczy struktur ortograficznych w nagłosie wyrazu.

Tabela 2.32 – Analiza słownika dotycząca diady *zi* w nagłosie wyrazu

Lp.	Struktura nagłosu	Formy podstawowe w słowniku *SJP.pl* zawierające strukturę z kolumny drugiej
1	*#zis*	*ziszczać, ziszczalność, ziszczalny*
2	*#zid*	*zidentyfikować, zideologizować, zidiociały, zidiocieć*
3	*#zig*	*zignorować, zigurat*
4	*#zil*	*zilustrować, Zile*
5	*#ził*	*Ził*
6	*#zin*	*zin, zindeksować, zindoktrynować, zindustrializować, zindywidualizować, zineb, zinfantylizować, zinfantylniały, zinfantylnieć, zinfiltrować, zinformatyzować, ziniantrop, zinjantrop, Zinowiew, zinowy, zinstrumentalizować, zinstrumentować, zinstytucjonalizować, zinstytucjonalizowany, zintegrować, zintegrowany, zintelektualizować, zintensyfikować, zintensywniały, zintensywnieć, zinterioryzować, zinternalizować, zinterpretować, zinwentaryzować*
7	*#zim*	*~~zima~~, ~~Zimak~~, Zimbabwe, Zimbabwejczyk, Zimbabwejka, zimbabwejski, Zimbabweńczyk, Zimbabweńka, zimbabweński, ~~Zimek~~, ~~Zimin~~, ~~Zimińska~~, ~~Zimiński~~, ~~zimnawy~~, ~~Zimniak~~, ~~zimnica~~, ~~Zimnicki~~, ~~zimniej~~, ~~zimniejszy~~, ~~Zimniewicz~~, ~~Zimnik~~, ~~zimnisko~~, ~~zimniusieńki~~, ~~zimniuteńki~~, ~~zimniutki~~, ~~zimno~~, ~~Zimnoch~~, ~~zimno~~, ~~zimnokrwistość~~, ~~zimnokrwisty~~, ~~zimnolubny~~, ~~zimnowałcowany~~, ~~zimnowodny~~, ~~zimnowojenny~~, ~~zimny~~, ~~Zimoch~~, ~~zimochów~~, ~~Zimoń~~, ~~zimoodporność~~, ~~zimoodporny~~, ~~zimorodek~~, ~~zimorodkowate~~, ~~Zimorowic~~, ~~zimotrwałość~~, ~~zimotrwały~~, ~~zimować~~, ~~zimowisko~~, ~~zimowit~~, ~~Zimowski~~, ~~zimowy~~, ~~zimoziele~~, ~~zimozielony~~, ~~zimoziół~~, ~~zimówka~~*
8	*#zir*	*zironizować, zirygować, zirytować, zirytowany*
9	*#ziś*	*ziścić*
10	*#dezid*	*dezideologizacja*
11	*#dezil*	*deziluzja, deziluzyjny*
12	*#dezin*	*dezinflacja, dezinflacyjny, dezinformacja, dezinformacyjny, dezinformować, dezintegracja, dezintegracyjny, dezintegralny, dezintegrator, dezintegrować, dezintegrująco, dezinwestycja, dezinwestycyjny*
13	*#rozis*	*roziskrzać, roziskrzony, roziskrzyć*
14	*#rozig*	*rozigrać, rozigrany*
15	*#rozin*	*rozindyczony, rozindyczyć*
16	*#rozir*	*rozirytować*
17	*#zdezin*	*zdezintegrować*
18	*#bezim*	*bezimienność, bezimienny, bezimplozyjny*
19	*#bezis*	*beziskrowy*
20	*#bezid*	*bezideowiec, bezideowość, bezideowy*
21	*#bezig*	*bezigłowy*
22	*#bezin*	*bezindukcyjny, bezinercyjny, bezinteresownie, bezinteresowność, bezinteresowny, bezinwazyjność, bezinwazyjny, bezinwestycyjność, bezinwestycyjny*
23	*#ubezim*	*ubezimiennić*

W kolumnie trzeciej w tabeli 2.32 wyszczególniono formy podstawowe wyrazów, które rozpoczynają się od struktur ujętych w kolumnie drugiej. Analiza wykazała, że w niemal wszystkich uwzględnionych strukturach nagłosu diada *zi* oznacza sekwencję fonemów /z.i/. Powinna być ona transkrybowana jako sekwencja fonemów /z'.i/, jeżeli jest częścią nagłosowej triady *#zim*. Triada ta występuje w formach wyrazów *zima* oraz *zimny* (i w wyrazach pokrewnych), a także w niektórych nazwiskach (np. *Zimoch*) – takie wyrazy w tabeli 2.32 zostały przekreślone. Poza tym w wypadku wyrazu *Zimbabwe* i wyrazów pokrewnych diada *zi* oznacza sekwencję fonemów /z.i/. Można je zidentyfikować na podstawie struktury: *#zimb*. Niekiedy nagłosowe konstrukcje, które zostały ujęte w kolumnie drugiej, mogą być poprzedzone przedrostkiem *nie*. Dotyczy to między innymi wyrazów:

nieziszczalność, nieziszczonych, niezidentyfikowane, niezideologizowany, niezignorowana, niezilustrowane, niezindeksowane, niezindoktrynowani, niezindustrializowane, niezinfantylizowani, niezinformatyzowane, niezintensyfikowany, niezironizowanego, niezirytowanego

Dla celów praktycznych można przyjąć, że każda struktura wymieniona w kolumnie drugiej w tabeli 2.32 może być poprzedzona przedrostkiem *nie*. Dlatego zbiór struktur ortograficznych umożliwiających identyfikację

wyrazów, w których diada *zi* powinna być transkrybowana jako sekwencja fonemów /z.i/ można powiększyć o kolejne elementy (np.: *#niezid, #niezin, #niezir, #nierozig* itp.).

2.5.7. Litera *ż*

Litera *ż* powinna być transkrybowana jako dźwięczny szczelinowy fonem /Z/. Wynik analizy dotyczącej transkrypcji w wariancie bezdźwięcznym (jako fonem /S/) został omówiony w *Asymilacji…* . Informacje dotyczące możliwości transkrypcji diady ortograficznej *dż* przedstawiono w części 2.5.8.

Reguła zamieszczona w tabeli 2.33 umożliwia przekształcanie litery *ż* w fonem /s'/. Takie przekształcenie dotyczy wyrazów typu: *najwyżsi, uboższi*.

Tabela 2.33 – Analiza słownika i korpusu dotycząca litery *ż*

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	X[*ż*]*si*=/s'/	32	*najświeższi, najdłuższi, chyższi*	17	*uboższi, najbliższi, najwyższi*

2.5.8. Diada ortograficzna *dż*

Diada ortograficzna *dż* w przyjętej transkrypcji podstawowej oznacza fonem /dZ/. Podstawowy problem związany z transkrypcją tego typu struktur wynika z faktu, że w określonych kontekstach oznaczają one sekwencję fonemów (zatem nie są wieloznakami). Poza tym mogą one występować w kontekście ubezdźwięczniającym. Analiza związana z możliwością transkrypcji diady *dż* jako fonem /tS/ została opisana w podrozdziale 2.8 w *Asymilacji…*.

Analizę ujętą w tabeli 2.34 oparto na drugiej części tabeli 25 w *Automatyzacji…* . Obejmuje ona wyrazy, w których diada ortograficzna *dż* powinna być transkrybowana jako sekwencja dwóch fonemów. Z analizy wynika, że w praktyce takie przypadki są spotykane niezwykle rzadko (za wyjątkiem wyrazów zawierających w nagłosie strukturę *#odż* lub *#podż*). Na podstawie reguł wyszczególnionych w tabeli 2.34 wyznaczono przedrostki, które zostały użyte w bardziej wnikliwej analizie (tabela 2.35).

Tabela 2.34 – Analiza słownika i korpusu dotycząca niestandardowej transkrypcji diady ortograficznej *dż*

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	*#na*[*d*]*ż*A=/d/	227	*nadżarcia, nadżerały, nadżółkła*	0	
2	*#o*[*d*]*ż*A=/d/	1005	*odżeglowała, odżałowałbyś, odżywiającej*	55	*odżył, odżegnywał, odżywiał*
3	*#prze*[*d*]*ż*A=/d/	10	*przedżołądek, przedżołądka, przedżołądkowi*	0	
4	*#po*[*d*]*ż*A=/d/	452	*podżarci, podżebrowy, podżegłem*	7	*podżegała, podżegaczy, podżyły*
5	*#niena*[*d*]*ż*A =/d/	100	*nienadżarty, nienadżarcia, nienadżółkłej*	0	
6	*#nieo*[*d*]*ż*A=/d/	365	*nieodżałowana, nieodżywianej, nieodżyłowanym*	1	*nieodżałowanej*
7	*#pona*[*d*]*ż*A =/d/	79	*ponadżerać, ponadżerano, ponadżeram*	0	
8	*#śró*[*d*]*ż*A=/d/	24	*śródżylny, śródżylną, śródżylne*	0	
9	*#prze*[*d*]*ż*M =/d/	11	*przedżniwny, przedżniwna, przedżniwną*	0	

Trzeba wspomnieć o istotnym problemie, który odnosi się do wybranych konstrukcji ortograficznych zawierających literę *d* na pierwszej pozycji (*dż, dz, dź, dzi*). W praktyce bifonematyczna realizacja takich struktur może ograniczać się do wymowy starannej i taką wymowę założono dla opisanych analiz. Przy wymowie mniej starannej fragment wyrazu złożony z głoski zwartej oraz z głoski szczelinowej może być realizowany jako głoska zwarto-szczelinowa ze względu na zjawisko asybilacji; omawiane konstrukcje mogą zatem być realizowane jako jedna głoska, pomimo obecności junktury morfologicznej.

Jest kwestią umowną, czy monofonematyczna interpretacja diady *dż* powinna zostać włączona do transkrypcji podstawowej. Każde połączenie litery *d* oraz litery ż, które wynika z procesu słowotwórczego i które oznacza sekwencję dwóch fonemów, można zatem uznać za wyjątkowe względem standardowej transkrypcji tej diady. W tabeli 2.35 zamieszczono wynik szczegółowej analizy – jej celem było wyznaczenie zbioru nagłosowych ciągów ortograficznych umożliwiających identyfikację wszystkich form fleksyjnych, w których diada *dż* oznacza sekwencję fonemów. Dzięki temu w projekcie konkretnego systemu transkrypcji można utworzyć jedną podstawową regułę dotyczącą diady *dż* (umożliwiającą transkrypcję tej diady jako fonem /dZ/) i jednocześnie można uwzględnić zbiór wyjątków. W wypadku niektórych wyrazów uznanych za takie wyjątki istnieje możliwość utworzenia nowych form poprzez dodanie przedrostka *nie* w nagłosie. Słownik *SJP.pl* zawiera wiele takich połączeń, np.: *nienadżarty, nienadżerający, nieodżałowanie, nieodżałowany*. Dla celów praktycznych związanych z funkcjonowaniem systemu transkrypcji przyjęto, że każdy wyraz zaliczony do zbioru wyjątków może być poprzedzony przedrostkiem *nie*. Takie rozwiązanie znacznie ułatwia konstruowanie reguł, ponieważ nie trzeba wyznaczać podzbioru wyrazów, które w rzeczywistości łączą się z tym przedrostkiem.

Część przedrostków zawiera w swojej strukturze ortograficznej krótszy przedrostek (na przykład przedrostek *nad* zawiera się w przedrostku *ponad*). Dlatego w niektórych wyrazach występują struktury ortograficzne, które pokrywają się jednocześnie z różnymi przedrostkami (np. *ponadżerać*). Czynnik ten jest istotny dla poprawnego wyznaczenia ostatecznego zbioru sekwencji spełniających pierwotny cel analizy. Wyrazy zawierające powtarzające się określone fragmenty oznaczono symbolem: * i w tabeli 2.35 zostały one umieszczone dwukrotnie (w wierszach dotyczących odpowiednich przedrostków).

Uwagi zawarte w trzech poprzednich akapitach odnoszą się również do analiz omówionych w kolejnych podrozdziałach, a dotyczących diad: *dz, dź* oraz triady *dzi*.

Tabela 2.35 – Szczegółowa analiza dotycząca dwóch możliwości transkrypcji ortograficznej diady *dż*

Lp.	Transkr.	Wykaz wyrazów
Wyrazy rozpoczynające się od sekwencji *nadż*		
1	/d.Z/	*nadżarty, nadżerać, nadżerka, nadżerkowy, nadżółkły, nadżółknąć, nadżółknięty, nadżarcie, nadżeranie, nadżółknięcie*
2	/dZ/	–
Wyrazy zawierające sekwencję *nadż* w śródgłosie		
3	/d.Z/	*ponadżerać*, ponadżeranie**
4	/dZ/	*menadżer, menadżerka, menadżerować, menadżerski, menadżerstwo, menadżeryzm*
Wyrazy rozpoczynające się od sekwencji *śród*		
5	/d.Z/	*śródżylny*
6	/dZ/	–
Wyrazy zawierające sekwencję *śródż* w śródgłosie		
7	/d.Z/	–
8	/dZ/	–
Wyrazy rozpoczynające się od sekwencji *odż*		
9	/d.Z/	*odżałować, odżałowanie, odżałowany, odżąć, odżęcie, odżeglować, odżeglowanie, odżegnać, odżegnanie, odżegnywanie, odżegnywać, odżelaziacz, odżelaziać, odżelazianie, odżelazić, odżużlacz, odżużlować, odżużlowanie, odżyłować, odżyłowywać, odżyłowany, odżyłowanie, odżyłowujący, odżyły, odżycie, odżynać, odżynający, odżynany, odżynanie, odżywać, odżywanie, odżywczość, odżywczy, odżywiać, odżywić, odżywka, odżywmy*
10	/dZ/	–

Wyrazy zawierające sekwencję *odż* w śródgłosie		
11	/d.Z/	*współodżywiać, współodżywiający, współodżywiać, podżarcie*, podżartowywać*, podżartowujący*, podżartowywanie*, podżebrowy*, podżebrze*, podżec*, podżegacz*, podżegaczka*, podżeganie*, podżegać*, podżegnąć*, podżegnięcie*, podżerać*, podżeranie*, podżuchwowy*, podżyły*, podżyrować*, podżyrowanie*, podżyrowany**
12	/dZ/	*Kambodża, Kambodżanin, Kambodżanka, kambodżański, niekambodżański, chodża, hodża, balkonolodżia, kilodżul, lodżia*
Wyrazy rozpoczynające się od sekwencji *przedż*		
13	/d.Z/	*przedżniwny, przedżołądek*
14	/dZ/	–
Wyrazy zawierające sekwencję *przedż* w śródgłosie		
15	/d.Z/	–
16	/dZ/	–
Wyrazy rozpoczynające się od sekwencji *podż*		
17	/d.Z/	*podżarcie*, podżartowywać*, podżartowujący*, podżartowywanie*, podżebrowy*, podżebrze*, podżec*, podżegacz*, podżegaczka*, podżeganie*, podżegać*, podżegnąć*, podżegnięcie*, podżerać*, podżeranie*, podżuchwowy*, podżyły*, podżyrować*, podżyrowanie*, podżyrowany**
18	/dZ/	–
Wyrazy zawierające sekwencję *podż* w śródgłosie		
19	/d.Z/	–
20	/dZ/	–
Wyrazy rozpoczynające się od sekwencji *ponad*		
21	/d.Z/	*ponadżerać*, ponadżeranie**
22	/dZ/	–
Wyrazy zawierające sekwencję *ponadż* w śródgłosie		
23	/d.Z/	–
24	/dZ/	–

Na podstawie omówionej analizy wyznaczono następujące nagłosowe struktury ortograficzne umożliwiające identyfikację wyrazów, w których diada *dż* powinna być transkrybowana jako sekwencja dwóch fonemów: *#nadż, #ponadż, #śródż, #odż, #współodż, #przedż, #podż, #ponadż, #nienadż, #nieponadż, #nieśródż, #nieodż, #niewspółodż, #nieprzedż, #niepodż, #nieponadż.*

2.5.9. Diada ortograficzna *dz*

W tej części zostały przedstawione wyniki analiz, które dotyczą diady ortograficznej *dz.* Zgodnie z przyjętą transkrypcją oznacza ona zwarto-szczelinowy dźwięczny fonem /dz/. Poza tym może być transkrybowana jako bezdźwięczny fonem /ts/ – analiza związana z taką możliwością została omówiona w podrozdziale 2.9 w *Asymilacji…* .

Diada *dz* może być uformowana w wyniku połączenia przedrostka z rdzeniem wyrazu. W takich wypadkach nie stanowi ona zatem digrafemu. W analizie (tabela 2.36) uwzględniono tylko te reguły, które nakazują transkrypcję diady *dz* jako sekwencję dwóch fonemów: /d.z/. Łatwo zauważyć, że większość omówionych struktur, które notuje słownik *SJP.pl*, w analizowanym korpusie występuje sporadycznie lub w ogóle. Wyjątek stanowią wyrazy zawierające ciąg ortograficzny *#odz*.

Tabela 2.36 – Analiza słownika i korpusu dotycząca niestandardowej transkrypcji diady ortograficznej *dz*

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	#*o*[*d*]*zy*=/d/	218	*odzysk, odzyskamy, odzywka*	472	*odzysk, odzywałam, odzywek*
2	#*nieo*[*d*]*zy*=/d/	66	*nieodzyskana, nieodzyskaną, nieodzyskane*	0	
3	#*na*[*d*]*z*(A–{*i,y*})=/d/	114	*nadzór, nadzoru, nadzorca*	16	*nadzorem, nadzór, nadzoruje*
4	#*o*[*d*]*z*(A–{*i,y*})=/d/	10	*odzew, odzewach, odzewami*	5	*odzew, odzewów*
5	#*prze*[*d*]*z*(A–{*i,y*})=/d/	68	*przedzachodni, przedzawałowym, przedzawodowej*	1	*przedzachodnie*
6	#*po*[*d*]*z*(A–{*i,y*})=/d/	203	*podzakres, podzelować, podzespołowi*	5	*podzamcze, podzelowane, podzamcza*
7	#*niena*[*d*]*z*(A–{*i,y*})=/d/	39	*nienadzorcza, nienadzorczą, nienadzorcze*	0	
8	#*nieo*[*d*]*z*(A–{*i,y*})=/d/	18	*nieodzowność, nieodzownie, nieodzownych*	11	*nieodzowną, nieodzownych, nieodzowne*
9	#*na*[*d*]*z*(D+M+R–*w*)=/d/	27	*nadzbiór, nadzmysłowościami, nadzmysłowemu*	0	
10	#*o*[*d*]*z*(D+M+R–*w*)=/d/	219	*odznaczaj, odzbroi, odzbroicie*	88	*odznaka, odznaczenie, odznaczył*
11	#*prze*[*d*]*z*(D+M+R–*w*)=/d/	42	*przedzlotowe, przedzgonny, przedzgonna*	0	
12	#*po*[*d*]*z*(D+M+R–*w*)=/d/	43	*podzbiornik, podzbiory, podzbiorze*	0	
13	#*niena*[*d*]*z*(D+M+R–*w*) =/d/	11	*nienadzmysłowa, nienadzmysłową, nienadzmysłowe*	0	
14	#*nieo*[*d*]*z*(D+M+R–*w*)=/d/	78	*nieodznaczająca, nieodzbrojeni, nieodzbrojenia*	0	
15	#*pona*[*d*]*z*(D+M+R–*w*) =/d/	22	*ponadzmysłowy, nieponadzmysłowa, nieponadzmysłową*	0	
16	#*prze*[*d*]*zw*{*a,o*}(X{*n,ń*}) =/d/	11	*przedzwojowy, przedzwojowa, przedzwojową*	0	

Na podstawie przytoczonych reguł wyznaczono zbiór przedrostków, które zostały użyte w bardziej szczegółowej analizie – wynik został zaprezentowany tabeli 2.37. Jej celem było wyznaczenie nagłosowych ciągów ortograficznych umożliwiających identyfikację wszystkich form fleksyjnych, w których diada *dz* oznacza sekwencję fonemów /d.z/. Wyznaczone ciągi liter nie mogą występować w nagłosie form zawierających diadę *dz*, która powinna być transkrybowana jako fonem /dz/. Uzyskano następujący zbiór: *#nadzb, #nadzm, #nadzor, #nadzór, #nadzw, #supernadz, #ponadz, #odz, #nieodzown, #podza, #podzb, #podzn, #podzwr, #przedza, #przedzg, #przedzj, #przedzm, #przedzwoj*. Zgodnie z wcześniejszymi uwagami dotyczącymi przedrostka *nie* ten zbiór zostaje powiększony o następujące ciągi znaków: *#nienadzb, #nienadzm, #nienadzor, #nienadzór, #nienadzw, #niesupernadz, #nieponadz, #nieodz, #niepodza, #niepodzb, #niepodzn, #niepodzwr, #nieprzedza, #nieprzedzg, #nieprzedzj, #nieprzedzm, #nieprzedzwoj*.

Tabela 2.37 – Szczegółowa analiza słownika dotycząca dwóch możliwości transkrypcji ortograficznej diady *dz*

Lp.	Transkr.	Wykaz wyrazów
Wyrazy rozpoczynające się od sekwencji *nadz*		
1	/d.z/	*nadzbiór, nadzmysłowość, nadzmysłowy, nadzorca, nadzorczy, nadzorczyni, nadzorować, nadzorowanie, nadzór, nadzwyczaj, nadzwyczajność, nadzwyczajny*
2	/dz/	*Nadzów, (w) Nadzowie, nadzy, nadze*
Wyrazy zawierające sekwencji *nadz* w śródgłosie		
3	/d.z/	*supernadzór*
4	/dz/	*Szewardnadze, nienadzy, knadze, tamilnadzki*
Wyrazy rozpoczynające się od sekwencji *ponadz*		
5	/d.z/	*ponadzmysłowy*
6	/dz/	–
Wyrazy zawierające sekwencję *ponadz* w śródgłosie		
7	/d.z/	–
8	/dz/	–
Wyrazy rozpoczynające się od sekwencji *odz*		
9	/d.z/	*odzbroić, odzew, odznaczać, odznaczenie, odznaczeniowy, odznaczyć, odznaka, odzwierciadlać, odzwierciadlić, odzwierciedlać, odzwierciedlenie, odzwierciedlić, odzwierzęcy, odzwyczaić, odzwyczajać, odzysk, odzyskać, odzyskiwać, odzysknica, odzywać, odzywka*
10	/dz/	–
Wyrazy zawierające sekwencję *odz* w śródgłosie		
11	/d.z/	*nieodzowność, nieodzowny, podzakres*, podzamcze*, podzamkowy*, podzapytanie*, podzastaw*, podzelować*, podzelowywać*, podzespołowy*, podzespół*, podzbiornik*, podzbiór*, podzbiórka*, podznaczenie*, podzwrotnikowy**
12	/dz/	– w niektórych formach fleksyjnych oraz wyrazach pochodnych od bezokoliczników zakończonych na: *-dzić*, np.: *chodzić (chodzę, chodzą), schodzić (schodzę), pochodzić (pochodzę, pochodzenie), rozpogodzić (rozpogodzę, rozpogodzenie), chłodzić (chłodzę, ochłodzenie), dowodzić (głównodowodzący);* – w mianowniku liczby mnogiej przymiotników zakończonych na: *-gi (srogi – srodzy, ubogi – ubodzy).* – w celowniku i miejscowniku rzeczowników zakończonych na: *-ga (droga – drodze, noga – nodze).* W formach fleksyjnych wyrazów: *Bodzak, Bodzanowice, Bodzanowo, Bodzanowski, Bodzanów, Bodzanówek, Bodzechowianin, Bodzechowianka, Bodzechowski, Bodzechów, Bodzek, Bodzentyn, bodzentynianin, bodzentynianka, bodzentyński, Bodzewko, Bodzęcin, bodzęciński, Bodzyniewo, modzel, Modzelewski, modzelowatość, modzelowaty, Nodzyński, dwurodzajowość, dwurodzajowy, grodza, innorodzajowy, jednorodzajowy, najurodzajniejszy, nieurodzaj, nieurodzajność, oburodzajowość, podrodzaj, rodzaj, rodzajnik, rodzajny, rodzajowość, rodzajowy, różnorodzajowy, urodzaj, urodzajniejszy, urodzajność, urodzajny, wielorodzajowy, wspólnorodzajowy, wysokourodzajny, rodzynek, rodzynka, rodzynki, rodzynkowy, wodza, wodzak, wodzarka, wodzostwo, wodzowski, Wodzyński, mokobodzki, świebodzki, dodzwaniać, dodzwonić, Przygodzki, wołogodzki, międzychodzki, Przychodzki, kłodzczanin, kłodzczanka, kłodzka, dębowokłodzki, dziadowokłodzki, inowłodzki, kłodzki, Kłodzko, kłodzkie, mrzygłodzki, kłodzka, kłodzkie, podzwaniać, podzwonić, podzwonne, antyciemnogrodzki, baligrodzki, białogrodzki, bielgorodzki, brodzki, carogrodzki, ciemnogrodzki, czerwonogrodzki, Grodzki, Jarodzki, jasnogrodzki, Kanabrodzki, krasnobrodzki, międzybrodzki, Nagrodzki, nowogrodzki, nowożmigrodzki, piotrogrodzki, podegrodzki, przygrodzki, rajgrodzki, rodzoniutki, siedmiogrodzki, Siemiątkowo Koziebrodzkie, Skrodzki, tarnogrodzki, Winogrodzki, wyszogrodzki, Zabrodzki, Zagrodzki, żmigrodzki, ciemnogrodztwo, Grodztwo, kazirodztwo, pierworodztwo, ciepłowodzki, czarnowodzki, Krasnowodzk, krasnowodzki, międzywodzki, napiwodzki, Petrozawodzk, petrozawodzki*
Wyrazy rozpoczynające się od sekwencji *podz*		
13	/d.z/	*podzakres*, podzamcze*, podzamkowy*, podzapytanie*, podzastaw*, podzelować*, podzelowywać*, podzespołowy*, podzespół*, podzbiornik*, podzbiór*, podzbiórka*, podznaczenie*, podzwrotnikowy**
14	/dz/	*podzwaniać, podzwonić, podzwonne*
Wyrazy zawierające sekwencję *podz* w śródgłosie		
15	/d.z/	–
16	/dz/	–
Wyrazy rozpoczynające się od sekwencji *przedz*		
17	/d.z/	*przedzachodni, przedzamcze, przedzamkowy, przedzawałowiec, przedzawałowy, przedzawodowy, przedzgonny, przedzjazdowy, przedzlotowy, przedzmrok, przedzwojowy*
18	/dz/	*przedzwonić, Przedzeń, przedzwaniać*

Wyrazy zawierające sekwencję *przedz* **w śródgłosie**		
19	/d.z/	–
20	/dz/	*poprzedzać, poprzedzający, uprzedzać, uprzedzający, uprzedzenie, uprzedzony, powyprzedzać, wyprzedzać, wyprzedzająco, wyprzedzenie, wyprzedzanie*

2.5.10. Diada ortograficzna *dź*

Diada ortograficzna *dź* oznacza dźwięczny zwarto-szczelinowy fonem /dz'/. Ten sam fonem częściej jest zapisywany przy użyciu triady *dzi* (przed literą oznaczającą samogłoskę inną niż *i*) lub przy użyciu diady *dz* (przed literą *i*, po której następuje litera oznaczająca spółgłoskę). W wypadku ortograficznej diady *dź* problem związany z transkrypcją bifonematyczną nie występuje (tabela 2.38).

Diada *dź* może być również transkrybowana jako bezdźwięczny fonem /ts'/, jeżeli znajduje się w kontekście ubezdźwięczniającym. Analiza dotycząca tego zagadnienia została przedstawiona w podrozdziale 2.10 w *Asymilacji…* .

Tabela 2.38 – Szczegółowa analiza słownika dotycząca dwóch możliwości transkrypcji ortograficznej diady *dź*

Lp.	Transkr.	Wykaz wyrazów
Wyrazy rozpoczynające się od sekwencji *nadź*		
1	/d.z'/	–
2	/dz'/	*nadźwiękawiać, nadźwiękawianie, nadźwigać*
Wyrazy zawierające sekwencję *nadź* w śródgłosie		
3	/d.z'/	–
4	/dz'/	*snadź*
Wyrazy rozpoczynające się od sekwencji *odź*		
5	/d.z'/	–
6	/dz'/	*odźwierna, odźwiernik, odźwierny*
Wyrazy zawierające sekwencję *odź* w śródgłosie		
7	/d.z'/	–
8	/dz'/	*antybodźcowy, bodźce, bodźcować, dochodźcie, bodźcowy, bodźcujący, przeciwbodźcowy, dodźwigać, łagodźmy, Chodźko, chodźżeż, inochodźcu, jednochodźca, złagodźcie, uchodźca, oswobodźcie, uchodźczy, uchodźczyni, uchodźstwo, wychodźca, wychodźczy, obchodźmy, wychodźstwo, pięciodźwięk, siedmiodźwiękowy, wielodźwięk, wielodźwiękowy, okołodźwiękowy, pochodźcie, pełnodźwięczny, wczesnodźwiękowy, podźgać, podźwięk, podźwigać, podźwignąć, podźwignięty, czterodźwięk, grodźczanin, grodźczanka, nowogrodźczanin, nowogrodźczanka, przechodźcie, rozchodźmy, uchodźca, szkodźcie, oblodźmy, głodźmy, okołodźwiękowe, posłodźmy, słodźcie, pełnodźwięcznych, podźgać, podźwigać, podźwignięcie, brodźmy, czterodźwięki, grodźmy, obrodźmy, rodźmy, smrodźcem*
Wyrazy rozpoczynające się od sekwencji *przedź*		
9	/d.z'/	–
10	/dz'/	*przedźwięczeć, przedźwięczenie, przedźwigać, przedźwigany, przedźwignąć, przedźwignięty*
Wyrazy zawierające sekwencję *przedź* w śródgłosie		
11	/d.z'/	–
12	/dz'/	*poprzedźcie, poprzedźcież, poprzedźmy, poprzedźmyż, poprzedźże, uprzedź, uprzedźcie, uprzedźcież, uprzedźmy, uprzedźmyż, uprzedźże, wyprzedź, wyprzedźcie, wyprzedźcież, wyprzedźmy, wyprzedźmyż, wyprzedźże*
Wyrazy rozpoczynające się od sekwencji *podź*		
13	/d.z'/	–
14	/dz'/	*podźgać, podźwięk, podźwigać, podźwignąć, podźwignięty*
Wyrazy zawierające sekwencję *podź* w śródgłosie		
15	/d.z'/	–
16	/dz'/	–

Wyrazy rozpoczynające się od sekwencji *ponadź*		
17	/d.z'/	–
18	/dz'/	–
Wyrazy zawierające sekwencję *ponadź* w śródgłosie		
19	/d.z'/	–
20	/dz'/	–
Wyrazy rozpoczynające się od sekwencji śródź		
21	/d.z'/	–
22	/dz'/	–
Wyrazy zawierające sekwencję *śródź* w śródgłosie		
23	/d.z'/	–
24	/dz'/	–

2.5.11. Triada ortograficzna *dzi*

Jeżeli triada ortograficzna *dzi* jest umiejscowiona przed literą samogłoskową (z wyłączeniem litery *i*), to stanowi ona trigrafem oznaczający fonem /dz'/. Natomiast przed literą oznaczającą spółgłoskę triada *dzi* oznacza sekwencję dwóch fonemów (/dz'.i/). Uzależnienie sposobu transkrypcji od kontekstu sprawiło, że triada *dzi* nie została włączona do transkrypcji podstawowej przyjętej na potrzeby tej publikacji (zob. podrozdz. 1.3).

Z triadą *dzi* związany jest omówiony już problem – może być ona uformowana w wyniku połączenia przedrostka zawierającego literę *d* na ostatniej pozycji z rdzeniem rozpoczynającym się od diady *zi*. Wynik analizy dotyczącej struktur tego typu zawarto tabeli 2.39. Użyto w niej reguł pochodzących z tabeli 25 w *Automatyzacji…* . Zaledwie trzy wiersze w tabeli (na pozycjach: 16, 17, 18) odnoszą się do kontekstu spółgłoskowego. Pozostałe wiersze dotyczą triady *dzi* w kontekście prawostronnym samogłoskowym.

Tabela 2.39 – Analiza słownika i korpusu dotycząca niestandardowej transkrypcji triady ortograficznej *dzi*

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	*#na*[*d*]*ziar*=/d/	10	*nadziarno, nadziaren, nadziarnie*	0	
2	*#o*[*d*]*ziar*=/d/	140	*odziarniać, odziarnię, odziarniony*	0	
3	*#po*[*d*]*ziar*=/d/	9	*podziarno, podziaren, podziarnie*	0	
4	*#nieo*[*d*]*ziar*=/d/	48	*nieodziarniająca, nieodziarniającą, nieodziarniające*	0	
5	*#na*[*d*]*ziem*=/d/	23	*nadziemny, nadziemna, nadziemskimi*	10	*nadziemskim, nadziemskiej, nadziemnych*
6	*#o*[*d*]*ziem*=/d/	32	*odziemny, odziemski, odziemscy*	0	
7	*#po*[*d*]*ziem*=/d/	20	*podziemie, podziemni, podziemnie*	321	*podziemny, podziemi, podziemnego*
8	*#niena*[*d*]*ziem*=/d/	23	*nienadziemna, nienadziemną, nienadziemne*	0	
9	*#nieo*[*d*]*ziem*=/d/	22	*nieodziemna, nieodziemną, nieodziemne*	0	
10	*#pona*[*d*]*ziem*=/d/	22	*ponadziemski, ponadziemskim, ponadziemskimi*	0	
11	*#śró*[*d*]*ziem*=/d/	105	*śródziemne, śródziemnego, śródziemnej*	5	*śródziemnego, śródziemnym, śródziemnomorskich*
12	*#na*[*d*]*ziom*=/d/	11	*nadziomek, nadziomka, nadziomkach*	0	
13	*#o*[*d*]*ziom*=/d/	21	*odziomek, odziomka, odziomkach*	0	
14	*#po*[*d*]*ziom*=/d/	10	*podziomek, podziomka, podziomkach*	0	

15	*#nieo*[*d*]*ziom*=/d/	11	*nieodziomkowa, nieodziomkową, nieodziomkowe*		
16	*#o*[*d*]*zip*=/d/	49	*odzipnąć, odzipną, odzipnął*	0	
17	*#nieo*[*d*]*zip*=/d/	8	*nieodzipnięcia, nieodzipnięciach, nieodzipnięciami*	0	
18	*#prze*[*d*]*zim*=/d/	29	*przedzimek, przedzimowy, przedzimowa*	1	*przedzimie*

Wykonano szczegółową analizę związaną z różnymi możliwościami transkrypcji triady *dzi* (w zależności od struktury morfologicznej wyrazu). Badanie zostało podzielone na dwa etapy. Pierwszy etap dotyczył kontekstu prawostronnego obejmującego spółgłoskę (tabela 2.40), w etapie drugim został uwzględniony kontekst samogłoskowy. Słownik *SJP.pl* zawiera tylko cztery formy podstawowe, w których po triadzie ortograficznej *dzi* występuje litera oznaczająca spółgłoskę i jednocześnie litery *d* oraz *z* oznaczają sekwencję dwóch fonemów (cała triada stanowi sekwencję trzech fonemów). Są to następujące wyrazy: *przedzimek, przedzimie, przedzimowy, odzipnąć*. A więc można wyznaczyć zbiór nagłosowych struktur ortograficznych umożliwiających identyfikację wszystkich form fleksyjnych, w których triada *dzi* nie powinna być transkrybowana jako ciąg fonemów /dz'.i/: *#przedzim, #odzip.* Ten zbiór można rozszerzyć o ciągi: *#nieprzedzim, #nieodzip.*

Tabela 2.40 – Szczegółowa analiza słownika dotycząca dwóch możliwości transkrypcji ortograficznej triady *dzi* umiejscowionej przed literą oznaczającą spółgłoskę

Lp.	Transkr.	Wykaz wyrazów
Wyrazy rozpoczynające się od sekwencji *nadzi*		
1	/d.z'.i/	–
2	/dz'.i/	*Nadziny* (należący do Nadii), *nadziwić*
Wyrazy zawierające sekwencję *nadzi* w śródgłosie		
3	/d.z'.i/	–
4	/dz'.i/	–
Wyrazy rozpoczynające się od sekwencji *odzi*		
5	/d.z'.i/	*odzipnąć*
6	/dz'.i/	–
Wyrazy zawierające sekwencję *odzi* w śródgłosie		
7	/d.z'.i/	–
8	/dz'.i/	*oswobodziciel, oswobodzicielka,* **Ś***wiebodzice, świebodzicki, świebodziczanin, świebodziczanka, oswobodzić, wyswobodzić, Świebodzin, świebodzinianin, świebodzinianka, Łobodziński, świebodziński, bodziszek, bodziszkowate, bodziszkowaty, jagodzica, Nagodzice, Przygodzice, przygodzicki, dogodzić, godzić, łagodzić, pogodzić, rozpogodzić, ugodzić, ułagodzić, wygodzić, wypogodzić, załagodzić, zgodzić, złagodzić, Godzik, Godzikowice, Godzilla, Godzimir, Godzimirostwo, Godzimirowie, Godzimirowy, amperogodzina, całogodzinny, cogodzinny, czterdziestodwugodzinny, czterdziestogodzinny, czterdziestojednogodzinny, czterdziestoośmiogodzinny, czternastogodzinny, czterogodzinny, czteroipółgodzinny, ćwierćgodzinny, dupogodziny, dwudziestoczterogodzinny, dwudziestodwugodzinny, dwudziestogodzinny, dwudziestojednogodzinny, dwudziestopięciogodzinny, dwudziestotrzygodzinny, dwugodzinny, dwugodzinowy, dwuipółgodzinny, dwunastogodzinny, dziesięciogodzinny, dziewięciogodzinny, dziewięcioipółgodzinny, dziewięćdziesięciogodzinny, dziewiętnastogodzinny, godzina, godzinami, godzineczka, godzinka, godzinki, godzinny, godzinowy, godzinówka, ilugodzinny, Jagodziny, jedenastogodzinny, jednogodzinny, jednogodzinowy, kilkogodzinny, kilkonastogodzinny, kilkugodzinny, kilkunastogodzinny, kilowatogodzina, koniogodzina, luksogodzina, lumenogodzina, Łagodzin, maszynogodzina, megawatogodzina, motogodzina, nadgodzina, nadgodzinny, nadgodzinowy, normogodzina, osiemdziesięciogodzinny, osiemnastogodzinny, osobogodzina, ośmiogodzinny, ośmioipółgodzinny, parogodzinny, pięciogodzinny, pięcioipółgodzinny, pięćdziesięciogodzinny, piętnastogodzinny, ponaddwugodzinny, ponaddwuipółgodzinny, ponadgodzinny, ponadpółgodzinny, półgodzina, półgodzinka, półgodzinny, półtoragodzinny, pracogodzina, roboczogodzina, robotnikogodzina, siedemdziesięciogodzinny, siedemnastogodzinny, siedmiogodzinny, siedmioipółgodzinny, szesnastogodzinny, sześciogodzinny, sześcioipółgodzinny, sześćdziesięciogodzinny, trzydziestogodzinny, trzygodzinny, trzyipółgodzinny, trzynastogodzinny, tylogodzinny, tylugodzinny, watogodzina, wielogodzinny, Godziński, Jagodziński, Łagodziński, Pogodziński, Przygodziński, Zgodziński, Godzisław,*

8	/dz'.i/	*Godzisławostwo, Godzisławowie, Godzisławowy, Godzisz, Godziszewo, Godziszewski, Godziszka, Godziszostwo, Godziszowie, godziszowski, Godziszowy, Godziszów, jagodzisko, godziwiej, godziwość, godziwszy, godziwy, najgodziwiej, najgodziwszy, najniegodziwiej, najniegodziwszy, niegodziwiec, niegodziwiej, niegodziwość, niegodziwszy, chodzić, dochodzić, nachodzić, nadchodzić, naschodzić, obchodzić, odchodzić, pochodzić, podchodzić, podochodzić, poodchodzić, poprzechodzić, poprzychodzić, porozchodzić, poschodzić, powchodzić, powschodzić, powychodzić, pozachodzić, przechodzić, przychodzić, rozchodzić, schodzić, uchodzić, wchodzić, wschodzić, wychodzić, zachodzić, znachodzić, chodzik, samochodzik, chodziłbych, samochodzina, Chodziński, samochodzisko, miodzik, jodzica, naszkodzić, przeszkodzić, szkodzić, uszkodzić, zaszkodzić, Szkodziński, lodzić, oblodzić, odlodzić, zalodzić, lodzik, Lodziński, młodziczka, Włodzice Małe, chłodzić, dosłodzić, głodzić, napłodzić, ochłodzić, odchłodzić, odmłodzić, osłodzić, płodzić, posłodzić, przechłodzić, przegłodzić, przesłodzić, przysłodzić, rozpłodzić, schłodzić, słodzić, spłodzić, wychłodzić, wygłodzić, wysłodzić, zagłodzić, zasłodzić, chłodzik, łodzik, łodzikowate, łodzikowaty, młodzik, młodzik, młodzik, młodzikowo, Młodzikówko, słodzik, Włodzimiera, Włodzimierka, włodzimierski, Włodzimierz, włodzimierzanin, włodzimierzanka, Włodzimierzostwo, Włodzimierzowie, Włodzimierzowy, Włodzimierzyny, Włodzimirów, kłodzina, słodziny, Włodziny, wysłodziny, Kłodziński, Kołodziński, Łodziński, Młodziński, kłodzisko, Włodzisław, Włodzisława, Włodzisławczyny, Włodzisławek, Włodzisławiny, włodzisławka, Włodzisławkowy, Włodzisławostwo, Włodzisławowie, Włodzisławowy, chłodziwo, młodziwo, smarochłodziwo, modzić, wymodzić, podziczały, podziczeć, spodzik, napodziwiać, podziw, podziwiać, podziwić, popodziwiać, Bogarodzica, Bogurodzica, grodzica, Grodzicki, grodziczeński, Grodziczno, odrodziciel, odrodzicielka, prarodzic, prarodzicielka, rodzic, rodzice, rodziciel, rodziciele, rodzicielka, rodzicielski, rodzicielsko-nauczycielski, rodzicielstwo, Zbrodzice, brodzić, dogrodzić, grodzić, nagrodzić, narodzić, nasmrodzić, obrodzić, odgrodzić, odrodzić, ogrodzić, pogrodzić, porodzić, przebrodzić, przegrodzić, przerodzić, rodzić, rozgrodzić, rozrodzić, smrodzić, urodzić, wrodzić, wygrodzić, wynagrodzić, wyrodzić, zagrodzić, zasmrodzić, zesmrodzić, zrodzić, brodzik, dobrodzika, Rodzik, starodzikowiecki, starodzikowski, rodzimak, rodzimek, rodzimiak, rodzimość, rodzimy, antyrodzinność, antyrodzinny, czterorodzinny, dwurodzinny, jednorodzinny, kilkorodzinny, kilkurodzinny, małżeńsko-rodzinny, nadrodzina, narodzinowy, narodziny, nierodzinność, podrodzina, porodziny, pozarodzinny, prorodzinny, pseudorodzina, rodzina, rodzinka, rodzinne, rodzinnokoleżeński, rodzinnoprawny, rodzinność, rodzinny, urodzinki, urodzinowoimieninowy, urodzinowy, urodziny, wewnątrzrodzinny, wielorodzinny, Brodziński, Grodziński, Rodziński, Brodzisław, Brodzisława, Brodzisławiny, Brodzisławostwo, Brodzisławowie, Brodzisławowy, Brodziszów, Grodzisk, grodziski, Grodzisko, grodziskowy, Grodzisław, Grodzisławostwo, Grodzisławowie, Grodzisławowy, grodziszczanin, grodziszczanka, grodziszcze, Grodziszczko, mazowieckogrodziski, najurodziwiej, najurodziwszy, nieurodziwość, retrodziwołąg, urodziwiec, urodziwiej, urodziwość, urodziwszy, urodziwy, Nawodzice, odwodziciel, przywodziciel, uwodziciel, uwodzicielka, uwodzicielski, uwodzicielstwo, wodzicielka, Wodzicki, wojewodzic, dowodzić, obwodzić, odwodzić, powodzić, pozawodzić, pozwodzić, przewodzić, przywodzić, rozwodzić, uwodzić, wodzić, wwodzić, wywodzi**ć**, zawodzić, zwodzić, wodzidło, dowodzik, przewodzik, wodzik, Wodzinek, Wodzinowski, wojewodzina, Wodziński, wodzirej, wodzirejka, Wodzisław, Wodzisława, Wodzisławek, wodzisławianin, wodzisławianka, Wodzisławiny, Wodzisławka, Wodzisławkowy, Wodzisławostwo, Wodzisławowie, Wodzisławowy, wodzisławski, epizodzik, przodzik, wrzodzik, epizodzista*
Wyrazy rozpoczynające się od sekwencji *przedzi*		
9	/d.z'.i/	*przedzimek, przedzimie, przedzimowy*
10	/dz'.i/	*przedziwniej, przedziwność, przedziwny*
Wyrazy zawierające sekwencję *przedzi* w śródgłosie		
11	/d.z'.i/	–
12	/dz'.i/	*poprzedzić, uprzedzić, wyprzedzić*
Wyrazy rozpoczynające się od sekwencji *podzi*		
13	/d.z'.i/	–
14	/dz'.i/	*podziczały, podziczeć, podziw, podziwiać, podziwić*
Wyrazy zawierające sekwencję *podzi* w śródgłosie		
15	/d.z'.i/	–
16	/dz'.i/	*napodziwiać, popodziwiać, spodzik*
Wyrazy rozpoczynające się od sekwencji *śródzi*		
17	/d.z'.i/	–
18	/dz'.i/	–
Wyrazy zawierające sekwencję *śródzi* w śródgłosie		
19	/d.z'.i/	–
20	/dz'.i/	–

Tabela 2.41 dotyczy różnych możliwości transkrypcji ortograficznej triady *dzi* w kontekście prawostronnym samogłoskowym. Można zauważyć, że często oznacza ona sekwencję fonemów (/d.z'/) ze względu na strukturę morfologiczną. Wyznaczenie nagłosowych ciągów ortograficznych, które umożliwiają identyfikację takich wyrazów, niekiedy jest utrudnione. Ma to związek z faktem występowania znacznej liczby wyrazów, w których triada *dzi* oznacza jeden fonem /dz'/. Dlatego ciągi powinny być na tyle długie, żeby na ich podstawie można było zidentyfikować tylko i wyłącznie formy fleksyjne, w których triada *dzi* oznacza sekwencję fonemów /d.z'/.

Na podstawie omówionej analizy wyznaczono następujący zbiór konstrukcji w nagłosie wyrazu: *#nadziar, #nadziem, #nadziom, #odziar, #odziem, #odziom, #odzip, #podziar, #podziem, #podziom, #podzięb, #ponadziem, #śródziem, #nadśródziem, #wschodniośródziem*. Poza tym ten zbiór można powiększyć o następujące struktury zawierające przedrostek *nie*: *#nienadziar, #nienadziem, #nienadziom, #nieodziar, #nieodziem, #nieodziom, #nieodzip, #niepodziar, #niepodziem, #niepodziom, #niepodzięb, #nieponadziem, #nieśródziem, #nienadśródziem, #niewschodniośródziem.*

Podsumowując część dotyczącą różnych możliwości transkrypcji triady *dzi*, trzeba zaznaczyć, że żaden kontekst nie jest dla niej ubezdźwięczniający (ze względu na literę *i* na ostatniej pozycji).

Tabela 2.41 – Szczegółowa analiza słownika dotycząca dwóch możliwości transkrypcji triady ortograficznej *dzi* umiejscowionej przed literą oznaczającą samogłoskę

Lp.	Transkr.	Wykaz wyrazów
Wyrazy rozpoczynające się od sekwencji *nadzi*		
1	/d.z'/	*nadziarno, nadziemny, nadziemski, nadziomek*
2	/dz'/	*Nadzia, nadziać, nadziak, nadział, nadziałka, nadziany, nadziąślak, nadzieja, nadzielać, nadzielić, nadzienie, nadziewacz, nadziewać, nadziewaka, nadziewany, nadzionko nadziewarka, nadziewka, nadziękować*
Wyrazy zawierające sekwencję *nadzi* w śródgłosie		
3	/d.z'/	*ponadziemski**
4	/dz'/	*esplanadzie, Granadzie, Kanadzie, cytrynadzie, nienadziana, żelazowanadzie, dziecinadzie, gonadzie, kanonadzie, monadzie, ponadzierać, ponadziewać, beznadziei, beznadziejność, beznadziejny*
Wyrazy rozpoczynające się od sekwencji *odzi*		
5	/d.z'/	*odziarnić, odziarniać, odziarnianie, odziarniarka, odziemek, odziemny, odziemski odziomek, odziomkowy*
6	/dz'/	*odziać, odziewać, odzienie, odziedziczyć, odziedziczać, odziedziczalność, odziedziczenie, odziedziczony, odzier, odzierać, odzierek, odziewać, odziewek, odzież, odzieżowiec, odzieżownictwo, odzieżowy, odzieżówka, odzierze*
Wyrazy zawierające sekwencję *odzi* w śródgłosie		
7	/d.z'/	*podziarno*, podziomek*, podziemie*, podziemny*, podziębić*, podziębiony**
8	/dz'/	*lód – lodzie, wschód – wchodzie, głód – głodzie* *bodziak, Bodzianowski, antybodziec, bodziec, Bodziejowice, chorobodzień, elektrobodziec, osobodzień, przeciwbodziec, Bodzio, Bodzioch, grubodzioby, grubodziób, codzienność, codzienny, niecodzienność, lododział, wododział, wododziałowy, dodzierać, każdodziennie, każdodzienny, przeodziać, godzianowski, Godzianów, jagodzian, jagodzianka, godzieski, Godziesze Wielkie, Godzieszów, długodzioby, długodziób, chodziarka, chodziarski, chodziarz, międzychodzianin, międzychodzianka, samochodziarz, chodzieski, Chodzież, chodzieżanin, chodzieżanka, inochodziec, jednochodziec, podschodzie, przechodzień, przychodzień, osiemdziesięciodziewięcioletni, pięciodzielny, średniodzietny, miodzio, rękodzielnictwo, rękodzielniczka, rękodzielniczy, rękodzielnik, rękodzieło, blaszkodziobe, blaszkodzioby, krótkodzióbkowy, miękkodziób, trzewikodziób, Lodzia, lodziara, lodziarka, lodziarnia, lodziarski, lodziarstwo, lodziarz, wielodziałowy, wielodziedzinowy, wielodzielny, wielodzietna, wielodzietność, wielodzietny, szablodziób, chłodziarka, chłodziarka-zamrażarka, chłodziarkozamrażarka, chłodziarstwo, chłodziarz, inowłodzianin, inowłodzianka, łodzianin, łodzianka, młodziak, młodzian, młodzianek, Młodzianowski, słodziak, Włodzia, całodzienny, chołodziec, kołodziej, Kołodziejak, Kołodziejczak, Kołodziejczyk, Kołodziejek, Kołodziejewo, kołodziejnia, Kołodziejski, kołodziejstwo, małodzietność, małodzietny, młodzie, młodziej, Młodziejewski, Młodziejowice, Młodziejów, młodzieniaszek, młodzieniec, młodzieńczość, młodzieńczy, Młodzieszyn, Młodzieszynek, młodzieszyński, młodzież, młodzieżowiec, młodzieżowość, młodzieżowy, młodzieżówa, młodzieżówka, najmłodziej, najsłodziej, słodziej, słodzieńki, Tułodziecki, Włodzienin, złodziej, złodziejaszek, złodziejka, złodziejski, złodziejstwo, łodziowy, piłodziobe, piłodzioby, Włodzio, Włodziowy, piłodziób, młodziuchny, młodziusieńki, młodziuśki, młodziuteńki, młodziutki, słodziuchny, słodziusieńki, słodziuśki, słodziuteńki, słodziutki, samodział, samodziałowy, najniesamodzielniej, najsamodzielniej, najsamodzielniejszy, niesamodzielność, samodzielniej, samodzielniejszy, samodzielność, samodzielny, samodzierżawca, samodzierżawie, samodzierżca, usamodzielniać, usamodzielnić, jednodzielny, jednodzietność, jednodzietny, kaznodzieja, kaznodziejka, kaznodziejski, kaznodziejstwo, telekaznodzieja, wczesnodziecięcy, wczesnodziejowy, czerwonodzioby,*

8	/dz'/	*czwórpodział, ekstraniespodzianka, niespodzianka, międzypodziałowy, niespodzianka, niespodziany, podziabać, podziać, podział, podziałać, podziałka, podziałowy, podziałówka, samopodział, trójpodział, zapodziać, nadspodziewany, najniespodziewaniej, najpodzielniejszy, niepodzielność, niespodziewaniej, niespodziewaność, niespodziewany, podzielać, podzielić, podzielnia, podzielnica, podzielniejszy, podzielnik, podzielność, podzielny, podziewać, spodziectwo, spodziewać, spodziewany, trójpodzielony, zapodziewać, podzięka, podziękować, podziękowanie, podzięce, podziobać, podzióbać, sierpodziób, podziurawić, podziurkować, baligrodzianin, baligrodzianka, Brodziak, brodzianin, brodzianka, ciemnogrodzianin, ciemnogrodzianka, czterodziałaniowy, krasnobrodzianin, krasnobrodzianka, nowogrodzianin, nowogrodzianka, nowożmigrodzianin, nowożmigrodzianka, podgrodzianin, podgrodzianka, rajgrodzianin, rajgrodzianka, siedmiogrodzianin, siedmiogrodzianka, środzianin, środzianka, tarnogrodzianin, tarnogrodzianka, wyszogrodzianin, wyszogrodzianka, żmigrodzianin, żmigrodzianka, borodziej, brodziec, brodziecki, czarodziej, czarodziejka, czarodziejski, czarodziejstwo, czterodzielny, Dnieprodzierżyńsk, dnieprodzierżyński, dobrodziej, dobrodziejka, dobrodziejstwo, dobrodzienianin, dobrodzienianka, Dobrodzień, dobrodzieński, Grodziec, grodziecki, grodzieński, Grodzieńszczyzna, Międzybrodzie Żywieckie, Międzybrodzie Bialskie, mikrodzielnica, Nowogrodziec, nowogrodziecki, Ogrodzieniec, ogrodzieniecki, ogrodzieńczanin, ogrodzieńczanka, Podegrodzie, podgrodzie, prodziekan, prodziekański, przedgrodzie, przygrodzie, Rodziewicz, Rodziewiczówna, smrodziec, smrodzieniec, smrodzieńcowy, starodzierzgoński, taborodzień, Zabrodzie, zagrodzieński, zarodziec, zbrodzień, grodziowy, ostrodziobe, ostrodzioby, zarodziowe, zarodziowy, ostrodziób, Borodziuk, Bartodziej, dłoniastodzielny, dwudziestodziewięcioletni, pierzastodzielny, trzydziestodziewięcioletni, żółtodziób, ciepłowodzianin, ciepłowodzianka, podwodziarz, powodzianin, powodzianka, wodzian, wodzianka, wodziarka, wojewodzianka, małowodzie, Międzywodzie, nadwodzie, podwodzie, przewodzień, wielowodzie, Wodzierady, wodzieradzki, brzytwodzioby, popowodziowy, powodziowy, przeciwpowodziowo, przeciwpowodziowy, wodzionka, brzytwodziób, krzywodziób, pochwodziób, przyodziać, przyodzienie, przyodziewać, przyodziewek, owrzodziały, wrzodzianka, dzisiejszodzienny, przodzie, owrzodzieć, roboczodzień, wrzodzieć, kaczodzioby, krzyżodziób, tekstylno-odzieżowy, osiemdziesięciodziewięcioletnim*
Wyrazy rozpoczynające się od sekwencji *przedzi*		
9	/d.z'/	–
10	/dz'/	*przedział, przedziałek, przedziałowość, przedziałowy, przedzie, przedzielać, przedzielić, przedzielenie, przedzierać, przedzierzgać, przedzierzgnąć, przedzierzgnięcie przedziesiątkować, przedziurawiać, przedziurawić, przedziurkować*
Wyrazy zawierające sekwencję *przedzi* w śródgłosie		
11	/d.z'/	–
12	/dz'/	*poprzedzielać, bezprzedziałowy*
Wyrazy rozpoczynające się od sekwencji *podzi*		
13	/d.z'/	*podziarno*, podziomek*, podziemie*, podziemny*, podziębić*, podziębiony**
14	/dz'/	*podziabać, podziać, podzianie, podział, podziałać, podziałka, podziałowy, podziałówka podzielać, podzielić, podzielnia, podzielnica, podzielniejszy, podzielnik, podzielność, podzielny, podziewać, podziobanie, podziębić, podziębiony, podzięka, podziękować, podziękowanie, podzięce, podziobać, podzióbać, podziurawić, podziurawienie, podziurawiony, podziurkować, podziurkowanie, podziurkowany*
Wyrazy zawierające sekwencję *podzi* w śródgłosie		
15	/d.z'/	–
16	/dz'/	*zapodziać, zapodziewać, niepodzielność, trójpodział, najpodzielniejszy, trójpodzielony, samopodział, czwórpodział, sierpodziób, ekstraniespodzianka, niespodzianka, niespodzianka, niespodziany, nadspodziewany, najniespodziewaniej, niespodziewaniej, niespodziewaność, niespodziewany, spodziectwo, spodziewać, spodziewany, międzypodziałowy, antypodzie*
Wyrazy rozpoczynające się od sekwencji *ponadzi*		
17	/d.z'/	*ponadziemski**
18	/dz'/	*ponadzierać, ponadziewać*
Wyrazy zawierające sekwencję *ponadzi* w śródgłosie		
19	/d.z'/	–
20	/dz'/	*tamponadzie*
Wyrazy rozpoczynające się od sekwencji *śródzi*		
21	/d.z'/	*śródziemne, śródziemnomorski, śródziemnomorskość*
22	/dz'/	–
Wyrazy zawierające sekwencję *śródzi* w śródgłosie		
23	/d.z'/	*nadśródziemnomorski, wschodniośródziemnomorski*
24	/dz'/	–

2.5.12. Diada ortograficzna *rz*

W przyjętej transkrypcji podstawowej diada ortograficzna *rz* oznacza dźwięczny szczelinowy fonem /rz/. Jednak częściej jest ona transkrybowana w wariancie bezdźwięcznym – jako fonem /S/. Obszerna analiza związana z tym zagadnieniem została omówiona w podrozdziale 2.12 w *Asymilacji…* .

Z transkrypcją diady *rz* powiązane jest istotne utrudnienie – w niektórych wyrazach nie jest ona digrafemem, tylko oznacza ciąg fonemów: /r.z/. W wypadku niektórych form fleksyjnych diada *rz* może oznaczać zarówno jeden, jak i dwa fonemy. Różnica między tymi możliwościami ma związek ze znaczeniem wyrazu. W dalszym ciągu przedstawiono wynik szczegółowej analizy tego problemu. Uwzględniono w niej zbiór struktur nagłosu, który został utworzony na podstawie informacji przedstawionych w *Automatyzacji…* . Badanie oparto przede wszystkim na słowniku *SJP.pl*, który zawiera informację o wszystkich formach fleksyjnych należących do poszczególnych form podstawowych wyrazów (związanych z określonymi znaczeniami). Dzięki temu można było wyznaczyć podzbiory form, w których diada ortograficzna *rz*:

– zawsze oznacza fonem /rz/;
– zawsze oznacza ciąg fonemów /r.z/;
– zawsze oznacza ciąg fonemów /r.z'/ (przed literą *i*);
– może oznaczać zarówno fonem /rz/, jak i sekwencję fonemów /r.z/ (w zależności od znaczenia wyrazu).

Na tej podstawie wyznaczono nagłosowe ciągi ortograficzne, które umożliwiają identyfikację wszystkich wyrazów należących do wymienionych podzbiorów. Są one niezbędne dla budowy reguł w systemie transkrypcji opartym na modelu wielowarstwowym. Taki sposób wyznaczania reguł stanowi alternatywę dla sposobu przedstawionego w *Automatyzacji…*, gdzie są one oparte na dodatkowej czynności wstawiania rozdzielaczy (znaczników) – w rozwiązaniach prezentowanych w niniejszej publikacji nie stosuje się takiego podejścia.

W opisach ujętych w częściach 2.5.12.1 do 2.5.12.16 zostały użyte definicje wyrazów pochodzące ze słownika *SJP.pl* oraz z internetowej wersji *Słownika języka polskiego PWN,* który powstał na podstawie *Słownika 100 tysięcy potrzebnych słów* pod red. J. Bralczyka [Bralczyk 2005]. Poza tym w tabelach 2.42–2.55 w kolumnie: *Transkrypcja diady rz* nie uwzględniono faktu, że w nielicznych wypadkach fonem /rz/ lub fonem /z/ (w sekwencji fonemów /r.z/) mogą znajdować się w kontekście ubezdźwięczniającym (na przykład w wygłosie absolutnym wyrazu).

2.5.12.1. Diada ortograficzna *rz* w strukturze: *#marz*

Ortograficzna diada *rz* umiejscowiona w nagłosowym ciągu *#marz* może oznaczać fonem /rz/ lub ciąg fonemów /r.z/. Słownik *SJP.pl* nie zawiera form fleksyjnych, które mogłyby być transkrybowane zgodnie z dwoma wymienionymi możliwościami. Diada *rz* powinna być transkrybowana jako fonem /rz/ w wyrazach: *marze* (miejscownik od: *mara* – „zjawa, widzenie senne" lub „gryzoń żyjący w Ameryce Południowej"); *marzenie, marzyć* i *marzyciel*; *marzanka*, *marzana*, *marzyca*, *marzymięta* (rośliny), *marzanna* („kukła lub roślina z rodziny marzanowatych"), *marzanowate* („rodzina roślin okrytonasiennych"); *Marzanna* („imię w mitologii słowiańskiej"), *Marzena i Marzenka* (imiona); *Marzęcice*, *Marzęcin* i *Marzęcino* (nazwy miejscowości); *Marzęcki* oraz *Marzyński* (nazwiska). Natomiast transkrypcja diady *rz* jako sekwencja fonemów /r.z/ jest właściwa dla wyrazu *marzłoć* („wieczna zmarzlina") oraz dla wyrazu *marznąć*. Można też wspomnieć o wyjątkowej transkrypcji wyrazu obcego: *marziale* – /m.a.r.ts.j.a.l.e/ (ten termin „odnosi się do tempa, rytmu lub też dynamiki utworu muzycznego").

Na podstawie analizy ujętej w tabeli 2.42 można wyznaczyć zbiór nagłosowych ciągów ortograficznych umożliwiających identyfikację form fleksyjnych, w których diada ortograficzna *rz* powinna być transkrybowana jako sekwencja fonemów /r.z/: *#marzł, #marzn*.

Tabela 2.42 – Analiza słownika dotycząca struktury: *#marz*

Transkrypcja diady *rz*	Wykaz form fleksyjnych
/rz/	*marze, marzę, marzana, marzan, marzanach, marzanami, marzaną, marzanę, marzanie, marzano, marzanom, marzany, marzanka, marzance, marzanek, marzankach, marzankami, marzanką, marzankę, marzanki, marzanko, marzankom, Marzanna, Marzann, Marzannach, Marzannami, Marzanną, Marzannę, Marzannie, Marzanno, Marzannom, Marzanny, Marzanniny, Marzannin, Marzannina, Marzanniną, Marzannine, Marzanninego, Marzanninej, Marzanninemu, Marzannini, Marzanninych, Marzanninym, Marzanninymi, marzannowate, marzannowatych, marzannowatym, marzannowatymi, marzannowaty, marzannowaci, marzannowata, marzannowatą, marzannowate, marzannowatego, marzannowatej, marzannowatemu, marzannowatych, marzannowatym, marzannowatymi, marzanowate, marzanowatych, marzanowatym, marzanowatymi, marzanowaty, marzanowaci, marzanowata, marzanowatą, marzanowate, marzanowatego, marzanowatej, marzanowatemu, marzanowatych, marzanowatym, marzanowatymi, marzanowiec, marzanowca, marzanowcach, marzanowcami, marzanowce, marzanowcem, marzanowcom, marzanowcowi, marzanowców, marzanowcu, marzący, marząca, marzącą, marzące, marzącego, marzącej, marzącemu, marząco, marzących, marzącym, marzącymi, marzec, marzec, Marzena, Marzen, Marzenach, Marzenami, Marzeną, Marzenę, Marzenie, Marzeno, Marzenom, Marzeny, Marzenie, marzenia, marzeniach, marzeniami, marzeniem, marzeniom, marzeniu, marzeń, Marzenin, Marzenina, Marzeninem, Marzeninie, Marzeninowi, Marzeniny, Marzenin, Marzenina, Marzeniną, Marzenine, Marzeninego, Marzeninej, Marzeninemu, Marzenini, Marzeninych, Marzeninym, Marzeninymi, marzeniowy, marzeniowa, marzeniową, marzeniowe, marzeniowego, marzeniowej, marzeniowemu, marzeniowi, marzeniowych, marzeniowym, marzeniowymi, Marzenka, Marzence, Marzenek, Marzenkach, Marzenkami, Marzenką, Marzenkę, Marzenki, Marzenko, Marzenkom, Marzęcice, Marzęcic, Marzęcicach, Marzęcicami, Marzęcicom, Marzęcin, Marzęcina, Marzęcinem, Marzęcinie, Marzęcinowi, Marzęcino, Marzęcina, Marzęcinem, Marzęcinie, Marzęcinu, Marzęcki, Marzęccy, Marzęcka, Marzęcką, Marzęckich, Marzęckie, Marzęckiego, Marzęckiej, Marzęckiemu, Marzęckim, Marzęckimi, Marzena, Marzeną, Marzenę, Marzenie, Marzeno, Marzeny, Marzenach, Marzenami, Marzenom, Marzen, marzyca, marzyc, marzycach, marzycami, marzycą, marzyce, marzycę, marzyco, marzycom, marzycy, marzyciel, marzyciela, marzycielach, marzycielami, marzyciele, marzycielem, marzycieli, marzycielom, marzycielowi, marzycielu, marzycielka, marzycielce, marzycielek, marzycielkach, marzycielkami, marzycielką, marzycielkę, marzycielki, marzycielko, marzycielkom, marzycielski, marzycielscy, marzycielska, marzycielską, marzycielskich, marzycielskie, marzycielskiego, marzycielskiej, marzycielskiemu, marzycielskim, marzycielskimi, marzycielsko, marzycielsku, marzycielskość, marzycielskości, marzycielskościach, marzycielskościami, marzycielskością, marzycielskościom, marzycielstwo, marzycielstw, marzycielstwa, marzycielstwach, marzycielstwami, marzycielstwem, marzycielstwie, marzycielstwom, marzycielstwu, marzyć, marz, marzą, marząc, marząca, marzącą, marzące, marzącego, marzącej, marzącemu, marzący, marzących, marzącym, marzącymi, marzcie, marzcież, marzenia, marzeniach, marzeniami, marzenie, marzeniem, marzeniom, marzeniu, marzeń, marzę, marzmy, marzmyż, marzono, marzy, marzycie, marzyli, marzyliby, marzylibyście, marzylibyśmy, marzyliście, marzyliśmy, marzył, marzyła, marzyłaby, marzyłabym, marzyłabyś, marzyłam, marzyłaś, marzyłby, marzyłbym, marzyłbyś, marzyłem, marzyłeś, marzyło, marzyłoby, marzyły, marzyłyby, marzyłybyście, marzyłybyśmy, marzyłyście, marzyłyśmy, marzymy, marzysz, marzże, marzymięta, marzymięcie, marzymięt, marzymiętach, marzymiętami, marzymiętą, marzymiętę, marzymięto, marzymiętom, marzymięty, Marzyński, Marzyńscy, Marzyńska, Marzyńską, Marzyńskich, Marzyńskie, Marzyńskiego, Marzyńskiej, Marzyńskiemu, Marzyńskim, Marzyńskimi*
/r.z/	*marzłoć, marzłoci, marzłociach, marzłociami, marzłocią, marzłocie, marzłociom, marzłota, marzłocie, marzłot, marzłotach, marzłotami, marzłotą, marzłotę, marzłoto, marzłotom, marzłoty, marznąć, marzli, marzliby, marzlibyście, marzlibyśmy, marzliście, marzliśmy, marzł, marzła, marzłaby, marzłabym, marzłabyś, marzłam, marzłaś, marzłby, marzłbym, marzłbyś, marzłem, marzłeś, marzło, marzłoby, marzły, marzłyby, marzłybyście, marzłybyśmy, marzłyście, marzłyśmy, marzną, marznąc, marznąca, marznącą, marznące, marznącego, marznącej, marznącemu, marznący, marznących, marznącym, marznącymi, marznął, marznąłby, marznąłbym, marznąłbyś, marznąłem, marznąłeś, marznę, marznęli, marznęliby, marznęlibyście, marznęlibyśmy, marznęliście, marznęliśmy, marznie, marzniecie, marzniemy, marzniesz, marznięcia, marznięciach, marznięciami, marznięcie, marznięciem, marznięciom, marznięciu, marznięć, marznięto, marznij, marznijcie, marznijcież, marznijmy, marznijmyż, marznijże, marziale*
/rz/ lub /r.z/	–

2.5.12.2. Diada ortograficzna *rz* w strukturze: *#domarz*

Sekwencja *#domarz* występuje w czasowniku *domarzać*, który ma dwa znaczenia różniące się wymową. Pierwsza możliwość to: „morzyć całkowicie, do reszty" – w tym znaczeniu wyrazu diada *rz* powinna być transkrybowana jako fonem /rz/. Znaczenie drugie to: „marznąć całkowicie, do reszty" – w tym wypadku właściwą transkrypcją diady *rz* jest ciąg fonemów /r.z/. Jako sekwencja fonemów /r.z/ diada *rz* powinna być transkrybowana także w wypadku wszystkich form wyrazów: *domarznąć, domarznięcie, domarznięty.*

Z danych zamieszczonych w tabeli 2.43 wynika, że w języku polskim nie występują formy, w których diada *rz* (w nagłosie *domarz*) oznaczałaby tylko fonem /rz/. Identyfikacja wszystkich wyrazów zawierających diadę

rz oznaczającą ciąg fonemów /r.z/ jest możliwa na podstawie ciągów ortograficznych: *#domarzn, #domarzł*. Formy, które można transkrybować zgodnie z dwoma omówionymi sposobami zawierają sekwencję *#domarza*.

Tabela 2.43 – Analiza słownika dotycząca struktury: *#domarz*

Transkrypcja diady *rz*	Wykaz form fleksyjnych
/rz/	–
/r.z/	*domarznij, domarznijcie, domarznijcież, domarznijmy, domarznijmyż, domarznijże, domarzli, domarzliby, domarzlibyście, domarzlibyśmy, domarzliście, domarzliśmy, domarzł, domarzła, domarzłaby, domarzłabym, domarzłabyś, domarzłam, domarzłaś, domarzłby, domarzłbym, domarzłbyś, domarzłem, domarzłeś, domarzło, domarzłoby, domarzły, domarzłyby, domarzłybyście, domarzłybyśmy, domarzłyście, domarzłyśmy, domarzną, domarznę, domarznie, domarzniecie, domarzniemy, domarzniesz, domarznięto, domarznięcia, domarznięciach, domarznięciami, domarznięcie, domarznięciem, domarznięciom, domarznięciu, domarznięć, domarzłszy, domarznął, domarznąłby, domarznąłbym, domarznąłbyś, domarznąłem, domarznąłeś, domarznąwszy, domarzniętego, domarzniętemu, domarzniętych, domarzniętym, domarzniętymi, domarznięci, domarznięta, domarzniętą, domarznięte, domarzniętej*
/rz/ lub /r.z/	*domarzać, domarzali, domarzaliby, domarzalibyście, domarzalibyśmy, domarzaliście, domarzaliśmy, domarzał, domarzała, domarzałaby, domarzałabym, domarzałabyś, domarzałam, domarzałaś, domarzałby, domarzałbym, domarzałbyś, domarzałem, domarzałeś, domarzało, domarzałoby, domarzały, domarzałyby, domarzałybyście, domarzałybyśmy, domarzałyście, domarzałyśmy, domarza, domarzacie, domarzają, domarzam, domarzamy, domarzasz, domarzano, domarzania, domarzaniach, domarzaniami, domarzanie, domarzaniem, domarzaniom, domarzaniu, domarzań, domarzaj, domarzajcie, domarzajcież, domarzajmy, domarzajmyż, domarzajże, domarzająca, domarzającą, domarzające, domarzającego, domarzającej, domarzającemu, domarzający, domarzających, domarzającym, domarzającymi, domarzając*

2.5.12.3. Diada ortograficzna *rz* w strukturze: *#pomarz*

Z informacji zamieszczonych w tabeli 2.44 wynika, że diada *rz* umiejscowiona w nagłosowej strukturze: *#pomarz* może być transkrybowana zgodnie z jedną z dwóch możliwości – jako fonem /rz/ lub jako sekwencja fonemów /r.z/. Diada *rz* powinna być transkrybowana jako fonem /rz/ we wszystkich formach wyrazów: *Pomarzanowice* (nazwa miejscowości), *pomarzyć*. Natomiast we wszystkich formach wyrazów: *pomarzły, pomarznąć*, *pomarznięty* diada *rz* oznacza sekwencję fonemów /r.z/. Identyfikacja takich wyrazów jest możliwa na podstawie ciągów: *#pomarzn, #pomarzł, #pomarzl*.

Tabela 2.44 – Analiza słownika dotycząca struktury: *#pomarz*

Transkrypcja diady *rz*	Wykaz form fleksyjnych
/rz/	*Pomarzanowice, Pomarzanowicach, Pomarzanowicami, Pomarzanowicom, Pomarzanowic, pomarzyć, pomarzyli, pomarzyliby, pomarzylibyście, pomarzylibyśmy, pomarzyliście, pomarzyliśmy, pomarzył, pomarzyła, pomarzyłaby, pomarzyłabym, pomarzyłabyś, pomarzyłam, pomarzyłaś, pomarzyłby, pomarzyłbym, pomarzyłbyś, pomarzyłem, pomarzyłeś, pomarzyło, pomarzyłoby, pomarzyły, pomarzyłyby, pomarzyłybyście, pomarzyłybyśmy, pomarzyłyście, pomarzyłyśmy, pomarzą, pomarzę, pomarzy, pomarzycie, pomarzymy, pomarzysz, pomarzono, pomarzenia, pomarzeniach, pomarzeniami, pomarzenie, pomarzeniem, pomarzeniom, pomarzeniu, pomarzeń, pomarz, pomarzcie, pomarzcież, pomarzmy, pomarzmyż, pomarzże, pomarzywszy*
/r.z/	*pomarznąć, pomarznij, pomarznijcie, pomarznijcież, pomarznijmy, pomarznijmyż, pomarznijże, pomarzli, pomarzliby, pomarzlibyście, pomarzlibyśmy, pomarzliście, pomarzliśmy, pomarzł, pomarzła, pomarzłaby, pomarzłabym, pomarzłabyś, pomarzłam, pomarzłaś, pomarzłby, pomarzłbym, pomarzłbyś, pomarzłem, pomarzłeś, pomarzło, pomarzłoby, pomarzły, pomarzłyby, pomarzłybyście, pomarzłybyśmy, pomarzłyście, pomarzłyśmy, pomarzną, pomarznę, pomarznie, pomarzniecie, pomarzniemy, pomarzniesz, pomarznięto, pomarznięcia, pomarznięciach, pomarznięciami, pomarznięcie, pomarznięciem, pomarznięciom, pomarznięciu, pomarznięć, pomarzłszy, pomarznął, pomarznąłby, pomarznąłbym, pomarznąłbyś, pomarznąłem, pomarznąłeś, pomarznąwszy, pomarzłego, pomarzłemu, pomarzłych, pomarzłym, pomarzłymi, pomarzłą, pomarzłe, pomarzłej, pomarzniętego, pomarzniętemu, pomarzniętych, pomarzniętym, pomarzniętymi, pomarznięci, pomarznięta, pomarzniętą, pomarznięte, pomarzniętej*
/rz/ lub /r.z/	–

2.5.12.4. Diada ortograficzna *rz* w strukturze: *#przemarz*

Wyraz *przemarzać* może być formą niedokonaną czasownika *przemarzyć* („spędzić pewien czas na marzeniach") lub formą niedokonaną czasownika *przemarznąć* („ulec uszkodzeniu lub zniszczeniu na skutek działania mrozu, bardzo zmarznąć"). W związku z tym znaczna liczba form fleksyjnych ujętych w tabeli 2.45 została zaliczona do kategorii wyrazów, w których diada *rz* może być transkrybowana jako fonem /rz/, a także jako sekwencja fonemów /r.z/. Poza tym we wszystkich formach wyrazów: *przemarzlina, przemarzły, przemarznąć, przemarznięcie, przemarznięty* diada *rz* powinna być transkrybowana jako ciąg fonemów /r.z/. Ich identyfikacja jest możliwa na podstawie struktur: *#przemarzl, #przemarzł, #przemarzn*. Identyfikacja form zawierających diadę *rz*, która może być transkrybowana zgodnie z dwiema uwzględnionymi możliwościami, jest możliwa na podstawie sekwencji: *#przemarzal, #przemarzał, #przemarza#, #przemarzam, #przemarzasz, #przemarzać#, #przemarzac, #przemarzano, #przemarzani, #przemarzań, #przemarzaj*.

Tabela 2.45 – Analiza słownika dotycząca struktury: *#przemarz*

Transkrypcja diady *rz*	Wykaz form fleksyjnych
/rz/	*przemarzana, przemarzaną, przemarzane, przemarzanego, przemarzanej, przemarzanemu, przemarzani, przemarzany, przemarzanych, przemarzanym, przemarzanymi, przemarzyć, przemarzeni, przemarzona, przemarzoną, przemarzone, przemarzonego, przemarzonej, przemarzonemu, przemarzony, przemarzonych, przemarzonym, przemarzonymi, przemarzyli, przemarzyliby, przemarzylibyście, przemarzylibyśmy, przemarzyliście, przemarzyliśmy, przemarzył, przemarzyła, przemarzyłaby, przemarzyłabym, przemarzyłabyś, przemarzyłam, przemarzyłaś, przemarzyłby, przemarzyłbym, przemarzyłbyś, przemarzyłem, przemarzyłeś, przemarzyło, przemarzyłoby, przemarzyły, przemarzyłyby, przemarzyłybyście, przemarzyłybyśmy, przemarzyłyście, przemarzyłyśmy, przemarzą, przemarzę, przemarzy, przemarzycie, przemarzymy, przemarzysz, przemarzono, przemarzenia, przemarzeniach, przemarzeniami, przemarzenie, przemarzeniem, przemarzeniom, przemarzeniu, przemarzeń, przemarz, przemarzcie, przemarzcież, przemarzmy, przemarzmyż, przemarzże, przemarzywszy*
/r.z/	*przemarzlina, przemarzliną, przemarzlinę, przemarzlinie, przemarzlino, przemarzliny, przemarzlinach, przemarzlinami, przemarzlinom, przemarzlin, przemarzłego, przemarzłemu, przemarzły, przemarzłych, przemarzłym, przemarzłymi, przemarzli, przemarzła, przemarzłą, przemarzłe, przemarzłej, przemarznąć, przemarznij, przemarznijcie, przemarznijcież, przemarznijmy, przemarznijmyż, przemarznijże, przemarzliby, przemarzlibyście, przemarzlibyśmy, przemarzliście, przemarzliśmy, przemarzł, przemarzłaby, przemarzłabym, przemarzłabyś, przemarzłam, przemarzłaś, przemarzłby, przemarzłbym, przemarzłbyś, przemarzłem, przemarzłeś, przemarzło, przemarzłoby, przemarzłyby, przemarzłybyście, przemarzłybyśmy, przemarzłyście, przemarzłyśmy, przemarzną, przemarznę, przemarznie, przemarzniecie, przemarzniemy, przemarzniesz, przemarznięto, przemarznięcia, przemarznięciach, przemarznięciami, przemarznięcie, przemarznięciem, przemarznięciom, przemarznięciu, przemarznięć, przemarzłszy, przemarznął, przemarznąłby, przemarznąłbym, przemarznąłbyś, przemarznąłem, przemarznąłeś, przemarznąwszy, przemarznięty, przemarzniętego, przemarzniętemu, przemarzniętych, przemarzniętym, przemarzniętymi, przemarznięci, przemarznięta, przemarzniętą, przemarznięte, przemarzniętej*
/rz/ lub /r.z/	*przemarzali, przemarzaliby, przemarzalibyście, przemarzalibyśmy, przemarzaliście, przemarzaliśmy, przemarzał, przemarzała, przemarzałaby, przemarzałabym, przemarzałabyś, przemarzałam, przemarzałaś, przemarzałby, przemarzałbym, przemarzałbyś, przemarzałem, przemarzałeś, przemarzało, przemarzałoby, przemarzały, przemarzałyby, przemarzałybyście, przemarzałybyśmy, przemarzałyście, przemarzałyśmy, przemarza, przemarzacie, przemarzają, przemarzam, przemarzamy, przemarzasz, przemarzać, przemarzano, przemarzania, przemarzaniach, przemarzaniami, przemarzanie, przemarzaniem, przemarzaniom, przemarzaniu, przemarzań, przemarzaj, przemarzajcie, przemarzajcież, przemarzajmy, przemarzajmyż, przemarzajże, przemarzająca, przemarzającą, przemarzające, przemarzającego, przemarzającej, przemarzającemu, przemarzający, przemarzających, przemarzającym, przemarzającymi, przemarzając*

2.5.12.5. Diada ortograficzna *rz* w strukturze: *#rozmarz*

Wyraz *rozmarzać* może być formą niedokonaną czasownika *rozmarznąć* („wrócić do stanu przed zamrożeniem, zamarznięciem") lub formą niedokonaną czasownika *rozmarzyć* („wprawić w stan marzenia"). Dlatego w wypadku niektórych form fleksyjnych ujętych w tabeli 2.46 diada *rz* może być transkrybowana zarówno jako fonem /rz/, jak i jako sekwencja fonemów /r.z/. Oprócz tego we wszystkich formach wyrazów: *rozmarzły, rozmarznięty* diada *rz* oznacza ciąg fonemów /r.z/, natomiast we wszystkich formach wyrazów: *rozmarzająco, rozmarzenie, rozmarzyć* oznacza ona fonem /rz/.

Na podstawie wyniku analizy przedstawionego w tabeli 2.46 zostały wyznaczone nagłosowe struktury ortograficzne umożliwiające identyfikację wszystkich form zawierających diadę *rz*, która oznacza sekwencję fonemów /r.z/: *#rozmarzł, #rozmarzl, #rozmarzn*. Identyfikacja wszystkich form zawierających diadę *rz*, która może być transkrybowana w dwóch rozpatrywanych wariantach, jest możliwa na podstawie struktur: *#rozmarzać#, #rozmarzal, #rozmarzał, #rozmarza#, #rozmarzac, #rozmarzają#, #rozmarzam, #rozmarzas, #rozmarzano, #rozmarzani, #rozmarzaj, #rozmarzań*.

Tabela 2.46 – Analiza słownika dotycząca struktury: *#rozmarz*

Transkrypcja diady *rz*	Wykaz form fleksyjnych
/rz/	*rozmarzana, rozmarzaną, rozmarzane, rozmarzanego, rozmarzanej, rozmarzanemu, rozmarzani, rozmarzany, rozmarzanych, rozmarzanym, rozmarzanymi, rozmarzyć, rozmarzeni, rozmarzona, rozmarzoną, rozmarzone, rozmarzonego, rozmarzonej, rozmarzonemu, rozmarzony, rozmarzonych, rozmarzonym, rozmarzonymi, rozmarzyli, rozmarzyliby, rozmarzylibyście, rozmarzylibyśmy, rozmarzyliście, rozmarzyliśmy, rozmarzył, rozmarzyła, rozmarzyłaby, rozmarzyłabym, rozmarzyłabyś, rozmarzyłam, rozmarzyłaś, rozmarzyłby, rozmarzyłbym, rozmarzyłbyś, rozmarzyłem, rozmarzyłeś, rozmarzyło, rozmarzyłoby, rozmarzyły, rozmarzyłyby, rozmarzyłybyście, rozmarzyłybyśmy, rozmarzyłyście, rozmarzyłyśmy, rozmarzną, rozmarzę, rozmarzy, rozmarzycie, rozmarzymy, rozmarzysz, rozmarzono, rozmarzenia, rozmarzeniach, rozmarzeniami, rozmarzenie, rozmarzeniem, rozmarzeniom, rozmarzeniu, rozmarzeń, rozmarz, rozmarzcie, rozmarzcież, rozmarzmy, rozmarzmyż, rozmarzże, rozmarzywszy*
/r.z/	*rozmarzły, rozmarzłego, rozmarzłemu, rozmarzłych, rozmarzłym, rozmarzłymi, rozmarzli, rozmarźli, rozmarzła, rozmarzłą, rozmarzłe, rozmarzłej, romarznąć, rozmarznij, rozmarznijcie, rozmarznijcież, rozmarznijmy, rozmarznijmyż, rozmarznijże, rozmarzliby, rozmarzlibyście, rozmarzlibyśmy, rozmarzliście, rozmarzliśmy, rozmarzł, rozmarzłaby, rozmarzłabym, rozmarzłabyś, rozmarzłam, rozmarzłaś, rozmarzłby, rozmarzłbym, rozmarzłbyś, rozmarzłem, rozmarzłeś, rozmarzło, rozmarzłoby, rozmarzłyby, rozmarzłybyście, rozmarzłybyśmy, rozmarzłyście, rozmarzłyśmy, rozmarzną, rozmarznę, rozmarznie, rozmarzniecie, rozmarzniemy, rozmarzniesz, rozmarznięto, rozmarznięcia, rozmarznięciach, rozmarznięciami, rozmarznięcie, rozmarznięciem, rozmarznięciom, rozmarznięciu, rozmarznięć, rozmarzłszy, rozmarznął, rozmarznąłby, rozmarznąłbym, rozmarznąłbyś, rozmarznąłem, rozmarznąłeś, rozmarznąwszy, rozmarznięty, rozmarzniętego, rozmarzniętemu, rozmarzniętych, rozmarzniętym, rozmarzniętymi, rozmarznięci, rozmarznięta, rozmarzniętą, rozmarznięte, rozmarzniętej*
/rz/ lub /r.z/	*rozmarzać, rozmarzali, rozmarzaliby, rozmarzalibyście, rozmarzalibyśmy, rozmarzaliście, rozmarzaliśmy, rozmarzał, rozmarzała, rozmarzałaby, rozmarzałabym, rozmarzałabyś, rozmarzałam, rozmarzałaś, rozmarzałby, rozmarzałbym, rozmarzałbyś, rozmarzałem, rozmarzałeś, rozmarzało, rozmarzałoby, rozmarzały, rozmarzałyby, rozmarzałybyście, rozmarzałybyśmy, rozmarzałyście, rozmarzałyśmy, rozmarza, rozmarzacie, rozmarzają, rozmarzam, rozmarzamy, rozmarzasz, rozmarzano, rozmarzania, rozmarzaniach, rozmarzaniami, rozmarzanie, rozmarzaniem, rozmarzaniom, rozmarzaniu, rozmarzań, rozmarzaj, rozmarzajcie, rozmarzajcież, rozmarzajmy, rozmarzajmyż, rozmarzajże, rozmarzająca, rozmarzającą, rozmarzające, rozmarzającego, rozmarzającej, rozmarzającemu, rozmarzający, rozmarzających, rozmarzającym, rozmarzającymi, rozmarzając*

2.5.12.6. Diada ortograficzna *rz* w strukturze: *#umarz*

Czasownik *umarzać* może oznaczać: „rezygnację z dalszego postępowania sądowego w określonej sprawie" lub „rezygnację całkowitą lub częściową ze ściągania jakichś należności pieniężnych". W takich znaczeniach diada *rz* powinna być transkrybowana jako fonem /rz/. Dotyczy to również przymiotnika *umarzalny* („o długu, kredycie itp.: dający się umorzyć"). Natomiast w innym znaczeniu wyrazu *umarzać*: „twardnieć na skutek mrozu, zmienić się w lód" diada *rz* oznacza sekwencję fonemów /r.z/. W wyrazie *umarznąć* („stwardnieć na skutek mrozu, zmienić się w lód"), a także w wyrazach: *umarzły, umarznięty* („taki, który umarzł") diada *rz* również oznacza sekwencję fonemów /r.z/. Identyfikacja form tych wyrazów jest możliwa na podstawie nagłosowych ciągów ortograficznych: *#umarzł, #umarzl, #umarzn*. Poza tym analiza przedstawiona w tabeli 2.47 wykazała, że identyfikacja form fleksyjnych zawierających diadę *rz*, która może być transkrybowana zarówno jako sekwencja fonemów /r.z/, jak i jako fonem /rz/, jest możliwa na podstawie następujących nagłosowych ciągów ortograficznych: *#umarzać, #umarzali, #umarzał, #umarza#, #umarzac, #umarzaj, #umarzam, #umarzani, #umarzano, #umarzań*.

Tabela 2.47 – Analiza słownika dotycząca struktury: *#umarz*

Transkrypcja diady *rz*	Wykaz form fleksyjnych
/rz/	*umarzana, umarzaną, umarzane, umarzanego, umarzanej, umarzanemu, umarzani, umarzany, umarzanych, umarzanym, umarzanymi, umarzalny, umarzalnego, umarzalnemu, umarzalnych, umarzalnym, umarzalnymi, umarzalni, umarzalna, umarzalną, umarzalne, umarzalnej*
/r.z/	*umarzły, umarzłego, umarzłemu, umarzłych, umarzłym, umarzłymi, umarzli, umarźli, umarzła, umarzłą, umarzłe, umarzłej, umarznąć, umarznij, umarznijcie, umarznijcież, umarznijmy, umarznijmyż, umarznijże, umarzliby, umarzlibyście, umarzlibyśmy, umarzliście, umarzliśmy, umarzł, umarzłaby, umarzłabym, umarzłabyś, umarzłam, umarzłaś, umarzłby, umarzłbym, umarzłbyś, umarzłem, umarzłeś, umarzło, umarzłoby, umarzłyby, umarzłybyście, umarzłybyśmy, umarzłyście, umarzłyśmy, umarzną, umarznę, umarznie, umarzniecie, umarzniemy, umarzniesz, umarznięto, umarznięcia, umarznięciach, umarznięciami, umarznięcie, umarznięciem, umarznięciom, umarznięciu, umarznięć, umarzłszy, umarznął, umarznąłby, umarznąłbym, umarznąłbyś, umarznąłem, umarznąłeś, umarznąwszy, umarznięty, umarzniętego, umarzniętemu, umarzniętych, umarzniętym, umarzniętymi, umarznięci, umarznięta, umarzniętą, umarznięte, umarzniętej*
/rz/ lub /r.z/	*umarzać, umarzali, umarzaliby, umarzalibyście, umarzalibyśmy, umarzaliście, umarzaliśmy, umarzał, umarzała, umarzałaby, umarzałabym, umarzałabyś, umarzałam, umarzałaś, umarzałby, umarzałbym, umarzałbyś, umarzałem, umarzałeś, umarzało, umarzałoby, umarzały, umarzałyby, umarzałybyście, umarzałybyśmy, umarzałyście, umarzałyśmy, umarza, umarzacie, umarzają, umarzam, umarzamy, umarzasz, umarzano, umarzania, umarzaniach, umarzaniami, umarzanie, umarzaniem, umarzaniom, umarzaniu, umarzań, umarzaj, umarzajcie, umarzajcież, umarzajmy, umarzajmyż, umarzajże, umarzająca, umarzającą, umarzające, umarzającego, umarzającej, umarzającemu, umarzający, umarzających, umarzającym, umarzającymi, umarzając*

2.5.12.7. Diada ortograficzna *rz* w strukturze: *#wymarz*

Czasownik *wymarzać* ma dwa znaczenia, które są związane z różnymi sposobami transkrypcji. Pierwsze znaczenie to: „ginąć od mrozu, wymierać z zimna, zamarzać". Znaczenie drugie to: „morząc głodem przez dłuższy czas osłabiać, wycieńczać". W pierwszej wymienionej możliwości diadzie *rz* odpowiada sekwencja fonemów /r.z/, natomiast w wypadku możliwości drugiej jest to fonem /rz/. Poza tym we wszystkich formach wyrazów: *wymarzły, wymarznąć, wmarznięty* diada *rz* powinna być transkrybowana jako /r.z/, a w wypadku form: *wymarzony, wymarzyć* odpowiednią transkrypcją jest fonem /rz/.

Analiza ujęta w tabeli 2.48 wykazała, że identyfikacja form fleksyjnych zawierających diadę *rz* oznaczającą tylko i wyłącznie sekwencję /r.z/ jest możliwa na podstawie nagłosowych struktur: *#wymarzł, #wymarzl, #wymarzn*. Identyfikacja wszystkich form zawierających diadę *rz*, którą można transkrybować na dwa sposoby, jest możliwa na podstawie ciągów: *#wymarzać, #wymarzal, #wymarzał, #wymarza#, #wymarzac, #wymarzam, #wymarzasz, #wymarzano, #wymarzani, #wymarzaj, #wymarzań*.

Tabela 2.48 – Analiza słownika dotycząca struktury: *#wymarz*

Transkrypcja diady *rz*	Wykaz form fleksyjnych
/rz/	*wymarzana, wymarzaną, wymarzane, wymarzanego, wymarzanej, wymarzanemu, wymarzani, wymarzany, wymarzanych, wymarzanym, wymarzanymi, wymarzony, wymarzonego, wymarzonemu, wymarzonych, wymarzonym, wymarzonymi, wymarzeni, wymarzona, wymarzoną, wymarzone, wymarzonej, wymarzyć, wymarzyli, wymarzyliby, wymarzylibyście, wymarzylibyśmy, wymarzyliście, wymarzyliśmy, wymarzył, wymarzyła, wymarzyłaby, wymarzyłabym, wymarzyłabyś, wymarzyłam, wymarzyłaś, wymarzyłby, wymarzyłbym, wymarzyłbyś, wymarzyłem, wymarzyłeś, wymarzyło, wymarzyłoby, wymarzyły, wymarzyłyby, wymarzyłybyście, wymarzyłybyśmy, wymarzyłyście, wymarzyłyśmy, wymarzą, wymarzę, wymarzy, wymarzycie, wymarzymy, wymarzysz, wymarzono, wymarzenia, wymarzeniach, wymarzeniami, wymarzenie, wymarzeniem, wymarzeniom, wymarzeniu, wymarzeń, wymarz, wymarzcie, wymarzcież, wymarzmy, wymarzmyż, wymarzże, wymarzywszy*

/r.z/	*wymarzły, wymarzłego, wymarzłemu, wymarzłych, wymarzłym, wymarzłymi, wymarzli, wymarźli, wymarzła, wymarzłą, wymarzłe, wymarzłej, wymarznąć, wymarznij, wymarznijcie, wymarznijcież, wymarznijmy, wymarznijmyż, wymarznijże, wymarzliby, wymarzlibyście, wymarzlibyśmy, wymarzliście, wymarzliśmy, wymarzł, wymarzłaby, wymarzłabym, wymarzłabyś, wymarzłam, wymarzłaś, wymarzłby, wymarzłbym, wymarzłbyś, wymarzłem, wymarzłeś, wymarzło, wymarzłoby, wymarzłyby, wymarzłybyście, wymarzłybyśmy, wymarzłyście, wymarzłyśmy, wymarzną, wymarznę, wymarznie, wymarzniecie, wymarzniemy, wymarzniesz, wymarznięto, wymarznięcia, wymarznięciach, wymarznięciami, wymarznięcie, wymarznięciem, wymarznięciom, wymarznięciu, wymarznięć, wymarzłszy, wymarznął, wymarznąłby, wymarznąłbym, wymarznąłbyś, wymarznąłem, wymarznąłeś, wymarznąwszy, wymarznięty, wymarzniętego, wymarzniętemu, wymarzniętych, wymarzniętym, wymarzniętymi, wymarznięci, wymarznięta, wymarzniętą, wymarznięte, wymarzniętej*
/rz/ lub /r.z/	*wymarzać, wymarzali, wymarzaliby, wymarzalibyście, wymarzalibyśmy, wymarzaliście, wymarzaliśmy, wymarzał, wymarzała, wymarzałaby, wymarzałabym, wymarzałabyś, wymarzałam, wymarzałaś, wymarzałby, wymarzałbym, wymarzałbyś, wymarzałem, wymarzałeś, wymarzało, wymarzałoby, wymarzały, wymarzałyby, wymarzałybyście, wymarzałybyśmy, wymarzałyście, wymarzałyśmy, wymarza, wymarzacie, wymarzają, wymarzam, wymarzamy, wymarzasz, wymarzano, wymarzania, wymarzaniach, wymarzaniami, wymarzanie, wymarzaniem, wymarzaniom, wymarzaniu, wymarzań, wymarzaj, wymarzajcie, wymarzajcież, wymarzajmy, wymarzajmyż, wymarzajże, wymarzająca, wymarzającą, wymarzające, wymarzającego, wymarzającej, wymarzającemu, wymarzający, wymarzających, wymarzającym, wymarzającymi, wymarzając*

2.5.12.8. Diada ortograficzna *rz* w strukturze: *#zamarz*

Istnieją różne możliwości transkrypcji struktury *#zamarz*. Czasownik *zamarzać* może oznaczać: „doprowadzać kogoś do stanu wycieńczenia lub do śmierci poprzez głodzenie lub złe odżywianie". W takim wariancie diada ortograficzna *rz* oznacza fonem /rz/. Inne znaczenia tego wyrazu to: „zamieniać się w lód lub umierać wskutek działania mrozu". W znaczeniu tym diada *rz* powinna być transkrybowana jako sekwencja fonemów /r.z/. Taki sposób transkrypcji jest właściwy również dla wyrazów: *zamarzalnia, zamarzły, zamarznąć, zamarznięty*. W wyrazie *zamarzyć* diada *rz* oznacza fonem /rz/.

Z analizy ujętej w tabeli 2.49 wynika, że identyfikacja wszystkich form fleksyjnych zawierających diadę ortograficzną *rz*, która może być transkrybowana tylko jako sekwencja fonemów /r.z/, jest możliwa na podstawie nagłosowych ciągów: *#zamarzaln, #zamarzl, #zamarzł, #zamarzn.* Identyfikacja form, dla których są dopuszczalne dwa warianty transkrypcji (/r.z/ lub /rz/), może być przeprowadzona na podstawie następujących struktur umiejscowionych w nagłosie: *#zamarzać, #zamarzal, #zamarzał, #zamarza#, #zamarzac, #zamarzaj, #zamarzam, #zamarzano, #zamarzani, #zamarzasz*.

Tabela 2.49 – Analiza słownika dotycząca struktury: *#zamarz*

Transkrypcja diady *rz*	Wykaz form fleksyjnych
/rz/	*zamarzana, zamarzaną, zamarzane, zamarzanego, zamarzanej, zamarzanemu, zamarzani, zamarzany, zamarzanych, zamarzanym, zamarzanymi, zamarzyć, zamarzyli, zamarzyliby, zamarzylibyście, zamarzylibyśmy, zamarzyliście, zamarzyliśmy, zamarzył, zamarzyła, zamarzyłaby, zamarzyłabym, zamarzyłabyś, zamarzyłam, zamarzyłaś, zamarzyłby, zamarzyłbym, zamarzyłbyś, zamarzyłem, zamarzyłeś, zamarzyło, zamarzyłoby, zamarzyły, zamarzyłyby, zamarzyłybyście, zamarzyłybyśmy, zamarzyłyście, zamarzyłyśmy, zamarzą, zamarzę, zamarzy, zamarzycie, zamarzymy, zamarzysz, zamarzono, zamarzenia, zamarzeniach, zamarzeniami, zamarzenie, zamarzeniem, zamarzeniom, zamarzeniu, zamarzeń, zamarz, zamarzcie, zamarzcież, zamarzmy, zamarzmyż, zamarzże, zamarzywszy*
/r.z/	*zamarzalnia, zamarzalnie, zamarzalni, zamarzalnią, zamarzalnię, zamarzalnio, zamarzalniach, zamarzalniami, zamarzalniom, zamarzalń, zamarzły, zamarzłego, zamarzłemu, zamarzłych, zamarzłym, zamarzłymi, zamarzli, zamarzła, zamarzłą, zamarzłe, zamarzłej, zamarznąć, zamarznij, zamarznijcie, zamarznijcież, zamarznijmy, zamarznijmyż, zamarznijże, zamarzliby, zamarzlibyście, zamarzlibyśmy, zamarzliście, zamarzliśmy, zamarzł, zamarzłaby, zamarzłabym, zamarzłabyś, zamarzłam, zamarzłaś, zamarzłby, zamarzłbym, zamarzłbyś, zamarzłem, zamarzłeś, zamarzło, zamarzłoby, zamarzłyby, zamarzłybyście, zamarzłybyśmy, zamarzłyście, zamarzłyśmy, zamarzną, zamarznę, zamarznie, zamarzniecie, zamarzniemy, zamarzniesz, zamarznięto, zamarznięcia, zamarznięciach, zamarznięciami, zamarznięcie, zamarznięciem, zamarznięciom, zamarznięciu, zamarznięć, zamarzłszy, zamarznął, zamarznąłby, zamarznąłbym, zamarznąłbyś, zamarznąłem, zamarznąłeś, zamarznąwszy, zamarznięty, zamarzniętego, zamarzniętemu, zamarzniętych, zamarzniętym, zamarzniętymi, zamarznięci, zamarznięta, zamarzniętą, zamarznięte, zamarzniętej*

/rz/ lub /r.z/	*zamarzać, zamarzali, zamarzaliby, zamarzalibyście, zamarzalibyśmy, zamarzaliście, zamarzaliśmy, zamarzał, zamarzała, zamarzałaby, zamarzałabym, zamarzałabyś, zamarzałam, zamarzałaś, zamarzałby, zamarzałbym, zamarzałbyś, zamarzałem, zamarzałeś, zamarzało, zamarzałoby, zamarzały, zamarzałyby, zamarzałybyście, zamarzałybyśmy, zamarzałyście, zamarzałyśmy, zamarza, zamarzacie, zamarzają, zamarzam, zamarzamy, zamarzasz, zamarzano, zamarzania, zamarzaniach, zamarzaniami, zamarzanie, zamarzaniem, zamarzaniom, zamarzaniu, zamarzań, zamarzaj, zamarzajcie, zamarzajcież, zamarzajmy, zamarzajmyż, zamarzajże, zamarzająca, zamarzającą, zamarzające, zamarzającego, zamarzającej, zamarzającemu, zamarzający, zamarzających, zamarzającym, zamarzającymi, zamarzając*

2.5.12.9. Diada ortograficzna *rz* w strukturze: *#omarz*

Nagłosowy ortograficzny ciąg *#omarz* występuje w wyrazach: *omarzać, omarzły, omarznąć, omarznięty*. We wszystkich formach tych wyrazów diada *rz* powinna być transkrybowana jako sekwencja fonemów /r.z/ (identyfikacja tych form jest możliwa na podstawie ciągów: *#omarzł, #omarzl, #omarza, #omarzn*). Istnieją tylko dwie formy fleksyjne (zawierające ciąg ortograficzny *#omarz*), w których diada *rz* oznacza fonem /rz/. Jest to miejscownik wyrazu *omar* oraz miejscownik wyrazu *Omar – omarze* i *Omarze*. Definicja wyrazu *omar* jest następująca: „dawna nazwa rdzy – rodzaju grzybów z rodziny rdzowatych, występujących w postaci pleśni na liściach rozmaitych roślin, źdźbłach traw, niekiedy na korze drzew (…)”. Wyraz *Omar* oznacza arabskie imię męskie. Ponieważ taki sposób transkrypcji dotyczy tylko dwóch form, zrezygnowano z prezentowania danych w formie tabeli.

2.5.12.10. Diada ortograficzna *rz* w strukturze: *#mierz*

Istnieje znaczna liczba form podstawowych zawierających ortograficzny ciąg *#mierz*: *mierzalność, mierzalny, mierzarka, mierzchnąć, mierzeja, Mierzejewo, Mierzejewski, Mierzeszyn, Mierzewo, Mierzęcice, mierzęcicki, Mierzęcin, mierznica, Mierzowice, mierzwa, Mierzwiak, mierzwiasty, mierzwić, Mierzwin, Mierzwiński, Mierzyce, mierzyć, mierzyn, mierzynek, Mierzynko, Mierzyński.* We wszystkich wymienionych wyrazach diada *rz* powinna być transkrybowana jako fonem /rz/. W tabeli 2.50 nie uwzględniono form wyrazu *mierzchnąć*, w którym diada *rz* oznacza fonem /S/ ze względu na zjawisko antycypacji artykulacyjnej. Poza tym w wyrazach *mierznąć* oraz mierznięcie *rz* oznacza sekwencję fonemów /r.z/, w wyrazie *mierzić* jest to sekwencja /r.z'/.

Z danych zamieszczonych w tabeli wynika, że nie występują formy fleksyjne zawierające nagłosowy ciąg *#mierz*, które mogłyby być transkrybowane na dwa różne sposoby w zależności od znaczenia. Wszystkie formy, w których diada *rz* oznacza sekwencję /r.z/, można zidentyfikować na podstawie nagłosowych ciągów: *#mierzną, #mierznę, #mierznie, #mierzl, #mierzł*. Formy, w których diada *rz* zawsze oznacza /r.z'/, można zidentyfikować na podstawie nagłosowego ciągu *#mierzi*.

Tabela 2.50 – Analiza słownika dotycząca struktury: *#mierz*

Transkrypcja diady *rz*	Wykaz form fleksyjnych
/rz/	*mierzalność, mierzalności, mierzalnością, mierzalnościach, mierzalnościami, mierzalnościom, mierzalny, mierzalnego, mierzalnemu, mierzalnych, mierzalnym, mierzalnymi, mierzalni, mierzalna, mierzalną, mierzalne, mierzalnej, mierzalnie, mierzarka, mierzarce, mierzarką, mierzarkę, mierzarki, mierzarko, mierzarkach, mierzarkami, mierzarkom, mierzarek, mierzeja, mierzei, mierzeją, mierzeję, mierzejo, mierzejach, mierzejami, mierzejom, mierzej, Mierzejewo, Mierzejewa, Mierzejewem, Mierzejewie, Mierzejewu, Mierzejewski, Mierzejewskich, Mierzejewskiego, Mierzejewskiemu, Mierzejewskim, Mierzejewskimi, Mierzejewscy, Mierzejewska, Mierzejewską, Mierzejewskie, Mierzejewskiej, Mierzeszyn, Mierzeszyna, Mierzeszynem, Mierzeszynie, Mierzeszynowi, Mierzewo, Mierzewa, Mierzewem, Mierzewie, Mierzewu, Mierzęcice, Mierzęcicach, Mierzęcicami, Mierzęcicom, mierzęcicki, mierzęcickich, mierzęcickiego, mierzęcickiemu, mierzęcickim, mierzęcickimi, mierzęciccy, mierzęcicku, mierzęcicka, mierzęcicką, mierzęcickie, mierzęcickiej, mierzęcicko, Mierzęcin, Mierzęcina, Mierzęcinem, Mierzęcinie, Mierzęcinowi, mierznica, mierznice, mierznicą, mierznicę, mierznico, mierznicy, mierznicach, mierznicami, mierznicom, mierznic, Mierzowice, Mierzowicach, Mierzowicami, Mierzowicom, Mierzowic, mierzwa, mierzwą, mierzwę, mierzwie, mierzwo, mierzwy, mierzwach, mierzwami, mierzwom, mierzw, Mierzwiak, Mierzwiakach, Mierzwiakami, Mierzwiakom, Mierzwiaka, Mierzwiakiem, Mierzwiakowi, Mierzwiaku, Mierzwiaków, Mierzwiakowie, Mierzwiaki, mierzwiasty, mierzwiastego, mierzwiastemu, mierzwiastych, mierzwiastym, mierzwiastymi, mierzwiaści, mierzwiasta, mierzwiastą, mierzwiaste, mierzwiastej, mierzwić, mierzwij, mierzwijcie, mierzwijcież, mierzwijmy, mierzwijmyż, mierzwijże, mierzwieni, mierzwiona, mierzwioną, mierzwione, mierzwionego, mierzwionej, mierzwionemu, mierzwiony, mierzwionych, mierzwionym, mierzwionymi, mierzwili, mierzwiliby, mierzwilibyście, mierzwilibyśmy, mierzwiliście, mierzwiliśmy, mierzwił, mierzwiła, mierzwiłaby, mierzwiłabym, mierzwiłabyś, mierzwiłam, mierzwiłaś, mierzwiłby, mierzwiłbym, mierzwiłbyś, mierzwiłem, mierzwiłeś, mierzwiło, mierzwiłoby, mierzwiły, mierzwiłyby, mierzwiłybyście, mierzwiłybyśmy, mierzwiłyście, mierzwiłyśmy, mierzwi, mierzwią, mierzwicie, mierzwię, mierzwimy, mierzwisz, mierzwiono, mierzwienia, mierzwieniach, mierzwieniami, mierzwienie, mierzwieniem, mierzwieniom, mierzwieniu, mierzwień, mierzwcie, mierzwcież, mierzwmy, mierzwmyż, mierzwże, mierzwiąca, mierzwiącą, mierzwiące, mierzwiącego, mierzwiącej, mierzwiącemu, mierzwiący, mierzwiących, mierzwiącym, mierzwiącymi, mierzwiąc, Mierzwin, Mierzwina, Mierzwinem, Mierzwinie, mierzwinowi, Mierzwiński, Mierzwińskich, Mierzwińskiego, Mierzwińskiemu, Mierzwińskim, Mierzwińskimi, Mierzwińscy, Mierzwińska, Mierzwińską, Mierzwińskie, Mierzwińskiej, Mierzyce, Mierzycach, Mierzycami, Mierzycom, Mierzyc, mierzyć, mierzeni, mierzona, mierzoną, mierzone, mierzonego, mierzonej, mierzonemu, mierzony, mierzonych, mierzonym, mierzonymi, mierzyli, mierzyliby, mierzylibyście, mierzylibyśmy, mierzyliście, mierzyliśmy, mierzył, mierzyła, mierzyłaby, mierzyłabym, mierzyłabyś, mierzyłam, mierzyłaś, mierzyłby, mierzyłbym, mierzyłbyś, mierzyłem, mierzyłeś, mierzyło, mierzyłoby, mierzyły, mierzyłyby, mierzyłybyście, mierzyłybyśmy, mierzyłyście, mierzyłyśmy, mierzą, mierzę, mierzy, mierzycie, mierzymy, mierzysz, mierzono, mierzenia, mierzeniach, mierzeniami, mierzenie, mierzeniem, mierzeniom, mierzeniu, mierzeń, mierz, mierzcie, mierzcież, mierzmy, mierzmyż, mierzże, mierząca, mierzącą, mierzące, mierzącego, mierzącej, mierzącemu, mierzący, mierzących, mierzącym, mierzącymi, mierząc, mierzyn, mierzynach, mierzynami, mierzynom, mierzyna, mierzynem, mierzynie, mierzynowi, mierzynów, mierzyny, mierzynek, Mierzynka, Mierzynkiem, mierzynkowi, Mierzynku, mierzynków, mierzynkach, mierzynkami, mierzynkom, mierzynki, Mierzynko, Mierzyński, Mierzyńskich, Mierzyńskiego, Mierzyńskiemu, Mierzyńskim, Mierzyńskimi, Mierzyńscy, Mierzyńska, Mierzyńską, Mierzyńskie, Mierzyńskiej*
/r.z/	*mierznąć, mierznij, mierznijcie, mierznijcież, mierznijmy, mierznijmyż, mierznijże, mierzli, mierzliby, mierzlibyście, mierzlibyśmy, mierzliście, mierzliśmy, mierzł, mierzła, mierzłaby, mierzłabym, mierzłabyś, mierzłam, mierzłaś, mierzłby, mierzłbym, mierzłbyś, mierzłem, mierzłeś, mierzło, mierzłoby, mierzły, mierzłyby, mierzłybyście, mierzłybyśmy, mierzłyście, mierzłyśmy, mierzną, mierznę, mierznie, mierzniecie, mierzniemy, mierzniesz, mierznięto, mierznięcia, mierznięciach, mierznięciami, mierznięcie, mierznięciem, mierznięciom, mierznięciu, mierznięć, mierznąca, mierznącą, mierznące, mierznącego, mierznącej, mierznącemu, mierznący, mierznących, mierznącym, mierznącymi, mierznąc*
/r.z'/	*mierzić, mierzili, mierziliby, mierzilibyście, mierzilibyśmy, mierziliście, mierziliśmy, mierził, mierziła, mierziłaby, mierziłabym, mierziłabyś, mierziłam, mierziłaś, mierziłby, mierziłbym, mierziłbyś, mierziłem, mierziłeś, mierziło, mierziłoby, mierziły, mierziłyby, mierziłybyście, mierziłybyśmy, mierziłyście, mierziłyśmy, mierzi, mierzicie, mierzimy, mierzisz*
/rz/ lub /r.z/	–

2.5.12.11. Diada ortograficzna *rz* w strukturze: *#obmierz*

Struktura *#obmierz* występuje w czasowniku *obmierzać*, który ma dwa znaczenia różniące się wymową. Pierwsze znaczenie to: „obrzydzać”. Natomiast znaczenie drugie to: „zmierzyć kogoś lub coś ze wszystkich stron”. W pierwszym wymienionym znaczeniu diada *rz* powinna być transkrybowana jako sekwencja fonemów /r.z/. W drugiej możliwości diada *rz* oznacza fonem /rz/. Poza tym transkrypcja jako /r.z/ jest właściwa

dla wyrazów: *obmierzłość, obmierzły, obmierznąć*. Z wyrazem *obmierzyć* (w aspekcie dokonanym) powiązana jest tylko jedna możliwość transkrypcji diady *rz* – jako fonem /rz/. W wypadku wyrazu *obmierzić* diadzie odpowiada sekwencja fonemów /r.z'/. Wszystkie formy fleksyjne tego typu można zidentyfikować na podstawie ortograficznego ciągu *#obmierzi*. Identyfikacja form, w których diada *rz* powinna być transkrybowana jako sekwencja fonemów /r.z/, jest możliwa na podstawie nagłosowych ciągów: *#obmierzn, #obmierzł, #obmierzle*. Formy zawierające diadę *rz*, która może być transkrybowana zarówno jako /r.z/, jak i jako /rz/, można zidentyfikować na podstawie nagłosowej struktury ortograficznej *#obmierza*.

Tabela 2.51 – Analiza słownika dotycząca struktury: *#obmierz*

Transkrypcja diady *rz*	Wykaz form fleksyjnych
/rz/	*obmierzeni, obmierzona, obmierzoną, obmierzone, obmierzonego, obmierzonej, obmierzonemu, obmierzony, obmierzonych, obmierzonym, obmierzonymi, obmierzyli, obmierzyliby, obmierzylibyście, obmierzylibyśmy, obmierzyliście, obmierzyliśmy, obmierzył, obmierzyła, obmierzyłaby, obmierzyłabym, obmierzyłabyś, obmierzyłam, obmierzyłaś, obmierzyłby, obmierzyłbym, obmierzyłbyś, obmierzyłem, obmierzyłeś, obmierzyło, obmierzyłoby, obmierzyły, obmierzyłyby, obmierzyłybyście, obmierzyłybyśmy, obmierzyłyście, obmierzyłyśmy, obmierzą, obmierzę, obmierzy, obmierzycie, obmierzymy, obmierzysz, obmierzono, obmierzenia, obmierzeniach, obmierzeniami, obmierzenie, obmierzeniem, obmierzeniom, obmierzeniu, obmierzeń, obmierz, obmierzcie, obmierzcież, obmierzmy, obmierzmyż, obmierzże, obmierzywszy*
/r.z/	*obmierznąć, obmierznij, obmierznijcie, obmierznijcież, obmierznijmy, obmierznijmyż, obmierznijże, obmierzli, obmierzliby, obmierzlibyście, obmierzlibyśmy, obmierzliście, obmierzliśmy, obmierzł, obmierzła, obmierzłaby, obmierzłabym, obmierzłabyś, obmierzłam, obmierzłaś, obmierzłby, obmierzłbym, obmierzłbyś, obmierzłem, obmierzłeś, obmierzło, obmierzłoby, obmierzły, obmierzłyby, obmierzłybyście, obmierzłybyśmy, obmierzłyście, obmierzłyśmy, obmierzną, obmierznę, obmierznie, obmierzniecie, obmierzniemy, obmierzniesz, obmierznięto, obmierźnięto, obmierznięcia, obmierznięciach, obmierznięciami, obmierznięcie, obmierznięciem, obmierznięciom, obmierznięciu, obmierznięć, obmierzłszy, obmierzłość, obmierzłości, obmierzłością, obmierzłościach, obmierzłościami, obmierzłościom, obmierzłego, obmierzłemu, obmierzłych, obmierzłym, obmierzłymi, obmierzłą, obmierzłe, obmierzłej, obmierzle*
/r.z'/	*obmierzić, obmierzili, obmierziliby, obmierzilibyście, obmierzilibyśmy, obmierziliście, obmierziliśmy, obmierził, obmierziła, obmierziłaby, obmierziłabym, obmierziłabyś, obmierziłam, obmierziłaś, obmierziłby, obmierziłbym, obmierziłbyś, obmierziłem, obmierziłeś, obmierziło, obmierziłoby, obmierziły, obmierziłyby, obmierziłybyście, obmierziłybyśmy, obmierziłyście, obmierziłyśmy, obmierzi, obmierzicie, obmierzimy, obmierzisz, obmierziwszy,*
/rz/ lub /r.z/	*obmierzać, obmierzana, obmierzaną, obmierzane, obmierzanego, obmierzanej, obmierzanemu, obmierzani, obmierzany, obmierzanych, obmierzanym, obmierzanymi, obmierzali, obmierzaliby, obmierzalibyście, obmierzalibyśmy, obmierzaliście, obmierzaliśmy, obmierzał, obmierzała, obmierzałaby, obmierzałabym, obmierzałabyś, obmierzałam, obmierzałaś, obmierzałby, obmierzałbym, obmierzałbyś, obmierzałem, obmierzałeś, obmierzało, obmierzałoby, obmierzały, obmierzałyby, obmierzałybyście, obmierzałybyśmy, obmierzałyście, obmierzałyśmy, obmierza, obmierzacie, obmierzają, obmierzam, obmierzamy, obmierzasz, obmierzano, obmierzania, obmierzaniach, obmierzaniami, obmierzanie, obmierzaniem, obmierzaniom, obmierzaniu, obmierzań, obmierzaj, obmierzajcie, obmierzajcież, obmierzajmy, obmierzajmyż, obmierzajże, obmierzająca, obmierzającą, obmierzające, obmierzającego, obmierzającej, obmierzającemu, obmierzający, obmierzających, obmierzającym, obmierzającymi, obmierzając*

2.5.12.12. Diada ortograficzna *rz* w strukturze: *#przemierz*

W *słowniku SJP.pl* występują następujące formy podstawowe wyrazów, które zawierają strukturę *#przemierz*: *przemierzać, przemierzyć, przemierzanie, przemierzły*. Niniejszy opis dotyczy drugiego wystąpienia diady *rz* w omawianej strukturze nagłosu – umiejscowiona w bezpośrednim sąsiedztwie litery *p* zawsze powinna być transkrybowana jako bezdźwięczny fonem /S/ ze względu na zjawisko perseweracji artykulacyjnej. W wyrazach: *przemierzać*, *przemierzyć* i *przemierzanie* diada *rz* oznacza fonem /rz/. W przymiotniku *przemierzły* właściwą transkrypcję stanowi sekwencja fonemów /r.z/. Identyfikacja form fleksyjnych jest możliwa na podstawie ciągów nagłosowych: *#przemierzł, #przemierzl*. Nie istnieją formy zawierające ortograficzną sekwencję *#przemierz*, które mogłyby być transkrybowane zgodnie z dwiema uwzględnionymi możliwościami.

Tabela 2.52 – Analiza słownika dotycząca struktury: *#przemierz*

Transkrypcja diady *rz*	Wykaz form fleksyjnych
/rz/	*przemierzać, przemierzana, przemierzaną, przemierzane, przemierzanego, przemierzanej, przemierzanemu, przemierzani, przemierzany, przemierzanych, przemierzanym, przemierzanymi, przemierzali, przemierzaliby, przemierzalibyście, przemierzalibyśmy, przemierzaliście, przemierzaliśmy, przemierzał, przemierzała, przemierzałaby, przemierzałabym, przemierzałabyś, przemierzałam, przemierzałaś, przemierzałby, przemierzałbym, przemierzałbyś, przemierzałem, przemierzałeś, przemierzało, przemierzałoby, przemierzały, przemierzałyby, przemierzałybyście, przemierzałybyśmy, przemierzałyście, przemierzałyśmy, przemierza, przemierzacie, przemierzają, przemierzam, przemierzamy, przemierzasz, przemierzano, przemierzania, przemierzaniach, przemierzaniami, przemierzanie, przemierzaniem, przemierzaniom, przemierzaniu, przemierzań, przemierzaj, przemierzajcie, przemierzajcież, przemierzajmy, przemierzajmyż, przemierzając, przemierzająca, przemierzającą, przemierzające, przemierzającego, przemierzającej, przemierzającemu, przemierzający, przemierzających, przemierzającym, przemierzającymi, przemierzając, przemierzyć, przemierzeni, przemierzona, przemierzoną, przemierzone, przemierzonego, przemierzonej, przemierzonemu, przemierzony, przemierzonych, przemierzonym, przemierzonymi, przemierzyli, przemierzyliby, przemierzylibyście, przemierzylibyśmy, przemierzyliście, przemierzyliśmy, przemierzył, przemierzyła, przemierzyłaby, przemierzyłabym, przemierzyłabyś, przemierzyłam, przemierzyłaś, przemierzyłby, przemierzyłbym, przemierzyłbyś, przemierzyłem, przemierzyłeś, przemierzyło, przemierzyłoby, przemierzyły, przemierzyłyby, przemierzyłybyście, przemierzyłybyśmy, przemierzyłyście, przemierzyłyśmy, przemierzą, przemierzę, przemierzy, przemierzycie, przemierzymy, przemierzysz, przemierzono, przemierzenia, przemierzeniach, przemierzeniami, przemierzenie, przemierzeniem, przemierzeniom, przemierzeniu, przemierzeń, przemierz, przemierzcie, przemierzcież, przemierzmy, przemierzmyż, przemierzże, przemierzywszy*
/r.z/	*przemierzły, przemierzłego, przemierzłemu, przemierzłych, przemierzłym, przemierzłymi, przemierzli, przemierzła, przemierzłą, przemierzłe, przemierzłej*
/rz/ lub /r.z/	–

2.5.12.13. Diada ortograficzna *rz* w strukturze: *#zmierz*

Wyraz *zmierzać* ma dwa znaczenia, które różnią się wymową. W wypadku znaczenia: „wzbudzać wstręt, odrazę, obrzydzać, mierzić" diada *rz* oznacza sekwencję fonemów /r.z/. Natomiast w znaczeniu: „poruszać się w jakimś kierunku" właściwą transkrypcję stanowi jeden fonem /rz/. W czasowniku *zmierznąć* („stać się dla kogoś ohydnym, budzącym niechęć, odrazę") oraz w przymiotniku *zmierzły* („obrzydliwy, obmierzły, wstrętny, ohydny") diada *rz* oznacza sekwencję fonemów.

Formy fleksyjne, w których diada *rz* może oznaczać tylko ciąg fonemów /r.z/, można zidentyfikować na podstawie struktur : *#zmierzł, #zmierzn, #zmierzl*. Formy, w których diada *rz* może być transkrybowana zarówno jako sekwencja fonemów /r.z/, jak i jako fonem /rz/ (w zależności od znaczenia), można zidentyfikować na podstawie nagłosowego ciągu *#zmierza*. W wyrazie *zmierzić* („wzbudzić w kimś wstręt, odrazę" – jest to aspekt dokonany wyrazu *zmierzać*) diada *rz* oznacza sekwencję /r.z'/ (wszystkie formy zawierają w nagłosie ciąg: *#zmierzi*).

Diada *rz* oznacza fonem /rz/ we wszystkich formach fleksyjnych wyrazów: *zmierzwiać, zmierzwić, zmierzwiony, zmierzyć*. Poza tym ortograficzna struktura *#zmierz* występuje w wyrazach: *zmierzch, zmierzchać, zmierzchnąć, zmierzchnica, zmierzchnikowiec, zmierzchowy.* We wszystkich formach tych wyrazów (nie wymieniono ich w tabeli 2.53) diada *rz* oznacza bezdźwięczny fonem /S/. Można je zidentyfikować na podstawie ortograficznego ciągu: *#zmierzch*.

Tabela 2.53 – Analiza słownika dotycząca struktury: *#zmierz*

Transkrypcja diady *rz*	Wykaz form fleksyjnych
/rz/	*zmierzwić, zmierzwieni, zmierzwiona, zmierzwioną, zmierzwione, zmierzwionego, zmierzwionej, zmierzwionemu, zmierzwiony, zmierzwionych, zmierzwionym, zmierzwionymi, zmierzwili, zmierzwiliby, zmierzwilibyście, zmierzwilibyśmy, zmierzwiliście, zmierzwiliśmy, zmierzwił, zmierzwiła, zmierzwiłaby, zmierzwiłabym, zmierzwiłabyś, zmierzwiłam, zmierzwiłaś, zmierzwiłby, zmierzwiłbym, zmierzwiłbyś, zmierzwiłem, zmierzwiłeś, zmierzwiło, zmierzwiłoby, zmierzwiły, zmierzwiłyby, zmierzwiłybyście, zmierzwiłybyśmy, zmierzwiłyście, zmierzwiłyśmy, zmierzwi, zmierzwią, zmierzwicie, zmierzwię, zmierzwimy, zmierzwisz, zmierzwiono, zmierzwienia, zmierzwieniach, zmierzwieniami, zmierzwienie, zmierzwieniem, zmierzwieniom, zmierzwieniu, zmierzwień, zmierzw, zmierzwcie, zmierzwcież, zmierzwmy, zmierzwmyż, zmierzwże, zmierzwiwszy, zmierzwiać, zmierzwiana, zmierzwianą, zmierzwiane, zmierzwianego, zmierzwianej, zmierzwianemu, zmierzwiani, zmierzwiany, zmierzwianych, zmierzwianym, zmierzwianymi, zmierzwiali, zmierzwialiby, zmierzwialibyście, zmierzwialibyśmy, zmierzwialiście, zmierzwialiśmy, zmierzwiał, zmierzwiała, zmierzwiałaby, zmierzwiałabym, zmierzwiałabyś, zmierzwiałam, zmierzwiałaś, zmierzwiałby, zmierzwiałbym, zmierzwiałbyś, zmierzwiałem, zmierzwiałeś, zmierzwiało, zmierzwiałoby, zmierzwiały, zmierzwiałyby, zmierzwiałybyście, zmierzwiałybyśmy, zmierzwiałyście, zmierzwiałyśmy, zmierzwia, zmierzwiacie, zmierzwiają, zmierzwiam, zmierzwiamy, zmierzwiasz, zmierzwiano, zmierzwiania, zmierzwianiach, zmierzwianiami, zmierzwianie, zmierzwianiem, zmierzwianiom, zmierzwianiu, zmierzwiań, zmierzwiaj, zmierzwiajcie, zmierzwiajcież, zmierzwiajmy, zmierzwiajmyż, zmierzwiajże, zmierzwiająca, zmierzwiającą, zmierzwiające, zmierzwiającego, zmierzwiającej, zmierzwiającemu, zmierzwiający, zmierzwiających, zmierzwiającym, zmierzwiającymi, zmierzwiając, zmierzyć, zmierzeni, zmierzona, zmierzoną, zmierzone, zmierzonego, zmierzonej, zmierzonemu, zmierzony, zmierzonych, zmierzonym, zmierzonymi, zmierzyli, zmierzyliby, zmierzylibyście, zmierzylibyśmy, zmierzyliście, zmierzyliśmy, zmierzył, zmierzyła, zmierzyłaby, zmierzyłabym, zmierzyłabyś, zmierzyłam, zmierzyłaś, zmierzyłby, zmierzyłbym, zmierzyłbyś, zmierzyłem, zmierzyłeś, zmierzyło, zmierzyłoby, zmierzyły, zmierzyłyby, zmierzyłybyście, zmierzyłybyśmy, zmierzyłyście, zmierzyłyśmy, zmierzą, zmierzę, zmierzy, zmierzycie, zmierzymy, zmierzysz, zmierzono, zmierzenia, zmierzeniach, zmierzeniami, zmierzenie, zmierzeniem, zmierzeniom, zmierzeniu, zmierzeń, zmierz, zmierzcie, zmierzcież, zmierzmy, zmierzmyż, zmierzże, zmierzywszy*
/r.z/	*zmierzły, zmierzłego, zmierzłemu, zmierzłych, zmierzłym, zmierzłymi, zmierzli, zmierzła, zmierzłą, zmierzłe, zmierzłej, zmierznąć, zmierznij, zmierznijcie, zmierznijcież, zmierznijmy, zmierznijmyż, zmierznijże, zmierzliby, zmierzlibyście, zmierzlibyśmy, zmierzliście, zmierzliśmy, zmierzł, zmierzłaby, zmierzłabym, zmierzłabyś, zmierzłam, zmierzłaś, zmierzłby, zmierzłbym, zmierzłbyś, zmierzłem, zmierzłeś, zmierzło, zmierzłoby, zmierzły, zmierzłyby, zmierzłybyście, zmierzłybyśmy, zmierzłyście, zmierzłyśmy, zmierzną, zmierznę, zmierznie, zmierzniecie, zmierzniemy, zmierzniesz, zmierznięto, zmierźnięto, zmierznięcia, zmierznięciach, zmierznięciami, zmierznięcie, zmierznięciem, zmierznięciom, zmierznięciu, zmierznięć, zmierzłszy*
/r.z'/	*zmierzić, zmierzili, zmierziliby, zmierzilibyście, zmierzilibyśmy, zmierziliście, zmierziliśmy, zmierził, zmierziła, zmierziłaby, zmierziłabym, zmierziłabyś, zmierziłam, zmierziłaś, zmierziłby, zmierziłbym, zmierziłbyś, zmierziłem, zmierziłeś, zmierziło, zmierziłoby, zmierziły, zmierziłyby, zmierziłybyście, zmierziłybyśmy, zmierziłyście, zmierziłyśmy, zmierzi, zmierzicie, zmierzimy, zmierzisz, zmierziwszy*
/rz/ lub /r.z/	*zmierzać, zmierzali, zmierzaliby, zmierzalibyście, zmierzalibyśmy, zmierzaliście, zmierzaliśmy, zmierzał, zmierzała, zmierzałaby, zmierzałabym, zmierzałabyś, zmierzałam, zmierzałaś, zmierzałby, zmierzałbym, zmierzałbyś, zmierzałem, zmierzałeś, zmierzało, zmierzałoby, zmierzały, zmierzałyby, zmierzałybyście, zmierzałybyśmy, zmierzałyście, zmierzałyśmy, zmierza, zmierzacie, zmierzają, zmierzam, zmierzamy, zmierzasz, zmierzano, zmierzania, zmierzaniach, zmierzaniami, zmierzanie, zmierzaniem, zmierzaniom, zmierzaniu, zmierzań, zmierzaj, zmierzajcie, zmierzajcież, zmierzajmy, zmierzajmyż, zmierzajże, zmierzająca, zmierzającą, zmierzające, zmierzającego, zmierzającej, zmierzającemu, zmierzający, zmierzających, zmierzającym, zmierzającymi, zmierzając*

2.5.12.14. Diada ortograficzna *rz* w strukturze: *#omierz*

Nagłosowa struktura *#omierz* występuje w wyrazie *omierzyć* („zmierzyć ze wszystkich stron”). W jego formach diada *rz* oznacza fonem /rz/. Natomiast w wyrazach: *omierzły* („taki, który budzi wstręt”), *omierznąć* („stać się odpychająco brzydkim”) oraz omierznięty *rz* oznacza sekwencję fonemów /r.z/. Takie formy można zidentyfikować na podstawie nagłosowych struktur: *#omierzł, #omierzl, #omierzn*.

Tabela 2.54 – Analiza słownika dotycząca struktury: *#omierz*

Transkrypcja diady *rz*	Wykaz form fleksyjnych
/rz/	*omierzeni, omierzona, omierzoną, omierzone, omierzonego, omierzonej, omierzonemu, omierzony, omierzonych, omierzonym, omierzonymi, omierzyli, omierzyliby, omierzylibyście, omierzylibyśmy, omierzyliście, omierzyliśmy, omierzył, omierzyła, omierzyłaby, omierzyłabym, omierzyłabyś, omierzyłam, omierzyłaś, omierzyłby, omierzyłbym, omierzyłbyś, omierzyłem, omierzyłeś, omierzyło, omierzyłoby, omierzyły, omierzyłyby, omierzyłybyście, omierzyłybyśmy, omierzyłyście, omierzyłyśmy, omierzą, omierzę, omierzy, omierzycie, omierzymy, omierzysz, omierzono, omierzenia, omierzeniach, omierzeniami, omierzenie, omierzeniem, omierzeniom, omierzeniu, omierzeń, omierz, omierzcie, omierzcież, omierzmy, omierzmyż, omierzże, omierzywszy*
/r.z/	*omierzłego, omierzłemu, omierzłych, omierzłym, omierzłymi, omierzli, omierzła, omierzłą, omierzłe, omierzłej, omierznij, omierznijcie, omierznijcież, omierznijmy, omierznijmyż, omierznijże, omierzli, omierzliby, omierzlibyście, omierzlibyśmy, omierzliście, omierzliśmy, omierzł, omierzła, omierzłaby, omierzłabym, omierzłabyś, omierzłam, omierzłaś, omierzłby, omierzłbym, omierzłbyś, omierzłem, omierzłeś, omierzło, omierzłoby, omierzły, omierzłyby, omierzłybyście, omierzłybyśmy, omierzłyście, omierzłyśmy, omierzną, omierznę, omierznie, omierzniecie, omierzniemy, omierzniesz, omierzną, omierznę, omierźnie, omierźniecie, omierźniemy, omierźniesz, omierznięto, omierźnięto, omierznięcia, omierznięciach, omierznięciami, omierznięcie, omierznięciem, omierznięciom, omierznięciu, omierznięć, omierzłszy, omierzniętego, omierzniętemu, omierzniętych, omierzniętym, omierzniętymi, omierznięci, omierznięta, omierzniętą, omierznięte, omierzniętej*

2.5.12.15. Diada ortograficzna *rz* w strukturze: *#namarz*

Ortograficzny ciąg *#namarz* może być częścią wyrazu *namarzyć się* („marzyć przez dłuższy czas, nasycić się marzeniami"), w którym diada *rz* oznacza fonem /rz/. W wyrazach: *namarzać* („pokryć się szronem"), *namarzły* („pokryty się szronem"), *namarznąć, namarznięty* diada *rz* oznacza sekwencję /r.z/. Formy tych wyrazów zawierają w nagłosie jeden z następujących ciągów ortograficznych: *#namarza, #namarzł, #namarzl, #namarzn.*

Tabela 2.55 – Analiza słownika dotycząca struktury: *#namarz*

Transkrypcja diady *rz*	Wykaz form fleksyjnych
/rz/	*namarzyli, namarzyliby, namarzylibyście, namarzylibyśmy, namarzyliście, namarzyliśmy, namarzył, namarzyła, namarzyłaby, namarzyłabym, namarzyłabyś, namarzyłam, namarzyłaś, namarzyłby, namarzyłbym, namarzyłbyś, namarzyłem, namarzyłeś, namarzyło, namarzyłoby, namarzyły, namarzyłyby, namarzyłybyście, namarzyłybyśmy, namarzyłyście, namarzyłyśmy, namarzą, namarzę, namarzy, namarzycie, namarzymy, namarzysz, namarzono, namarz, namarzcie, namarzcież, namarzmy, namarzmyż, namarzże, namarząca, namarzącą, namarzące, namarzącego, namarzącej, namarzącemu, namarzący, namarzących, namarzącym, namarzącymi, namarząc*
/r.z/	*namarzłego, namarzłemu, namarzłych, namarzłym, namarzłymi, nienamarzłego, namarzli, namarźli, namarzła, namarzłą, namarzłe, namarzłej, namarzali, namarzaliby, namarzalibyście, namarzalibyśmy, namarzaliście, namarzaliśmy, namarzał, namarzała, namarzałaby, namarzałabym, namarzałabyś, namarzałam, namarzałaś, namarzałby, namarzałbym, namarzałbyś, namarzałem, namarzałeś, namarzało, namarzałoby, namarzały, namarzałyby, namarzałybyście, namarzałybyśmy, namarzałyście, namarzałyśmy, namarza, namarzacie, namarzają, namarzam, namarzamy, namarzasz, namarzano, namarzania, namarzaniach, namarzaniami, namarzanie, namarzaniem, namarzaniom, namarzaniu, namarzań, namarzaj, namarzajcie, namarzajcież, namarzajmy, namarzajmyż, namarzajże, namarzająca, namarzającą, namarzające, namarzającego, namarzającej, namarzającemu, namarzający, namarzających, namarzającym, namarzającymi, namarzając, namarznij, namarznijcie, namarznijcież, namarznijmy, namarznijmyż, namarznijże, namarzli, namarzliby, namarzlibyście, namarzlibyśmy, namarzliście, namarzliśmy, namarzł, namarzłaby, namarzłabym, namarzłabyś, namarzłam, namarzłaś, namarzłby, namarzłbym, namarzłbyś, namarzłem, namarzłeś, namarzło, namarzłoby, namarzły, namarzłyby, namarzłybyście, namarzłybyśmy, namarzłyście, namarzłyśmy, namarzną, namarznę, namarznie, namarzniecie, namarzniemy, namarzniesz, namarznięto, namarznięcia, namarznięciach, namarznięciami, namarznięcie, namarznięciem, namarznięciom, namarznięciu, namarznięć, namarzłszy, namarznął, namarznąłby, namarznąłbym, namarznąłbyś, namarznąłem, namarznąłeś, namarznąwszy, namarzniętego, namarzniętemu, namarzniętych, namarzniętym, namarzniętymi, namarznięci, namarznięta, namarzniętą, namarznięte, namarzniętej*

2.5.12.16. Diada ortograficzna *rz* zawsze oznaczająca sekwencję fonemów /r.z/

Istnieje zbiór wyrazów, w których ortograficzna diada *rz* zawsze oznacza sekwencję fonemów /r.z/. Wszystkie formy fleksyjne tych wyrazów można zidentyfikować na podstawie nagłosowych ciągów ortograficznych: *#obmarz, #odmarz, #przymarz, #nadmarz, #podmarz, #zmarz, #wmarz*. W tabeli 2.56 przedstawiono formy podstawowe wyrazów, które można zidentyfikować na podstawie wymienionych struktur.

Tabela 2.56 – Formy podstawowe wyrazów, w których diada *rz* zawsze oznacza sekwencję fonemów /r.z/

Lp.	Struktura nagłosu	Wykaz form podstawowych
1	*#obmarz*	*obmarzać, obmarzły, obmarznąć, obmarznięty*
2	*#odmarz*	*odmarzać, odmarzły, odmarznąć, odmarznięty*
3	*#przymarz*	*przymarzać, przymarzły, przymarznąć, przymarznięty*
4	*#nadmarz*	*nadmarzać, nadmarzły, nadmarznąć, nadmarznięty*
5	*#podmarz*	*podmarzać, podmarzły, podmarznąć, podmarznięty*
6	*#zmarz*	*zmarzlak, zmarzlina, zmarzlinowy, zmarzluch, zmarzłoć, zmarzłość, zmarzły, zmarznąć, zmarznięty*
7	*#wmarz*	*wmarzać, wmarzły, wmarznąć, wmarznięty*

2.6. Transkrypcja liter i diad ortograficznych oznaczających spółgłoski właściwe bezdźwięczne

2.6.1. Litera *c*

Litera *c* oznacza pojedynczy fonem /ts/. Może też wchodzić w skład digrafemów: *cz* lub *ci* (zostały one omówione oddzielnie). Analiza oparta na regułach zawartych w tabeli 28 w *Automatyzacji…* wykazała, że w słowniku w stopniu marginalnym występują wyrazy, w których fonem odpowiadający literze *c* nie jest realizowany, natomiast w korpusie takie wyrazy nie wystąpiły w ogóle (tabela 2.57). Również rzadko są spotykane wyrazy, w których litera *c* oznacza dźwięczny fonem /dz/. Analiza dotycząca umiejscowienia litery *c* w kontekście udźwięczniającym została omówiona podrozdziale 3.1 w *Asymilacji…* .

Tabela 2.57 – Analiza słownika i korpusu dotycząca litery *c*

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	X[*c*](T–*h*)DT=1	2	*zeświecczcie,zeświecczcież*	0	

2.6.2. Litera *ć*

Reguły przytoczone w tabeli 2.58 dotyczą wyrazów, w których fonem odpowiadający literze *ć* nie jest realizowany (np.: *sześćset, sześćdziesiąt*). Inne reguły zawarte w pierwszej oraz w drugiej części tabeli 29 w *Automatyzacji…* są związane z wymową dźwięczną; analiza oparta na tych regułach została omówiona w podrozdziale 3.2 w *Asymilacji…* .

Tabela 2.58 – Analiza słownika i korpusu dotycząca litery *ć*

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	*ś*[*ć*]*s*=1	51	*sześćset, sześćsetlecia, sześćsetnym*	24	*sześćset*
2	*ś*[*ć*]*c*=1	387	*bezcześćcie, doczyśćcie, mieśćcie*	26	*puśćcie, dopuśćcie, czyśćcu*
3	(X–#-*a*–*r*–*ś*)[*ć*]DD =1	874	*dziewięćdziesiąt, dziewięćdziesięcioletni, dziewięćdziesiątym*	287	*pięćdziesiąt, pięćdziesiątym, pięćdziesiątej*
4	*ś*[*ć*]DD=1	295	*sześćdziesiąt, sześćdziesięciolatek, sześćdziesięcioletni*	73	*sześćdziesięciu, sześćdziesiąt, sześćdziesiąte*

2.6.3. Litera *h*

Transkrypcja litery *h* jest oparta na jednej regule: X[*h*]X=/x/ (tabela 30 w *Automatyzacji…*). Litera *h* powinna być transkrybowana jako fonem /x/, chyba że jest częścią digrafemu *ch*, który również jest transkrybowany jako fonem /x/. Wynika to z faktu, że ewentualne dźwięczne głoski są alofonami fonemu /x/ [Nawrocki, Gonet 2004; Nawrocki 2004].

2.6.4. Litera *f*

Reguła zawarta w tabeli 2.59 dotyczy redukcji w transkrypcji diady ortograficznej *ff* (redukcja nie występuje w wypadku wewnątrzwyrazowej diady *ff* umiejscowionej przed literą oznaczającą samogłoskę). Autorka *Automatyzacji…* podkreśliła, że takie uproszczenie jest opcjonalne. Fonem odpowiadający literze *f* może występować w kontekście udźwięczniającym; stosowna analiza związana z tym zagadnieniem została zawarta w podrozdziale 3.4 w *Asymilacji…* .

Tabela 2.59 – Analiza słownika i korpusu dotycząca litery *f*

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	*f*[*f*]TA=1	104	*fotooffset, offsetu, offsetobiorca*	0	

2.6.5. Litera *k*

W tabeli 2.60 zamieszczono wynik analizy wykonanej na podstawie reguł pochodzących z tabeli 32 w *Automatyzacji…* (z pierwszej oraz z drugiej części). Przyjęto założenie, że /c/ jest wyodrębnionym fonemem. Dwa początkowe wiersze dotyczą litery *k* umiejscowionej przed literą *i* lub przed literą *j*. W takim kontekście litera *k* powinna być transkrybowana jako fonem /c/. Reguły ujęte w wierszach 3–5 dotyczą wymowy geminaty ortograficznej *kk* w wyrazach: *lekki, lekko, miękki, miękko*. Zgodnie z regułą zapisaną w wierszu trzecim pierwsza litera *k* w wyrazie *lekki* jest transkrybowana jako fonem /c/ (to oznacza, że redukcja nie występuje). W wyrazach *miękki* oraz *miękko* fonem odpowiadający pierwszej literze *k* jest pomijany. W wierszach tabeli 2.60 podano dodatkowe warianty reguł, które obowiązują przy założeniu, że głoska [k̀] (AS) jest alofonem fonemu /k/, zatem fonem /c/ nie jest wyodrębniony. Analizę dotyczącą litery *k* w kontekście udźwięczniającym omówiono w podrozdziale 3.5 w *Asymilacji…* .

Tabela 2.60 – *A*naliza słownika i korpusu dotycząca litery *k*

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	X[*k*]*i*=/c/ lub X[*k*]*i*=/k/	269887	*dolnołużycki, skiszonego, suwalskiego*	92324	*szklanki, cukierki, czarodziejskiego*
2	X[*k*]*j*=/c/ lub X[*k*]*j*=/k/	78	*skjerem, Łukjanience, skjorami*	0	
3	(X–*ę*)[*k*]*ki*=/c/ lub (X–*ę*)[*k*]*ki*=/k/	184	*lekkich, lekkim, lekkimi*	278	*lekki, Mekki, lekkiego*
4	*ę*[*k*]*k*(V–*i*)=1	245	*miękkawy, ultramiękką, miękkopuchy*	111	*miękka, miękkość, miękkopióra*
5	*ę*[*k*]*ki*=1	55	*miękki, ultramiękkie, półmiękkie*	238	*miękkich, miękkiego, miękkim*

2.6.6. Litera *p*

Litera *p* oznacza zwarty, bezdźwięczny fonem /p/. Może być ona transkrybowana jako dźwięczny fonem /b/ – analiza została omówiona w podrozdziale 3.6 w *Asymilacji…* .

2.6.7. Litera *s*

W przyjętej transkrypcji podstawowej litera *s* oznacza bezdźwięczny, szczelinowy fonem /s/. Możliwość transkrypcji tej litery jako fonem dźwięczny została omówiona w podrozdziale 3.7 w *Asymilacji…* . Poza tym w tabeli 2.61 zawarto dwie reguły pochodzące z tabeli 34 w *Automatyzacji…*, dotyczące transkrypcji litery *s* jako fonem /s'/ w określonych kontekstach: przed diadą *si* oraz przed triadą *mie* (ten wariant dotyczy wymowy mniej starannej).

Tabela 2.61 – Analiza słownika i korpusu dotycząca litery *s*

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	X[*s*]*si*=/s'/	652	*nassijmy, bessie, wyssijmy*	165	*bessie, Barbarossie*
2	X[*s*]*mie*O=/s'/	16	*basmie, Kosmie, pasmie*	0	

W dalszym ciągu przedstawiono wynik analizy dotyczącej ortograficznej diady *si*, która zgodnie z zasadami ortofonicznymi języka polskiego powinna być transkrybowana jako fonem /s'/ (przed literą oznaczającą samogłoskę inną niż *i*) lub jako sekwencja fonemów /s'.i/ (przed literą oznaczającą spółgłoskę lub półsamogłoskę). Reguły przytoczone w tabeli 2.62 pochodzą z drugiej części tabeli 34 w *Automatyzacji…* . Dotyczą one wyrazów, w których diada ortograficzna *si* oznacza ciąg fonemów /s.i/. Taki sposób transkrypcji można zatem uznać za wyjątkowy. Wymowa niektórych nazw własnych może się cechować wariantywnością lub może być trudna do określenia [Łobacz 1999; Madelska 2005].

Tabela 2.62 – Analiza słownika i korpusu dotycząca diady *si* (część 1)

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	#[*s*]*i*O=/s/	1	*si*	15	*si*
2	#[*s*]*ialic*=/s/	11	*sialiczny, sialiczna, sialiczną*	0	
3	#[*s*]*ial*(X–*i*) =/s/	30	*sialach, sialami, sialowego*	0	
4	#[*s*]*ido*=/s/	83	*sidol, ~~Sidorkiewicz~~, ~~Sidorska~~*	0	
5	#[*s*]*ilik*=/s/	93	*silikat, silikon, silikoza*	0	
6	#[*s*]*il*(A-*i*) =/s/	203	*silan, silosem, ~~silenie~~*	57	*~~siłą~~, ~~siłąc~~, ~~siłę~~*
7	#[*s*]*il*O=/s/	1	*sil*	1	*sil*
8	#[*s*]*im*=/s/	316	*simir, simkach, ~~Simińskich~~*	4	*simkę*, *Simba*
9	#[*s*]*inan*=/s/	20	*sinantrop*	0	
10	#[*s*]*inol*=/s/	33	*sinolog, sinologiczne*	0	
11	#[*s*]*inu*=/s/	64	*sinus, sinusoidalna, sinuśnica*	0	
12	#[*s*]*is*=/s/	44	*sistrum, sistrami*	0	
13	#*mak*[*s*]*i*O=/s/	1	*maksi*	0	
14	#*plek*[*s*]*iO*=/s/	1	*pleksi*	0	
15	#*mak*[*s*]*im*=/s/	50	*Maksim, Maksima, Maksimowicz*	2	*Maksimowicz, maksimum*
16	#*elik*[*s*]*i*(X–O–{*a,d,l,m,n,s*})=/s/	10	*eliksir, eliksirach, eliksirami*	21	*eliksirów, eliksir, eliksiru*
17	#*mak*[*s*]*i*(X–O–{*a,d,l,m,n,s*})=/s/	11	*~~Maksie~~, ~~Maksiem~~*	0	
18	#*plek*[*s*]*i*(X–O–{*a,d,l,m,n,s*})=/s/	31	*pleksiglas, pleksigum, pleksiglasie*	1	*pleksiglasowej*

Analizując dane zawarte w tabeli 2.62, łatwo zauważyć, że przytoczone reguły wymagają aktualizacji. Niektóre dotyczą wyrazów, w których diada *si* nie powinna być transkrybowana jako sekwencja fonemów /s.i/. Na podstawie uzupełniającej analizy słownika wyznaczono zbiór form fleksyjnych, których wymowa jest niezgodna z tymi regułami. W kolumnie drugiej w tabeli 2.63 ponownie przytoczono reguły z tabeli 34 w *Automatyzacji…* (został użyty tylko pierwszy człon reguł; niektóre struktury zmodyfikowano). W kolumnie trzeciej są wymienione tylko te wyrazy, w których diada ortograficzna *si* oznacza ciąg fonemów /s.i/. W kolumnie czwartej ujęto formy podstawowe wyrazów (w niektórych wypadkach również formy fleksyjne), które pokrywają się ze strukturami wymienionymi w kolumnie drugiej i w których diada *si* oznacza sekwencję fonemów /s'.i/. Ostatnia kolumna w tabeli 2.63 zawiera nagłosowe struktury ortograficzne umożliwiające identyfikację form fleksyjnych, w których diada *si* powinna być transkrybowana jako sekwencja fonemów /s.i/. Te struktury nie mogą występować w wyrazach, w których diada *si* oznacza fonemy /s'.i/. Przy projektowaniu systemu transkrypcji można zatem przyjąć następującą transkrypcję domyślną diady *si*: sekwencja fonemów /s'.i/ przed literą oznaczającą spółgłoskę lub półsamogłoskę oraz fonem /s'/ przed literą oznaczającą samogłoskę (inną niż *i*). Uzyskany zbiór struktur nagłosowych z ostatniej kolumny może być użyty do identyfikacji wyjątków. W wypadku form: *sil, sili, sile* dopuszczalne są dwie możliwości transkrypcji w zależności od znaczenia.

Tabela 2.63 – Analiza słownika i korpusu dotycząca diady *si* (część 2)

Lp.	Struktura nagłosu według M.S.-B.	Transkrypcja diady ortograficzne *si* jako /s.i/	Transkrypcja diady ortograficzne *si* jako /s'.i/	Nagłosowe struktury ortograficzne umożliwiające identyfikację wyjątków
1	*#si*O	*si*		*#si#*
2	*#sialic*	*sialiczny*		*#sial*
3	*#sial*(X–*i*)	*sialma, sialowy, sial*		*#sial* (jw.)
4	*#sido*	*sidol*	*Sidorkiewicz, Sidorski, Sidor, Sidorówka*	*#sidol*
5	*#silik*	*silikat, silikatowy, silikatyzacja, silikażel, silikon, silikonowy, silikotermia, silikoza*		*#silik*
6	*#sil*(A–*i*)	*silan, silanol, silentium, Silesia, Siloe, silos, silosokombajn, silosować, silosowy, silumin, siluminowy,* formy wyrazu *sil*: *silach, silami, silom, silem, silowi, silu, silów, sile*	niektóre formy wyrazu *silić,* np.: *silą, siląc, siląca, silącą, silące, silącego, silącej, silącemu, silący, silących, silącym, silącymi, silenia, sileniach, sileniami, silenie, sileniem, sileniom, sileniu, sileń, silę;* wyrazy: *sile* (forma wyrazu: *siła*)*, Silec*	*#sila, #silen, #siles, #silo, #silu, #siló, #silem*
7	*#sili*(X–*k*)	*sili* (forma od: *sil*), *silifikacja, silimanit, silit*	niektóre formy wyrazu *silić, np.*: *sili, sililiby, silicie, silimy, siliłem, silisz*	*#silif, #silim, #silit*
8	*#sil*O	*sil*	*sil* (forma od: *silić*)	
9	*#sim*	*SIM, sima, simental, simentale, simentalerski, simentalski, simir, Simbierowicz, Simka, Simonow, simpleks, simpleksowy, Simplus, Simson, Simula, Simba*	*Simiński*	*#sim#, #sima, #simą, #simę, #simie, #simy, #sime, #simir, #simb, #simk, #simo, #simp, #sims, #simul*
10	*#sin*(X–*i*)	*sinfonietta, singel, singleton, singelton, singlista, sinantrop, sinantropus, sine, sinolog, sinologia, sinologiczny, sinto, sinus, sinusoida, sinusoidalny, sinusowy, sinuśnica*		*#sinf, #sing, #sina, #sine, #sino, #sint, #sinu*
11	*#sis*	*sisal, sistrum*	*Sisice, Sisicki*	*#sisa, #sist*
12	*#eliksi*O			
13	*#maksi*O	*maksi*		*#maksi#*
14	*#pleksi*O	*pleksi*		*#pleksi#*
15	*#eliksim*			
16	*#maksim*	*Maksim, maksimafilia, Maksimow, Maksimowicz, maksimum*		*#maksim*
17	*#pleksim*			
18	#eliksi(X–O–{*a,d,l,m,n,s*})	*eliksir*		*#eliksir*
19	*#maksi*(X–O–{*a,d,l,m,n,s*})		*Maksio, Maksiu, Maksie*	
20	*#pleksi*(X–O–{*a,d,l,m,n,s*})	*pleksiglas, pleksiglasowy, pleksigum*		*#pleksig*

2.6.8. Litera *ś*

W transkrypcji podstawowej litera *ś* oznacza bezdźwięczny szczelinowy fonem /s'/. Litera *ś* może być również transkrybowana w wariancie dźwięcznym (jako fonem /z'/). Wszystkie reguły zawarte w pierwszej oraz drugiej części tabeli 35 w *Automatyzacji…* dotyczą tej możliwości; pełna analiza na nich oparta została omówiona w podrozdziale 3.8 w *Asymilacji…* .

2.6.9. Litera *t*

Transkrypcję litery *t* komplikuje fakt, że występuje ona w triadach ortograficznych *trz* oraz *trw*. Poza tym litera *t*, podobnie jak inne litery oznaczające spółgłoski bezdźwięczne, może być transkrybowana w wariancie dźwięcznym. Ta możliwość została dokładnie omówiona w podrozdziale 3.9 w *Asymilacji…* .

Reguła *n*[*t*]Rw=/d/ pochodzi z drugiej części tabeli 36 w *Automatyzacji…* . Dotyczy ona transkrypcji triady ortograficznej *trw*. Litera *t* oznacza w niej fonem /t/ (na przykład w wyrazach: *wytrwały, trwoga*), za wyjątkiem wyrazów złożonych, w których pierwszy człon stanowi prefiks *kontr*, a drugi człon rozpoczyna się od litery *w.* W takich konstrukcjach litera *w* jest umiejscowiona za junkturą morfologiczną, co sprawia, że jest transkrybowana w wariancie dźwięcznym. Również litera *t* jest wtedy transkrybowana w wariancie dźwięcznym. Słownik *SJP.pl* zawiera następujące formy podstawowe wyrazów zawierających omawianą konstrukcję: *kontrwalacja, kontrwynalazca, kontrwywiad, kontrwywiadowca, kontrwywiadowczy*.

W dalszym ciągu omówiono możliwości transkrypcji triady ortograficznej *trz*. W tabeli 2.64 przedstawiono podsumowanie (na podstawie reguł zawartych w *Automatyzacji…*), które ma następujący układ: w kolumnie drugiej wyszczególniono różne możliwości kontekstu prawostronnego triady *trz*, a w kolumnie trzeciej oraz w kolumnie czwartej podano informacje dotyczące dwóch możliwości transkrypcji. Przy pierwszej możliwości *trz* jest częścią wyrazu *wewnątrz* (kolumna trzecia), natomiast możliwość druga dotyczy pozostałych wypadków (kolumna czwarta).

Jeżeli ortograficzna triada *trz* należy do członu złożenia: *wewnątrz*, to przy ustalaniu statusu dźwięczności można odnieść się do zasad właściwych dla fonetyki międzywyrazowej. Jeżeli w kontekście prawostronnym znajduje się litera bądź diada ortograficzna oznaczające spółgłoskę właściwą, to status dźwięczności tej spółgłoski decyduje o ewentualnej sonoryzacji *trz*. Jeżeli w kontekście prawostronnym znajduje się samogłoska, półsamogłoska lub spółgłoska sonorna, to sonoryzacja jest prawdopodobna tylko w wymowie krakowskopoznańskiej.

Jeżeli triada *trz* nie jest częścią członu *wewnątrz* (kolumna czwarta tabeli 2.64), to zasady dotyczące modyfikacji statusu dźwięczności są podobne w obu typach wymowy i pokrywają się z ogólnymi zasadami dotyczącymi modyfikacji statusu dźwięczności w ramach wewnątrzwyrazowych grup spółgłoskowych. Warto zauważyć, że triada *trz* przed samogłoską jest transkrybowana jako sekwencja fonemów /t.S/ – wynika to z reguł zawartych w *Automatyzacji…* .

Tabela 2.64 – Podsumowanie dotyczące możliwości transkrypcji ortograficznej triady *trz*

Lp.	Kontekst prawostronny	Transkrypcja triady *trz* w wyrazie *wewnątrz*	Transkrypcja triady *trz* w wyrazie innym niż *wewnątrz*
1	Spółgłoska właściwa bezdźwięczna	/tS/ (WKP i WW) np. *wewnątrzsieciowy*	/tS/ (WKP i WW) np. *trzcinowy*
2	Spółgłoska właściwa dźwięczna	/dZ/ (WKP i WW) np. *wewnątrzgrupowy*	/dZ/ (WKP i WW) np. *dopatrzże*
3	Samogłoska	/dZ/ (WKP) lub /tS/ (WW) np. *wewnątrzatomowy*	/t.S/ (WKP i WW) np. *jątrzysz*
4	Spółgłoska sonorna z wyłączeniem wyrazów zawierających sufiks *my*	/dZ/ (WKP) lub /tS/ (WW) np. *wewnątrzmolekularny*	/tS/ (WKP i WW) np.: *trzmiel, wywietrznik*
5	Sufiks *my*		/dZ/ (WKP) lub /tS/ (WW) np. *popatrzmy*

Z informacji zamieszczonych w tabeli 2.64 wynika, że w wymowie krakowskopoznańskiej triada *trz* umiejscowiona w członie złożenia: *wewnątrz* ulega zarówno sonoryzacji, jak i asybilacji (jest transkrybowana jako fonem /dZ/) pod warunkiem, że jest umiejscowiona przed literą oznaczającą: spółgłoskę właściwą dźwięczną, samogłoskę lub spółgłoskę sonorną. W wymowie warszawskiej dźwięczny wariant transkrypcji jest możliwy jedynie przed literą oznaczającą spółgłoskę właściwą dźwięczną.

W *Automatyzacji…* zostały omówione dwie reguły, które dotyczą redukcji fonemu odpowiadającego literze *t*: X[*t*]*ci*=1 oraz X[*t*]*t*O=1. Pierwsza przytoczona reguła odnosi się do triady *tci*, na przykład w wyrazach: *Lisetcie, Lotcie, risotcie* (formy od: *Lisetta, Lotta, risotto*). Natomiast reguła X[*t*]*t*O=1 dotyczy redukcji w ramach diady

ortograficznej *tt* umiejscowionej w wygłosie wyrazu, na przykład w wyrazach: *watt* (pas wybrzeża), *Wiolett* i *pitt* (formy od: *Wioletta, pitta*).

2.6.10. Diada orograficzna *sz*

Diada ortograficzna *sz* może być transkrybowana jako bezdźwięczny, szczelinowy fonem /S/ (wariant zgodny z transkrypcją podstawową) lub w wariancie dźwięcznym jako fonem /Z/. Analiza dotycząca drugiej możliwości została omówiona w podrozdziale 3.10 w *Asymilacji…* .

2.6.11. Diada ortograficzna *cz*

Diada ortograficzna *cz* domyślnie oznacza bezdźwięczny, zwarto-szczelinowy fonem /tS/. Może być również transkrybowana w wariancie dźwięcznym – analiza dotycząca tej możliwości została omówiona w podrozdziale 3.11 w *Asymilacji…* .

2.6.12. Diady ortograficzne: *ci, si* oraz *zi* (przed literą oznaczającą samogłoskę)

W tej części został poszerzony opis zawarty w punkcie f w części 2.2.6. Dotyczy on struktury: {*s,c,z*}*i*A, w której litera *i* jest tylko znakiem miękkości, zatem jest ona komponentem digrafemu. Diady: *ci, si* oraz *zi* są digrafemami tylko w kontekście samogłoskowym (z wyłączeniem litery *i,* która jednak nie występuje po takich konstrukcjach). Dlatego nie zostały włączone do transkrypcji podstawowej (zob. podrozdz. 1.3). Można im jednak przypisać domyślny status tzw. dźwięczności ortograficznej (digrafemy: *ci* oraz *si* są bezdźwięczne, natomiast *zi* jest dźwięczny). Tabela 2.65 jest zmodyfikowaną wersją tabeli 2.14 i dotyczy omawianych diad ortograficznych.

Tabela 2.65 – Analiza słownika i korpusu dotycząca diad ortograficznych: *si, ci* oraz *zi*

Lp.	Reguła według M.S.-B.	Liczba wyrazów w *SJP.pl*	Przykładowe wyrazy z *SJP.pl*	Liczba wyrazów w korp.	Przykładowe wyrazy z korpusu
1	[*si*](A–*i*)=/s'/	64152	*łosiem, księżyc, Sierackiego*	127810	*posiada, miesiąc, sąsiednie*
2	[*ci*](A–*i*)=/ts'/	299412	*odległościami, napięciowej, niecierpiącego*	84288	*dwanaście, wreszcie, pociąg*
3	(X–*d*)[*zi*](A–*i*)=/z'/	21863	*ziarnowemu, zapyziałą, uziemionym*	12570	*wziął, gałęziami, jeziorko*

3. Podstawy wielowarstwowego modelu transkrypcji

3.1. Wprowadzenie

W rozdziale tym omówiono założenia dotyczące wielowarstwowego modelu transkrypcji (dalej: WMT). Został on zaprojektowany na podstawie rozwiązań zaproponowanych przez Marię Steffen-Batogową, choć znacznie różni się od swojego pierwowzoru. Występują jednak pewne cechy wspólne – przede wszystkim możliwość określania kontekstu lewostronnego oraz kontekstu prawostronnego w regułach czy też możliwość definiowania i stosowania zbiorów segmentów. Niektóre reguły pochodzące z *Automatyzacji…* mogą być przeniesione do projektów opartych na modelu wielowarstwowym bez konieczności wprowadzania zmian, inne natomiast wymagają modyfikacji. Istnieją też reguły, które w ogóle nie są odpowiednie dla omawianego modelu. Z punktu widzenia terminologii i rozwiązań właściwych dla WMT rozwiązanie przedstawione przez Marię Steffen-Batogową w *Automatyzacji…* można uznać za system jednowarstwowy.

U podstaw zaprezentowanego modelu wielowarstwowego leży wieloetapowe przetwarzanie wyrazów ortograficznych. W konsekwencji wyrazy stają się transkrypcją stopniowo. Ma to związek z różnymi problemami, takimi jak: wydzielanie digrafemów i trigrafemów, zmiana statusu dźwięczności, czy uproszczenia w grupach spółgłoskowych. Model wielowarstwowy umożliwia przyporządkowywanie poszczególnych procesów (na przykład procesów asymilacyjnych) i innych przekształceń do różnych warstw. Dzięki temu zbiór reguł staje się bardziej przejrzysty. Poza tym reguły związane z określonym procesem mogą być modyfikowane niezależnie od reguł dotyczących innych przekształceń, które zostały przypisane do innych warstw. Z tych właściwości wynikają podstawowe zalety wielowarstwowego modelu transkrypcji – można mówić o elastyczności oraz o przejrzystości projektów na nim opartych. Nie ma założenia dotyczącego liczby warstw, ich funkcji czy też liczby reguł zdefiniowanych w poszczególnych warstwach. Specjalny mechanizm realizujący koncepcję WMT umożliwia projekcję poszczególnych znaków ortograficznych na ostateczny zapis fonemów lub głosek (osiągany po przetworzeniu w ostatniej warstwie systemu opartego na WMT).

W WMT nie ma założenia dotyczącego sposobu zapisywania wyrazów w czasie przetwarzania w poszczególnych warstwach. Oznacza to, że można przyjąć dowolny (umowny) sposób zapisu. W przykładowym projekcie omówionym w rozdziale czwartym użyto alfabetu SAMPA (zob. podrozdz. 1.1). Problem nazewnictwa segmentów w czasie przetwarzania wyrazów przy użyciu rozwiązań opartych na WMT został już poruszony we *Wprowadzeniu*.

3.2. Podstawowe pojęcia

Istnieje kilka podstawowych pojęć, które mają kluczowe znaczenie dla opisu budowy i funkcjonowania WMT. Należy do nich wspomniane już pojęcie warstwy. Warstwa w WMT jest samodzielnym mechanizmem mającym własny zestaw reguł i służącym do przetwarzania ciągów segmentów. Pojedyncza warstwa jest przetwornikiem, do którego są przekazywane dane tekstowe i na wyjściu którego są umieszczane dane tekstowe przetworzone. Sposób przetwarzania jest sprecyzowany przez reguły zdefiniowane w tej warstwie. Istotny jest fakt, że poszczególne warstwy w systemie transkrypcji opartym na WMT funkcjonują niezależnie od siebie. Trzeba jednak zaznaczyć, że reguły należące do określonej warstwy muszą być skonstruowane w ten sposób, żeby możliwe było przetwarzanie zapisu wyrazów otrzymywanych z warstwy poprzedniej. Inaczej mówiąc: reguły związane z kolejną warstwą muszą „rozumieć” sposób zapisywania wyrazów przetworzonych w warstwie poprzedniej.

Kolejne istotne pojęcia elementarne to: segment oraz indeks. Segment jest fragmentem przetwarzanego wyrazu. Podział danego wyrazu na segmenty funkcjonuje na dowolnym etapie jego przetwarzania (w ramach dowolnej warstwy). W WMT występują dwa rodzaje segmentów: pojedyncze i wielokrotne. Segmentami pojedynczymi mogą być podstawowe jednostki związane z danym systemem zapisu. W zapisie ortograficznym są to litery, natomiast w transkrypcji fonologicznej pojęcie to odnosi się do fonemów. Mogą to również być jednostki umowne. Oto przykłady segmentów pojedynczych stosowanych w projekcie omówionym w niniejszej monografii (zob. rozdz. 4):

– *a* (segment obejmujący jedną literę);
– /a/ (segment obejmujący jeden fonem);
– /ts/ (segment obejmujący jeden fonem).

Zapisywanie kilku jednostek, które z określonych względów stanowią wyodrębnioną całość, można zrealizować w segmentach wielokrotnych. Poszczególne składniki segmentu wielokrotnego (na przykład fonemy) łączy się przy użyciu znaku: &. Oto dwa przykłady segmentów wielokrotnych:
– /e/&/n/ (segment wielokrotny obejmujący ciąg fonemów);
– /o/&/w~/ (segment wielokrotny obejmujący ciąg fonemów).

Indeks jest numerem segmentu w wyrazie. Trzeba podkreślić, że w momencie przekazywania wyrazu do pierwszej warstwy liczba indeksów jest identyczna jak liczba znaków w tym wyrazie powiększona o 2 ze względu na znaczniki początku i końca wyrazu. Tabela 3.1 obrazuje tę zależność w przykładowym wyrazie *wszędzie*.

Tabela 3.1 – Początkowy stan przypisania indeksów do segmentów w wyrazie: *wszędzie*

Segment	#	*w*	*s*	*z*	*ę*	*d*	*z*	*i*	*e*	#
Indeks	0	1	2	3	4	5	6	7	8	9

Przypisanie indeksów do poszczególnych segmentów może się zmieniać w miarę przetwarzania wyrazu w kolejnych warstwach. Istotny jest fakt, że początkowy zbiór indeksów jest identyczny jak zbiór indeksów w wyrazie po przetworzeniu w ostatniej warstwie danego rozwiązania opartego na WMT (w podanym przykładzie jest to zbiór dziesięcioelementowy). W czasie przetwarzania wyrazu mogą występować zmiany dotyczące przypisania poszczególnych indeksów do segmentów oraz zmiany dotyczące liczby oraz treści poszczególnych segmentów. Zmiany zachodzą w wyniku operacji, które są zdefiniowane w regułach związanych z kolejnymi warstwami. Liczba indeksów w przetwarzanym wyrazie może ulec zmianie tylko w jednym wypadku i trzeba podkreślić, że zawsze jest to zmiana tymczasowa stosowana w określonej warstwie. Ma to związek z etykietą, która może być wygenerowana w wyniku przetwarzania następnego wyrazu (wyrazy są przetwarzane od strony prawej do strony lewej). Etykieta może być dołączona do danego wyrazu jako oddzielny segment (za segmentem zawierającym znacznik końca wyrazu). Ten mechanizm został szczegółowo omówiony w podrozdziale 5.6.

Warto zaznaczyć, że w omówionym przykładowym projekcie systemu transkrypcji opartym na WMT (zob. rozdz. 4) w wyniku przetwarzania wyrazów w pierwszej warstwie zwracany jest zapis w przyjętej transkrypcji podstawowej z niezbędnymi modyfikacjami dotyczącymi przede wszystkim kontekstu samogłoskowego. W rozwiązaniu tym zapis ortograficzny funkcjonuje więc tylko do momentu przekazania wyrazu do pierwszej warstwy.

3.3. Struktura reguł WMT

Każda reguła składa się z dwóch zasadniczych części – pierwszy człon reguły wskazuje na fragment wyrazu, dla którego będzie ona stosowana (przetwarzany wyraz musi zawierać ten fragment, żeby reguła mogła być użyta). Natomiast w członie drugim są zdefiniowane wszelkie modyfikacje, które dotyczą fragmentu wyrazu wskazanego w pierwszym członie. Poza tym do danej reguły może być dołączona lista wyjątków lub lista włączeń (zob. podrozdz. 3.5). Obejmuje ona nagłosowe, śródgłosowe lub wygłosowe ciągi ortograficzne, umożliwiające identyfikację określonych form fleksyjnych.

Oba człony reguły są podzielone na taką samą liczbę mniejszych, analogicznych względem siebie elementów, które odpowiadają poszczególnym segmentom stanowiącym fragmenty przetwarzanego wyrazu. Schemat budowy reguły WMT został przedstawiony w tabeli 3.2.

Tabela 3.2 – Schemat budowy reguły WMT

Pierwszy człon reguły			Drugi człon reguły		
element 1	element 2	element n	element 1*	element 2*	element m

W pierwszym członie reguły należącej do pierwszej warstwy określonego rozwiązania (systemu transkrypcji) opartego na WMT każdy element odpowiada dokładnie jednej literze (zob. podrozdz. 3.1). Znaczenie elementów reguł w kolejnych warstwach wynika z zastosowanych wcześniej reguł oraz z przyjętego sposobu zapisywania segmentów – nie jest to z góry zdefiniowane. Poszczególne elementy są rozdzielane znakiem kropki. Jeden element pierwszego członu reguły może odnosić się do konkretnego segmentu wyrazu lub może być to zapis określający zbiór segmentów, na przykład:

– [A]–*ą*–*ę* (zbiór A bez liter: *ą* oraz *ę*);
– [A]+*ł*+*j* (zbiór A powiększony o litery: *ł* oraz *j*);
– *p*+*d*+*t*+*d*+*k*+*g* (sześcioelementowy zbiór liter).

Wszystkie używane oznaczenia zbiorów muszą być najpierw zdefiniowane, na przykład definicja zbioru A mogłaby wyglądać następująco: [A]={*a, ą, e, ę, o, ó, u, i, y*}. Poniższa lista zawiera przykładowe konstrukcje pierwszego członu reguły:

– *p.i.*[A]–*i* (zapis właściwy dla wyrazów: *pianino, opieszały, pieniądze*);
– *r.ł.b* (zapis właściwy dla wyrazów: *dotarłby, nażarłby, wsparłbyś*);
– *a.i* (zapis właściwy dla wyrazów: *zaistnieć, zaiskrzyć*).

Te trzy przykłady są odpowiednie dla pierwszej warstwy rozwiązania opartego na WMT (dotyczą one wyrazów ortograficznych). Trzeci element w pierwszym przykładzie obejmuje zbiór dopuszczalnych segmentów (tutaj: zbiór liter). Symboli zbiorów używa się w celu określenia dopuszczalnego kontekstu dla elementu umiejscowionego w pierwszym członie reguły, któremu towarzyszy definicja modyfikacji zapisana na analogicznej pozycji w drugim członie tej samej reguły. Dla elementu w pierwszym członie reguły, który obejmuje pewien zbiór segmentów, nie powinna być definiowana modyfikacja w drugim członie tej reguły. Wyjątek od tej zasady dotyczy funkcjonowania list odwzorowań. Ten mechanizm umożliwia definiowanie zbioru przekształceń dla różnych wariantów segmentu, przy czym z góry nie wiadomo, jaki segment będzie częścią konkretnego przetwarzanego wyrazu. Mechanizm został szczegółowo omówiony w części poświęconej zagadnieniom zaawansowanym związanym z WMT (zob. podrozdz. 5.2).

Informacje przedstawione w dalszym ciągu dotyczą konstrukcji drugiego członu w regułach WMT. Liczba elementów w obu członach jest identyczna, więc poszczególne elementy w drugim członie danej reguły odnoszą się do analogicznych elementów w pierwszym członie tej reguły, które z kolei odnoszą się do segmentów w przetwarzanych wyrazach. Charakter i funkcje elementów w drugim członie są jednak inne. Każdy taki element może zawierać zapis związany z jedną z trzech możliwości:

– przepisanie w niezmienionej postaci segmentu wyrazu wskazanego przez analogiczny element umiejscowiony w pierwszym członie reguły;
– przepisanie w postaci zmodyfikowanej segmentu wyrazu wskazanego przez analogiczny element umiejscowiony w pierwszym członie reguły;
– użycie specjalnego polecenia dotyczącego operacji, która ma być przeprowadzona na segmencie wyrazu wskazanym przez analogiczny element umiejscowiony w pierwszym członie reguły.

Ostatnia opcja wymieniona na powyższej liście wymaga dodatkowego wyjaśnienia. W drugim członie reguł można używać następujących poleceń:

– !NC (ang. no change) – to polecenie może być stosowane w każdym wypadku, a w szczególności dla elementów, które w pierwszym członie reguły zostały użyte tylko dla określenia kontekstu;
– !ML (ang. move left) – w wyniku użycia tego polecenia następuje usunięcie danego segmentu w przetwarzanym wyrazie, a indeks, który był przypisany do tego segmentu, zostaje dołączony do indeksu segmentu poprzedzającego dany segment (a więc pojedynczy segment może być powiązany z kilkoma indeksami);
– !RM (ang. remove) – to polecenie usuwa segment, ale pozostawia indeks na swoim miejscu – po zastosowaniu tego polecenia dany indeks zostaje przypisany do specjalnego segmentu: EMP (ang. empty).

Operacje na wyrazach przetwarzanych w danej warstwie mogą być wykonywane między innymi przy użyciu wymienionych poleceń. Trzeba zaznaczyć, że określona reguła jest stosowana w odniesieniu do przetwarzanego wyrazu pod warunkiem, że zawiera on ciąg segmentów zgodny ze strukturą zdefiniowaną w pierwszym członie tej reguły. Omówiona struktura pierwszego członu reguły: *p.i.*[A]–*i* występuje między innymi w wyrazach: *pianino* i *dopiekać*. W wyrazie *pianino* wskazane trzy elementy pierwszego członu reguły pokrywają się z segmentami (literami), do których są przypisane indeksy: 1, 2, 3 (przy założeniu, że do segmentu zawierającego znak # jest przyporządkowany indeks 0). W wyrazie *dopiekać* wspomniana struktura pierwszego członu reguły pokrywa się z segmentami (literami), do których są przypisane indeksy: 3, 4, 5. A więc w pierwszym wypadku byłby przetwarzany ciąg segmentów (liter): *pia* (wraz z przypisanymi indeksami: 1, 2, 3), natomiast w drugim byłby

to ciąg *pie* (wraz z indeksami: 3, 4, 5). Podane przykłady dotyczą przetwarzania w ramach pierwszej warstwy, a więc każdy segment jest jedną literą i liczba indeksów jest równa liczbie liter w wejściowym wyrazie ortograficznym powiększonej o 2 (ze względu na użycie znaczników: #).

W dalszym ciągu rozdziału omówiono operacje (czynności), które mogą być zdefiniowane w drugim członie reguły. Większość przykładów nawiązuje do projektu systemu transkrypcji omówionego w rozdziale czwartym. Istotny jest również fakt, że treść każdego segmentu jest zapisywana przy użyciu ciągów znaków, a dowolna zmiana struktury danego ciągu oznacza zmianę segmentu.

3.4. Operacje wykonywane przy użyciu reguł WMT

3.4.1. Zachowanie segmentu i indeksu (brak zmiany)

Najprostsza operacja polega na zachowaniu segmentu w niezmienionej postaci oraz na pozostawieniu na tej samej pozycji indeksu, który jest przypisany do tego segmentu. Jeżeli dany element w ramach pierwszego członu reguły jest konkretnym segmentem, to dla osiągnięcia zamierzonego celu można użyć polecenia !NC. Inny sposób polega na przepisaniu identycznej treści segmentu w analogicznym elemencie drugiego członu reguły. Oto przykłady reguł obrazujących te dwie możliwości:

/b/./s/./t/./v/>/p/.!NC.!NC./f/;
/b/./s/./t/./v/>/p/./s/./t/./f/;

Obie reguły są odpowiednie dla przetwarzania wyrazu *hrabstwo* zapisanego przy użyciu transkrypcji podstawowej (/x/./r/./a/./b/./s/./t/./v/./o/). Taki zapis byłby uzyskany po przetworzeniu wyrazu ortograficznego w pierwszej warstwie przykładowego systemu transkrypcji opartego na WMT (zob. rozdz. 4). Musiałyby być użyte następujące reguły:

h>/x/;
r>/r/;
a>/a/;
b>/b/;
s>/s/;
t>/t/;
w>/v/;
o>/o/;

Zatem dwie omawiane reguły nie są właściwe dla warstwy pierwszej (wejściowej), do której są przekazywane wyłącznie wyrazy ortograficzne. W transkrypcji podstawowej zostaje zachowany status dźwięczności przypisany do znaków ortograficznych (zob. podrozdz. 1.4). W tabeli 3.3 oraz w tabeli 3.4 przedstawiono schemat działania omawianych reguł na przykładzie przetwarzania wyrazu /x/./r/./a/./b/./s/./t/./v/./o/.

Tabela 3.3 – Schemat działania reguły: /b/./s/./t/./v/>/p/.!NC.!NC./f/;

Indeksacja początkowa	0	1	2	3	4	5	6	7	8	9
Wyraz przed zast. reguły	#	/x/	/r/	/a/	/b/	/s/	/t/	/v/	/o/	#
I człon reguły					/b/	/s/	/t/	/v/		
II człon reguły					/p/	!NC	!NC	/f/		
Wyraz po zast. reguły	#	/x/	/r/	/a/	/p/	/s/	/t/	/f/	/o/	#
Indeksacja końcowa	0	1	2	3	4	5	6	7	8	9

Tabela 3.4 – Schemat działania reguły: /b/./s/./t/./v/>/p/./s/./t/./f/;

Indeksacja początkowa	0	1	2	3	4	5	6	7	8	9
Wyraz przed zast. reguły	#	/x/	/r/	/a/	/b/	/s/	/t/	/v/	/o/	#
I człon reguły					/b/	/s/	/t/	/v/		
II człon reguły					/p/	/s/	/t/	/f/		
Wyraz po zast. reguły	#	/x/	/r/	/a/	/p/	/s/	/t/	/f/	/o/	#
Indeksacja końcowa	0	1	2	3	4	5	6	7	8	9

Różnica między tymi dwiema możliwościami jest istotna. Jeżeli dany segment zostanie przepisany w odpowiednim elemencie drugiego członu reguły w niezmienionej postaci (jak w wypadku segmentów: /s/ i /t/ w przykładzie podanym w tabeli 3.4), to w przetwarzanym wyrazie możliwość jego dalszej modyfikacji zostanie zablokowana. Oznacza to, że przy użyciu innych reguł zdefiniowanych w tej warstwie nie będzie można zmienić tego segmentu (dotyczy to reguł stosowanych później). Użycie polecenia !NC również sprawia, że dany segment nie zostanie zmieniony, jednak nadal będzie istniała możliwość zamiany tego segmentu na podstawie innych reguł zdefiniowanych w tej samej warstwie.

Mechanizm działa w następujący sposób: najpierw są sprawdzane wszystkie reguły zdefiniowane w danej warstwie pod kątem możliwości ich użycia w aktualnie przetwarzanym wyrazie (sprawdzanie rozpoczyna się od reguł najdłuższych). Reguły odpowiednie dla tego wyrazu są na bieżąco rejestrowane (w czasie sprawdzania) i jednocześnie na podstawie ich struktury są blokowane odpowiednie segmenty. Zablokowanie poszczególnych segmentów oznacza, że nie mogą być one modyfikowane na podstawie reguł sprawdzanych później. Po sprawdzeniu wszystkich reguł zarejestrowane zmiany są wykonywane, następnie segmenty zostają odblokowane i wyraz zostaje przekazany do przetwarzania w kolejnej warstwie.

Jeżeli element pierwszego członu reguły jest zbiorem segmentów, to w celu uniknięcia zmian w analogicznym elemencie drugiego członu można użyć tylko polecenia !NC (ponieważ z góry nie wiadomo, który segment z tego zbioru będzie częścią przetwarzanego wyrazu). Oto przykład reguły zawierającej omawianą strukturę:

[WB2]./v/>!NC./f/;

W przykładowym projekcie systemu transkrypcji (zob. rozdz. 4) zbiór WB2 obejmuje wszystkie segmenty drugiego stopnia, które są spółgłoskami właściwymi bezdźwięcznymi. W tabeli 3.5 zobrazowano działanie podanej reguły na przykładzie wyrazu /g/./w/./u/./p/./s/./t/./v/./o/. Taki zapis zostałby uzyskany po przetworzeniu wyrazu *głupstwo* w pierwszej warstwie. W tym celu zostałyby użyte następujące reguły:

g>/g/;
ł>/w/;
u>/u/;
p>/p/;
s>/s/;
t>/t/;
w>/v/;
o>/o/;

Dzięki użyciu polecenia !NC w odniesieniu do dowolnego segmentu ze zbioru WB2 (tutaj jest to: /t/) zostaje zachowana możliwość modyfikowania tego elementu przy użyciu innych reguł zdefiniowanych w tej samej warstwie.

Tabela 3.5 – Schemat działania reguły: [WB2]./v/>!NC./f/;

Indeksacja początkowa	0	1	2	3	4	5	6	7	8	9
Wyraz przed zast. reguły	#	/g/	/w/	/u/	/p/	/s/	/t/	/v/	/o/	#
I człon reguły							[WB2]	/v/		
II człon reguły							!NC	/f/		
Wyraz po zast. reguły	#	/g/	/w/	/u/	/p/	/s/	/t/	/f/	/o/	#
Indeksacja końcowa	0	1	2	3	4	5	6	7	8	9

3.4.2. Modyfikacja segmentu z zachowaniem jego dotychczasowego indeksu

Można wyróżnić następujące operacje modyfikacji segmentu w przetwarzanym wyrazie, które nie powodują zmiany indeksu przypisanego do tego segmentu:
– zamiana segmentu pojedynczego na inny segment pojedynczy;
– zamiana segmentu pojedynczego na segment wielokrotny;
– zamiana segmentu wielokrotnego na inny segment wielokrotny;
– zamiana segmentu wielokrotnego na segment pojedynczy.

Operacja polegająca na zamianie segmentu pojedynczego na inny segment pojedynczy polega na zastąpieniu ciągu znaków przypisanego do danego indeksu innym ciągiem znaków. Ten rodzaj operacji został już uwzględniony w opisie w części 3.4.1. Segment /v/ (drugiego stopnia) był zastępowany jego bezdźwięcznym odpowiednikiem – segmentem /f/. Przykładowa reguła, która umożliwia zastąpienie segmentu pojedynczego segmentem wielokrotnym, wygląda następująco:
ę.t>/e/&/n/.!NC;

Sprawia ona, że indeks pierwotnie przypisany do segmentu *ę* zostaje powiązany z segmentem wielokrotnym obejmującym /e/ oraz /n/. Działanie podanej reguły zobrazowano na przykładzie ujętym w tabeli 3.6.

Tabela 3.6 – Schemat działania reguły: *ę.t*>/e/&/n/.!NC;

Indeksacja początkowa	0	1	2	3	4	5	6
Wyraz przed zast. reguły	#	*p*	*ę*	*t*	*l*	*a*	#
I człon reguły			*ę*	*t*			
II człon reguły			/e/&/n/	!NC			
Wyraz po zast. reguły	#	*p*	/e/&/n/	*t*	*l*	*a*	#
Indeksacja końcowa	0	1	2	3	4	5	6

Kolejny rodzaj wyszczególnionej operacji to zastąpienie segmentu wielokrotnego innym segmentem wielokrotnym. Taka konieczność mogłaby wystąpić w wypadku zaprojektowania dodatkowej warstwy w konkretnym projekcie opartym na WMT, której celem byłoby przekształcanie uzyskiwanej transkrypcji fonologicznej na transkrypcję fonetyczną. Dla ostatniego wyszczególnionego wariantu operacji, czyli dla zamiany segmentu wielokrotnego na segment pojedynczy, trudno znaleźć analogię w języku polskim.

3.4.3. Usunięcie segmentu (polecenie: !RM)

Użycie polecenia !RM (ang. remove) powoduje usunięcie segmentu oraz pozostawienie indeksu związanego z tym segmentem na tej samej pozycji w przetwarzanym wyrazie. Po zastosowaniu tego polecenia dany indeks zostaje przypisany do specjalnego segmentu EMP (ang. empty). Jeżeli z usuniętym segmentem była powiązana większa liczba indeksów (większa niż jeden), to te wszystkie indeksy zostają przypisane do segmentu EMP. Użycie polecenia !RM modyfikuje indeksowanie wyrazu w tym sensie, że bezpowrotnie wyłącza określony indeks (lub indeksy) z dalszego przetwarzania. Polecenie !RM może być używane dla przetwarzania fragmentów wyrazu, które nie są realizowane.

W tabeli 3.7 przedstawiono schemat przetwarzania wyrazu /p/./o/./r/./ts/./j/./i/. Taki zapis byłby uzyskany po przetworzeniu wyrazu *porcji* w pierwszej warstwie przykładowego systemu transkrypcji (zob. rozdz. 4). Zostałyby użyte następujące reguły:

p>/p/;
o>/o/;
r>/r/;
c>/ts/;
j>/j/;
i>/i/;

W drugiej warstwie tego rozwiązania nie zostałby wprowadzone żadne zmiany. W warstwie trzeciej zostałaby użyta reguła zawierająca polecenie !RM:

/ts/./j/./i/.#>!NC.!RM.!NC.!NC;

Sprawia ona, że indeks o numerze 5, który był przypisany do segmentu /j/, zostaje powiązany z segmentem EMP.

Tabela 3.7 – Schemat działania reguły: /ts/./j/./i/.#>!NC.!RM.!NC.!NC;

Indeksacja początkowa	0	1	2	3	4	5	6	7
Wyraz przed zast. reguły	#	/p/	/o/	/r/	/ts/	/j/	/i/	#
I człon reguły					/ts/	/j/	/i/	#
II człon reguły					!NC	!RM	!NC	!NC
Wyraz po zast. reguły	#	/p/	/o/	/r/	/ts/	EMP	/i/	#
Indeksacja końcowa	0	1	2	3	4	5	6	7

3.4.4. Usunięcie segmentu z jednoczesnym przesunięciem indeksu w lewo (polecenie: !ML)

Użycie polecenia !ML (ang. move left) sprawia, że dany segment zostaje usunięty, a indeks powiązany z tym segmentem zostaje dołączony do indeksu segmentu poprzedzającego dany segment. To polecenie służy przede wszystkim do wydzielania digrafemów i trigrafemów w wyrazach ortograficznych. Użycie polecenia !ML można zademonstrować na przykładzie przetwarzania wyrazu ortograficznego *szafa*. W tym wyrazie do litery *s* początkowo jest przyporządkowany indeks 1, do litery *z* jest przypisany indeks 2 (przy uwzględnieniu indeksacji znacznika #). Diada ortograficzna *sz* musi być ostatecznie zapisana jako jeden segment /S/. Do niego zostają przypisane dwa indeksy, które pierwotnie były przypisane do segmentów: *s* oraz *z*. Można to uzyskać przy użyciu następującej reguły:

s.z>/S/.!ML;

Oprócz użycia polecenia !ML trzeba zadbać również o to, żeby segment, do którego zostaną przypisane dwa indeksy, został odpowiednio zmodyfikowany. Jeżeli w wyniku przetwarzania wyrazów ortograficznych miałby być uzyskiwany zapis fonologiczny, to ortograficzny segment *s* należałoby zastąpić segmentem /S/. W tabeli 3.8 przedstawiono schemat działania powyższej reguły w odniesieniu do ortograficznego wyrazu *szafa*. Po takiej operacji indeksy: 1 i 2 zostałyby przypisane do segmentu /S/.

Tabela 3.8 – Schemat działania reguły *s.z*>/S/.!ML;

Indeksacja początkowa	0	1	2	3	4	5	6
Wyraz przed zast. reguły	#	*s*	*z*	*a*	*f*	*a*	#
I człon reguły		*s*	*z*				
II człon reguły		/S/	!ML				
Wyraz po zast. reguły	#	/S/	–	*a*	*f*	*a*	#
Indeksacja końcowa	0	1,2	–	3	4	5	6

Trzeba wspomnieć o kilku istotnych ograniczeniach związanych ze stosowaniem polecenia !ML. Nie można go używać bezpośrednio po elemencie umiejscowionym w drugim członie reguły, który zawiera polecenie !RM. Nie można go również stosować w odniesieniu do segmentu, do którego jest przypisany indeks 1, a także w odniesieniu do segmentu zawierającego etykietę EMP lub znak #.

Można natomiast użyć polecenia !ML w kilku kolejnych elementach drugiego członu reguły (umiejscowionych w swoim bezpośrednim sąsiedztwie). Może to mieć zastosowanie dla wydzielania trigrafemów. Przykładowa reguła zawierająca sekwencję dwóch poleceń !ML wygląda następująco:

d.z.i.a>/dz'/.!ML.!ML.!NC;

W tabeli 3.9 przedstawiono schemat działania tej reguły w odniesieniu do ortograficznego wyrazu *wydział*.

Tabela 3.9 – Schemat działania reguły *d.z.i.a*>/dz'/.!ML.!ML.!NC;

Indeksacja początkowa	0	1	2	3	4	5	6	7	8
Wyraz przed zast. reguły	#	*w*	*y*	*d*	*z*	*i*	*a*	*ł*	#
I człon reguły			*y*	*d*	*z*	*i*	*a*		
II człon reguły			!NC	/dz'/	!ML	!ML	!NC		
Wyraz po zast. reguły	#	*w*	*y*	/dz'/	–	–	*a*	*ł*	#
Indeksacja końcowa	0	1	2	3,4,5	–	–	6	7	8

3.5. Listy wyjątków i listy włączeń

Ostatnie zagadnienie omówione w tym rozdziale obejmuje listy wyjątków oraz listy włączeń. Listy te mogą być dołączane do reguł. Do jednej reguły może być dołączona jedna lista zawierająca ciągi znaków ortograficznych umożliwiających identyfikację określonych form fleksyjnych. W zapisie tych struktur nie są używane kropki, ale może być używany znak # w celu identyfikacji określonych form ortograficznych na podstawie fragmentów zawartych w nagłosie lub w wygłosie wyrazu. Reguła zawierająca listę wyjątków ma następującą strukturę:

pierwszy_człon_reguły>drugi_człon_reguły>!EXC:wyjątek_1,wyjątek_2,...,wyjatek_n;

Oto przykład reguły zgodnej z podanym schematem:

d.ż>/dZ/.!ML>!EXC:*#nadż,#ponadż,#śródż,#odż,#współodż,#przedż,#podż,#ponadż,#nienadż,#nieponadż,#-nieśródż,#nieodż,#niewspółodż,#nieprzedż,#niepodż,#nieponadż*;

Poszczególne elementy na liście są rozdzielone tylko znakiem przecinka (bez odstępu). Powyższa reguła zawiera listę wyjątków (lub inaczej: listę wyłączeń – ang. exclusions). Oznacza to, że każda forma fleksyjna, która może być zidentyfikowana na podstawie dowolnego ciągu ortograficznego włączonego do podanej listy (w tym przykładzie są to sekwencje nagłosowe), nie jest objęta tą regułą – czyli podana reguła nie może być dla niej użyta.

Listy włączeń umożliwiają identyfikację zbioru form fleksyjnych, dla których ma być zastosowana dana reguła. Nie może ona być użyta dla przetwarzania wyrazów spoza zbioru pokrywającego się z określoną listą włączeń. Różnica między zapisem listy włączeń a zapisem listy wyjątków dotyczy tylko używanego polecenia: !INC (ang. inclusions) zamiast polecenia !EXC.

4. Przykładowy projekt oparty na WMT

4.1. Wprowadzenie

W poprzednich rozdziałach omówiono podstawowe pojęcia oraz założenia dotyczące WMT. Przedstawiono też wnikliwe analizy związane z konwersją tekstu ortograficznego na transkrypcję fonologiczną. Ten rozdział dotyczy przykładowego systemu transkrypcji opartego na modelu wielowarstwowym.

Opisywany projekt jest rozwiązaniem statycznym. Oznacza to, że dla określonego zbioru reguł przetwarzanie danego tekstu ortograficznego zawsze daje taki sam rezultat. Nie jest brane pod uwagę prawdopodobieństwo poszczególnych wariantów wymowy. Drugie założenie dotyczy redukcji liczby reguł w poszczególnych warstwach. Wynikają z niego określone konsekwencje – zmniejszanie liczby reguł wiąże się z ich uogólnianiem. W regułach ogólnych struktura wyrazów często jest wyrażona przy użyciu szerokich klas dźwięków, a pojedyncze reguły dotyczą większej liczby wyrazów. Trzeba jednak podkreślić, że WMT umożliwia tworzenie reguł szczegółowych bez konieczności modyfikowania reguł ogólnych. Reguły szczegółowe najczęściej są dłuższe niż reguły ogólne (w sensie liczby elementów), mają zatem wyższy priorytet – w ten sposób funkcjonuje mechanizm przesłaniania reguł. Kolejne założenie dotyczy uwzględnionego typu wymowy. Jest to wymowa warszawska, normatywna dla języka polskiego i jako związana z głównym ośrodkiem kulturalnym uznawana za wzorzec ogólnopolski. Pozostaje jeszcze założenie związane z faktem, że wszystkie zaprojektowane reguły są przeznaczone dla przetwarzania małych liter. Każdy wyraz przed przetworzeniem musi zostać zapisany małymi literami.

W rozdziale piątym omówiono zaawansowane zagadnienia związane z WMT. Na podstawie tych informacji można rozbudować projekt opisany w niniejszym rozdziale. Wyszczególniono w nim trzy warstwy, które umownie nazwano: warstwą transkrypcji podstawowej, warstwą statusu dźwięczności oraz warstwą modyfikacyjną. Prezentowane rozwiązanie jest systemem transkrypcji fonologicznej, jednak istnieje możliwość jego rozbudowy do systemu transkrypcji fonetycznej (zgodnie z informacjami zawartymi we *Wprowadzeniu*).

4.2. Warstwa transkrypcji podstawowej

Podstawowy cel związany z przetwarzaniem wyrazów w pierwszej warstwie prezentowanego projektu polega na wydzielaniu liter, diad i triad ortograficznych odnoszących się do segmentów fonologicznych. To wstępne przekształcenie jest oparte przede wszystkim na transkrypcji podstawowej, dlatego ta warstwa została nazwana warstwą transkrypcji podstawowej. Wyodrębnianie ciągów znaków odpowiadających poszczególnym fonemom wymaga uwzględnienia faktu wielofunkcyjności litery *i*. Jeżeli ta litera jest tylko znakiem miękkości, to jest traktowana jako komponent wieloznaku (zob. podrozdz. 1.2 i 1.3).

Zgodnie z podstawową ideą prezentowaną w tej publikacji zapis wyrazów uzyskiwany po przetworzeniu w warstwie transkrypcji podstawowej nie jest ostateczną transkrypcją. Jest to zapis pośredni, który opiera się na transkrypcji podstawowej. Uwzględniono w nim również modyfikacje segmentów wynikające z występowaniem litery *i w* kontekście prawostronnym. Jednostki w wyrazach przekazywanych do drugiej warstwy omawianego projektu są umownie nazywane segmentami drugiego stopnia. Zapis wyrazów otrzymywany w wyniku przetworzenia w warstwie transkrypcji podstawowej nie zawiera modyfikacji związanych z asymilacją dźwięczności w ramach grup spółgłoskowych niejednorodnych pod względem statusu tzw. dźwięczności ortograficznej (zob. podrozdz. 1.4). Nie zawiera też uproszczeń grup spółgłoskowych oraz kilku innych istotnych modyfikacji – reguły związane z wymienionymi zagadnieniami zostały umieszczone w warstwach wyższych. Więcej informacji dotyczących inwentarza segmentów używanego do zapisywania wyrazów przetworzonych w pierwszej warstwie prezentowanego rozwiązania przedstawiono we wstępie do opisu warstwy drugiej.

Jednostki przekazywane (w wyrazach) do pierwszej warstwy każdego projektu opartego na WMT są zdefiniowane przez sam model – zawsze są to pojedyncze znaki ortograficzne (litery). W opisanym projekcie są one umownie nazywane segmentami pierszego stopnia. W tabeli 4.1 przedstawiono inwentarz tych segmentów. W zapisie przykładowych wyrazów poszczególne litery rozdzielono kropką.

Tabela 4.1 – Inwentarz segmentów w wyrazach przekazywanych do warstwy transkrypcji podstawowej

Lp.	Litera	Przykładowy wyraz	Lp.	Litera	Przykładowy wyraz
1	*a*	*m.a.k*	17	*f*	*f.u.t.r.o*
2	*e*	*z.e.r.o*	18	*w*	*w.i.a.t.r*
3	*i*	*n.i.c*	19	*s*	*w.y.s.o.k.i*
4	*o*	*s.o.w.a*	20	*z*	*r.z.e.k.a*
5	*y*	*b.y.k*	21	*ż*	*ż.a.b.a*
6	*u*	*j.u.t.r.o*	22	*ś*	*ś.m.i.e.c.h*
7	*ó*	*ó.s.m.y*	23	*ź*	*ź.r.e.b.a.k*
8	*ę*	*j.ę.z.y.k*	24	*h*	*h.o.r.y.z.o.n.t*
9	*ą*	*p.a.j.ą.k*	25	*p*	*p.a.l.e.c*
10	*ł*	*p.u.ł.k.a*	26	*b*	*b.u.d.a*
11	*j*	*j.e.d.e.n*	27	*t*	*t.a.m.a*
12	*l*	*w.i.e.l.e*	28	*d*	*d.o.m*
13	*r*	*r.y.b.a*	29	*k*	*p.o.k.ó.j*
14	*m*	*m.o.r.z.e*	30	*g*	*g.o.ś.ć*
15	*n*	*m.o.n.e.t.a*	31	*c*	*c.z.e.k*
16	*ń*	*k.o.ń*	32	*ć*	*ć.m.a*

Ostatnia kwestia dotyczy zdefiniowanych zbiorów jednostek pierwszego stopnia. Mogą one zawierać tylko elementy wyszczególnione w tabeli 4.1. Sposób definiowania i używania zbiorów segmentów w projektach opartych na WMT został omówiony w rozdziale trzecim. Poniżej przedstawiono definicje zbiorów, które są używane tylko w regułach należących do warstwy transkrypcji podstawowej (stąd cyfra 1 w nazwach tych zbiorów). Definicja zbioru X1, który obejmuje wszystkie litery, przyjmuje następującą postać:

[X1]={*a, e, i, o, y, u, ó, ę, ą, ł, j, l, r, m, n, ń, f, w, s, z, ż, ś, ź, h, p, b, t, d, k, g, c, ć*};

Zbiór SA1 obejmuje litery oznaczające samogłoski:

[SA1]={*a, e, i, o, y, u, ó, ę, ą*};

Zbiór SP1 stanowi różnicę zbiorów: X1 i SA1 – obejmuje on litery oznaczające spółgłoski i półsamogłoski:

[SP1]={*ł, j, l, r, m, n, ń, f, w, s, z, ż, ś, ź, h, p, b, t, d, k, g, c, ć*};

4.2.1. Reguły transkrypcji podstawowej

Zgodnie z informacjami zawartymi w rozdziale pierwszym transkrypcja podstawowa nie jest częścią WMT, jednak może ona ułatwić użycie tego modelu. Reguły transkrypcji podstawowej dotyczące poszczególnych liter są jednoelementowe, więc są przesłaniane przez reguły co najmniej dwuelementowe, należące do tej samej warstwy. Zatem reguły dłuższe są traktowane jako wyjątki w stosunku do reguł krótszych. Dzięki takiemu podejściu dla danej litery lub diady ortograficznej umiejscowionych w określonym kontekście ortograficznym nie trzeba tworzyć żadnych dodatkowych reguł (poza regułami ustalonymi w ramach transkrypcji podstawowej), jeżeli właściwa transkrypcja tych struktur pokrywa się z transkrypcją podstawową. Trzeba podkreślić, że rozważania w tym miejscu dotyczą pierwszej warstwy projektu opartego na WMT. Odnoszą się więc do przekształcania wyrazów ortograficznych na oczekiwaną transkrypcję pośrednią (nieostateczną). Oto zestaw reguł przeznaczonych dla przetwarzania pojedynczych liter:

a>/a/;
b>/b/;
c>/ts/;
ć>/ts'/;
d>/d/;
e>/e/;
f>/f/;
g>/g/;

h>/x/;
i>/i/;
j>/j/;
k>/k/;
l>/l/;
ł>/w/;
m>/m/;
n>/n/;
ń>/n'/;
o>/o/;
ó>/u/;
p>/p/;
r>/r/;
s>/s/;
ś>/s'/;
t>/t/;
u>/u/;
w>/v/;
y>/y/;
z>/z/;
ż>/Z/;
ź>/z'/;

Utworzono też kilka reguł związanych z przetwarzaniem diad ortograficznych, które zostały włączone do transkrypcji podstawowej. Są to reguły dwuelementowe. Mają więc wyższy priorytet niż wymienione reguły jednoelementowe. Poniższa lista obejmuje reguły, które nie zawierają wyjątków:

c.h>/x/.!ML;
c.z>/tS/.!ML;
d.ź>/dz'/.!ML;
s.z>/S/.!ML;

W podrozdziale 2.5.8 przedstawiono analizę dotyczącą diady ortograficznej *dż*. Wynika z niej, że w określonych wyrazach *dż* nie stanowi digrafemu (oznacza sekwencję fonemów /d.Z/). Na podstawie analizy ujętej w tabeli 2.35 utworzono listę ciągów nagłosowych, które umożliwiają identyfikacje takich wyrazów. Musi ona być uwzględniona w regułach. Pierwsza możliwość polega na utworzeniu tylko jednej reguły, do której jest dołączona lista ciągów ortograficznych dla identyfikacji wyjątków:

d.ż>/dZ/.!ML>!EXC:*#nadż,#ponadż,#śródż,#odż,#współodż,#przedż,#podż,#ponadż,#nienadż,#nieponadż,#nieśródż,#nieodż,#niewspółodż,#nieprzedż,#niepodż,#nieponadż*;

Dla form fleksyjnych, które zostały wykluczone z podanej reguły, zastosowanie mają reguły jednoelementowe przeznaczone dla przetwarzania litery *d* oraz litery *ż*. Druga możliwość polega na utworzeniu reguły podstawowej:

d.ż>/dZ/.!ML;

Dla każdego ciągu ortograficznego umożliwiającego identyfikację określonego wyjątku trzeba utworzyć oddzielną regułę, która musi być dłuższa niż reguła podstawowa (będzie więc miała wyższy priorytet). Na przykład dla nagłosowego ciągu *nadż* można utworzyć następującą regułę:

#.n.a.d.ż>!NC.!NC.!NC./d/./Z/;

Na jej podstawie litery w przetwarzanych wyrazach: *d* i *ż* są przekształcane w segmenty drugiego stopnia: /d/ i /Z/. Powoduje to ich zablokowanie dla modyfikacji wynikających z innych reguł zdefiniowanych w omawianej warstwie. Podana wcześniej reguła podstawowa nie mogłaby być użyta.

W podobny sposób można zaprojektować reguły dla ortograficznej diady *dz* umiejscowionej w kontekście prawostronnym litery innej niż *i*. Z analizy przedstawionej w podrozdziale 2.59 uzyskano listę ciągów nagłoso-

wych umożliwiających identyfikację wyrazów, w których diada *dz* powinna być transkrybowana jako sekwencja fonemów /d.z/. Jako pierwszą możliwość zakłada się włączenie tej listy wyjątków do reguły:

d.z.[X1]–*i*>/dz/.!ML.!NC>!EXC:*#nadzb,#nadzm,#nadzor,#nadzór,#nadzw,#supernadz,#ponadz,#odz,#nieodzown,#podza,#podzb,#podzn,#podzwr,#przedza,#przedzg,#przedzj,#przedzm,#przedzwoj,#nienadzb,#nienadzm,#nienadzor,#nienadzór,#nienadzw,#niesupernadz,#nieponadz,#nieodz,#niepodza,#niepodzb,#niepodzn,#niepodzwr,#nieprzedza,#nieprzedzg,#nieprzedzj,#nieprzedzm,#nieprzedzwoj*;

Druga możliwość polega na utworzeniu reguły podstawowej:

d.z.[X1]–*i*>/dz/.!ML.!NC;

Jednocześnie dla każdego elementu z listy wyjątków trzeba utworzyć odpowiednią regułę przesłaniającą regułę podstawową. Na przykład reguła utworzona dla nagłosowego ciągu *nadzor* ma następującą postać:

#.n.a.d.z.o.r>!NC.!NC.!NC./d/./z/.!NC.!NC;

Z analiz omówionych w podrozdziale 2.5.12 uzyskano wykaz nagłosowych ciągów umożliwiających identyfikację wszystkich wyrazów, w których diada *rz* oznacza sekwencję fonemów /r.z/ lub /r.z'/. Pierwsza możliwość polega na dołączeniu tej listy do jednej reguły:

r.z>/rz/.!ML>!EXC:*#marzł,#marzn,#domarzn,#domarzł,#pomarzn,#pomarzł,#pomarzl,#przemarzl,#przemarzł,#przemarzn,#rozmarzł,#rozmarzl,#rozmarzn,#umarzł,#umarzl,#umarzn,#wymarzł,#wymarzl,#wymarzn,#zamarzaln,#zamarzl,#zamarzł,#zamarzn,#omarzł,#omarzl,#omarza,#omarzn,#mierzną,#mierznę,#mierznie,#mierzl,#mierzł,#obmierzn,#obmierzł,#obmierzle,#przemierzł,#przemierzl,#zmierzł,#zmierzn,#zmierzl,#omierzł,#omierzl,#omierzn,#namarza,#namarzł,#namarzl,#namarzn,#obmarz,#odmarz,#przymarz,#nadmarz,#podmarz,#zmarz,#wmarz,#mierzi,#obmierzi,#zmierzi,#niemarzł,#niemarzn,#niedomarzn,#niedomarzł,#niepomarzn,#niepomarzł,#niepomarzl,#nieprzemarzl,#nieprzemarzł,#nieprzemarzn,#nierozmarzł,#nierozmarzl,#nierozmarzn,#nieumarzł,#nieumarzl,#nieumarzn,#niewymarzł,#niewymarzl,#niewymarzn,#niezamarzaln,#niezamarzl,#niezamarzł,#niezamarzn,#nieomarzł,#nieomarzl,#nieomarza,#nieomarzn,#niemierzną,#niemierznę,#niemierznie,#niemierzl,#niemierzł,#nieobmierzn,#nieobmierzł,#nieobmierzle,#nieprzemierzł,#nieprzemierzl,#niezmierzł,#niezmierzn,#niezmierzl,#nieomierzł,#nieomierzl,#nieomierzn,#nienamarza,#nienamarzł,#nienamarzl,#nienamarzn,#nieobmarz,#nieodmarz,#nieprzymarz,#nienadmarz,#niepodmarz,#niezmarz,#niewmarz,#niemierzi,#nieobmierzi,#niezmierzi*;

W powyższej regule lista wyjątków obejmuje również wyrazy rozpoczynające się od przedrostka *nie*. Dla celów praktycznych związanych z transkrypcją tekstu można założyć, że wszystkie wyznaczone ciągi ortograficzne mogą występować w takim kontekście. Druga możliwość polega na utworzeniu reguły podstawowej:

r.z>/rz/.!ML;

Może ona być przesłaniana przez reguły dotyczące wyjątków, na przykład:

#.m.a.r.z.ł>!NC.!NC.!NC./r/./z/.!NC;

Z ortograficzną diadą *rz* jest powiązany problem, który został omówiony w podrozdziale 2.5.12. Wyznaczono zbiór struktur nagłosowych umożliwiających identyfikację wszystkich form fleksyjnych, w których diada *rz* może być transkrybowana zarówno jako fonem /rz/, jak i jako sekwencja fonemów /r.z/. Trudno wskazać optymalne rozwiązanie tego problemu. Można przyjąć jeden wariant transkrypcji przy jednoczesnym wprowadzeniu specjalnych oznaczeń, które umożliwiłyby dalsze przetwarzanie wyrazów przez zewnętrzny system analizy semantycznej. Inny sposób, który nie wymaga angażowania dodatkowych narzędzi, polega na przyjęciu jednego – bardziej prawdopodobnego sposobu transkrypcji poszczególnych wyrazów. W omawianym projekcie przyjęto notację bifonematyczną z dodatkowym oznaczeniem wariantywności transkrypcji przy użyciu znaku zapytania. Pierwsza opcja polega na utworzeniu reguły z listą włączeń (zob. podrozdz. 3.5). Umożliwia ona przetwarzanie form fleksyjnych, które da się zidentyfikować na podstawie dołączonej listy ciągów ortograficznych. Sama reguła musi być co najmniej trójelementowa – po to, żeby przesłaniała elementarną regułę dwuelementową związaną z przetwarzaniem diady *rz* (*r.z*>/rz/.!ML;). Wykorzystano fakt, że wszystkie formy, które muszą być uwzględnione, zawierają triadę ortograficzną *rza*. Tak powstała reguła o następującej konstrukcji:

r.z.a>/r?/./z?/.!NC>!INC:*#domarza,#przemarzal,#przemarzał,#przemarza#,#przemarzam,#przemarzasz,#przemarzać#,#przemarzac,#przemarzano,#przemarzani,#przemarzań,#przemarzaj,#rozmarzać#,#rozmarzal,#rozmarzał,#rozmarza#,#rozmarzac,#rozmarzają#,#rozmarzam,#rozmarzas,#rozmarzano,#rozmarzani,#rozmarzaj,#rozmarzań,#umarzać,#umarzali,#umarzał,#umarza#,#umarzac,#umarzaj,#umarzam,#umarzani,#umarzano,#umarzań,#wymarzać,#wymarzal,#wymarzał,#wymarza#,#wymarzac,#wymarzam,#wymarzasz,#wymarzano,#wymarzani,#wymarzaj,#wymarzań,#zamarzać,#zamarzal,#zamarzał,#zamarza#,#zamarzac,#zamarzaj,#zamarzam,#zamarzano,#zamarzani,#zamarzasz,#obmierza,#zmierza,#niedomarza,#nieprzemarzal,#nieprzemarzał,#nieprzemarza#,#nieprzemarzam,#nieprzemarzasz,#nieprzemarzać#,#nieprzemarzac,#nieprzemarzano,#nieprzemarzani,#nieprzemarzań,#nieprzemarzaj,#nierozmarzać#,#nierozmarzal,#nierozmarzał,#nierozmarza#,#nierozmarzac,#nierozmarzają#,#nierozmarzam,#nierozmarzas,#nierozmarzano,#nierozmarzani,#nierozmarzaj,#nierozmarzań,#nieumarzać,#nieumarzali,#nieumarzał,#nieumarza#,#nieumarzac,#nieumarzaj,#nieumarzam,#nieumarzani,#nieumarzano,#nieumarzań,#niewymarzać,#niewymarzal,#niewymarzał,#niewymarza#,#niewymarzac,#niewymarzam,#niewymarzasz,#niewymarzano,#niewymarzani,#niewymarzaj,#niewymarzań,#niezamarzać,#niezamarzal,#niezamarzał,#niezamarza#,#niezamarzac,#niezamarzaj,#niezamarzam,#niezamarzano,#niezamarzani,#niezamarzasz,#nieobmierza,#niezmierza*;

Druga opcja polega na utworzeniu oddzielnej reguły dla każdej struktury ortograficznej umożliwiającej identyfikację form, w których diada *rz* może być transkrybowana zgodnie z dwoma uwzględnionymi wariantami. Przykład reguły utworzonej dla nagłosowego ciągu *#domarza* jest następujący:

#.d.o.m.a.r.z.a>!NC.!NC.!NC.!NC.!NC./r?/./z?/.!NC;

4.2.2. Reguły przetwarzania struktur ortograficznych niewłączonych do transkrypcji podstawowej

W tej części omówiono reguły dotyczące struktur ortograficznych, które są wieloznakami tylko w określonych kontekstach i z tego powodu nie zostały włączone do transkrypcji podstawowej (zob. podrozdz. 1.3). Muszą one być zawarte w pierwszej warstwie prezentowanego projektu, ponieważ umożliwiają poprawne wyodrębnianie fragmentów związanych z określonymi segmentami fonologicznymi.

Ortograficzna konstrukcja {*s,c,z*}*i* umiejscowiona przed literą oznaczającą samogłoskę inną niż *i* stanowi digrafem (zob. podrozdz. 2.6.12). Dla przetwarzania takich struktur utworzono następujące reguły:

s.i.[SA1]–*i*>/s'/.!ML.!NC;
c.i.[SA1]–*i*>/ts'/.!ML.!NC;
z.i.[SA1]–*i*>/z'/.!ML.!NC;

Polecenie !ML umożliwia przesunięcie indeksu związanego z literą *i* do poprzedniego segmentu. Po jego użyciu do pojedynczego segmentu zostają przypisane dwa indeksy oraz transkrypcja właściwa dla omawianej warstwy. Wynika to z faktu, że w tych konstrukcjach litera *i* jest tylko znakiem miękkości. Pierwsza z trzech wymienionych reguł wymaga utworzenia reguły uzupełniającej ze względu na nietypową wymowę ortograficznego wyrazu *sial*:

#.s.i.a.l>!NC./s/./i/.!NC.!NC;

Odnosi się ona również do form fleksyjnych oraz wyrazów pochodnych (np.: *sialów, sialiczny*). Poza tym jest to reguła pięcioelementowa, czyli przesłania podaną regułę trójelementową.

W dalszym ciągu omówiono reguły związane z przetwarzaniem diad: *ci, zi* oraz *si* przed literą oznaczającą spółgłoskę lub w wygłosie wyrazu. W takich strukturach litera *i* jest znakiem miękkości i jednocześnie oznacza fonem /i/. Dla przetwarzania diady *ci* można zastosować następującą regułę:

c.i.[SP1]+#>/ts'/./i/.!NC;

Diady: *zi* oraz *si* umiejscowione przed spółgłoską w niektórych wyrazach mogą oznaczać sekwencję fonemów: /z.i/ lub /s.i/ (zob. tabela 2.32 oraz tabela 2.63). Istnieją dwie możliwości konstrukcji odpowiednich reguł. Są one analogiczne do możliwości przedstawionych w części 4.2.1. Pierwsza polega na utworzeniu jednej reguły zawierającej listę ciągów ortograficznych, które umożliwiają identyfikację wszystkich form fleksyjnych stanowiących wyjątki:

z.i.[SP1]+#>/z'/./i/.!NC>!EXC:*#zis,#zid,#zig,#zil,#ził,#zin,#zimb,#zir,#ziś,#dezid,#dezil,#dezin,#rozis,#rozig,#rozin,#rozir,#zdezin,#bezim,#bezis,#bezid,#bezig,#bezin,#ubezim,#niezis,#niezid,#niezig,#niezil,#nieził,#niezin,#niezimb,#niezir,#nieziś,#niedezid,#niedezil,#niedezin,#nierozis,#nierozig,#nierozin,#nierozir,#niezdezin,#niebezim,#niebezis,#niebezid,#niebezig,#niebezin,#nieubezim*;

Druga możliwość polega na utworzeniu reguły podstawowej oraz oddzielnych reguł uzupełniających dla wyjątków. Oto reguła podstawowa:

z.i.[SP1]+#>/z'/./i/.!NC;

Przykładowa reguła uzupełniająca, która dotyczy nagłosowej struktury *#zis* (jak w wyrazie *ziszczyć*), ma następującą konstrukcję (jest to reguła czteroelementowa):

#.z.i.s>!NC./z/./i/.!NC;

Dla przetwarzania diady *si* umiejscowionej przed literą oznaczającą spółgłoskę lub w wygłosie wyrazu można utworzyć jedną regułę zawierającą listę sekwencji ortograficznych dla identyfikacji wyjątków:

s.i.[SP1]+#>/s'/./i/.!NC>!EXC:*#sidol,#silik,#sila,#silen,#siles,#silo,#silu,#siló,#silem,#silif,#silim,#silit,#sim#,#sima,#simą,#simę,#simie,#simy,#sime,#simir,#simb,#simk,#simo,#simp,#sims,#simul,#sinf,#sing,#sina,#sine,#sino,#sint,#sinu,#sisa,#sist,#maksi#,#pleksi#,#maksim,#eliksir,#pleksig*;

Podejście alternatywne polega na utworzeniu trójelementowej reguły podstawowej:

s.i.[SP1]+#>/s'/./i/.!NC;

Muszą jej towarzyszyć dłuższe reguły przesłaniające, na przykład:

#.s.i.l.i.k>!NC./s/./i/.!NC.!NC.!NC;

Oddzielną regułę trzeba utworzyć dla wyrazu *si*:

#.s.i.#>!NC./s/./i/.!NC;

W wypadku ortograficznej triady *dzi* trzeba uwzględnić dodatkową komplikację, o której wspomniano w podrozdziale 2.5.11. Jest ona związana z możliwością wystąpienia junktury morfologicznej wewnątrz tej konstrukcji – ten problem dotyczy zarówno kontekstu spółgłoskowego, jak i kontekstu samogłoskowego. Triada *dzi* w kontekście spółgłoskowym lub w wygłosie wyrazu nie stanowi trigrafemu, tylko oznacza sekwencję fonemów /dz'.i/. Na podstawie analizy ujętej w tabeli 2.40 utworzono listę wyjątków, która została włączona do odpowiedniej reguły:

d.z.i.[SP1]+#>/dz'/.!ML.!NC.!NC>!EXC:*#przedzim,#odzip,#nieprzedzim,#nieodzip*;

Użycie podanej reguły sprawia, że indeks początkowo przypisany do litery *z* zostaje przesunięty w lewo i powiązany z segmentem /dz'/. Druga możliwość polega na utworzeniu reguły podstawowej:

d.z.i.[SP1]+#>/dz'/.!ML.!NC.!NC;

Przy takiej konstrukcji niezbędne jest utworzenie reguły uzupełniającej dla każdej pozycji ujętej na liście wyjątków, na przykład:

#.p.r.z.e.d.z.i.m>!NC.!NC.!NC.!NC.!NC./d/./z'/.!NC.!NC;

W kontekście samogłoskowym (z wyjątkiem litery *i*) triada *dzi* stanowi trigrafem. Na podstawie analizy ujętej w tabeli 2.41 utworzono listę wyjątków, które zostały uwzględnione w regule:

d.z.i.[SA1]–*i*>/dz'/.!ML.!ML.!NC>!EXC:*#nadziar,#nadziem,#nadziom,#odziar,#odziem,#odziom,#podziar,#podziem,#podziom,#podzięb,#ponadziem,#śródziem,#nadśródziem,#wschodniośródziem,#nienadziar,#nienadziem,#nienadziom,#nieodziar,#nieodziem,#nieodziom,#niepodziar,#niepodziem,#niepodziom,#niepodzięb,#nieponadziem,#nieśródziem,#nienadśródziem,#niewschodniośródziem*;

Dwukrotnie występuje w niej polecenie !ML. Użycie reguły sprawia, że do segmentu /dz'/ zostają przypisane łącznie trzy indeksy, które pierwotnie były przypisane do liter: *d*, *z* oraz *i*, przy czym powiązanie indeksu litery *d* z /dz'/ jest domyślne i wynika z pozycji tej litery w regule. Przypisanie indeksów pierwotnie powiązanych z literami: *z* oraz *i* do segmentu /dz'/ wynika z użycia polecenia !ML.

Druga możliwość polega na utworzeniu reguły podstawowej oraz reguł uzupełniających związanych z wyjątkami. Reguła podstawowa ma następującą budowę:

d.z.i.[SA1]–*i*>/dz'/.!ML.!ML.!NC;

Przykładowa reguła uzupełniająca dotyczy nagłosowej struktury *#nadziem*:

#.n.a.d.z.i.e.m>!NC.!NC.!NC./d/./z'/.!ML.!NC.!NC;

Punkt g w podrozdziale 2.2.6 dotyczy transkrypcji struktury *ni*A. Litera *i* w wyrazach rodzimych jest tylko znakiem miękkości. Odpowiednia reguła podstawowa może mieć następującą postać:

n.i.[SA1]–*i*>/n'/.!ML.!NC;

Trzeba uwzględnić liczne wyrazy obce, które mają wymowę asynchroniczną. Dla każdego wyjątku można utworzyć oddzielną regułę, która ze względu na długość będzie przesłaniać podaną regułę podstawową, na przykład:

k.i.n.o.m.a.n.i>!NC.!NC.!NC.!NC.!NC.!NC./n'/./j/>!EXC:*manij#*;

Reguła dotyczy wszystkich form fleksyjnych rzeczownika *kinomania* z wyłączeniem mało prawdopodobnej, ale dopuszczalnej formy *kinomanij*.

Zbiór wyjątków od reguły podstawowej obejmuje wszystkie formy fleksyjne i wyrazy pochodne od wyrazu *mania*. Zawiera on także formy fleksyjne wyrazów zakończonych w mianowniku na *fonia, frenia, genia, gonia, tonia, linia* (należy uwzględnić wyjątki podane w opisie). Jego elementami są również wyrazy zakończone na: *nium, nion, nior* (oprócz wyrazów: *junior, senior, Pinior* i *Konior),* a także wszystkie wyrazy wymienione w tabeli 2.15, dla których przyjęto wymowę asynchroniczną (nie są przekreślone). Przykładowe reguły mają następującą postać:

#.m.a.n.i.ę.#>!NC.!NC.!NC./n'/./j/.!NC.!NC;
#.m.a.n.i.a.k.a.l.n>!NC.!NC.!NC./n'/./j/.!NC.!NC.!NC.!NC.!NC;
#.k.o.s.m.o.g.o.n.i>!NC.!NC.!NC.!NC.!NC.!NC.!NC.!NC./n'/./j/>!EXC:*gonij#*;
#.c.y.r.k.o.n.i>!NC.!NC.!NC.!NC.!NC.!NC./n'/./j/>!EXC:*#cyrkonij*;
n.i.o.r.#>/n'/./j/.!NC.!NC.!NC>!EXC:*#junior,#senior,#pinior,#konior*;
#.u.r.a.n.i.a.n>!NC.!NC.!NC.!NC./n'/./j/.!NC.!NC;

Diada *ni* w kontekście spółgłoskowym, półsamogłoskowym lub w wygłosie wyrazu oznacza sekwencję fonemów:

n.i.[SP1]+# >/n'/./i/.!NC;

4.2.3. Reguły przetwarzania litery *y*

Podstawowa reguła dotycząca przetwarzania litery *y* została włączona do zbioru reguł transkrypcji podstawowej (zob. podrozdz. 4.2.1). Trzeba utworzyć dodatkowe reguły dla kontekstów omówionych w podrozdziale 2.2.2. Litera *y* powinna być w nich transkrybowana jako fonem /j/. Można przenieść na grunt modelu wielowarstwowego reguły przytoczone z tabeli 2.1 w *Automatyzacji…* . Ich zapis po modyfikacji i przy uwzględnieniu definicji przyjętych zbiorów jest następujący:

#.y.[SA1]>!NC./j/.!NC;
[SA1].*y.*[SA1]>!NC./j/.!NC;
[SA1].*y.*[X1]–[SA1]–*j*>!NC./j/.!NC;

4.2.4. Reguły przetwarzania litery *u*

W podrozdziale 2.25 stwierdzono, że litera *u* w ortograficznej diadzie *eu* w zasadzie powinna być transkrybowana jako fonem /w/ z wyjątkiem wyrazów: *Pentateuch, reunion, teurgia.* Poza tym diada *eu* oznacza ciąg fonemów /e.u/ w wyrazach zawierających struktury ortograficzne: *ieu*, *rzeu*, *eusz*, *eus*, *eum#*. Można zatem utworzyć regułę podstawową:

e.u>!NC./w/;

Wyjątki trzeba uwzględnić w dodatkowych regułach, które ze względu na długość będą przesłaniać podaną regułę podstawową:

i.e.u>!NC.!NC./u/;
r.z.e.u>!NC.!NC.!NC./u/;
e.u.s.z>!NC./u/.!NC.!NC;
e.u.s>!NC./u/.!NC;
e.u.m.#>!NC./u/.!NC.!NC;
#.p.e.n.t.a.t.e.u.c.h>!NC.!NC.!NC.!NC.!NC.!NC.!NC.!NC./u/.!NC.!NC;
#.r.e.u.n.i.o.n>!NC.!NC.!NC./u/.!NC.!NC.!NC.!NC;
#.t.e.u.r.g>!NC.!NC.!NC./u/.!NC.!NC;

W ortograficznej diadzie *au* literze *u* odpowiada fonem /w/. Można utworzyć ogólną regułę podstawową:

a.u>!NC./w/;

Konsekwentnie trzeba utworzyć dłuższe reguły umożliwiające poprawną transkrypcję wszystkich form wyrazów: *laurka, laurkowy, asauł, auł, esauł, saksauł*:

#.l.a.u.r.k>!NC.!NC.!NC./u/.!NC.!NC;
#.l.a.u.r.c>!NC.!NC.!NC./u/.!NC.!NC;
#.a.s.a.u.ł>!NC.!NC.!NC.!NC./u/.!NC;
#.a.u.ł>!NC.!NC./u/.!NC;
#.e.s.a.u.ł>!NC.!NC.!NC.!NC./u/.!NC;
#.s.a.k.s.a.u.ł>!NC.!NC.!NC.!NC.!NC.!NC./u/.!NC;

Zbiór wyjątków od reguły podstawowej obejmuje również wyrazy zawierające jeden z następujących ciągów ortograficznych: *pozau*, *prau*, *dnau*, *unau*, *enau*, *ezau*, *onau*, *anau, ynau*. Zgodnie z informacjami podanymi w *Automatyzacji…* litera *u* powinna w nich być transkrybowana jako fonem /u/. Szczegółowa analiza omówiona w podrozdziale 2.25 wykazała, że istnieją odstępstwa od tej zasady. Wyrazy, które zostały przekreślone w tabeli 2.9 (np.: *metopozaury, tezurus*), stanowią wyjątki od wyjątków. Można utworzyć reguły, które ze względu na długość będą przesłaniać dwuelementową regułę podstawową, ale w tych regułach również trzeba uwzględnić wspomniane wyjątki:

p.o.z.a.u>!NC.!NC.!NC.!NC./u/>!EXC:*#metopozaur*;
p.r.a.u>!NC.!NC.!NC./u/;
d.n.a.u>!NC.!NC.!NC./u/;
u.n.a.u>!NC.!NC.!NC./u/;
e.n.a.u>!NC.!NC.!NC./u/>!EXC:*#adenauer,#birkenau*;
e.z.a.u>!NC.!NC.!NC./u/>!EXC:*#stezaur,#tezaur*;
o.n.a.u>!NC.!NC.!NC./u/>!EXC:*#aeronaut,#argonaut,#astroaeronaut,#astronaut,#bioastronaut,#biokosmonaut,#hiponaut,#infonaut,#kosmonaut,#lunonaut,#paleoastronaut,#selenonaut,#taikonaut,#ufonaut*;
a.n.a.u>!NC.!NC.!NC./u/>!EXC:*#akwanaut*;
y.n.a.u>!NC.!NC.!NC./u/;

Analogiczną strukturę reguł można zastosować dla wyrazów zawierających triadę *nau* w nagłosie. W takich wypadkach litera *u* powinna być transkrybowana jako fonem /u/, za wyjątkiem form wyrazów: *nauplius, Nautilus, nautofon, nautologia, nautologiczny, nautyczny, nautyka, naumachia, Nauranka, Naurańczyk, naurański, Nauru, nautykwariat, Nauzykaa.* Odpowiednia reguła ma następującą konstrukcję:

#.n.a.u>!NC.!NC.!NC./u/>!EXC:*#nauplius,#nautilus,#nautofon,#nautolog,#nautyczn,#nautyk,#naumach,#nauran,#naurań,#nauru#,#nautykwar,#nauzyk*;

Również dla wyrazów ortograficznych zawierających w nagłosie strukturę *#zau*(X–*t*) (w których litera *u* powinna być transkrybowana jako fonem /u/) można utworzyć regułę, która będzie przesłaniać podstawową regułę dwuelementową. Nie obejmuje ona wyrazów *zautomatyzować i zautonomizować*:

#.z.a.u.[X1]–*t*>!NC.!NC.!NC./u/.!NC;

4.2.5. Reguły przetwarzania litery *i*

Włączenie reguł dotyczących litery *i* do pierwszej warstwy projektu ma istotne znaczenie, ponieważ zapewniają one odpowiednie przyporządkowywanie znaków ortograficznych do segmentów fonologicznych. Ma to związek z wielofunkcyjnością litery *i*.

Reguły zawarte w tabeli 2.10 pokrywają się z transkrypcją podstawową (nie ma więc konieczności tworzenia reguł dodatkowych). Odnoszą się one do konstrukcji typu: *stoi, naiwny, naigrywać*. Poza tym na podstawie tabeli 2.11 można utworzyć następujące reguły:

p.i.[SA1]–*i*>!NC./j/.!NC;
b.i.[SA1]–*i* >!NC./j/.!NC;
f.i.[SA1]–*i*>!NC./j/.!NC;
w.i.[SA1]–*i*>!NC./j/.!NC;
m.i.[SA1]–*i*>!NC./j/.!NC;
t.i.[SA1]–*i*>!NC./j/.!NC;
d.i.[SA1]–*i*>!NC./j/.!NC;
h.i.[SA1]–*i*>!NC./j/.!NC;
l.i.[SA1]–*i*>!NC./j/.!NC;
r.i.[SA1]–*i*>!NC./j/.!NC;

Dotyczą one struktur {*p, b, f, w, m, t, d, h, l, r*}*i* przed literą oznaczającą samogłoskę (oprócz *i*); są to zatem struktury umiejscowione w nagłosie lub w śródgłosie wyrazu. Natomiast tabela 2.12 dotyczy struktur {*p, b, f, w, m, t, d, h, l, r*}*ii* umiejscowionych w wygłosie. Istnieją dwie możliwości transkrypcji takich triad ortograficznych. Pierwsza możliwość jest zgodna z transkrypcją podstawową. Przyjęto rozwiązanie polegające na uwzględnieniu w pierwszej warstwie wariantu pokrywającego się z transkrypcją podstawową (co w praktyce oznacza brak konieczności tworzenia dodatkowych reguł), natomiast możliwość ewentualnego odniesienia do wariantu drugiego została uwzględniona w warstwie modyfikacyjnej.

Zgodnie z *Automatyzacją*... we wszystkich wyrazach pochodnych od wyrazów: *ekspiacja, kwietyzm, klient, miastenia, patriarcha, patriota* literze *i* odpowiada sekwencja fonemów /i.j/. Do warstwy transkrypcji podstawowej można włączyć następujące reguły odnoszące się do wszystkich form fleksyjnych podanych wyrazów:

#.e.k.s.p.i.a.c.j>!NC.!NC.!NC.!NC.!NC./i/&/j/.!NC.!NC.!NC;
#.k.w.i.e.t.y.z.m>!NC.!NC.!NC./i/&/j/.!NC.!NC.!NC.!NC.!NC;
#.k.l.i.e.n.t>!NC.!NC.!NC./i/&/j/.!NC.!NC.!NC;
#.m.i.a.s.t.e.n>!NC.!NC./i/&/j/.!NC.!NC.!NC.!NC.!NC;
#.p.a.t.r.i.a.r.c.h>!NC.!NC.!NC.!NC.!NC./i/&/j/.!NC.!NC.!NC.!NC;
#.p.a.t.r.i.o.t>!NC.!NC.!NC.!NC.!NC./i/&/j/.!NC.!NC;

Zgodnie z informacjami ujętymi w podrozdziale 2.2.6 (w punkcie d) w strukturze {*#,a*}{*t,p*}*r*[*i*]A literze *i* odpowiada sekwencja fonemów /i.j/. Można więc zaproponować odpowiednie konstrukcje reguł dla pierwszej warstwy omawianego projektu:

#.t.r.i.[SA1]>!NC.!NC.!NC./i/&/j/.!NC;
#.p.r.i.[SA1]>!NC.!NC.!NC./i/&/j/.!NC;
a.t.r.i.[SA1].[X1]–#>!NC.!NC.!NC./i/&/j/.!NC.!NC;
a.p.r.i.[SA1].[X1]–#>!NC.!NC.!NC./i/&/j/.!NC.!NC;

Na podstawie przedrostków: *bi, mili, poli, mini* (umiejscowionych przed rdzeniem zawierającym samogłoskę na pierwszej pozycji) utworzono reguły, które ze względu na rozmiar przesłaniają reguły podane wcześniej (odnoszące się do tabeli 2.11). Litera *i* przed junkturą morfologiczną powinna być transkrybowana jako fonem /i/. Na podstawie informacji zamieszczonych w podrozdziale 2.2.6 (w punkcie e) skonstruowano następujący zestaw reguł:

#.b.i.a.n.d.r>!NC.!NC./i/.!NC.!NC.!NC.!NC;
#.m.i.l.i.a.m.p>!NC.!NC.!NC.!NC./i/.!NC.!NC.!NC;
#.p.o.l.i.a>!NC.!NC.!NC.!NC./i/.!NC;
#.p.o.l.i.e>!NC.!NC.!NC.!NC./i/.!NC;
#.p.o.l.i.i>!NC.!NC.!NC.!NC./i/./i/;
#.p.o.l.i.u>!NC.!NC.!NC.!NC./i/.!NC;
#.p.o.l.i.o.c>!NC.!NC.!NC.!NC./i/.!NC.!NC;
#.m.i.n.i.[SA1]>!NC.!NC.!NC.!NC./i/.!NC>!EXC:*#miniator,#miniatur,#minier,#minięc,#minion,#miniow,#miniów,#minięć#,#minier#,#minia#,#minią#*;

Zgodnie z *Automatyzacją…* litera *i* w strukturze *gi*A oznacza fonem /j/, jeżeli kontekst jest inny niż litera *e*. Zakładając, że głoska [ǵ] (AS) jest alofonem fonemu /J/, można skonstruować odpowiednią regułę:

g.i.[SA1]–*e*>/J/./j/.!NC;

Przy założeniu, że pierwsza litera *i* w triadzie *gii* umiejscowionej w wygłosie wyrazu oznacza fonem /i/, trzeba utworzyć regułę czteroelementową, która będzie przesłaniać podaną regułę trójelementową:

g.i.i.#>/J/./i/.!NC.!NC;

Ortograficzna triada *gie* jest wymawiana w sposób synchroniczny za wyjątkiem wyraz *higiena*. Można zaproponować dwie reguły:

g.i.e>/J/.!ML.!NC;
h.i.g.i.e.n>!NC.!NC./J/./j/.!NC.!NC;

Reguła druga ma wyższy priorytet ze względu na długość. Triada *gie* umiejscowiona w wygłosie najczęściej jest wymawiana w sposób asynchroniczny. Ma to związek z występowaniem znacznej liczby wyrazów obcych. Liczba wyrazów rodzimych realizowanych synchronicznie w wygłosie jest stosunkowo niewielka. W punkcie h w podrozdziale 2.2.6 wymieniono ciągi ortograficzne umożliwiające ich identyfikację. Te struktury zostały użyte w konstrukcji następujących reguł:

g.i.e.#>/J/./j/.!NC.!NC>!EXC:*nogie#,rogie#,mózgie#,brzegie#,chędogie#,wargie#,długie#,drugie#,nagie#,połogie#,tęgie#,ubogie#,wielgie#,pręgie#*;
g.i.e.#>/J/.!ML.!NC.!NC>!INC:*nogie#,rogie#,mózgie#,brzegie#,chędogie#,wargie#,długie#,drugie#,nagie#,połogie#,tęgie#,ubogie#,wielgie#,pręgie#*;

Pierwsza reguła dotyczy wymowy asynchronicznej w wyrazach obcych. Lista wyjątków odnosi się do wyrazów rodzimych wymawianych synchronicznie. Natomiast reguła druga dotyczy tych samych wyrazów rodzimych, które są identyfikowane przy użyciu listy włączeń.

W wypadku wyodrębnienia fonemu /J/ (wiąże się to z obecnością wymowy synchronicznej) do omawianej warstwy projektu trzeba włączyć regułę dla odpowiedniej transkrypcji diady *gi* umiejscowionej przed spółgłoską, półsamogłoską lub w wygłosie wyrazu:

g.i.[SP1]+#>/J/.!NC.!NC;

Możliwa jest również bardziej współczesna interpretacja zakładająca pełną asynchronię wymowy. W takim wariancie fonem /J/ nie zostaje wyodrębniony i wystarczy utworzyć jedną regułę, która przyjmuje następujący kształt:

g.i.[SA1]>/g/./j/.!NC;

Zostały utworzone dwie reguły dla struktury *ki*A. Zgodnie z nimi ortograficzna triada *kie* ma wymowę synchroniczną, a pozostałe konstrukcje mają wymowę asynchroniczną:

k.i.[SA1]–*e*>/c/./j/.!NC;
k.i.e>/c/.!ML.!NC;

Obecność wymowy synchronicznej wiąże się z wyodrębnieniem fonemu /c/. A więc do warstwy transkrypcji podstawowej trzeba włączyć regułę dla transkrypcji diady *ki* umiejscowionej przed spółgłoską, półsamogłoską lub w wygłosie wyrazu:

k.i.[SP1]+#>/c/.!NC.!NC;

Można również odnieść się do bardziej współczesnej interpretacji, która zakłada pełną asynchronię wymowy (w tym wariancie fonem /c/ nie zostaje wyodrębniony):

k.i.[SA1]=/k/./j/.!NC;

W warstwie transkrypcji podstawowej dla struktur *gi*A oraz *ki*A można przyjąć wymowę w pełni asynchroniczną. Omówione reguły związane z wyodrębnianiem fonemu /J/ oraz fonemu /c/ można przenieść do warstwy modyfikacyjnej. Przy takim podejściu wymowa synchroniczna może być traktowana jako opcja.

4.2.6. Reguły przetwarzania liter: *ę* oraz *ą*

W prezentowanym projekcie reguły dotyczące transkrypcji litery *ą* oraz litery *ę* zostały przesunięte do warstwy modyfikacyjnej (jest to warstwa trzecia). Do pierwszej warstwy projektu włączono reguły sprawiające, że litery te są przekształcane na tymczasową notację, która nie ma charakteru bifonematycznego: /ą/ oraz /ę/. Taki sposób zapisu zostaje zachowany w niezmienionej postaci aż do momentu przekazania wyrazu do warstwy modyfikacyjnej. Odpowiednie reguły mają postać:

ą>/ą/;
ę>/ę/;

4.3. Warstwa statusu dźwięczności

Reguły włączone do drugiej warstwy prezentowanego rozwiązania umożliwiają modyfikację statusu dźwięczności przetwarzanych segmentów. Wyrazy przekazywane do tej warstwy mają strukturę wynikającą z modyfikacji wykonywanych w ramach warstwy transkrypcji podstawowej. Zgodnie z przyjętą terminologią składają się one z segmentów drugiego stopnia i nie można ich traktować jako pełnoprawnych (zdefiniowanych) jednostek lingwistycznych. Inwentarz segmentów drugiego stopnia wynika z konstrukcji reguł włączonych do warstwy transkrypcji podstawowej. Zdecydowana większość elementów w tym inwentarzu pokrywa się z przyjętą transkrypcją podstawową, ponieważ przede wszystkim na tej transkrypcji opiera się funkcjonowanie pierwszej warstwy omawianego projektu. Poza tym reguły w omawianej warstwie mają związek z wielofunkcyjnością litery *i* oraz z tymi strukturami ortograficznymi, które nie należą do przyjętej transkrypcji podstawowej. W tabeli 4.2 zostały wymienione wszystkie segmenty należące do inwentarza drugiego stopnia. Do warstwy statusu dźwięczności mogą być przekazywane wyrazy zapisane przy użyciu tylko tych jednostek (znakiem * oznaczono segmenty opcjonalne).

Tabela 4.2 – Inwentarz segmentów drugiego stopnia

Lp.	Segment drugiego stopnia	Przykładowy wyraz	Dźwięczność	Lp.	Segment drugiego stopnia	Przykładowy wyraz	Dźwięczność
1	/a/	/m.**a**.k/	tak	21	/rz/	/**rz**.e.k.a/	tak
2	/e/	/z.**e**.r.o/	tak	22	/Z/	/**Z**.a.b.a/	tak
3	/i/	/n'.**i**.ts/	tak	23	/s'/	/**s'**.m.j.e.x/	nie
4	/o/	/s.**o**.v.a/	tak	24	/z'/	/**z'**.a.r.n.o/	tak
5	/y/	/b.**y**.k/	tak	25	/x/	/k.u.**x**.n'.a/	nie
6	/u/	/j.**u**.t.r.o/	tak	26	/p/	/**p**.a.l.e.ts/	nie
7	/w/	/p.u.**w**.k.a/	tak	27	/b/	/**b**.u.d.a/	tak
8	/j/	/**j**.e.d.e.n/	tak	28	/t/	/**t**.a.m.a/	nie
9	/ę/	/j.**ę**.z.y.k/	tak	29	/d/	/**d**.o.m/	tak
10	/ą/	/p.a.j.**ą**.k/	tak	30	/k/	/p.o.**k**.u.j/	nie
11	/l/	/v.j.e.**l**.e/	tak	31	/c/*	/**c**.i.n.o/	nie
12	/r/	/**r**.y.b.a/	tak	32	/g/	/**g**.o.s'.ts'/	tak
13	/m/	/**m**.o.Z.e/	tak	33	/J/*	/**J**.i.t.a.r.a/	tak
14	/n/	/m.o.**n**.e.t.a/	tak	34	/ts/	/**ts**.y.r.k/	nie
15	/n'/	/k.o.**n'**/	tak	35	/dz/	/**dz**.b.a.n/	tak
16	/f/	/**f**.u.t.r.o/	nie	36	/tS/	/**tS**.a.s/	nie
17	/v/	/**v**.j.a.t.r/	tak	37	/dZ/	/**dZ**.u.m.a/	tak
18	/s/	/v.y.**s**.o.k.i/	nie	38	/ts'/	/k.o.**ts'**.o.w/	nie
19	/z/	/k.o.**z**.a/	tak	39	/dz'/	/**dz'**.a.w.k.a/	tak
20	/S/	/m.a.**S**.t/	nie				

Podstawowy cel omawianej warstwy ma związek z modyfikacją statusu dźwięczności segmentów w wyrazach przetworzonych w warstwie transkrypcji podstawowej. Zmiana statusu dźwięczności dotyczy spółgłosek właściwych – każdy segment drugiego stopnia będący spółgłoską właściwą bezdźwięczną (z wyjątkiem segmentu /x/) ma swój dźwięczny odpowiednik, który charakteryzuje się takim samym miejscem i sposobem artykulacji. Trzeba podkreślić, że status dźwięczności segmentów drugiego stopnia wynika wprost z transkrypcji podstawowej, czyli jest to tak zwana dźwięczność ortograficzna (zob. podrozdz. 1.4). Odzwierciedla ona status dźwięczności domyślnie przypisany do liter, diad oraz triad w pierwotnym zapisie ortograficznym. Na przykład wyraz *gorzki* po przetworzeniu w warstwie transkrypcji podstawowej ma zapis: /g.o.rz.k.i/. W tym zapisie segment /rz/ jest dźwięczny i wynika to ze struktury ortograficznej wyrazu *gorzki* oraz z transkrypcji podstawowej, zgodnie z którą diada *rz* oznacza dźwięczny szczelinowy fonem /rz/. W tabeli 4.2 informację o statusie dźwięczności segmentów drugiego stopnia podano w kolumnie *Dźwięczność*.

Modyfikacja statusu dźwięczności obejmuje tylko te nagłosowe i śródgłosowe grupy spółgłoskowe, które w pierwotnym zapisie ortograficznym są niejednorodne pod względem tzw. dźwięczności ortograficznej. Polega ona na ujednoliceniu statusu dźwięczności segmentów drugiego stopnia. Niektóre segmenty dźwięczne mogą być więc zastępowane przez ich bezdźwięczne odpowiedniki lub segmenty bezdźwięczne mogą być zastępowane przez odpowiedniki dźwięczne. Używane w tej publikacji sformułowanie: „modyfikacja statusu dźwięczności segmentów" oznacza zastępowanie określonych segmentów drugiego stopnia ich dźwięcznymi lub bezdźwięcznymi odpowiednikami.

W tej części przedstawiono również reguły przeznaczone do modyfikacji dźwięczności, które wynikają z zależności międzywyrazowych. Realizacja tych procesów ma związek z typem wymowy i jest warunkowana zjawiskiem antycypacji. Oznacza to, że ewentualne zmiany statusu dźwięczności w obrębie spółgłoskowego wygłosu danego wyrazu są uzależnione od struktury nagłosu wyrazu następnego – sytuacja odwrotna nie jest dozwolona. Modyfikacja statusu dźwięczności może obejmować pojedynczą spółgłoskę lub grupę spółgłoskową umiejscowioną w wygłosie wyrazu. Dotyczy to zarówno grup jednorodnych, jak i grup niejednorodnych pod względem tzw. dźwięczności ortograficznej.

W dalszym ciągu został zawarty szczegółowy opis dotyczący dopuszczalnych modyfikacji statusu dźwięczności segmentów. Dla każdego segmentu drugiego stopnia (spółgłoski właściwej) wyznaczono podstawową regułę modyfikacji statusu dźwięczności, która odnosi się do minimalnego kontekstu determinującego taką zmianę. Wszystkie reguły podstawowe są dwuelementowe i zgodnie z przyjętymi założeniami mogą być przesłaniane przez reguły dłuższe, które zostały zdefiniowane w ramach tej samej warstwy.

Trzeba podkreślić, że reguły dotyczące modyfikacji statusu dźwięczności są oparte na obszernych badaniach omówionych w rozdziałach 2–5 w *Asymilacji…* . Sposób zapisywania grup spółgłoskowych w tamtych

analizach pokrywa się ze sposobem zapisywania grup w wyrazach przetworzonych w pierwszej warstwie. Szczególne znaczenie ma metoda zapisywania nagłosowych i śródgłosowych grup spółgłoskowych, która została użyta w rozdziale drugim w *Asymilacji…* . Uwzględniono w niej nie tylko transkrypcję podstawową, ale również ewentualne modyfikacje wynikające z kontekstu samogłoskowego. W rezultacie uzyskiwany jest zapis, w którym są wydzielone struktury pokrywające się z fonemami, jednak ich dźwięczność odzwierciedla tzw. dźwięczność ortograficzną liter (zob. podrozdz. 1.4). Zapis uzyskiwany w wyniku przetwarzania wyrazów w pierwszej warstwie prezentowanego tutaj projektu ma te same cechy. Dlatego wyniki analiz z *Asymilacji…* można wykorzystać dla projektowania reguł w warstwie statusu dźwięczności.

Trzeba zdefiniować zbiory używane w regułach należących do warstwy statusu dźwięczności. Zbiory te mogą obejmować tylko segmenty drugiego stopnia. Z takich jednostek są zbudowane wyrazy przetworzone w pierwszej warstwie rozwiązania. Zostały one wyszczególnione w tabeli 4.2. W nazwach zbiorów cyfra 2 oznacza, że są one zdefiniowane w warstwie drugiej. Zbiór WD2 obejmuje spółgłoski właściwe dźwięczne. Definicja tego zbioru jest następująca:

[WD2]={/z/,/Z/,/z'/,/b/,/d/,/g/,/J/,/dz/,/dZ/,/dz'/};

Definicja zbioru WB2 obejmuje segmenty drugiego stopnia będące spółgłoskami właściwymi bezdźwięcznymi:

[WB2]={/f/,/s/,/S/,/s'/,/x/,/p/,/t/,/k/,/c/,/ts/,/tS/,/ts'/};

Ostatnia definicja obejmuje segmenty drugiego stopnia stanowiące spółgłoski sonorne. Do tego zbioru włączono również półsamogłoski (na podstawie ustalenia umownego):

[SO2]={/w/,/j/,/l/,/r/,/m/,/n/,/n'/};

Ze zbioru WD2 zostały wykluczone segmenty: /v/ oraz /rz/ drugiego stopnia. W przyjętej transkrypcji podstawowej odpowiadają one literze *w* oraz diadzie *rz*, a więc podlegają zjawisku perseweracji artykulacyjnej – mogą być ubezdźwięczniane na podstawie kontekstu poprzedzającego. Ewentualne włączenie tych segmentów do zbioru WD2 spowodowałoby, że przykładowa reguła: /ts'/.[WD2]>/dz'/.!NC; oraz reguły podobne nie mogłyby być używane. Powodowałaby ona niepoprawne dźwięczne przekształcenie nagłosu w przykładowych wyrazach: *ćwierć, ćwiczyć*. W omawianym rozwiązaniu jest to również przyczyna używania oddzielnej notacji fonemu /Z/ (/rz/) (zob. podrozdz. 1.1).

4.3.1. Reguły modyfikacji statusu dźwięczności segmentu /v/ drugiego stopnia

Zmiana statusu dźwięczności segmentu /v/ drugiego stopnia może wynikać zarówno ze struktury kontekstu następującego (ma to związek ze zjawiskiem antycypacji artykulacyjnej), jak i ze struktury kontekstu poprzedzającego (dotyczy to zjawiska perseweracji artykulacyjnej). Na podstawie podsumowania przedstawionego w tabeli 2.10 w *Asymilacji…* można utworzyć dwie podstawowe reguły dwuelementowe przetwarzania segmentu /v/ drugiego stopnia w ramach warstwy statusu dźwięczności:

[WB2]./v/>!NC./f/;
/v/.[WB2]>/f/.!NC;

Istnieje znaczna liczba wewnątrzwyrazowych grup spółgłoskowych, w których kontekst ubezdźwięczniający obejmuje więcej niż jeden segment drugiego stopnia. W regułach utworzonych dla przetwarzania takich grup nie zostały użyte oznaczenia szerokich klas dźwięków (symbole nazw zbiorów). Wynika to z faktu, że reguły te są powtarzane w opisach dotyczących różnych segmentów (w niniejszym rozdziale). Aby ostatecznie uniknąć powtarzania tych samych reguł, wykonano analizę podsumowującą (zob. podrozdz. 4.3.22), która dotyczy wszystkich grup omawianego typu. Dzięki temu, że nie zostały użyte symbole zbiorów, możliwe było wykluczenie reguł występujących powtórnie.

Listę reguł związanych z modyfikacją statusu dźwięczności segmentu /v/ drugiego stopnia (w grupach spółgłoskowych, w których występują co najmniej dwa segmenty drugiego stopnia w kontekście ubezdźwięczniającym) podzielono na kilka części. Pierwsza część dotyczy zjawiska perseweracji artykulacyjnej:

/rz/./k/./v/>/S/.!NC./f/;
/rz/./k/./v/./j/>/S/.!NC./f/.!NC;
/z/./t/./v/>/s/.!NC./f/;
/z/./s/./v/>/s/.!NC./f/;
/z/./S/./v/>/s/.!NC./f/;
/z/./k/./v/>/s/.!NC./f/;
/b/./x/./v/>/p/.!NC./f/;
/d/./x/./v/>/t/.!NC./f/;
/d/./s'/./v/>/t/.!NC./f/;
/d/./tS/./v/>/t/.!NC./f/;
/d/./k/./v/>/t/.!NC./f/;
/d/./t/./v/>/t/.!NC./f/;
/dz/./t/./v/>/ts/.!NC./f/;
/z/./x/./v/./j/>/s/.!NC./f/.!NC;
/z/./s'/./v/./j/>/s/.!NC./f/.!NC;
/z/./x/./v/>/s/.!NC./f/;
/z/./ts'/./v/./j/>/s/.!NC./f/.!NC;
/z/./k/./v/./j/>/s/.!NC./f/.!NC;
/z/./t/./v/./j/>/s/.!NC./f/.!NC;
/d/./s'/./v/./j/>/t/.!NC./f/.!NC;
/d/./k/./v/./j/>/t/.!NC./f/.!NC;
/dz/./t/./v/./j/>/ts/.!NC./f/.!NC;
/b/./s/./t/./v/>/p/.!NC.!NC./f/;
/dz'/./s/./t/./v/>/ts'/.!NC.!NC./f/;
/b/./s/./t/./v/./j/>/p/.!NC.!NC./f/.!NC;
/dz'/./s/./t/./v/./j/>/ts'/.!NC.!NC./f/.!NC;
/b/./p/./s/./t/./v/>/p/.!NC.!NC.!NC./f/;
/b/./p/./s/./t/./v/./j/>/p/.!NC.!NC.!NC./f/.!NC;

Warto zauważyć, że wszystkie reguły zamieszczone na tej liście są co najmniej trójelementowe, czyli mają wyższy priorytet niż podstawowe reguły dwuelementowe związane z przetwarzaniem segmentu /v/ drugiego stopnia. Poniższe reguły również dotyczą zjawiska perseweracji, jednak segment /v/ jest w nich poprzedzony spółgłoską sonorną. To sprawia, że jego ubezdźwięcznienie jest mniej prawdopodobne:

/k/./r/./v/>!NC.!NC./f/;
/t/./r/./v/>!NC.!NC./f/;
/p/./l/./v/>!NC.!NC./f/;
/z/./k/./r/./v/>/s/.!NC.!NC./f/;
/z/./t/./r/./v/>/s/.!NC.!NC./f/;
/d/./s/./k/./r/./v/./j/>/t/.!NC.!NC.!NC./f/.!NC;

Oddzielna reguła została utworzona dla przetwarzania wyrazu *wszechwładca*. W wyrazie tym segment /v/ drugiego stopnia nie podlega perseweracji artykulacyjnej ze względu na obecność junktury morfologicznej. Przepisanie tego segmentu w drugim członie reguły sprawia, że jego status dźwięczności nie może być zmieniony (przy użyciu krótszej reguły podstawowej):

/S/./e/./x/./v/./w/>!NC.!NC.!NC./v/.!NC;

Kolejne prezentowane reguły mają związek z prawostronnym kontekstem ubezdźwięczniającym dla segmentu /v/ drugiego stopnia (w tych grupach zjawisko desonoryzacji obejmuje co najmniej dwa segmenty):

/rz/./v/./ts'/>/S/./f/.!NC;
/v/./rz/./k/>/f/./S/.!NC;
/v/./rz/./c/>/f/./S/.!NC;
/v/./rz/./k/./j/>/f/./S/.!NC.!NC;
/v/./v/./s/./t/./rz/>/f/./f/.!NC.!NC./S/;

/v/./v/./s/./k/>/f/./f/.!NC.!NC;
/v/./x/./rz/>/f/.!NC./S/;
/v/./k/./rz/>/f/.!NC./S/;
/v/./p/./rz/>/f/.!NC./S/;
/v/./t/./rz/>/f/.!NC./S/;
/v/./s/./k/./rz/>/f/.!NC.!NC./S/;
/v/./s/./t/./rz/>/f/.!NC.!NC./S/;
/v/./dz'/./ts'/>/f/./ts'/.!NC;
/j/./v/./s/./t/./rz/>!NC./f/.!NC.!NC./S/;
/z'/./v/./ts/>/s'/./f/.!NC;
/z'/./v/./ts'/>/s'/./f/.!NC;
/d/./v/./s'/>/t/./f/.!NC;
/d/./v/./S/>/t/./f/.!NC;
/d/./v/./tS/>/t/./f/.!NC;
/z/./v/./s/./t/>/s/./f/.!NC.!NC;
/z/./v/./s'/./ts'/>/s/./f/.!NC.!NC;
/d/./v/./s/./t/>/t/./f/.!NC.!NC;

Niektóre wymienione reguły odnoszą się zarówno do zjawiska antycypacji, jak i perseweracji artykulacyjnej (na przykład reguły utworzone dla grup: /v.x.rz/, /v.k.rz/).

Utworzono jedną regułę, która dotyczy przetwarzania form wyrazu *sewrski*. W wyrazie tym prawostronny kontekst ubezdźwięczniający segmentu /v/ drugiego stopnia jest poprzedzony spółgłoską sonorną:

/v/./r/./s/>/f/.!NC.!NC;

Istnieje kilka grup umiejscowionych w nagłosie wyrazu, które są złożone z segmentu /v/ drugiego stopnia oraz z innego segmentu (najczęściej z segmentu /rz/), który również podlega desonoryzacji. Dla przetwarzania takich grup utworzono następujące reguły:

#./v/./x/./rz/>!NC./f/.!NC./S/;
#./v/./k/./rz/>!NC./f/.!NC./S/;
#./v/./p/./rz/>!NC./f/.!NC./S/;
#./v/./t/./rz/>!NC./f/.!NC./S/;
#./v/./s/./k/./rz/>!NC./f/.!NC.!NC./S/;
#./v/./s/./t/./rz/>!NC./f/.!NC.!NC./S/;
#./v/./z/./S/>!NC./f/./s/.!NC;

W wypadku grup nagłosowych, w których tylko segment /v/ znajduje się w kontekście ubezdźwięczniającym (lewostronnym lub prawostronnym), zastosowanie mają dwie podstawowe reguły dwuelementowe, a więc nie ma potrzeby tworzenia oddzielnych reguł.

4.3.2. Reguły modyfikacji statusu dźwięczności segmentu /b/ drugiego stopnia

Na podstawie tabeli 2.18 w *Asymilacji…* można utworzyć jedną regułę obejmującą różne konteksty ubezdźwięczniające segmentu /b/ drugiego stopnia:

/b/.[WB2]>/p/.!NC;

Pierwszy jej człon występuje również w bardziej złożonych strukturach wewnątrzwyrazowych grup spółgłoskowych, np.: /b/.[WB2].[SO2], /b/.[WB2].[WB2].[SO2]. Można także utworzyć jedną regułę dla przetwarzania ciągu złożonego z segmentu /b/, ze spółgłoski sonornej (lub półsamogłoski) oraz ze spółgłoski właściwej bezdźwięcznej:

/b/.[SO2].[WB2]>/p/.!NC.!NC;

Utworzono oddzielną regułę dla przetwarzania wyrazu *gibbsyt* (regułę związaną z redukcją włączono do warstwy modyfikacyjnej):

/b/./b/./s/>/p/./p/.!NC;

Pozostałe reguły dotyczą struktur spółgłoskowych, w których desonoryzacja obejmuje więcej segmentów drugiego stopnia (więcej niż jeden segment /b/). Te reguły nie zawierają oznaczeń klas dźwięków (symboli zdefiniowanych zbiorów):

/b/./t/./rz/>/p/.!NC./S/;
/b/./x/./v/>/p/.!NC./f/;
/b/./s/./t/./v/>/p/.!NC.!NC./f/;
/b/./s/./t/./v/./j/>/p/.!NC.!NC./f/.!NC;
/b/./p/./s/./t/./v/>/p/.!NC.!NC.!NC./f/;
/b/./p/./s/./t/./v/./j/>/p/.!NC.!NC.!NC./f/.!NC;
/b/./s/./t/./rz/>/p/.!NC.!NC./S/;
/rz/./b/./ts'/>/S/./p/.!NC;
/rz/./b/./ts/>/S/./p/.!NC;
/b/./rz/./ts'/>/p/./S/.!NC;
/z'/./b/./ts/>/s'/./p/.!NC;
/z'/./b/./ts'/>/s'/./p/.!NC;

4.3.3. Reguły modyfikacji statusu dźwięczności segmentu /g/ drugiego stopnia

Na podstawie wyników zamieszczonych w tabeli 2.22 w *Asymilacji…* można utworzyć podstawową regułę dwuelementową, która pozwala na zmianę statusu dźwięczności segmentu /g/ drugiego stopnia:

/g/.[WB2]>/k/.!NC;

Ponadto dla przetwarzania wyrazów zawierających w swojej strukturze diadę ortograficzną *gg* utworzono następujące reguły trójelementowe:

/g/./g/./s/>/k/./k/.!NC;
/g/./g/./s'/>/k/./k/.!NC;

W kilku grupach spółgłoskowych pomiędzy segmentem /g/ drugiego stopnia a spółgłoską właściwą bezdźwięczną jest umiejscowiona spółgłoska sonorna (lub półsamogłoska). Dotyczy to grup: /g.r.ts/, /g.w.S/, /l.g.w.S/. Dla takich struktur można utworzyć jedną regułę:

/g/.[SO2].[WB2]>/k/.!NC.!NC;

Półsamogłoska poprzedzająca kontekst ubezdźwięczniający występuje również w grupie /Z.g.w.S/. Dla transkrypcji tej grupy utworzono osobną regułę, ponieważ zawiera ona dwa segmenty dźwięczne drugiego stopnia, których status dźwięczności jest zmieniany:

/Z/./g/./w/./S/>/S/./k/.!NC.!NC;

Ubezdźwięczniający kontekst dla co najmniej dwóch segmentów drugiego stopnia (łącznie z segmentem /g/) jest obecny także w kilku innych grupach. Dla przetwarzania tych grup utworzono następujące reguły:

/z/./g/./rz/./ts'/>/s/./k/./S/.!NC;
/g/./rz/./ts/>/k/./S/.!NC;
/g/./dz/./k/>/k/./ts/.!NC;
/g/./dz/./ts/>/k/./ts/.!NC;

4.3.4. Reguły modyfikacji statusu dźwięczności segmentu /z'/ drugiego stopnia

Podstawowa reguła dwuelementowa, która umożliwia transkrypcję segmentu /z'/ umiejscowionego w kontekście ubezdźwięczniającym, to:

/z'/.[WB2]>/s'/.!NC;

Utworzono dwie reguły dotyczące struktur, w których segment /z'/ oraz kontekst ubezdźwięczniający są rozdzielone spółgłoska sonorną:

/z'/./n'/./ts'/>/s'/.!NC.!NC;
/z'/./l/./ts'/>/s'/.!NC.!NC;

Można je zastąpić jedną uogólnioną regułą:

/z'/.[SO2].[WB2]>/s'/.!NC.!NC;

Bez użycia symboli zbiorów utworzono reguły związane ze zmianą statusu dźwięczności dwóch segmentów drugiego stopnia:

/z'/./v/./ts'/>/s'/./f/.!NC;
/z'/./v/./ts/>/s'/./f/.!NC;
/z'/./b/./ts/>/s'/./p/.!NC;
/z'/./b/./ts'/>/s'/./p/.!NC;
/z'/./dz'/./ts/>/s'/./ts'/.!NC;
/z'/./dz'/./ts'/>/s'/./ts'/.!NC;
/z'/./dz'/./tS/>/s'/./ts'/.!NC;

4.3.5. Reguły modyfikacji statusu dźwięczności segmentu /d/ drugiego stopnia

Podstawowa reguła związana z modyfikacją statusu dźwięczności segmentu /d/ dotyczy kilku struktur wewnątrzwyrazowych grup spółgłoskowych, np.: /d/.[WB2].[SO2].[SO2], /d/.[WB2].[SO2], /d/.[WB2].[WB2]. Ma ona następujący kształt:

/d/.[WB2]>/t/.!NC;

Z tabeli 2.30 w *Asymilacji…* wynika, że istnieje znaczna liczba różnych struktur wewnątrzwyrazowych grup spółgłoskowych, w których między segmentem /d/ drugiego stopnia a spółgłoską właściwą bezdźwięczną jest umiejscowiona spółgłoska sonorna (bądź półsamogłoska). Dla przetwarzania takich struktur można utworzyć jedną regułę:

/d/.[SO2].[WB2]>/t/.!NC.!NC;

Poza tym istnieje kilka grup, w których co najmniej dwa segmenty dźwięczne drugiego stopnia są oddzielone od kontekstu ubezdźwięczniającego spółgłoską sonorną. Utworzono dla nich następujące reguły:

/z/./d/./r/./s/./k/>/s/./t/.!NC.!NC.!NC;
/z/./d/./r/./s/./ts/>/s/./t/.!NC.!NC.!NC;
/z/./d/./r/./ts/>/s/./t/.!NC.!NC;

Oprócz tego istnieje liczny zbiór grup w wyrazach zwracanych przez pierwszą warstwę, w których co najmniej dwa segmenty dźwięczne (łącznie z segmentem /d/ drugiego stopnia) są umiejscowione w kontekście ubezdźwięczniającym. Oto reguły utworzone dla modyfikacji statusu dźwięczności w tych grupach:

/d/./rz/./ts'/>/t/./S/.!NC;
/d/./v/./s'/>/t/./f/.!NC;
/d/./v/./S/>/t/./f/.!NC;
/d/./v/./tS/>/t/./f/.!NC;
/d/./v/./s/./t/>/t/./f/.!NC.!NC;
/d/./x/./rz/>/t/.!NC./S/;
/d/./k/./rz/>/t/.!NC./S/;
/d/./p/./rz/>/t/.!NC./S/;
/d/./t/./rz/>/t/.!NC./S/;

/d/./k/./rz/./t/>/t/.!NC./S/.!NC;
/d/./x/./v/>/t/.!NC./f/;
/d/./s'/./v/>/t/.!NC./f/;
/d/./tS/./v/>/t/.!NC./f/;
/d/./k/./v/>/t/.!NC./f/;
/d/./t/./v/>/t/.!NC./f/;
/d/./s'/./v/./j/>/t/.!NC./f/.!NC;
/d/./k/./v/./j/>/t/.!NC./f/.!NC;
/d/./s/./k/./rz/>/t/.!NC.!NC./S/;
/d/./s/./t/./rz/>/t/.!NC.!NC./S/;
/d/./s/./k/./r/./v/./j/>/t/.!NC.!NC.!NC./f/.!NC;
/n/./d/./rz/./ts'/>!NC./t/./S/.!NC;
/z/./d/./ts/>/s/./t/.!NC;
/z/./d/./k/>/s/./t/.!NC;
/z/./d/./rz/./ts'/>/s/./t/./S/.!NC;

Wykaz obejmuje kilka grup zawierających sekwencję /d.rz/ lub /t.rz/. Podane reguły dotyczą tylko modyfikacji statusu dźwięczności. Redukcja związana ze zjawiskiem asybilacji jest realizowana w czasie przetwarzania w warstwie modyfikacyjnej (zob. podrozdz. 4.4.9 i 4.4.10).

4.3.6. Reguły modyfikacji statusu dźwięczności segmentu /z/ drugiego stopnia

Podstawowa reguła dotycząca kontekstu ubezdźwięczniającego segmentu /z/ drugiego stopnia jest dwuelementowa. Może ona być stosowana również dla bardziej złożonych struktur spółgłoskowych, np.: /z/.[WB2].[SO2], /z/.[WB2].[WB2]. Konstrukcja tej reguły jest następująca:

/z/.[WB2]>/s/.!NC;

Na podstawie informacji zawartych w podrozdziale 2.6 w *Asymilacji…* można utworzyć ogólną regułę dotyczącą prawostronnego kontekstu ubezdźwięczniającego poprzedzonego spółgłoską sonorną lub półsamogłoską:

/z/.[SO2].[WB2]>/s/.!NC.!NC;

Występują trzy grupy, w których przed spółgłoską sonorną umiejscowiona jest sekwencja /z.d/. Są to więc grupy zawierające więcej niż jeden segment drugiego stopnia, których status dźwięczności jest modyfikowany. Oto reguły przetwarzania tych grup:

/z/./d/./r/./ts/>/s/./t/.!NC.!NC;
/z/./d/./r/./s/./k/>/s/./t/.!NC.!NC.!NC;
/z/./d/./r/./s/./ts/>/s/./t/.!NC.!NC.!NC;

Poniższe reguły dotyczą grup, w których zjawisko desonoryzacji obejmuje segmenty: /z/ i /v/ drugiego stopnia. Segment /v/ jest poprzedzony spółgłoską sonorną, dlatego jego ubezdźwięcznienie jest mniej prawdopodobne:

/z/./k/./r/./v/>/s/.!NC.!NC./f/;
/z/./t/./r/./v/>/s/.!NC.!NC./f/;

Poza tym istnieje stosunkowo liczny zbiór innych grup spółgłoskowych zawierających jeszcze jeden segment drugiego stopnia (obok segmentu /z/), który podlega desonoryzacji, jednak kontekst ubezdźwięczniający jest umiejscowiony w ich bezpośrednim sąsiedztwie. Ostania reguła na poniższej liście odnosi się do nagłosu wyrazu:

/z/./v/./s/./t/>/s/./f/.!NC.!NC;
/z/./v/./s'/./ts'/>/s/./f/.!NC.!NC;
/z/./x/./rz/>/s/.!NC./S/;
/z/./k/./rz/>/s/.!NC./S/;
/z/./p/./rz/>/s/.!NC./S/;
/z/./t/./rz/>/s/.!NC./S/;

/z/./x/./v/>/s/.!NC./f/;
/z/./s/./v/>/s/.!NC./f/;
/z/./k/./v/>/s/.!NC./f/;
/z/./t/./v/>/s/.!NC./f/;
/z/./x/./v/./j/>/s/.!NC./f/.!NC;
/z/./s'/./v/./j/>/s/.!NC./f/.!NC;
/z/./ts'/./v/./j/>/s/.!NC./f/.!NC;
/z/./k/./v/./j/>/s/.!NC./f/.!NC;
/z/./t/./v/./j/>/s/.!NC./f/.!NC;
/z/./s/./k/./rz/>/s/.!NC.!NC./S/;
/z/./s/./p/./rz/>/s/.!NC.!NC./S/;
/z/./s/./t/./rz/>/s/.!NC.!NC./S/;
/z/./S/./v/>/s/.!NC./f/;
/z/./g/./rz/./ts'/>/s/./k/./S/.!NC;
/z/./d/./ts/>/s/./t/.!NC;
/z/./d/./k/>/s/./t/.!NC;
/z/./d/./rz/./ts'/>/s/./t/./S/.!NC;
/z/./dz/./ts/>/s/./ts/.!NC;
/z/./dz/./k/>/s/./ts/.!NC;
#./v/./z/./S/>!NC./f/./s/.!NC;

4.3.7. Reguły modyfikacji statusu dźwięczności segmentu /Z/ drugiego stopnia

Dla przetwarzania segmentu /Z/ drugiego stopnia utworzono następującą regułę podstawową:

/Z/.[WB2]>/S/.!NC;

W wypadku dwóch grup spółgłoskowych kontekst ubezdźwięczniający jest poprzedzony spółgłoską sonorną bądź półsamogłoską. Specjalnie dla tych grup utworzono reguły:

/Z/./n'/./ts'/>/S/.!NC.!NC;
/Z/./g/./w/./S/>/S/./k/.!NC.!NC;

Druga wymieniona reguła dotyczy grupy zawierającej dwa dźwięczne segmenty drugiego stopnia, których status dźwięczności musi być zmieniony. Poza tym istnieją jeszcze trzy inne grupy spółgłoskowe, w których występuje drugi segment (poza /Z/) wrażliwy na kontekst ubezdźwięczniający. Oto reguły utworzone dla przetwarzania tych grup:

/Z/./dZ/./ts/>/S/./tS/.!NC;
/Z/./dZ/./ts'/>/S/./tS/.!NC;
/Z/./dZ/./k/>/S/./tS/.!NC;

4.3.8. Reguły modyfikacji statusu dźwięczności segmentu /dZ/ drugiego stopnia

Dla modyfikacji statusu dźwięczności segmentu /dZ/ drugiego stopnia utworzono regułę analogiczną do innych reguł omówionych w tym rozdziale:

/dZ/.[WB2]>/tS/.!NC;

Na podstawie informacji zamieszczonych w tabeli 2.40 w *Asymilacji…* utworzono trzy reguły dla przetwarzania grup wewnątrzwyrazowych, zawierających w kontekście ubezdźwięczniającym dwa segmenty:

/Z/./dZ/./k/>/S/./tS/.!NC;
/Z/./dZ/./ts'/>/S/./tS/.!NC;
/Z/./dZ/./ts/>/S/./tS/.!NC;

4.3.9. Reguły modyfikacji statusu dźwięczności segmentu /dz/ drugiego stopnia

Podstawowa reguła umożliwiająca przetwarzanie segmentu /dz/ drugiego stopnia ma strukturę podobną do innych reguł dwuelementowych dotyczących kontekstu ubezdźwięczniającego:
/dz/.[WB2]>/ts/.!NC;

Utworzono sześć reguł (na podstawie tabeli 2.42 w *Asymilacji…*) dotyczących wewnątrzwyrazowych grup spółgłoskowych, w których zjawisko desonoryzacji obejmuje również inny segment drugiego stopnia (poza /dz/). Mogą to być segmenty: /v/, /z/ lub /g/. Dla grupy /dz.t.v.j/ utworzono oddzielną regułę, pomimo że odpowiednia byłaby reguła utworzona dla grupy /dz.t.v/. Redukcja ilościowa reguł związanych z desonoryzacją została omówiona w części 4.3.22, która stanowi podsumowanie niniejszego opisu. Oto lista reguł omówionych w niniejszym akapicie:

/dz/./t/./v/>/ts/.!NC./f/;
/dz/./t/./v/./j/>/ts/.!NC./f/.!NC;
/z/./dz/./k/>/s/./ts/.!NC;
/z/./dz/./ts/>/s/./ts/.!NC;
/g/./dz/./k/>/k/./ts/.!NC;
/g/./dz/./ts/>/k/./ts/.!NC;

4.3.10. Reguły modyfikacji statusu dźwięczności segmentu /dz'/ drugiego stopnia

Na podstawie analizy zamieszczonej w tabeli 2.45 w *Asymilacji…* można utworzyć podstawową regułę dwuelementową, która pozwala na zmianę statusu dźwięczności segmentu /dz'/ drugiego stopnia:

/dz'/.[WB2]>/ts'/.!NC;

Zgodnie z przyjętą konwencją w regułach nie używa się symboli zbiorów, jeżeli desonoryzacja obejmuje co najmniej dwa segmenty drugiego stopnia. Dotyczy to następujących reguł:

/dz'/./s/./t/./v/>/ts'/.!NC.!NC./f/;
/dz'/./s/./t/./v/./j/>/ts'/.!NC.!NC./f/.!NC;
/v/./dz'/./ts'/>/f/./ts'/.!NC;
/z'/./dz'/./ts/>/s'/./ts'/.!NC;
/z'/./dz'/./tS/>/s'/./ts'/.!NC;
/z'/./dz'/./ts'/>/s'/./ts'/.!NC;

4.3.11. Reguły modyfikacji statusu dźwięczności segmentu /rz/ drugiego stopnia

Zmiana statusu dźwięczności segmentu /rz/ drugiego stopnia może mieć związek zarówno ze zjawiskiem antycypacji artykulacyjnej, jak i ze zjawiskiem perseweracji artykulacyjnej. Zatem przy konstruowaniu reguł musi być brana pod uwagę zarówno struktura kontekstu następującego, jak i struktura kontekstu poprzedzającego. Na podstawie wyniku analizy, który został ujęty w tabeli 2.50 w *Asymilacji…* utworzono dwie podstawowe reguły dwuelementowe:

[WB2]./rz/>!NC./S/;
/rz/.[WB2]>/S/.!NC;

Pierwsza reguła ma związek ze zjawiskiem perseweracji artykulacyjnej, natomiast reguła druga odnosi się do zjawiska antycypacji artykulacyjnej. Poza tym istnieje stosunkowo liczny zbiór grup spółgłoskowych, w których zjawisko desonoryzacji obejmuje też inny segment lub inne segmenty drugiego stopnia (oprócz /rz/). Dla przetwarzania takich grup utworzono następujące reguły:

/v/./v/./s/./t/./rz/>/f/./f/.!NC.!NC./S/;
/v/./x/./rz/>/f/.!NC./S/;
/v/./k/./rz/>/f/.!NC./S/;

/v/./p/./rz/>/f/.!NC./S/;
/v/./t/./rz/>/f/.!NC./S/;
/v/./s/./k/./rz/>/f/.!NC.!NC./S/;
/v/./s/./t/./rz/>/f/.!NC.!NC./S/;
/j/./v/./s/./t/./rz/>!NC./f/.!NC.!NC./S/;
/z/./x/./rz/>/s/.!NC./S/;
/z/./k/./rz/>/s/.!NC./S/;
/z/./p/./rz/>/s/.!NC./S/;
/z/./t/./rz/>/s/.!NC./S/;
/d/./x/./rz/>/t/.!NC./S/;
/b/./t/./rz/>/p/.!NC./S/;
/d/./k/./rz/>/t/.!NC./S/;
/d/./p/./rz/>/t/.!NC./S/;
/d/./t/./rz/>/t/.!NC./S/;
/z/./s/./k/./rz/>/s/.!NC.!NC./S/;
/z/./s/./p/./rz/>/s/.!NC.!NC./S/;
/z/./s/./t/./rz/>/s/.!NC.!NC./S/;
/b/./s/./t/./rz/>/p/.!NC.!NC./S/;
/d/./s/./k/./rz/>/t/.!NC.!NC./S/;
/d/./s/./t/./rz/>/t/.!NC.!NC./S/;

Wszystkie wymienione reguły dotyczą zjawiska perseweracji artykulacyjnej segmentu /rz/. Utworzono następujące reguły dla grup, w których /rz/ ulega ubezdźwięcznieniu w wyniku antycypacji artykulacyjnej i w których zjawisko desonoryzacji obejmuje również inny segment drugiego stopnia:

/rz/./v/./ts'/>/S/./f/.!NC;
/rz/./k/./v/>/S/.!NC./f/;
/rz/./k/./v/./j/>/S/.!NC./f/.!NC;
/rz/./b/./ts/>/S/./p/.!NC;
/rz/./b/./ts'/>/S/./p/.!NC;
/v/./rz/./k/>/f/./S/.!NC;
/v/./rz/./c/>/f/./S/.!NC;
/v/./rz/./k/./j/>/f/./S/.!NC.!NC;
/n/./d/./rz/./ts'/>!NC./t/./S/.!NC;
/b/./rz/./ts'/>/p/./S/.!NC;
/d/./rz/./ts'/>/t/./S/.!NC;
/g/./rz/./ts/>/k/./S/.!NC;
/d/./k/./rz/./t/>/t/.!NC./S/.!NC;
/z/./d/./rz/./ts'/>/s/./t/./S/.!NC;
/z/./g/./rz/./ts'/>/s/./k/./S/.!NC;

Została utworzona oddzielna reguła dla przewarzania mało prawdopodobnej formy *zwichrzże*. W drugim członie tej reguły został przepisany segment /rz/. Zapobiega to zmianie jego statusu dźwięczności na podstawie reguły krótszej związanej ze zjawiskiem perseweracji artykulacyjnej:

/x/./rz/./Z/>/x/./rz/.!NC;

Kolejna lista obejmuje reguły dotyczące nagłosowych grup spółgłoskowych, w których segment /rz/ drugiego stopnia powinien być przekształcony w wariant bezdźwięczny (w wyniku zjawiska perseweracji artykulacyjnej). Poza tym te grupy zawierają jeszcze segment /v/ drugiego stopnia, który również powinien być ubezdźwięczniony. Jeżeli nagłosowe grupy spółgłoskowe nie zawierają drugiego segmentu umiejscowionego w kontekście ubezdźwięczniającym (poza /rz/), to dla ich przetwarzania mogą być użyte podstawowe reguły dwuelementowe: [WB2]./rz/>!NC./S/ lub /rz/.[WB2]>/S/.!NC. Oto ostatni wykaz reguł dotyczących zmiany statusu dźwięczności segmentu /rz/ drugiego stopnia:

#./v/./x/./rz/>!NC./f/.!NC./S/;
#./v/./k/./rz/>!NC./f/.!NC./S/;
#./v/./p/./rz/>!NC./f/.!NC./S/;
#./v/./t/./rz/>!NC./f/.!NC./S/;
#./v/./s/./t/./rz/>!NC./f/.!NC.!NC./S/;
#./v/./s/./k/./rz/>!NC./f/.!NC.!NC./S/;

4.3.12. Reguły modyfikacji statusu dźwięczności segmentu /ts/ drugiego stopnia

Opis zawarty w dalszym ciągu tego rozdziału dotyczy kontekstu udźwięczniającego dla bezdźwięcznych segmentów drugiego stopnia. Zostały przedstawione reguły odnoszące się do grup spółgłoskowych umiejscowionych w śródgłosie wyrazów. Ewentualna sonoryzacja segmentów drugiego stopnia może wynikać tylko i wyłącznie ze zjawiska antycypacji artykulacyjnej (nigdy ze zjawiska perseweracji artykulacyjnej). A więc kluczowe znaczenie ma struktura prawostronnego kontekstu poszczególnych segmentów.

Analizując dane zamieszczone w tabeli 3.2 w *Asymilacji…* można zauważyć, że kontekst mający wpływ na zmianę statusu dźwięczności segmentu /ts/ drugiego stopnia zawsze jest spółgłoską zwartą dźwięczną. Poza tym niemal we wszystkich wymienionych grupach sonoryzacja obejmuje tylko segment /ts/. Wystarczy zatem utworzyć jedną regułę:

/ts/.[WD2]>/dz/.!NC;

Występuje tylko jedna wewnątrzwyrazowa grupa spółgłoskowa, w której sonoryzacja obejmuje też drugi segment drugiego stopnia. Dla przetwarzania tej grupy utworzono następującą regułę:

/ts/./tS/./Z/>/dz/./dZ/.!NC;

Segment /ts/ drugiego stopnia ulega sonoryzacji w formach wyrazu *szwarcwaldzki* ze względu na strukturę morfologiczną. Utworzono regułę o następującej postaci:

/a/./r/./ts/./v/./a/>!NC.!NC./dz/./v/.!NC;

4.3.13. Reguły modyfikacji statusu dźwięczności segmentu /ts'/ drugiego stopnia

Reguły dotyczące kontekstu udźwięczniającego segmentu /ts'/ drugiego stopnia są oparte na tabeli 3.4 w *Asymilacji…* . Trzeba zaznaczyć, że część reguł ujętych w tej tabeli dotyczy tylko wymowy krakowskopoznańskiej, dlatego w niniejszym opisie nie zostały one użyte. Można utworzyć następującą regułę podstawową:

/ts'/.[WD2]>/dz'/.!NC;

W niektórych grupach po segmencie /ts'/ drugiego stopnia występuje segment /v/, który nie jest ubezdźwięczniany w wyniku perseweracji ze względu na strukturę morfologiczną wyrazu. Umiejscowienie junktury bezpośrednio przed nim (na przykład w wyrazach: ćwierć*wałek, ćwierćwiecze*) sprawia, że można odnieść się do zasad fonetyki międzywyrazowej przy ustalaniu statusu dźwięczności. W takich wypadkach segment /ts'/ powinien być przekształcany w dźwięczny segment /dz'/. Dla omawianych struktur wystarczy utworzyć jedną regułę:

/r/./ts'/./v/>!NC./dz'/./v/;

Jest to reguła trójelementowa; ma więc wyższy priorytet niż reguła dwuelementowa: [WB2]./v/>!NC./f/. Umieszczenie w drugim członie reguły zapisu: /v/ (zamiast polecenia !NC) sprawia natomiast, że ten segment nie może być zmodyfikowany na podstawie podanej reguły dwuelementowej (zob. podrozdz. 3.4.1). W wypadku kilku grup zjawisko sonoryzacji obejmuje dwa segmenty. Dla ich przetwarzania odpowiednie są następujące reguły:

/s'/./ts'/./Z/>/z'/./dz'/.!NC;
/p/./ts'/./Z/>/b/./dz'/.!NC;
/s'/./ts'/./dz'/>/z'/./dz'/.!NC;

4.3.14. Reguły modyfikacji statusu dźwięczności segmentu /f/ drugiego stopnia

Z informacji ujętych w tabeli 3.6 w *Asymilacji…* wynika, że minimalny kontekst udźwięczniający segmentu /f/ drugiego stopnia obejmuje najczęściej spółgłoskę zwartą dźwięczną umiejscowioną w bezpośrednim kontekście prawostronnym. Można zatem skonstruować następującą regułę uogólnioną:

/f/.[WD2]>/v/.!NC;

Trzeba utworzyć kilka reguł uzupełniających ze względu na strukturę wyrazów: *Wolfsburg* i *Softbank*:

/l/./f/./s/./b/>!NC./v/./z/.!NC;
/f/./t/./b/>/v/./d/.!NC;

Również w wyrazie *efwupowski* segment odpowiadający literze *w* nie podlega perseweracji artykulacyjnej ze względu na konstrukcję morfologiczną. W tym wypadku stanowi on kontekst udźwięczniający dla segmentu /f/ drugiego stopnia. Można to ująć w regule:

/e/./f/./v/>!NC./v/./v/;

Ze względu na długość podana reguła ma wyższy priorytet niż podstawowa reguła dwuelementowa dotycząca desonoryzacji segmentu /v/ drugiego stopnia (zob. podrozdz. 4.3.1).

4.3.15. Reguły modyfikacji statusu dźwięczności segmentu /k/ drugiego stopnia

Z tabeli 3.8 w *Asymilacji…* wynika, że dla segmentu /k/ drugiego stopnia minimalnym kontekstem udźwięczniającym jest spółgłoska właściwa dźwięczna. Reguła podstawowa ma następującą postać:

/k/.[WD2]>/g/.!NC;

W kilku wyrazach (np. *milkłby*) kontekst udźwięczniający jest poprzedzony półsamogłoską /w/. Oto reguła utworzona dla przetwarzania takich struktur:

/k/./w/./b/>/g/.!NC.!NC;

Sonoryzacja obejmuje również przedrostek *eks* przed spółgłoską właściwą dźwięczną. Odpowiednia reguła umożliwia zmianę statusu dźwięczności dwóch segmentów drugiego stopnia:

/k/./s/.[WD2]>/g/./z/.!NC;

Zjawisko sonoryzacji występuje również w wypadku grup spółgłoskowych umiejscowionych w następujących wyrazach: *dyskdżokej*, *iskrzże, powiększże, zakomplekśże, zmiękczże* oraz *ekswojskowy*. Oto reguły utworzone dla transkrypcji podanych grup:

/s/./k/./dZ/>/z/./g/.!NC;
/s/./k/./rz/./Z/>/z/./g/./rz/.!NC;
/k/./S/./Z/>/g/./Z/.!NC;
/k/./s'/./Z/>/g/./z'/.!NC;
/k/./tS/./Z/>/g/./dZ/.!NC;
/e/./k/./s/./v/>!NC./g/./z/./v/;

4.3.16. Reguły modyfikacji dźwięczności segmentu /p/ drugiego stopnia

Na podstawie wyniku analizy zamieszczonego w tabeli 3.10 w *Asymilacji…* można utworzyć następującą regułę podstawową pozwalającą na modyfikację statusu dźwięczności segmentu /p/ drugiego stopnia:

/p/.[WD2]>/b/.!NC;

W dwóch grupach pomiędzy segmentem /p/ a spółgłoską właściwą dźwięczną jest umiejscowiona półsamogłoska lub spółgłoska płynna. Dla takich grup utworzono następujące reguły:

/p/./w/./b/>/b/.!NC.!NC;
/p/./l/./Z/>/b/.!NC.!NC;

Cztery grupy zawierają dwa segmenty drugiego stopnia (łącznie z /p/), które są umiejscowione w kontekście udźwięczniającym. Oto reguły utworzone dla przetwarzania tych grup:

/p/./tS/./Z/>/b/./dZ/.!NC;
/p/./ts'/./Z/>/b/./dz'/.!NC;
/p/./S/./Z/>/b/./Z/.!NC;
/p/./s/./b/>/b/./z/.!NC;

Segment /p/ drugiego stopnia ulega sonoryzacji w formach wyrazu *topwanta ze* względu na strukturę morfologiczną. Utworzono następującą regułę:

/o/./p/./v/./a/>!NC./b/./v/.!NC;

4.3.17. Reguły modyfikacji statusu dźwięczności segmentu /s/ drugiego stopnia

Na podstawie informacji zawartych w tabeli 3.12 w *Asymilacji…* można utworzyć podstawową regułę przeznaczoną do modyfikacji statusu dźwięczności segmentu /s/ drugiego stopnia:

/s/.[WD2]>/z/.!NC;

W wypadku dwóch struktur spółgłoskowych kontekst udźwięczniający jest umiejscowiony za półsamogłoską, natomiast jedna z tych struktur obejmuje dwa segmenty drugiego stopnia, które ulegają sonoryzacji:

/s/./w/./b/>/z/.!NC.!NC;
/s/./t/./w/./b/>/z/./d/.!NC.!NC;

Modyfikacja statusu dźwięczności dotyczy również wyrazów złożonych zawierających w swojej strukturze prefiks *eks* lub prefiks *post*. Dla transkrypcji takich wyrazów utworzono reguły:

/k/./s/.[WD2]>/g/./z/.!NC;
/s/./t/.[WD2]>/z/./d/.!NC;

Reguła dotycząca przetwarzania prefiksu *eks* jest właściwa również dla złożenia *folksdojcz*. Reguła dotycząca prefiksu *post* jest odpowiednia również dla wyrazu *natomiastby*. Utworzono oddzielną regułę dla przetwarzania wyrazu *ekswojskowy* (zbiór WD2 nie obejmuje segmentu /v/ drugiego stopnia):

/e/./k/./s/./v/>!NC./g/./z/./v/;

W wypadku niektórych wewnątrzwyrazowych grup spółgłoskowych zjawisko sonoryzacji obejmuje dwa segmenty drugiego stopnia (łącznie z /s/). Utworzono dla nich następujące reguły:

/s/./k/./dZ/>/z/./g/.!NC;
/s/./k/./rz/./Z/>/z/./g/./rz/.!NC;
/r/./s/./t/./v/./Z/>!NC./z/./d/./v/.!NC;
/p/./s/./b/>/b/./z/.!NC;
/l/./f/./s/./b/>!NC./v/./z/.!NC;

4.3.18. Reguły modyfikacji statusu dźwięczności segmentu /s'/ drugiego stopnia

Analiza słownika *SJP.pl* (tabela 3.14 w *Asymilacji…*) wykazała, że w języku polskim istnieje niewielka liczba grup spółgłoskowych, w których segment /s'/ drugiego stopnia jest umiejscowiony w kontekście udźwięczniającym. Taka grupa występuje w wyrazie *prośba*. Można utworzyć regułę uogólnioną:

/s'/.[WD2]>/z'/.!NC;

W jednej grupie kontekst udźwięczniający jest poprzedzony spółgłoską płynną. Oto reguła utworzona dla niej:

/s'/./l/./Z/>/z'/.!NC.!NC;

W trzech grupach spółgłoskowych sonoryzacja obejmuje również inny segment drugiego stopnia:

/s'/./ts'/./Z/>/z'/./dz'/.!NC;
/k/./s'/./Z/>/g/./z'/.!NC;
/s'/./ts'/./dz'/>/z'/./dz'/.!NC;

4.3.19. Reguły modyfikacji statusu dźwięczności segmentu /t/ drugiego stopnia

Na podstawie danych zawartych w tabeli 3.16 w *Asymilacji…* można utworzyć podstawową regułę pozwalającą na modyfikację statusu dźwięczności segmentu /t/ drugiego stopnia umiejscowionego w kontekście udźwięczniającym:

/t/.[WD2]>/d/.!NC;

Dla przetwarzania wyrazów zawierających w swojej strukturze prefiks *kontr* lub prefiks *post* odpowiednie są następujące reguły:

/t/./r/.[WD2]>/d/.!NC.!NC;
/s/./t/.[WD2]>/z/./d/.!NC;

Druga wymieniona reguła umożliwia modyfikację dwóch segmentów bezdźwięcznych drugiego stopnia. Istnieje kilka innych grup zawierających dwa segmenty bezdźwięczne drugiego stopnia (łącznie z /t/) w kontekście udźwięczniającym. Dla ich przetwarzania utworzono reguły:

/f/./t/./b/>/v/./d/.!NC;
/t/./tS/./Z/>/d/./dZ/.!NC;
/r/./s/./t/./v/./Z/>!NC./z/./d/./v/.!NC;
/s/./t/./w/./b/>/z/./d/.!NC.!NC;

Utworzone oddzielne reguły dla przetwarzania wyrazów złożonych z członu *wewnątrz* oraz z wyrazu rozpoczynającego się od spółgłoski właściwej dźwięcznej (np.: *wewnątrzgrupowy, wewnątrzzakładowy*). Zamiana na fonem /dZ/ jest realizowana w warstwie modyfikacyjnej, natomiast poniższe reguły odnoszą się jedynie do ujednolicenia statusu dźwięczności sekwencji segmentów /t/ i /rz/ drugiego stopnia:

/t/./rz/.[WD2]>/d/./rz/.!NC;
/t/./rz/./v/>/d/./rz/./v/;

W tych regułach segmenty: /rz/ oraz /v/ zostały przepisane w drugim członie (zamiast użycia polecenia !NC). Dzięki temu nie mogą być zmieniane na podstawie reguł krótszych (mających związek ze zjawiskiem perseweracji artykulacyjnej). Pierwsza wymieniona reguła odnosi się również do formy *przypatrzże*.

4.3.20. Reguły modyfikacji statusu dźwięczności segmentu /S/ drugiego stopnia

Dla modyfikacji statusu dźwięczności segmentu /S/ drugiego stopnia utworzono regułę podstawową (zob. tabela 3.21 w *Asymilacji…*):

/S/.[WD2]>/Z/.!NC;

W wyrazach słownika *SJP.pl* występuje jedna wewnątrzwyrazowa grupa spółgłoskowa, w której kontekst udźwięczniający segmentu /S/ drugiego stopnia jest poprzedzony spółgłoską płynną. Oto reguła uwzględniająca strukturę tej grupy:

/S/./l/./Z/>/Z/.!NC.!NC;

W czterech grupach spółgłoskowych sonoryzacja obejmuje dwa segmenty drugiego stopnia. Utworzono następujące reguły związane ze zmianą statusu dźwięczności:

/k/./S/./Z/>/g/./Z/.!NC;
/p/./S/./Z/>/b/./Z/.!NC;
/S/./tS/./b/>/Z/./dZ/.!NC;
/S/./tS/./Z/>/Z/./dZ/.!NC;

4.3.21. Reguły modyfikacji statusu dźwięczności segmentu /tS/ drugiego stopnia

Utworzono podstawową regułę dla modyfikacji statusu dźwięczności segmentu /tS/ drugiego stopnia (zob. tabela 3.23 w *Asymilacji…*):

/tS/.[WD2]>/dZ/.!NC;

Odnosi się ona do następujących wewnątrzwyrazowych grup spółgłoskowych: /tS.Z/, /tS.b/, /tS.d/. Poza tym trzeba uwzględnić grupy, w których kontekst udźwięczniający obejmuje również drugi segment:

/S/./tS/./b/>/Z/./dZ/.!NC;
/p/./tS/./Z/>/b/./dZ/.!NC;
/S/./tS/./Z/>/Z/./dZ/.!NC;
/k/./tS/./Z/>/g/./dZ/.!NC;
/ts/./tS/./Z/>/dz/./dZ/.!NC;
/t/./tS/./Z/>/d/./dZ/.!NC;

Oddzielana reguła została utworzona dla przetwarzania wyrazu *doświadczże*:

/d/./tS/./Z/>/d/./dZ/.!NC;

Przepisanie segmentu /d/ drugiego stopnia w drugim członie reguły sprawia, że jego status dźwięczności nie może być zmieniany na podstawie reguły podstawowej.

4.3.22. Podsumowanie dotyczące reguł w warstwie statusu dźwięczności

Status dźwięczności segmentów drugiego stopnia pokrywa się z tzw. dźwięcznością ortograficzną znaków, diad oraz triad ortograficznych. Omówione reguły umożliwiają zmianę statusu dźwięczności segmentów drugiego stopnia w trakcie przetwarzania wyrazów w warstwie statusu dźwięczności. Niniejsze podsumowanie podzielono na dwie zasadnicze części. Pierwsza część dotyczy modyfikacji polegającej na zastępowaniu segmentów dźwięcznych drugiego stopnia ich bezdźwięcznymi odpowiednikami, część druga podsumowania dotyczy operacji analogicznej i odwrotnej. Informacje zawarte w częściach 4.3.1 do 4.3.21 odnoszą się tylko do nagłosu i śródgłosu wyrazów. Opis dotyczący modyfikacji statusu dźwięczności w wygłosie wyrazu znajduje się w części 4.3.23.

Najwięcej reguł włączonych do warstwy statusu dźwięczności dotyczy grup spółgłoskowych, w których modyfikacja obejmuje co najmniej dwa segmenty drugiego stopnia. Każda reguła tego typu była wymieniana w co najmniej dwóch opisach. Aby zredukować liczbę reguł oraz wykluczyć reguły powtarzające się, wykonano specjalne podsumowanie (tabela 4.3). Uwzględniono w nim wszystkie utworzone reguły związane z desonoryzacją przynajmniej dwóch segmentów (kolumna druga tabeli). W każdej komórce kolumny trzeciej wymienione zostały segmenty dźwięczne drugiego stopnia, których dotyczy dana reguła. W kolumnie czwartej podano reguły

alternatywne względem wymienionych w kolumnie drugiej. W większości wypadków jedną regułę wieloelementową można zastąpić dwiema podstawowymi regułami dwuelementowymi, które użyte razem umożliwiają wykonywanie takich samych przekształceń. Ponieważ wymienione w kolumnie czwartej reguły dwuelementowe zostały już zdefiniowane (oznaczono je etykietą: *zdefiniowane*), to odpowiednie reguły w kolumnie drugiej można wykluczyć z ostatecznego zbioru reguł należących do warstwy statusu dźwięczności. W niektórych regułach ujętych w kolumnie czwartej użyto mechanizmu list odwzorowań (zob. podrozdz. 5.2). Dzięki temu można było znacznie zredukować ostateczną liczbę reguł. Jeżeli w kolumnie czwartej nie ma żadnej informacji, to oznacza, że nie ma żadnej alternatywy dla odpowiedniej reguły zamieszczonej w kolumnie drugiej. Reguła ta nie może być usunięta z ostatecznego zbioru wszystkich reguł należących do drugiej warstwy projektu.

Tabela 4.3 – Podsumowanie dotyczące reguł związanych z desonoryzacją w grupach śródgłosowych

Lp.	Reguły ujęte w podrozdziałach 4.3.1–4.3.21	Segmenty drugiego stopnia	Reguły alternatywne
1	/b/./b/./s/>/p/./p/.!NC;	/b/, /b/	[WD2].[WD2].[WB2]>!LVL.!LVL.!NC;
2	/b/./x/./v/>/p/.!NC./f/;	/b/, /v/	/b/.[WB2]>/p/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
3	/b/./p/./s/./t/./v/./j/>/p/.!NC.!NC.!NC./f/.!NC;	/b/, /v/	/b/.[WB2]>/p/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
4	/b/./p/./s/./t/./v/>/p/.!NC.!NC.!NC./f/;	/b/, /v/	/b/.[WB2]>/p/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
5	/b/./rz/./ts'/>/p/./S/.!NC;	/b/, /rz/	[WD2]./rz/.[WB2]>!LVL./S/.!NC;
6	/b/./s/./t/./rz/>/p/.!NC.!NC./S/;	/b/, /rz/	/b/.[WB2]>/p/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
7	/b/./s/./t/./v/./j/>/p/.!NC.!NC./f/.!NC;	/b/, /v/	/b/.[WB2]>/p/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
8	/b/./s/./t/./v/>/p/.!NC.!NC./f/;	/b/, /v/	/b/.[WB2]>/p/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
9	/b/./t/./rz/>/p/.!NC./S/;	/b/, /rz/	/b/.[WB2]>/p/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
10	/d/./x/./v/>/t/.!NC./f/;	/d/, /v/	/d/.[WB2]>/t/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
11	/d/./k/./rz/./t>/t/.!NC./S/.!NC;	/d/, /rz/	/d/.[WB2]>/t/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
12	/d/./k/./rz/>/t/.!NC./S/;	/d/, /rz/	/d/.[WB2]>/t/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
13	/d/./k/./v/./j/>/t/.!NC./f/.!NC;	/d/, /v/	/d/.[WB2]>/t/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
14	/d/./k/./v/>/t/.!NC./f/;	/d/, /v/	/d/.[WB2]>/t/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
15	/d/./p/./rz/>/t/.!NC./S/;	/d/, /rz/	/d/.[WB2]>/t/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
16	/d/./rz/./ts'/>/t/./S/.!NC;	/d/, /rz/	[WD2]./rz/.[WB2]>!LVL./S/.!NC;
17	/d/./s/./k/./r/./v/./j/>/t/.!NC.!NC.!NC./f/.!NC;	/d/, /v/	
18	/d/./s/./k/./rz/>/t/.!NC.!NC./S/;	/d/, /rz/	/d/.[WB2]>/t/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
19	/d/./s/./t/./rz/>/t/.!NC.!NC./S/;	/d/, /rz/	/d/.[WB2]>/t/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
20	/d/./s'/./v/./j/>/t/.!NC./f/.!NC;	/d/, /v/	/d/.[WB2]>/t/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
21	/d/./s'/./v/>/t/.!NC./f/;	/d/, /v/	/d/.[WB2]>/t/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)

Lp.	Reguły ujęte w podrozdziałach 4.3.1–4.3.21	Segmenty drugiego stopnia	Reguły alternatywne
22	/d/./t/./rz/>/t/.!NC./S/;	/d/, /rz/	/d/.[WB2]>/t/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
23	/d/./t/./v/>/t/.!NC./f/;	/d/, /v/	/d/.[WB2]>/t/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
24	/d/./tS/./v/>/t/.!NC./f/;	/d/, /v/	/d/.[WB2]>/t/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
25	/d/./v/./s/./t/>/t/./f/.!NC.!NC;	/d/, /v/	[WD2]./v/.[WB2]>!LVL./f/.!NC;
26	/d/./v/./S/>/t/./f/.!NC;	/d/, /v/	[WD2]./v/.[WB2]>!LVL./f/.!NC;
27	/d/./v/./s'/>/t/./f/.!NC;	/d/, /v/	[WD2]./v/.[WB2]>!LVL./f/.!NC;
28	/d/./v/./tS/>/t/./f/.!NC;	/d/, /v/	[WD2]./v/.[WB2]>!LVL./f/.!NC;
29	/d/./x/./rz/>/t/.!NC./S/;	/d/, /rz/	/d/.[WB2]>/t/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
30	/dz'/./s/./t/./v/./j/>/ts'/.!NC.!NC./f/.!NC;	/dz'/, /v/	/dz'/.[WB2]>/ts'/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
31	/dz'/./s/./t/./v/>/ts'/.!NC.!NC./f/;	/dz'/, /v/	/dz'/.[WB2]>/ ts'/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
32	/dz/./t/./v/./j/>/ts/.!NC./f/.!NC;	/dz/, /v/	/dz/.[WB2]>/ts/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
33	/dz/./t/./v/>/ts/.!NC./f/;	/dz/, /v/	/dz/.[WB2]>/ts/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
34	/g/./dz/./k/>/k/./ts/.!NC;	/dz/, /g/	[WD2].[WD2].[WB2]>!LVL.!LVL.!NC;
35	/g/./dz/./ts/>/k/./ts/.!NC;	/dz/, /g/	[WD2].[WD2].[WB2]>!LVL.!LVL.!NC;
36	/g/./rz/./ts/>/k/./S/.!NC;	/g/, /rz/	[WD2]./rz/.[WB2]>!LVL./S/.!NC;
37	/j/./v/./s/./t/./rz/>!NC./f/.!NC.!NC./S/;	/v/, /rz/	/v/.[WB2]>/f/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
38	/n/./d/./rz/./ts'/>!NC./t/./S/.!NC;	/d/, /rz/	[WD2]./rz/.[WB2]>!LVL./S/.!NC;
39	/rz/./b/./ts/>/S/./p/.!NC;	/b/, /rz/	/rz/.[WD2].[WB2]>/S/.!LVL.!NC;
40	/rz/./b/./ts'/>/S/./p/.!NC;	/b/, /rz/	/rz/.[WD2].[WB2]>/S/.!LVL.!NC;
41	/rz/./k/./v/./j/>/S/.!NC./f/.!NC;	/v/, /rz/	/rz/.[WB2]>/S/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
42	/rz/./k/./v/>/S/.!NC./f/;	/v/, /rz/	/rz/.[WB2]>/S/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
43	/rz/./v/./ts'/>/S/./f/.!NC;	/v/, /rz/	/rz/./v/.[WB2]>/S/./f/.!NC;
44	/v/./dz'/./ts'/>/f/./ts'/.!NC;	/dz'/, /v/	/v/.[WD2].[WB2]>/f/.!LVL.!NC;
45	/v/./x/./rz/>/f/.!NC./S/;	/rz/, /v/	/v/.[WB2]>/f/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
46	/v/./k/./rz/>/f/.!NC./S/;	/v/, /rz/	/v/.[WB2]>/f/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
47	/v/./p/./rz/>/f/.!NC./S/;	/rz/, /v/	/v/.[WB2]>/f/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
48	/v/./rz/./k/./j/>/f/./S/.!NC.!NC;	/v/, /rz/	/v/./rz/.[WB2]>/f/./S/.!NC;
49	/v/./rz/./k/>/f/./S/.!NC;	/v/, /rz/	/v/./rz/.[WB2]>/f/./S/.!NC;
50	/v/./rz/./c/>/f/./S/.!NC;	/v/, /rz/	/v/./rz/.[WB2]>/f/./S/.!NC;
51	/v/./s/./k/./rz/>/f/.!NC.!NC./S/;	/v/, /rz/	/v/.[WB2]>/f/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
52	/v/./s/./t/./rz/>/f/.!NC.!NC./S/;	/v/, /rz/	/v/.[WB2]>/f/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
53	/v/./t/./rz/>/f/.!NC./S/;	/v/, /rz/	/v/.[WB2]>/f/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
54	/v/./v/./s/./t/./rz/>/f/./f/.!NC.!NC./S/;	/v/, /rz/	
55	/v/./v/./s/./k/>/f/./f/.!NC.!NC;	/v/,/v/	

Lp.	Reguły ujęte w podrozdziałach 4.3.1–4.3.21	Segmenty drugiego stopnia	Reguły alternatywne
56	/z'/./b/./ts/>/s'/./p/.!NC;	/b/, /z'/	[WD2].[WD2].[WB2]>!LVL.!LVL.!NC;
57	/z'/./b/./ts'/>/s'/./p/.!NC;	/b/, /z'/	[WD2].[WD2].[WB2]>!LVL.!LVL.!NC;
58	/z/./d/./k/>/s/./t/.!NC;	/d/, /z/	[WD2].[WD2].[WB2]>!LVL.!LVL.!NC;
59	/z/./d/./r/./s/./k/>/s/./t/.!NC.!NC.!NC;	/d/, /z/	
60	/z/./d/./r/./s/./ts/>/s/./t/.!NC.!NC.!NC;	/d/, /z/	
61	/z/./d/./r/./ts/>/s/./t/.!NC.!NC;	/d/, /z/	
62	/z/./d/./rz/./ts'/>/s/./t/./S/.!NC;	/d/, /rz/, /z/	
63	/z/./d/./ts/>/s/./t/.!NC;	/d/, /z/	[WD2].[WD2].[WB2]>!LVL.!LVL.!NC;
64	/z/./dz/./k/>/s/./ts/.!NC;	/dz/, /z/	[WD2].[WD2].[WB2]>!LVL.!LVL.!NC;
65	/Z/./dZ/./k/>/S/./tS/.!NC;	/dZ/, /Z/	[WD2].[WD2].[WB2]>!LVL.!LVL.!NC;
66	/Z/./dZ/./ts/>/S/./tS/.!NC;	/dZ/, /Z/	[WD2].[WD2].[WB2]>!LVL.!LVL.!NC;
67	/z/./dz/./ts/>/s/./ts/.!NC;	/dz/, /z/	[WD2].[WD2].[WB2]>!LVL.!LVL.!NC;
68	/Z/./dZ/./ts'/>/S/./tS/.!NC;	/dZ/, /Z/	[WD2].[WD2].[WB2]>!LVL.!LVL.!NC;
69	/z'/./dz'/./ts/>/s'/./ts'/.!NC;	/dz'/, /z'/	[WD2].[WD2].[WB2]>!LVL.!LVL.!NC;
70	/z'/./dz'/./tS/>/s'/./ts'/.!NC;	/dz'/, /z'/	[WD2].[WD2].[WB2]>!LVL.!LVL.!NC;
71	/z'/./dz'/./ts'/>/s'/./ts'/.!NC;	/dz'/, /z'/	[WD2].[WD2].[WB2]>!LVL.!LVL.!NC;
72	/z/./g/./rz/./ts'/>/s/./k/./S/.!NC;	/g/, /rz/, /z/	
73	/Z/./g/./w/./S/>/S/./k/.!NC.!NC;	/g/, /Z/	
74	/z/./x/./rz/>/s/.!NC./S/;	/rz/, /z/	/z/.[WB2]>/s/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
75	/z/./k/./r/./v/>/s/.!NC.!NC./f/;	/v/, /z/	
76	/z/./k/./rz/>/s/.!NC./S/;	/rz/, /z/	/z/.[WB2]>/s/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
77	/z/./k/./v/./j/>/s/.!NC./f/.!NC;	/v/, /z/	/z/.[WB2]>/s/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
78	/z/./k/./v/>/s/.!NC./f/;	/v/, /z/	/z/.[WB2]>/s/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
79	/z/./p/./rz/>/s/.!NC./S/;	/rz/, /z/	/z/.[WB2]>/s/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
80	/z/./s/./k/./rz/>/s/.!NC.!NC./S/;	/rz/, /z/	/z/.[WB2]>/s/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
81	/z/./s/./p/./rz/>/s/.!NC.!NC./S/;	/rz/, /z/	/z/.[WB2]>/s/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
82	/z/./s/./t/./rz/>/s/.!NC.!NC./S/;	/rz/, /z/	/z/.[WB2]>/s/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
83	/z/./s'/./v/./j/>/s/.!NC./f/.!NC;	/v/, /z/	/z/.[WB2]>/s/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
84	/z/./s/./v/>/s/.!NC./f/;	/v/, /z/	/z/.[WB2]>/s/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
85	/z/./S/./v/>/s/.!NC./f/;	/v/, /z/	/z/.[WB2]>/s/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
86	/z/./t/./r/./v/>/s/.!NC.!NC./f/;	/v/, /z/	
87	/z/./t/./rz/>/s/.!NC./S/;	/rz/, /z/	/z/.[WB2]>/s/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
88	/z/./t/./v/./j/>/s/.!NC./f/.!NC;	/v/, /z/	/z/.[WB2]>/s/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
89	/z/./t/./v/>/s/.!NC./f/;	/v/, /z/	/z/.[WB2]>/s/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
90	/z/./ts'/./v/./j/>/s/.!NC./f/.!NC;	/v/, /z/	/z/.[WB2]>/s/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
91	/z/./v/./s/./t/>/s/./f/.!NC.!NC;	/v/, /z/	[WD2]./v/.[WB2]>!LVL./f/.!NC;

Lp.	Reguły ujęte w podrozdziałach 4.3.1–4.3.21	Segmenty drugiego stopnia	Reguły alternatywne
92	/z/./v/./s'/./ts'/>/s/./f/.!NC.!NC;	/v/, /z/	[WD2]./v/.[WB2]>!LVL./f/.!NC;
93	/z'/./v/./ts/>/s'/./f/.!NC;	/v/, /z'/	[WD2]./v/.[WB2]>!LVL./f/.!NC;
94	/z'/./v/./ts'/>/s'/./f/.!NC;	/v/, /z'/	[WD2]./v/.[WB2]>!LVL./f/.!NC;
95	/z/./x/./v/./j/>/s/.!NC./f/.!NC;	/v/, /z/	/z/.[WB2]>/s/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)
96	/z/./x/./v/>/s/.!NC./f/;	/v/, /z/	/z/.[WB2]>/s/.!NC; (zdefiniowana) [WB2]./v/>!NC./f/; (zdefiniowana)

Wykonano oddzielne podsumowanie reguł związanych z desonoryzacją w nagłosowych grupach spółgłoskowych. Rezultat tego podsumowania został ujęty w tabeli 4.4. Z informacji w niej zamieszczonych wynika, że wszystkie reguły przetwarzania nagłosowych grup spółgłoskowych zawierają co najmniej dwa segmenty drugiego stopnia, które w procesie transkrypcji powinny być zastąpione ich bezdźwięcznymi odpowiednikami. Sześć reguł w kolumnie drugiej można zastąpić parą reguł dwuelementowych, które zostały zdefiniowane dla przetwarzania grup śródgłosowych. Regułę zapisaną w wierszu siódmym można zastąpić regułą zdefiniowaną w tabeli 4.3. Wszystkie reguły zamieszczone w kolumnie drugiej w tabeli 4.4 można usunąć z ostatecznego zbioru reguł należących do warstwy statusu dźwięczności. To samo dotyczy reguł umożliwiających przetwarzanie nagłosowych grup zawierających tylko jeden segment dźwięczny drugiego stopnia, którego status dźwięczności musi być zmieniony.

Tabela 4.4 – Podsumowanie dotyczące reguł związanych z desonoryzacją w grupach nagłosowych

Lp.	Reguły ujęte w podrozdziałach 4.3.1–4.3.21	Segmenty drugiego stopnia	Reguły alternatywne
1	#./v/./x/./rz/>!NC./f/.!NC./S/;	/v/, /rz/	/v/.[WB2]>/f/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
2	#./v/./k/./rz/>!NC./f/.!NC./S/;	/v/, /rz/	/v/.[WB2]>/f/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
3	#./v/./p/./rz/>!NC./f/.!NC./S/;	/v/, /rz/	/v/.[WB2]>/f/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
4	#./v/./t/./rz/>!NC./f/.!NC./S/;	/v/, /rz/	/v/.[WB2]>/f/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
5	#./v/./s/./k/./rz/>!NC./f/.!NC.!NC./S/;	/v/, /rz/	/v/.[WB2]>/f/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
6	#./v/./s/./t/./rz/>!NC./f/.!NC.!NC./S/;	/v/, /rz/	/v/.[WB2]>/f/.!NC; (zdefiniowana) [WB2]./rz/>!NC./S/; (zdefiniowana)
7	#./v/./z/./S/>!NC./f/./s/.!NC;	/v/, /z/	/v/.[WD2].[WB2]>/f/.!LVL.!NC; (zdefiniowana w tabeli 4.3)

Na podstawie informacji przedstawionych w tej części można wyznaczyć ostateczny zbiór reguł dla warstwy statusu dźwięczności. Reguły te dotyczą modyfikacji związanych ze zjawiskiem desonoryzacji w nagłosowych oraz śródgłosowych grupach spółgłoskowych:

[WB2]./v/>!NC./f/;
/v/.[WB2]>/f/.!NC;
/b/.[WB2]>/p/.!NC;
/b/.[SO2].[WB2]>/p/.!NC.!NC;
/g/.[WB2]>/k/.!NC;
/g/./g/./s/>/k/./k/.!NC;

/g/./g/./s'/>/k/./k/.!NC;
/g/.[SO2].[WB2]>/k/.!NC.!NC;
/x/./rz/./Z/>/x/./rz/.!NC;
/z'/.[WB2]>/s'/.!NC;
/z'/.[SO2].[WB2]>/s'/.!NC.!NC;
/d/.[WB2]>/t/.!NC;
/d/.[SO2].[WB2]>/t/.!NC.!NC;
/k/./r/./v/>!NC.!NC./f/;
/p/./l/./v/>!NC.!NC./f/;
/t/./r/./v/>!NC.!NC./f/;
/S/./e/./x/./v/./w/>!NC.!NC.!NC./v/.!NC;
/v/./r/./s/>/f/.!NC.!NC;
/z/.[WB2]>/s/.!NC;
/z/.[SO2].[WB2]>/s/.!NC.!NC;
/Z/.[WB2]>/S/.!NC;
/Z/./n'/./ts'/>/S/.!NC.!NC;
/dZ/.[WB2]>/tS/.!NC;
/dz/.[WB2]>/ts/.!NC;
/dz'/.[WB2]>/ts'/.!NC;
[WB2]./rz/>!NC./S/;
/rz/.[WB2]>/S/.!NC;
[WD2]./rz/.[WB2]>!LVL./S/.!NC;
[WD2]./v/.[WB2]>!LVL./f/.!NC;
[WD2].[WD2].[WB2]>!LVL.!LVL.!NC;
/rz/.[WD2].[WB2]>/S/.!LVL.!NC;
/rz/./v/.[WB2]>/S/./f/.!NC;
/v/.[WD2].[WB2]>/f/.!LVL.!NC;
/d/./s/./k/./r/./v/./j/>/t/.!NC.!NC.!NC./f/.!NC;
/v/./rz/.[WB2]>/f/./S/.!NC;
/v/./v/./s/./t/./rz/>/f/./f/.!NC.!NC./S/;
/v/./v/./s/./k/>/f/./f/.!NC.!NC;
/z/./d/./r/./s/./k/>/s/./t/.!NC.!NC.!NC;
/z/./d/./r/./s/./ts/>/s/./t/.!NC.!NC.!NC;
/z/./d/./r/./ts/>/s/./t/.!NC.!NC;
/z/./d/./rz/./ts'/>/s/./t/./S/.!NC;
/z/./g/./rz/./ts'/>/s/./k/./S/.!NC;
/Z/./g/./w/./S/>/S/./k/.!NC.!NC;
/z/./k/./r/./v/>/s/.!NC.!NC./f/;
/z/./t/./r/./v/>/s/.!NC.!NC./f/;

Znaczna liczba reguł związanych ze zjawiskiem sonoryzacji umożliwia modyfikację statusu dźwięczności dwóch segmentów drugiego stopnia. Reguły takie były powtarzane w opisach (zob. podrozdz. 4.3.12 do 4.3.21). Zostały one wymienione w kolumnie drugiej tabeli 4.5. W kolumnie trzeciej podano segmenty drugiego stopnia, których dotyczą te reguły. W kolumnie czwartej zawarto reguły alternatywne dla większości reguł z kolumny drugiej. Został w nich użyty mechanizm list odwzorowań (zob. podrozdz. 5.2). W ten sposób można zredukować zbiór wszystkich reguł z kolumny drugiej.

Tabela 4.5 – Podsumowanie dotyczące reguł związanych z sonoryzacją w grupach śródgłosowych

Lp.	Reguły omówione	Segmenty drugiego stopnia	Reguły alternatywne
1	/d/./tS/./Z/>/d/./dZ/.!NC;	/d/, /tS/	
2	/f/./t/./b/>/v/./d/.!NC;	/f/, /t/	[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;
3	/k/./S/./Z/>/g/./Z/.!NC;	/k/, /S/	[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;
4	/k/./s'/./Z/>/g/./z'/.!NC;	/k/, /s'/	[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;
5	/k/./s/.[WD2]>/g/./z/.!NC;	/k/, /s/	[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;
6	/e/./k/./s/./v/>NC!./g/./z/./v/;	/k/, /s/	
7	/k/./tS/./Z/>/g/./dZ/.!NC;	/k/, /tS/	[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;
8	/l/./f/./s/./b/>!NC./v/./z/.!NC;	/f/, /s/	[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;
9	/p/./s/./b/>/b/./z/.!NC;	/p/, /s/	[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;
10	/p/./S/./Z/>/b/./Z/.!NC;	/p/, /S/	[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;
11	/p/./tS/./Z/>/b/./dZ/.!NC;	/p/, /tS/	[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;
12	/p/./ts'/./Z/>/b/./dz'/.!NC;	/p/, /ts'/	[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;
13	/r/./s/./t/./v/./Z/>!NC./z/./d/./v/.!NC;	/s/, /t/	
14	/s'/./ts'/./Z/>/z'/./dz'/.!NC;	/s'/, /ts'/	[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;
15	/s/./k/./dZ/>/z/./g/.!NC;	/k/, /s/	[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;
16	/s/./k/./rz/./Z/>/z/./g/./rz/.!NC;	/s/, /k/	
17	/s/./t/.[WD2]>/z/./d/.!NC;	/s/, /t/	[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;
18	/s/./t/./w/./b/>/z/./d/.!NC.!NC;	/s/, /t/	
19	/S/./tS/./b/>/Z/./dZ/.!NC;	/S/, /tS/	[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;
20	/s'/./ts'/./dz'/>/z'/./dz'/.!NC;	/s'/, /ts'/	[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;
21	/S/./tS/./Z/>/Z/./dZ/.!NC;	/S/, /tS/	[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;
22	/t/./tS/./Z/>/d/./dZ/.!NC;	/tS/, /t/	[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;
23	/ts/./tS/./Z/>/dz/./dZ/.!NC;	/tS/, /ts/	[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;

Po wykonaniu redukcji można utworzyć ostateczną listę, która obejmuje reguły odnoszące się do zjawiska sonoryzacji w nagłosowych i śródgłosowych grupach spółgłoskowych:

/ts/.[WD2]>/dz/.!NC;
/ts'/.[WD2]>/dz'/.!NC;
/r/./ts'/./v/>!NC./dz'/./v/;
/f/.[WD2]>/v/.!NC;
/e/./f/./v/>!NC./v/./v/;
/k/.[WD2]>/g/.!NC;
/k/./w/./b/>/g/.!NC.!NC;
/p/.[WD2]>/b/.!NC;
/p/./w/./b/>/b/.!NC.!NC;
/p/./l/./Z/>/b/.!NC.!NC;
/r/./s/./t/./v/./Z/>!NC./z/./d/./v/.!NC;
/s/.[WD2]>/z/.!NC;
/s/./w/./b/>/z/.!NC.!NC;
/s/./t/./w/./b/>/z/./d/.!NC.!NC;
/s'/.[WD2]>/z'/.!NC;
/s'/./l/./Z/>/z'/.!NC.!NC;
/t/.[WD2]>/d/.!NC;
/t/./r/.[WD2]>/d/.!NC.!NC;
/S/.[WD2]>/Z/.!NC;
/S/./l/./Z/>/Z/.!NC.!NC;

/tS/.[WD2]>/dZ/.!NC;
/d/./tS/./Z/>/d/./dZ/.!NC;
e/./k/./s/./v/>NC!./g/./z/./v /;
/o/./p/./v/./a/>!NC./b/./v/.!NC;
/s/./k/./rz/./Z/>/z/./g/./rz/.!NC;
/t/./rz/.[WD2]>/d/./rz/.!NC;
/t/./rz/./v/>/d/./rz/./v/;
/a/./r/./ts/./v/./a/>!NC.!NC./dz/./v/.!NC;
[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;

4.3.23. Reguły modyfikacji statusu dźwięczności w wygłosie wyrazu

Informacje przedstawione w częściach 4.3.1 do 4.3.22 dotyczą grup spółgłoskowych umiejscowionych w nagłosie lub w śródgłosie wyrazu. W tej części zostały omówione reguły, które umożliwiają modyfikację statusu dźwięczności w wygłosowych grupach spółgłoskowych zapisanych w transkrypcji drugiego stopnia. Ewentualna decyzja o tego typu modyfikacjach jest podejmowana na podstawie etykiety uzyskanej z przetwarzania następnego wyrazu (jest on przetwarzany w pierwszej kolejności). Na potrzeby omawianego projektu ustalono następujące etykiety: UBDZ oraz UDZ. Oznaczają one: międzywyrazowy kontekst ubezdźwięczniający oraz międzywyrazowy kontekst udźwięczniający (zob. podrozdz. 5.6). Odpowiednia etykieta zostaje dołączona do wyrazu w trakcie jego przetwarzania w konkretnej warstwie. Stanowi ona oddzielny segment (za segmentem oznaczającym granicę wyrazu). Może być używana w regułach, jednak reguły nie mogą modyfikować tej etykiety (w drugim członie reguły do elementu jej odpowiadającego należy przypisać polecenie !NC).

Elementarne reguły umożliwiające modyfikację statusu dźwięczności pojedynczych segmentów spółgłoskowych drugiego stopnia, które są umiejscowione w wygłosie (na przykład w wyrazach: /k.o.t/, /p.o.d/), mają następującą postać:

[WB2].#.UDZ>!LVO.!NC.!NC;
[WD2]+/rz/+/v/.#.UBDZ>!LVL.!NC.!NC;

Pierwsza wymieniona reguła dotyczy sonoryzacji segmentu ze zbioru WB2 umiejscowionego w wygłosie wyrazu. Reguła druga dotyczy ubezdźwięcznienia segmentu ze zbioru WD2 (powiększonego o elementy: /rz/ oraz /v/). W obu wypadkach zastosowano mechanizm list odwzorowań. Polecenie !LVO powoduje przekształcenie dowolnego segmentu drugiego stopnia będącego spółgłoską właściwą bezdźwięczną na jego dźwięczny odpowiednik. Działanie polecenia !LVL jest analogiczne i odwrotne (definicja obu poleceń znajduje się w części 5.2). Trzeba zaznaczyć, że te reguły są trójelementowe, więc będą przesłaniane przez reguły dłuższe – dzięki temu można konstruować reguły dla grup spółgłoskowych umiejscowionych w wygłosie wyrazu.

W dalszym ciągu zostały omówione reguły dotyczące trzech kategorii grup spółgłoskowych umiejscowionych w wygłosie wyrazu. Zostały one ujęte w podrozdziale 5.2 w *Asymilacji…* (kategoria czwarta obejmuje grupy złożone wyłącznie ze spółgłosek sonornych, dlatego nie jest tutaj omawiana). W drugim członie każdej wymienionej reguły są przepisywane również segmenty drugiego stopnia, które nie są zmieniane. Ten zabieg zapewnia, że struktura wygłosowych grup nie może być modyfikowana na podstawie reguł krótszych ustalonych dla grup śródgłosowych (w rzeczywistości ten problem dotyczy tylko niektórych grup wygłosowych).

Grupy spółgłoskowe zaliczone do pierwszej kategorii zawierają co najmniej jedną spółgłoskę właściwą bezdźwięczną i nie zawierają żadnych spółgłosek właściwych dźwięcznych (w transkrypcji drugiego stopnia). Grupy takie mogą ulegać sonoryzacji pod warunkiem, że do przetwarzanego wyrazu zostanie dołączony segment zawierający etykietę UDZ. Na podstawie analizy przedstawionej w tabelach 5.1 i 5.2 w *Asymilacji…* utworzono listę reguł, które umożliwiają zastępowanie segmentów bezdźwięcznych drugiego stopnia ich dźwięcznymi odpowiednikami. Został w nich użyty mechanizm list odwzorowań. Dla czterech wygłosowych grup spółgłoskowych utworzono reguły bez użycia symboli zbiorów – ze względu na złożoną i specyficzną strukturę tych grup:

[SO2].[WB2].#.UDZ>!NC.!LVO.!NC.!NC;
[SO2].[WB2].[SO2].#.UDZ>!NC.!LVO.!NC.!NC.!NC;

[SO2].[WB2].[WB2].#.UDZ>!NC.!LVO.!LVO.!NC.!NC;
[WB2].[SO2].#.UDZ>!LVO.!NC.!NC.!NC;
[WB2].[WB2].#.UDZ>!LVO.!LVO.!NC.!NC;
[WB2].[WB2].[SO2].#.UDZ>!LVO.!LVO.!NC.!NC.!NC;
/r/./n/./s/./t/.#.UDZ>/r/./n/./z/./d/.!NC.!NC;
/j/./s/./t/./r/.#.UDZ>/j/./z/./d/./r/.!NC.!NC;
/p/./s/./k/.#.UDZ>/b/./z/./g/.!NC.!NC;
/k/./s/./t/./r/.#.UDZ>/g/./z/./d/./r/.!NC.!NC;

Druga kategoria (w opisie zawartym w części 5.2 w *Asymilacji…*) obejmuje grupy umiejscowione w wygłosie, które zawierają co najmniej jedną spółgłoskę właściwą dźwięczną i niezawierające żadnej spółgłoski właściwej bezdźwięcznej (w transkrypcji drugiego stopnia). Grupy te ulegają desonoryzacji, jeżeli do przetwarzanego wyrazu zostaje dołączona etykieta UBDZ. Na podstawie analizy, której wyniki przedstawiono w tabelach 5.3 i 5.4 w *Asymilacji…*, utworzono następujące reguły dotyczące omawianej kategorii grup wygłosowych:

[SO2]./v/.#.UBDZ>!NC./f/.!NC.!NC;
[SO2].[WD2]+/rz/+/v/.#.UBDZ>!NC.!LVL.!NC.!NC;
[SO2].[WD2]+/rz/+/v/.[SO2].#.UBDZ>!NC.!LVL.!NC.!NC.!NC;
[WD2]+/rz/+/v/./v/.#.UBDZ>!LVL./f/.!NC.!NC;
[WD2]+/rz/+/v/.[SO2].#.UBDZ>!LVL.!NC.!NC.!NC;
[WD2] +/rz/+/v/.[WD2] +/rz/+/v/.#.UBDZ>!LVL.!LVL.!NC.!NC;
/rz/./v/.#.UBDZ>/S/./f/.!NC.!NC;
/rz/./m/.#.UBDZ>/S/./m/.!NC.!NC;
/rz/./w/.#.UBDZ>/S/./w/.!NC.!NC;
/rz/./b/.#.UBDZ>/S/./p/.!NC.!NC;
/rz/./g/.#.UBDZ>/S/./k/.!NC.!NC;
/v/./n'/.#.UBDZ>/f/./n'/.!NC.!NC;
/v/./r/.#.UBDZ>/f/./r/.!NC.!NC;
/v/./d/.#.UBDZ>/f/./t/.!NC.!NC;
/v/./dz'/.#.UBDZ>/f/./ts'/.!NC.!NC;
/m/./rz/.#.UBDZ>/m/./S/.!NC.!NC;
/j/./rz/.#.UBDZ>/j/./S/.!NC.!NC;
/n/./g/./w/.#.UBDZ>/n/./k/./w/.!NC.!NC;
/r/./d/./v/.#.UBDZ>/r/./t/./f/.!NC.!NC;
/n/./d/./rz/.#.UBDZ>/n/./t/./S/.!NC.!NC;
/m/./b/./d/.#.UBDZ>/m/./p/./t/.!NC.!NC;
/b/./rz/.#.UBDZ>/p/./S/.!NC.!NC;
/d/./rz/.#.UBDZ>/t/./S/.!NC.!NC;
/d/./r/./n'/.#.UBDZ>/t/./r/./n'/.!NC.!NC;
/z/./d/./rz/.#.UBDZ>/s/./t/./S/.!NC.!NC;
/z/./g/./rz/.#.UBDZ>/s/./k/./S/.!NC.!NC;
/z/./d/./r/.#.UBDZ>/s/./t/./r/.!NC.!NC;

Tylko sześć reguł na tej liście zawiera symbole zbiorów. Został w nich użyty mechanizm list odwzorowań. Pozostałe reguły dotyczą konkretnych grup (mają one zróżnicowaną strukturę, zatem tworzenie reguł uogólnionych nie jest uzasadnione).

Kolejna omawiana kategoria to tzw. grupy mieszane. Zapisane w transkrypcji podstawowej zawierają co najmniej jedną spółgłoskę właściwą dźwięczną oraz co najmniej jedną spółgłoskę właściwą bezdźwięczną. Z informacji zamieszczonych w podrozdziale 5.2 w *Asymilacji…* wynika, że istnieją dwie możliwości modyfikacji statusu dźwięczności w takich grupach – są one związane z obecnością międzywyrazowego kontekstu ubezdźwięczniającego lub międzywyrazowego kontekstu udźwięczniającego. Ze względu na zróżnicowaną strukturę oraz stosunkowo niewielką liczbę grup należących do omawianej kategorii w regułach nie zastosowano uogólnień struktury (poprzez użycie symboli zdefiniowanych zbiorów). Oto pierwszy wykaz omawianych reguł (związanych z kontekstem ubezdźwięczniającym):

/b/./ts'/.#.UBDZ>/p/./ts'/.!NC.!NC;
/b/./p/./s/./t/./v/.#.UBDZ>/p/./p/./s/./t/./f/.!NC.!NC;
/b/./r/./tS/.#.UBDZ>/p/./r/./tS/.!NC.!NC;
/b/./s/.#.UBDZ>/p/./s/.!NC.!NC;
/b/./s/./k/.#.UBDZ>/p/./s/./k/.!NC.!NC;
/b/./s/./t/./v/.#.UBDZ >/p/./s/./t/./f/.!NC.!NC;
/x/./d/.#.UBDZ>/x/./t/.!NC.!NC;
/x/./rz/.#.UBDZ>/x/./S/.!NC.!NC;
/x/./rz/./t/.#.UBDZ>/x/./S/./t/.!NC.!NC;
/x/./v/.#.UBDZ>/x/./f/.!NC.!NC;
/ts/./t/./v/.#.UBDZ>/ts/./t/./f/.!NC.!NC;
/tS/./b/.#.UBDZ>/tS/./p/.!NC.!NC;
/d/./tS/.#.UBDZ>/t/./tS/.!NC.!NC;
/d/./s/./k/.#.UBDZ>/t/./s/./k/.!NC.!NC;
/d/./s'/.#.UBDZ>/t/./s'/.!NC.!NC;
/dz/./k/.#.UBDZ>/ts/./k/.!NC.!NC;
/dz/./t/./v/.#.UBDZ>/ts/./t/./f/.!NC.!NC;
/dz'/./s/./t/./v/.#.UBDZ>/ts'/./s/./t/./f/.!NC.!NC;
/g/./f/.#.UBDZ>/k/./f/.!NC.!NC;
/j/./s/./t/./v/.#.UBDZ>/j/./s/./t/./f/.!NC.!NC;
/k/./rz/.#.UBDZ>/k/./S/.!NC.!NC;
/k/./v/.#.UBDZ>/k/./f/.!NC.!NC;
/l/./s/./t/./v/.#.UBDZ>/l/./s/./t/./f/.!NC.!NC;
/m/./s/./t/./v/.#.UBDZ>/m/./s/./t/./f/.!NC.!NC;
/n/./ts/./t/./v/.#.UBDZ>/n/./ts/./t/./f/.!NC.!NC;
/n'/./s/./t/./v/.#.UBDZ>/n'/./s/./t/./f/.!NC.!NC;
/p/./rz/.#.UBDZ>/p/./S/.!NC.!NC;
/p/./s/./t/./v/.#.UBDZ>/p/./s/./t/./f/.!NC.!NC;
/r/./b/./s/./t/.#.UBDZ>/r/./p/./s/./t/.!NC.!NC;
/r/./s/./t/./v/.#.UBDZ>/r/./s/./t/./f/.!NC.!NC;
/r/./t/./v/.#.UBDZ>/r/./t/./f/.!NC.!NC;
/rz/./x/.#.UBDZ>/S/./x/.!NC.!NC;
/rz/./x/./w/.#.UBDZ>/S/./x/./w/.!NC.!NC;
/s/./k/./rz/.#.UBDZ>/s/./k/./S/.!NC.!NC;
/s/./k/./v/.#.UBDZ>/s/./k/./f/.!NC.!NC;
/s/./t/./rz/.#.UBDZ>/s/./t/./S/.!NC.!NC;
/s/./t/./v/.#.UBDZ>/s/./t/./f/.!NC.!NC;
/S/./tS/./b/.#.UBDZ>/S/./tS/./p/.!NC.!NC;
/S/./v/.#.UBDZ>/S/./f/.!NC.!NC;
/s'/./b/.#.UBDZ>/s'/./p/.!NC.!NC;
/t/./rz/.#.UBDZ>/t/./S/.!NC.!NC;
/t/./v/.#.UBDZ>/t/./f/.!NC.!NC;
/v/./ts/.#.UBDZ>/f/./ts/.!NC.!NC;
/v/./ts'/.#.UBDZ>/f/./ts'/.!NC.!NC;
/v/./s/./k/.#.UBDZ>/f/./s/./k/.!NC.!NC;
/v/./s/./t/./v/.#.UBDZ>/f/./s/./t/./f/.!NC.!NC;
/z'/./ts'/.#.UBDZ>/s'/./ts'/.!NC.!NC;
/Z/./S/.#.UBDZ>/S/./S/.!NC.!NC;

Druga lista jest analogiczna do listy przedstawionej powyżej i zawiera reguły właściwe dla międzywyrazowego kontekstu udźwięczniającego:

/b/./ts'/.#.UDZ>/b/./dz'/.!NC.!NC;
/b/./p/./s/./t/./v/.#.UDZ>/b/./b/./z/./d/./v/.!NC.!NC;
/b/./r/./tS/.#.UDZ>/b/./r/./dZ/.!NC.!NC;

/b/./s/.#.UDZ>/b/./z/.!NC.!NC;
/b/./s/./k/.#.UDZ>/b/./z/./g/.!NC.!NC;
/b/./s/./t/./v/.#.UDZ>/b/./z/./d/./v/.!NC.!NC;
/x/./d/.#.UDZ>/x/./d/.!NC.!NC;
/x/./rz/.#.UDZ>/x/./rz/.!NC.!NC;
/x/./rz/./t/.#.UDZ>/x/./rz/./d/.!NC.!NC;
/x/./v/.#.UDZ>/x/./v/.!NC.!NC;
/ts/./t/./v/.#.UDZ>/dz/./d/./v/.!NC.!NC;
/tS/./b/.#.UDZ>/dZ/./b/.!NC.!NC;
/d/./tS/.#.UDZ>/d/./dZ/.!NC.!NC;
/d/./s/./k/.#.UDZ>/d/./z/./g/.!NC.!NC;
/d/./s'/.#.UDZ>/d/./z'/.!NC.!NC;
/dz/./k/.#.UDZ>/dz/./g/.!NC.!NC;
/dz/./t/./v/.#.UDZ>/dz/./d/./v/.!NC.!NC;
/dz'/./s/./t/./v/.#.UDZ>/dz'/./z/./d/./v/.!NC.!NC;
/g/./f/.#.UDZ>/g/./v/.!NC.!NC;
/j/./s/./t/./v/.#.UDZ>/j/./z/./d/./v/.!NC.!NC;
/k/./rz/.#.UDZ>/g/./rz/.!NC.!NC;
/k/./v/.#.UDZ>/g/./v/.!NC.!NC;
/l/./s/./t/./v/.#.UDZ>/l/./z/./d/./v/.!NC.!NC;
/m/./s/./t/./v/.#.UDZ>/m/./z/./d/./v/.!NC.!NC;
/n/./ts/./t/./v/.#.UDZ>/n/./dz/./d/./v/.!NC.!NC;
/n'/./s/./t/./v/.#.UDZ>/n'/./z/./d/./v/.!NC.!NC;
/p/./rz/.#.UDZ>/b/./rz/.!NC.!NC;
/p/./s/./t/./v/.#.UDZ>/b/./z/./d/./v/.!NC.!NC;
/r/./b/./s/./t/.#.UDZ>/r/./b/./z/./d/.!NC.!NC;
/r/./s/./t/./v/.#.UDZ>/r/./z/./d/./v/.!NC.!NC;
/r/./t/./v/.#.UDZ>/r/./d/./v/.!NC.!NC;
/rz/./x/.#.UDZ>/rz/./x/.!NC.!NC;
/rz/./x/./w/.#.UDZ>/rz/./x/./w/.!NC.!NC;
/s/./k/./rz/.#.UDZ>/z/./g/./rz/.!NC.!NC;
/s/./k/./v/.#.UDZ>/z/./g/./v/.!NC.!NC;
/s/./t/./rz/.#.UDZ>/z/./d/./rz/.!NC.!NC;
/s/./t/./v/.#.UDZ>/z/./d/./v/.!NC.!NC;
/S/./tS/./b/.#.UDZ>/Z/./dZ/./b/.!NC.!NC;
/S/./v/.#.UDZ>/Z/./v/.!NC.!NC;
/s'/./b/.#.UDZ>/z'/./b/.!NC.!NC;
/t/./rz/.#.UDZ>/d/./rz/.!NC.!NC;
/t/./v/.#.UDZ>/d/./v/.!NC.!NC;
/v/./ts/.#.UDZ>/v/./dz/.!NC.!NC;
/v/./ts'/.#.UDZ>/v/./dz'/.!NC.!NC;
/v/./s/./k/.#.UDZ>/v/./z/./g/.!NC.!NC;
/v/./s/./t/./v/.#.UDZ>/v/./z/./d/./v/.!NC.!NC;
/z'/./ts'/.#.UDZ>/z'/./dz'/.!NC.!NC;
/Z/./S/.#.UDZ>/Z/./Z/.!NC.!NC;

4.4. Warstwa modyfikacyjna

Funkcje ostatniej (trzeciej) warstwy są związane z różnymi modyfikacjami w zapisie transkrypcyjnym. Autor starał się tak zaprojektować układ warstw, aby warstwa modyfikacyjna obejmowała przekształcenia, które mogą być przeprowadzone w różnorodny sposób. Dzięki takiemu podejściu można tworzyć różne wersje trzeciej warstwy i stosować te wersje zamiennie. Podobną funkcjonalność mają systemy generatywne (zob. podrozdz. 5.4). Nie wszystkie reguły, które mogą przybierać różnorodny kształt, włączono do warstwy modyfikacyjnej. Dotyczy

to przede wszystkim reguł związanych z modyfikacją statusu dźwięczności – zostały one włączone do warstwy drugiej. Niektóre reguły omówione w tej części są związane z przekształceniami obligatoryjnymi (na przykład z przekształceniami tymczasowej notacji litery *ę* oraz litery *ą* na transkrypcję właściwą). Inne mają charakter fakultatywny. Dotyczą na przykład uproszczeń w ramach niektórych grup spółgłoskowych.

Wzorem opisów dotyczących reguł włączonych do dwóch początkowych warstw projektu, w tym miejscu również musi być ustalony inwentarz segmentów, z których są zbudowane wyrazy przekazywane do warstwy modyfikacyjnej. Poza tym muszą być zdefiniowane zbiory używane w regułach. Inwentarz segmentów trzeciego stopnia wygląda identycznie jak inwentarz drugiego stopnia (zob. tab. 4.2), dlatego nie jest on tutaj ponownie prezentowany. Istnieje jednak istotna różnica związana ze sposobem korzystania z tych inwentarzy. W wyrazach zapisanych przy użyciu inwentarza drugiego stopnia status dźwięczności segmentów odzwierciedla domyślną dźwięczność znaków, diad i triad ortograficznych. W transkrypcji drugiego stopnia przykładowych wyrazów: *podkowa, władca, nadfiolet, fałdka, odpiorą* zostałby użyty segment /d/. Tymczasem w wyrazach zapisanych przy użyciu inwentarza trzeciego stopnia status dźwięczności segmentów wynika z konstrukcji reguł włączonych do drugiej warstwy projektu. Jest to zatem docelowy status dźwięczności, który z założenia odzwierciedla faktyczne cechy artykulacyjne mowy. W zapisie wymienionych wyrazów zostałby użyty segment /t/.

Definicje zbiorów ustalonych dla warstwy modyfikacyjnej są podobne do definicji omówionych w podrozdziale 4.3. Inne są symbole tych zbiorów. Nie ma potrzeby wykluczania segmentów: /v/ oraz /rz/ ze zbioru WD3. Poza tym ustalono zbiór SA3 obejmujący samogłoski. Oto definicje zbiorów włączone do warstwy modyfikacyjnej:

```
[WD3]={/v/,/z/,/rz/,/Z/,/z'/,/b/,/d/,/g/,/dz/,/dZ/,/dz'/};
[WB3]={/f/,/s/,/S/,/s'/,/x/,/p/,/t/,/k/,/ts/,/tS/,/ts'/};
[SO3]={/w/,/j/,/l/,/r/,/m/,/n/,/n'/};
[SA3]={/a/,/e/,/o/,/u/,/i/,/y/,/ą/,/ę/};
```

4.4.1. Reguły przetwarzania segmentu /ą/ trzeciego stopnia

W omawianym projekcie właściwe reguły dotyczące przetwarzania litery *ą* zostały włączone do warstwy trzeciej ze względu na różne możliwości transkrypcji. W warstwie transkrypcji podstawowej zawarto regułę, która umożliwia przekształcenie litery *ą* na umowny zapis pośredni: /ą/. W takiej postaci jest on przekazywany najpierw do warstwy statusu dźwięczności, a następnie do warstwy modyfikacyjnej (zob. podrozdz. 4.2.6).

Konstrukcja reguł w warstwie modyfikacyjnej została oparta na regułach pochodzących z *Automatyzacji…* (zob. podrozdz. 2.2.3). Wymagały one gruntownej modyfikacji, ponieważ pierwotnie został uwzględniony kontekst ortograficzny. W warstwie modyfikacyjnej są przetwarzane wyrazy zapisane w sposób wynikający z przetwarzania w dwóch początkowych warstwach. Oto zestaw reguł związanych z transkrypcją pośredniej notacji litery *ą*:

```
/ą/./l/+/w/+/m/>/o/.!NC;
/ą/./p/+/b/>/o/&/m/.!NC;
/ą/./t/+/d/+/k/+/g/>/o/&/n/.!NC;
/ą/./ts/+/dz/>/o/&/n/.!NC;
/ą/./ts'/+/dz'/>/o/&/n'/.!NC;
/ą/./s'/+/z'/+/f/+/v/+/s/+/z/+/S/+/Z/+/x/>/o/&/w~/.!NC;
/ą/.#>/o/&/w~/.!NC;
```

Zastosowano w nich modyfikację polegającą na zastąpieniu segmentu jednokrotnego segmentem wielokrotnym.

4.4.2. Reguły przetwarzania segmentu /ę/ trzeciego stopnia

Również /ę/ jest umownym zapisem pośrednim używanym w warstwie transkrypcji podstawowej. W warstwie modyfikacyjnej zawarto następujące reguły dotyczące przetwarzania tej notacji (są one oparte na tabeli 2.5):

/ę/./l/+/w/>/e/.!NC;
/ę/.#>/e/.!NC;
/ę/./p/+/b/>/e/&/m/.!NC;
/ę/./t/+/d/+/k/+/g/>/e/&/n/.!NC;
/ę/./ts/+/dz/>/e/&/n/.!NC;
/ę/./ts'/+/dz'/>/e/&/n'/.!NC;
/ę/./s'/+/z'/+/f/+/v/+/s/+/z/+/S/+/Z/+/rz/+/x/>/e/&/w~/.!NC;

4.4.3. Reguły przetwarzania segmentu /i/ trzeciego stopnia

W części 2.2.6 wspomniano o wygłosowej konstrukcji: *{p, b, f, w, m, t, d, h, l, r}ii* (występuje ona między innymi w wyrazach: *kopii, komedii, monarchii*). Zgodnie z tym opisem pierwsza litera w diadzie ortograficznej *ii* może oznaczać fonem /i/ lub fonem /j/. W wypadku pierwszej możliwości nie ma potrzeby tworzenia dodatkowych reguł. Do warstwy modyfikacyjnej można włączyć reguły, które są zgodne z drugą podaną opcją, jednak trzeba w nich uwzględnić inwentarz trzeciego stopnia (konstrukcja oryginalnych reguł jest oparta na zapisie ortograficznym):

/p/./i/./i/.#>!NC./j/.!NC.!NC;
/b/./i/./i/.#>!NC./j/.!NC.!NC;
/f/./i/./i/.#>!NC./j/.!NC.!NC;
/v/./i/./i/.#>!NC./j/.!NC.!NC;
/m/./i/./i/.#>!NC./j/.!NC.!NC;
/t/./i/./i/.#>!NC./j/.!NC.!NC;
/d/./i/./i/.#>!NC./j/.!NC.!NC;
/x/./i/./i/.#>!NC./j/.!NC.!NC;
/l/./i/./i/.#>!NC./j/.!NC.!NC;
/r/./i/./i/.#>!NC./j/.!NC.!NC;

4.4.4. Reguły przetwarzania segmentu /j/ trzeciego stopnia

Kolejne prezentowane reguły dotyczą redukcji segmentu odpowiadającego literze *j* wchodzącej w skład wygłosowej struktury: *ji#*. Reguły przytoczone w tabeli 2.19 trzeba dostosować do zapisu właściwego dla warstwy modyfikacyjnej. Użyto w nich polecenia !RM. Uwzględniono tylko takie struktury, które występują w słowniku *SJP*.pl:

/ts/./j/./i/.#>!NC.!RM.!NC.!NC;
/s/./j/./i/.#>!NC.!RM.!NC.!NC;
/t/./j/./i/.#>!NC.!RM.!NC.!NC;
/z/./j/./i/.#>!NC.!RM.!NC.!NC;

4.4.5. Reguły przetwarzania segmentu /m/ trzeciego stopnia

Do warstwy modyfikacyjnej włączono reguły dotyczące alternatywnej transkrypcji segmentu odpowiadającego literze *m*. Są one oparte są na regułach ujętych w tabeli 2.24 (przy założeniu, że głoska [ũ̯] (AS) należy do fonemu /w~/):

/m/./f/>/w~/.!NC;
/m/./v/>/w~/.!NC;

4.4.6. Reguły przetwarzania segmentu /n/ trzeciego stopnia

Litera *n* umiejscowiona przed literami lub przed diadami ortograficznymi oznaczającymi fonemy: /s/, /z/ oraz /Z/ (/rz/) jest transkrybowana jako fonem /w~/ (zob. podrozdz. 2.4.2). Natomiast przed literami, diadami

lub triadami ortograficznymi oznaczającymi fonemy: /i/, /ts'/, /dz'/, /s'/, /z'/, /n'/ jest notowana jako fonem /n'/. Te reguły obowiązują przy założeniu, że głoska [ũ] (AS) jest alofonem fonemu /w~/:

/n/./s/>/w~/.!NC;
/n/./z/>/w~/.!NC;
/n/./Z/>/w~/.!NC;
/n/./rz/>/w~/.!NC;
/n/./i/>/n'/.!NC;
/n/./ts'/>/n'/.!NC;
/n/./dz'/>/n'/.!NC;
/n/./s'/>/n'/.!NC;
/n/./z'/>/n'/.!NC;
/n/./n'/>/n'/.!NC;

4.4.7. Reguły przetwarzania segmentu /n'/ trzeciego stopnia

Reguły omówione w tej części dotyczą alternatywnej transkrypcji segmentu odpowiadającego literze *ń*. Zgodnie z informacjami podanymi w części 2.4.3 każdorazowa transkrypcja jako fonem /n'/ jest właściwa tylko dla wymowy starannej. Wariant taki pokrywa się z przyjętą transkrypcją podstawową, jest zatem stosowany bezwarunkowo na podstawie reguł włączonych do pierwszej warstwy omawianego projektu (w każdym kontekście). Do warstwy modyfikacyjnej włączono reguły związane z wymową mniej staranną. W ich konstrukcji trzeba uwzględnić sposób zapisywania wyrazów przetworzonych w dwóch początkowych warstwach. W warstwie modyfikacyjnej segment /n'/ trzeciego stopnia jest zastępowany segmentem wielokrotnym /j/&/m/ lub /j/&/n/ – wynika to z następujących reguł:

/n'/./b/>/j/&/m/.!NC;
/n'/./t/>/j/&/n/.!NC;
/n'/./d/>/j/&/n/.!NC;
/n'/./k/>/j/&/n/.!NC;
/n'/./ts/>/j/&/n/.!NC;
/n'/./tS/>/j/&/n/.!NC;

4.4.8. Reguły przetwarzania segmentu /w/ trzeciego stopnia

Podstawowa reguła dotycząca transkrypcji litery *ł* została przytoczona w części 2.3.2. Jest ona zgodna z przyjętą transkrypcją podstawową, więc jest stosowana bezwarunkowo w czasie przetwarzania wyrazów w pierwszej warstwie omawianego projektu. Uzyskana transkrypcja w niezmienionej postaci jest przekazywana do warstwy modyfikacyjnej. W tabeli 2.22 ujęto reguły dotyczące geminaty ortograficznej *łł* umiejscowionej przed *b* lub przed *sz*. Zgodnie z nimi fonem odpowiadający drugiej literze *ł* nie jest realizowany. Reguły w warstwie modyfikacyjnej dostosowane do inwentarza trzeciego stopnia mają następującą konstrukcję:

/w/./w/./b/>!NC.!RM.!NC;
/w/./w/./S/>!NC.!RM.!NC;

Kolejna kwestia dotyczy transkrypcji półsamogłoskowego fonemu /w/ umiejscowionego między spółgłoskami. Zgodnie z informacjami zamieszczonymi w podrozdziale 2.3.2 taki fonem może nie być realizowany. Oto zestaw odpowiednich reguł (na podstawie tabeli 2.22), których struktura została dostosowana do inwentarza segmentów trzeciego stopnia oraz do wykonywanych w poprzedniej warstwie modyfikacji statusu dźwięczności niektórych spółgłosek właściwych (zob. podrozdz. 4.3):

/p/./w/./k/>!NC.!RM.!NC;
/t/./w/./k/>!NC.!RM.!NC;
/s/./w/./k/>!NC.!RM.!NC;
/t/./w/./ts/>!NC.!RM.!NC;

/r/./w/./S/>!NC.!RM.!NC;
/r/./w/./b/>!NC.!RM.!NC;
/x/./w/./S/>!NC.!RM.!NC;
/k/./w/./S/>!NC.!RM.!NC;
/p/./w/./S/>!NC.!RM.!NC;
/s/./w/./S/>!NC.!RM.!NC;
/t/./w/./S/>!NC.!RM.!NC;
/z/./w/./b/>!NC.!RM.!NC;
/g/./w/./b/>!NC.!RM.!NC;
/x/./w/./b/>!NC.!RM.!NC;
/d/./w/./b/>!NC.!RM.!NC;
/b/./w/./b/>!NC.!RM.!NC;
/p/./w/./tS/>!NC.!RM.!NC;
/t/./w/./S/>!NC.!RM.!NC;
/k/./w/./S/>!NC.!RM.!NC;
/s/./w/./S/>!NC.!RM.!NC;
/g/./w/./b/>!NC.!RM.!NC;
/d/./w/./b/>!NC.!RM.!NC;
/z/./w/./b/>!NC.!RM.!NC;
/b/./w/./b/>!NC.!RM.!NC;

Ostatni prezentowany w tej części zestaw reguł ma związek z redukcją fonemu odpowiadającego literze *ł* umiejscowionej w wygłosie wyrazu. Pierwsza reguła na poniższej liście dotyczy geminaty *łł*:

/w/./w/.#>!NC.!RM.!NC;
/r/./w/.#>!NC.!RM.!NC;
/x/./w/.#>!NC.!RM.!NC;
/k/./w/.#>!NC.!RM.!NC;
/p/./w/.#>!NC.!RM.!NC;
/s/./w/.#>!NC.!RM.!NC;
/t/./w/.#>!NC.!RM.!NC;
/b/./w/.#>!NC.!RM.!NC;
/d/./w/.#>!NC.!RM.!NC;
/g/./w/.#>!NC.!RM.!NC;
/z/./w/.#>!NC.!RM.!NC;

4.4.9. Reguły przetwarzania triady ortograficznej *drz*

Z reguł przytoczonych w tabeli 2.31 wynika, że jeżeli ortograficzna triada *drz* nie jest umiejscowiona przed literą oznaczającą samogłoskę, to powinna być transkrybowana jako fonem /dZ/ lub jako fonem /tS/ (jeśli kontekst jest ubezdźwięczniający). W wypadku ortograficznej triady *drz* (oraz omówionej w następnej części triady *trz*) nabywanie właściwych cech transkrypcyjnych przebiega stopniowo i odbywa się w czasie przetwarzania wyrazu we wszystkich warstwach omawianego projektu. W warstwie transkrypcji podstawowej zostają wydzielone dwa segmenty: *d* oraz *rz*. Następuje też ich przekształcenie w odpowiednie segmenty drugiego stopnia: /d/ oraz /rz/. W czasie przetwarzania w warstwie statusu dźwięczności może nastąpić zamiana na ich bezdźwięczne odpowiedniki: /t/ oraz /S/ (jeżeli kontekst prawostronny jest ubezdźwięczniający). W warstwie modyfikacyjnej podane segmenty są zastępowane jednym fonemem zwarto-szczelinowym (ma to związek ze zjawiskiem asybilacji). W tym celu zostaje użyte polecenie !ML:

/d/./rz/.[WD3]+[SO3]>/dZ/.!ML.!NC;
/t/./S/.[WB3]>/tS/.!ML.!NC;

4.4.10. Reguły przetwarzania triady ortograficznej *trz*

Przetwarzanie ortograficznej triady *trz* jest realizowane w etapach. W warstwie transkrypcji podstawowej zostają wydzielone segmenty drugiego stopnia: /t/ oraz /rz/. W trakcie przetwarzania w warstwie drugiej następuje ujednolicenie statusu dźwięczności tych segmentów. W zależności od kontekstu uzyskiwana jest jedna z następujących sekwencji segmentów trzeciego stopnia: /t.S/ lub /d.Z/. W warstwie modyfikacyjnej następuje ich przekształcenie w segmenty pojedyncze (w określonych kontekstach). Odzwierciedla to istotę zjawiska asybilacji. Na podstawie podsumowania zawartego w tabeli 2.62 utworzono reguły dla warstwy modyfikacyjnej. Nie ma potrzeby tworzenia reguł dla sekwencji /d.rz/ oraz /t.S/ umiejscowionych przed spółgłoską właściwą, ponieważ reguły omówione w części 4.4.9 (reguły związane z transkrypcją ortograficznej triady *drz*) są odpowiednie również dla triady *trz* przetworzonej w dwóch początkowych warstwach:

/d/./rz/.[WD3]+[SO3]>/dZ/.!ML.!NC;
/t/./S/.[WB3]>/tS/.!ML.!NC;

Poza tym do warstwy modyfikacyjnej trzeba włączyć regułę związaną z kontekstem obejmującym spółgłoski sonorne:

/t/./S/.[SO3]>/tS/.!ML.!NC;

Nie ma potrzeby tworzenia dodatkowych reguł dla kontekstu samogłoskowego. Wyjątek stanowi triada ortograficzna *trz* w członie *wewnątrz* umiejscowionym przed literą oznaczającą samogłoskę. Można zatem zaproponować odpowiednią regułę:

/t/./S/.[SA3]>/tS/.!ML.!NC>!INC:*#wewnątrz*;

W konstrukcji tej reguły została uwzględniona jednoelementowa lista włączeń, która obejmuje nagłosową sekwencję ortograficzną *#wewnątrz*.

4.4.11. Reguły związane z wyodrębnieniem fonemów: /J/ i /c/

Można przyjąć rozwiązanie polegające na włączeniu do warstwy transkrypcji podstawowej reguł dostosowanych do pełnej asynchronii wymowy konstrukcji *gi*A oraz *ki*A:

g.i.[SA1]>/g/./j/.!NC;
k.i.[SA1]>/k/./j/.!NC;

Przy takim podejściu wymowę synchroniczną i wyodrębnienie fonemów: /J/ oraz /c/ można traktować jako opcję. Zatem odpowiednie reguły z podrozdziału 4.2.5 mogłyby zostać przesunięte do warstwy modyfikacyjnej. Po niezbędnych modyfikacjach przybrałyby one następujący kształt:

/g/./j/.[SA3]–/e/>/J/.!NC.!NC;
/g/./j/./e/>/J/.!ML.!NC;
/x/./i/./g/./j/./e/./n/>!NC.!NC./J/.!NC.!NC.!NC;
/g/./j/./e/.#>/J/.!NC.!NC.!NC>!EXC:*nogie#,rogie#,mózgie#,brzegie#,chędogie#,wargie#,długie#,drugie#,nagie#,połogie#,tęgie#,ubogie#,wielgie#,pręgie#*;
/g/./j/./e/.#>/J/.!ML.!NC.!NC>!INC:*nogie#,rogie#,mózgie#,brzegie#,chędogie#,wargie#,długie#,drugie#,nagie#,połogie#,tęgie#,ubogie#,wielgie#,pręgie#*;
/k/./j/.[SA3]–/e/>/c/.!NC.!NC;
/k/./j/./e/>/c/.!ML.!NC;

W wypadku wyodrębnienia fonemów /J/ i /c/ dopiero w warstwie modyfikacyjnej muszą zostać utworzone reguły dla odpowiedniej transkrypcji diad ortograficznych: *gi* oraz *ki* umiejscowionych przed spółgłoską, półsamogłoską lub w wygłosie wyrazu. Reguły w omawianej warstwie mają następującą postać:

/g/./i/.[WD3]+[WB3]+[SO3]+#>/J/.!NC.!NC;
/k/./i/.[WD3]+[WB3]+[SO3]+#>/c/.!NC.!NC;

Poza tym można uwzględnić reguły dotyczące mniej starannego wariantu wymowy triad: *kii* oraz *gii* umiejscowionych w wygłosie wyrazu. Mają one związek z redukcją segmentu odpowiadającego pierwszej literze *i*. Ich ostateczny kształt zależy od reguł włączonych do warstwy transkrypcji podstawowej.

4.4.12. Pozostałe reguły włączone do warstwy modyfikacyjnej

Reguły omówione w tej części dotyczą przekształceń (najczęściej elizji) różnych segmentów trzeciego stopnia. Poniższa reguła odnosi się do redukcji głoski odpowiadającej literze *w* w wyrazach typu: *krakowski, krakowscy, Beniowski, Poniatowscy, warszawska* (zob. podrozdz. 2.5.1). Użycie listy włączeń sprawia, że reguła może być użyta tylko w odniesieniu do wyrazów posiadających określoną konstrukcję wygłosu:

[SA3]./f/./s/>!NC.!RM.!NC>!INC:*wski#,wskich#,wskiego#,wskiemu#,wskim#,wskimi#,wscy#,wsku#,wska#,wską#,wskiej#,wsko#*;

Jednocześnie trzeba uwzględnić formy wyrazu *szewski* oraz *szewstwo*, w których segment /f/ trzeciego stopnia nie jest redukowany. W poniższej regule czteroelementowej jest on przepisywany w drugim członie, a więc nie może być usunięty na podstawie omówionej reguły trójelementowej:

/S/./e/./f/./s/>!NC.!NC./f/.!NC;

Kolejna omawiana reguła dotyczy redukcji segmentu /f/ trzeciego stopnia w formach i w wyrazach pokrewnych wyrazu *pierwszy*. Na podstawie reguł przytoczonych w podrozdziale 2.5.1 można zaproponować następującą konstrukcję:

/r/./f/./s'/+/S/.[SA3]>!NC.!RM.!NC.!NC;

W poniższych regułach uwzględniono konteksty omówione w podrozdziale 2.6. Niektóre oryginalne reguły zostały zmodyfikowane na podstawie analizy słownika *SJP*.pl:

/ts/./tS/./ts'/>!RM.!NC.!NC;
/dz'/./dz'/>!RM.!NC>!INC:*ięćdzies,#sześćdzies*;
/o/./f/./f/./s/>!NC.!NC.!RM.!NC;
/k/./k/>!RM.!NC>>!INC:*#miękk*;

W podrozdziale 2.6.7 poruszono problem transkrypcji jako /s'/ pierwszej litery w strukturach ortograficznych: *ssi* oraz *smie*. Jest to konsekwencja upodobnienia pod względem miejsca artykulacji (miękkości). Analiza słownika wykazała, że podane tam reguły nie są odpowiednie dla przetwarzania niektórych wyrazów obcych. Skonstruowano następujące reguły dla warstwy modyfikacyjnej:

/s/./s'/.[SA3]=/s'/./!NC/./!NC/>!INC:*#bessie,#nassij,#wyssij,#zassij,#barbarossie,#ekspressie,#odessie*;
/s/./m/./j/./e/.#=/s'/.!NC.!NC.!NC;

Do warstwy statusu dźwięczności włączono reguły, których użycie powoduje wygenerowanie ciągów identycznych segmentów trzeciego stopnia, na przykład:

/b/./b/./s/>/p/./p/.!NC;
/g/./g/./s/>/k/./k/.!NC;
/g/./g/./s'/>/k/./k/.!NC;

Z dużym prawdopodobieństwem w takich wypadkach realizowana jest tylko jedna głoska. A więc do warstwy modyfikacyjnej należałoby włączyć reguły umożliwiające redukcję, na przykład:

/p/./p/./s/>!RM.!NC.!NC;
/k/./k/./s/>!RM.!NC.!NC;
/k/./k/./s'/>!RM.!NC.!NC;

Często tego typu konstrukcje mają związek ze strukturą morfologiczną wyrazów, a dwa takie same segmenty trzeciego stopnia są rozdzielone junkturą (np.: *przeciwwstrząsowy, rozskubywać, rozstrzygać, Hongkong,*

odtwarzać, bezsprzeczny, trynkgeld, szlupbelka, kwintdecyma, postdatować). Przy wyraźnej junkturze morfologicznej prawdopodobieństwo redukcji maleje. Natomiast czynnikiem, który ma wpływ na zwiększenie tego prawdopodobieństwa jest zmniejszenie staranności wymowy (np.: *przewyższa, bądźcie, flamandzcy*). To zagadnienie wymaga dodatkowych badań.

Do warstwy modyfikacyjnej przesunięto regułę związaną z redukcją fonemu odpowiadającego drugiej literze *l* w diadzie ortograficznej *ll* umiejscowionej przed literą oznaczającą głoskę inną niż samogłoska (na podstawie tabeli 18 alfa w *Automatyzacji…*). Strukturę reguły dostosowano do inwentarza trzeciego stopnia i użyto w niej polecenia !RM:

/l/./l/.[WD3]+[WB3]+[SO3]>!NC.!RM.!NC;

4.5. Przetwarzanie przykładowych wyrazów ortograficznych

W tej części dokładnie omówiono działanie WMT na przykładach przetwarzania konkretnych wyrazów ortograficznych. Zostały użyte reguły opisane w częściach 4.2 do 4.4. Najpierw zaprezentowano przykłady stosunkowo nieskomplikowane, w których przetwarzanie ogranicza się do zmiany etykiet powiązanych z poszczególnymi indeksami. W kolejnych zastosowano polecenia umożliwiające przesuwanie indeksów lub wiązanie ich z segmentem pustym (EMP). Z przykładów omówionych w częściach 4.5.1–4.5.13 wykluczono czynnik fonetyki międzywyrazowej. W opisie użyto następujących skrótów dotyczących nazw poszczególnych warstw: WTP (warstwa transkrypcji podstawowej), WSD (warstwa statusu dźwięczności), WM (warstwa modyfikacyjna). Poszczególne etapy przetwarzania wyrazów są przedstawione w formie tabelarycznej, jednak ze względów praktycznych zrezygnowano z numerowania tych tabel oraz z umieszczania nad nimi nagłówków. Wynika to z faktu, że te tabele pełnią rolę arkuszy, które umożliwiają prezentację uporządkowanego zapisu stanu segmentów oraz indeksów przypisanych do tych segmentów. W omówionych przykładach pominięto fakt, że przed przetworzeniem zapis ortograficzny każdego wyrazu jest zamieniany na małe litery (wszystkie wyrazy zostały zapisane przy użyciu tylko małych liter).

4.5.1. Przetwarzanie wyrazu: *jeden*

Najpierw ortograficzny wyraz *jeden* zostaje podzielony na litery. Po obu stronach wyrazu zostaje dołączony segment zawierający znacznik granicy wyrazu. Przypisanie indeksów do segmentów przed przekazaniem wyrazu do WTP jest następujące:

Segment	#	*j*	*e*	*d*	*e*	*n*	#
Indeks	0	1	2	3	4	5	6

W czasie przetwarzania w WTP są stosowane reguły jednoelementowe:

j>/j/;
e>/e/;
d>/d/;
e>/e/;
n>/n/;

WTP nie zawiera dłuższych reguł, które mogłyby być użyte w wypadku tego wyrazu. Reguły zdefiniowane w WTP (oraz w każdej innej warstwie) nie są stosowane natychmiast. Najpierw sprawdzane są wszystkie reguły włączone do danej warstwy pod kątem możliwości ich użycia dla obecnie przetwarzanego wyrazu (reguły są sprawdzane w kolejności od najdłuższych do najkrótszych). W czasie sprawdzania poszczególnych reguł zostają zablokowane segmenty, które mają być zmodyfikowane. Segment zablokowany w czasie sprawdzania danej reguły nie może być zmodyfikowany w wyniku użycia reguł sprawdzanych później (w ramach jednej warstwy). Po sprawdzeniu wszystkich reguł należących do danej warstwy następuje wykonanie zmian wynikających z odpowiednich reguł. Stan wyrazu *jeden* po sprawdzeniu wszystkich reguł należących do WTP i przed zastosowaniem pięciu wymienionych reguł jest następujący:

Segment	#	*j*	*e*	*d*	*e*	*n*	#
Indeks	0	1^	2^	3^	4^	5^	6

W zapisie tym znak ^ oznacza, że segmenty powiązane z poszczególnymi indeksami zostały zablokowane. Po sprawdzeniu wszystkich reguł wykonywane są odpowiednie zmiany oraz zostają usunięte znaczniki blokady:

Segment	#	/j/	/e/	/d/	/e/	/n/	#
Indeks	0	1	2	3	4	5	6

W takiej postaci wyraz jest przekazywany najpierw do WSD, a następnie do WM. Warstwy te nie zawierają jednak żadnej reguły, która mogłaby być użyta. Dlatego podana transkrypcja jest ostateczna. Jej projekcja na początkowy zapis ortograficzny (przy zachowaniu indeksacji tego zapisu) jest następująca:

Segment	#	/j/	/e/	/d/	/e/	/n/	#
Indeks	0	1	2	3	4	5	6
Zapis ort.		*j*	*e*	*d*	*e*	*n*	

4.5.2. Przetwarzanie wyrazu: *szukaj*

Początkowe przypisanie indeksów do segmentów wygląda następująco (w takiej postaci wyraz *szukaj* jest przekazywany do WTP):

Segment	#	*s*	*z*	*u*	*k*	*a*	*j*	#
Indeks	0	1	2	3	4	5	6	7

W ramach WTP występuje jedna reguła dwuelementowa (właściwa dla przetwarzanego wyrazu), która ma wyższy priorytet niż reguły jednoelementowe:

s.z>/S/.!ML;

Najpierw rejestrowana jest zmiana oczekująca:

Segment przed zmianą	*s*	*z*
Indeks przed zmianą	1	2
Segment po zmianie	/S/	–
Indeks po zmianie	1,2	–

Podana reguła nie jest stosowana natychmiast – przedtem muszą zostać sprawdzone inne reguły pod kątem możliwości użycia w przetwarzanym wyrazie (w pierwszej kolejności są sprawdzane pozostałe reguły o takiej samej długości, a następnie reguły krótsze). Rejestracja zmiany oczekującej powoduje zablokowanie odpowiednich segmentów w przetwarzanym wyrazie:

Segment	#	*s*	*z*	*u*	*k*	*a*	*j*	#
Indeks	0	1^	2^	3	4	5	6	7

Zablokowane segmenty nie mogą być już modyfikowane w ramach tej warstwy, jednak ciągle mogą być używane dla określania kontekstu (w innych regułach może być używane polecenie !NC w odniesieniu do takich segmentów). W ramach WTP dla przetwarzania segmentów, które jeszcze nie zostały zablokowane, mogą być użyte następujące reguły jednoelementowe:

u>/u/;
k>/k/;
a>/a/;
j>/j/;

Obecność każdej wymienionej reguły powoduje rejestrację nowej zmiany oczekującej, na przykład schemat rejestracji zmiany związanej z pierwszą wymienioną regułą wygląda następująco:

Segment przed zmianą	*u*
Indeks przed zmianą	3
Segment po zmianie	/u/
Indeks po zmianie	3

Po zarejestrowaniu wszystkich zmian w ramach WTP stan przetwarzanego wyrazu jest następujący:

Segment	#	*s*	*z*	*u*	*k*	*a*	*j*	#
Indeks	0	1^	2^	3^	4^	5^	6^	7

Po sprawdzeniu wszystkich możliwości użycia reguł należących do WTP oraz po zarejestrowaniu odpowiednich zmian oczekujących zmiany te zostają wykonane i następuje odblokowanie segmentów. Po tej operacji segmenty w przetwarzanym wyrazie oraz przypisanie indeksów do tych segmentów są następujące:

Segment	#	/S/	/u/	/k/	/a/	/j/	#
Indeks	0	1,2	3	4	5	6	7

Wyraz w tej postaci jest przekazywany najpierw do WSD, a następnie do WM. Warstwy te nie zawierają jednak żadnych reguł, które mogłyby być użyte w tym wypadku. Nie są zatem wykonywane żadne kolejne zmiany, a ostateczna transkrypcja wyrazu *szukaj* wraz z projekcją na pierwotny zapis ortograficzny jest następująca:

Segment	#	/S/	/u/	/k/	/a/	/j/	#
Indeks	0	1,2	3	4	5	6	7
Zapis ort.		*s,z*	*u*	*k*	*a*	*j*	

4.5.3. Przetwarzanie wyrazu: *praugrofiński*

Kolejny omawiany przykład dotyczy wyrazu *praugrofiński,* który jest przekazywany do WTP w następującej postaci:

Segment	#	*p*	*r*	*a*	*u*	*g*	*r*	*o*	*f*	*i*	*ń*	*s*	*k*	*i*	#
Indeks	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14

W zbiorze reguł w ramach WTP znajduje się jedna odpowiednia reguła czteroelementowa, która ma większy priorytet niż reguły krótsze:

p.r.a.u>!NC.!NC.!NC./u/;

Obecność tej reguły sprawia, że zostaje zarejestrowana następująca zmiana:

Segment przed zmianą	*u*
Indeks przed zmianą	4
Segment po zmianie	/u/
Indeks po zmianie	4

Rejestracja tej zmiany powoduje zablokowanie segmentu o indeksie 4:

Segment	#	*p*	*r*	*a*	*u*	*g*	*r*	*o*	*f*	*i*	*ń*	*s*	*k*	*i*	#
Indeks	0	1	2	3	4^	5	6	7	8	9	10	11	12	13	14

Obecność tej blokady sprawia, że nie mogą być zastosowane następujące krótsze reguły zdefiniowane w WTP:

a.u>!NC./w/;
u>/u/;

WTP zawiera też reguły jednoelementowe, które mogą być użyte:

p>/p/;
r>/r/;
a>/a/;
g>/g/;
r>/r/;
o>/o/;
f>/f/;
i>/i/;
ń>/n'/;
s>/s/;
k>/k/;
i>/i/;

Po zarejestrowaniu wszystkich zmian oczekujących stan przetwarzanego wyrazu jest następujący:

Segment	#	*p*	*r*	*a*	*u*	*g*	*r*	*o*	*f*	*i*	*ń*	*s*	*k*	*i*	#
Indeks	0	1^	2^	3^	4^	5^	6^	7^	8^	9^	10^	11^	12^	13^	14

W wyrazie zostają wprowadzone wszystkie zmiany oczekujące oraz zostają zniesione blokady segmentów. Wyraz opuszcza WTP w następującej postaci:

Segment	#	/p/	/r/	/a/	/u/	/g/	/r/	/o/	/f/	/i/	/n'/	/s/	/k/	/i/	#
Indeks	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14

Następnie jest on przekazywany do WSD, która nie zawiera żadnych odpowiednich reguł, podobnie jak WM. Zatem podana transkrypcja jest ostateczna. Projekcja tej transkrypcji na początkowy zapis ortograficzny jest następująca:

Segment	#	/p/	/r/	/a/	/u/	/g/	/r/	/o/	/f/	/i/	/n'/	/s/	/k/	/i/	#
Indeks	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Zapis ort.		*p*	*r*	*a*	*u*	*g*	*r*	*o*	*f*	*i*	*ń*	*s*	*k*	*i*	

4.5.4. Przetwarzanie wyrazu: *klientem*

Wyraz *klientem* zostaje przekazany do WTP w następującej postaci:

Segment	#	*k*	*l*	*i*	*e*	*n*	*t*	*e*	*m*	#
Indeks	0	1	2	3	4	5	6	7	8	9

WTP zawiera regułę siedmioelementową, która jest właściwa dla tego wyrazu:

#.k.l.i.e.n.t>!NC.!NC.!NC./i/&/j/.!NC.!NC.!NC;

Ta reguła umożliwia zamianę ortograficznego segmentu *i* na segment wielokrotny /i/&/j/:

Segment przed zmianą	*i*
Indeks przed zmianą	3
Segment po zmianie	/i/&/j/
Indeks po zmianie	3

Po rejestracji tej zmiany stan wyrazu zmienia się:

Segment	#	*k*	*l*	*i*	*e*	*n*	*t*	*e*	*m*	#
Indeks	0	1	2	3^	4	5	6	7	8	9

Poza tym WTP zawiera sześć odpowiednich reguł jednoelementowych:

k>/k/;
l>/l/;
e>/e/;
n>/n/;
t>/t/;
m>/m/;

Na podstawie tych reguł rejestrowane są odpowiednie zmiany oraz następuje blokada kolejnych segmentów:

Segment	#	*k*	*l*	*i*	*e*	*n*	*t*	*e*	*m*	#
Indeks	0	1^	2^	3^	4^	5^	6^	7^	8^	9

Wszystkie zarejestrowane zmiany zostają wykonane i wyraz przybiera następujący kształt:

Segment	#	/k/	/l/	/i/&/j/	/e/	/n/	/t/	/e/	/m/	#
Indeks	0	1	2	3	4	5	6	7	8	9

Następnie wyraz jest przetwarzany w WSD oraz WM. W warstwach tych nie ma jednak żadnych reguł, które mogłyby być zastosowane, dlatego ostateczna transkrypcja wraz z projekcją na zapis ortograficzny wygląda następująco:

Segment	#	/k/	/l/	/i/&/j/	/e/	/n/	/t/	/e/	/m/	#
Indeks	0	1	2	3	4	5	6	7	8	9
Zapis ort.		*k*	*l*	*i*	*e*	*n*	*t*	*e*	*m*	

4.5.5. Przetwarzanie wyrazu: *linia*

Ortograficzny wyraz *linia* jest przekazywany do WTP z następującym przypisaniem indeksów:

Segment	#	*l*	*i*	*n*	*i*	*a*	#
Indeks	0	1	2	3	4	5	6

Zgodnie z treścią podrozdziału 4.2.2 zakłada się, że WTP zawiera następującą regułę siedmioelementową:

l.i.n.i.a.#>!NC.!NC./n'/./j/.!NC.!NC;

Na podstawie tej reguły rejestrowana jest zmiana:

Segment przed zmianą	*n*	*i*
Indeks przed zmianą	3	4
Segment po zmianie	/n'/	/j/
Indeks po zmianie	3	4

Rejestracja ta sprawia, że nie może być zastosowana krótsza reguła przeznaczona dla transkrypcji wyrazów rodzimych:

n.i.[SA1]–*i*>/n'/.!ML.!NC;

Wynika to z faktu, że segmenty powiązane z indeksami: 3 i 4 zostały zablokowane:

Segment	#	*l*	*i*	*n*	*i*	*a*	#
Indeks	0	1	2	3^	4^	5	6

Dla wyrazu w takiej postaci mogą być zastosowane następujące reguły jednoelementowe:

l>/l/;
i>/i/;
a>/a/;

Rejestracja zmian związanych z tymi regułami powoduje blokadę kolejnych segmentów:

Segment	#	*l*	*i*	*n*	*i*	*a*	#
Indeks	0	1^	2^	3^	4^	5^	6

Następnie wszystkie zarejestrowane zmiany są wykonywane, a wyraz przyjmuje następującą postać:

Segment	#	/l/	/i/	/ n'/	/j/	/a/	#
Indeks	0	1	2	3	4	5	6

Wyraz zostaje przekazany najpierw do WSD, a następnie do WM. Ze względu jednak na brak reguł, które mogłyby być użyte, ostateczna transkrypcja omawianego wyrazu wraz z projekcją na zapis ortograficzny wygląda następująco:

Segment	#	/l/	/i/	/ n'/	/j/	/a/	#
Indeks	0	1	2	3	4	5	6
Zapis ort.		*l*	*i*	*n*	*i*	*a*	

4.5.6. Przetwarzanie wyrazu: *marzłem*

Wyraz *marzłem* po rozdzieleniu na litery oraz po przypisaniu indeksów zostaje przekazany do WTP:

Segment	#	*m*	*a*	*r*	*z*	*ł*	*e*	*m*	#
Indeks	0	1	2	3	4	5	6	7	8

Zakłada się, że dla konwersji diady *rz* zdefiniowano jedną regułę podstawową:

r.z>/rz/.!ML;

Zgodnie z tym założeniem są tworzone reguły uzupełniające odnoszące się do wyrazów, w których diada *rz* oznacza sekwencję fonemów /r.z/ (zob. podrozdz. 4.2.1). W wypadku wystąpienia struktury *#marzł* odpowiednia reguła ma następującą postać:

#.m.a.r.z.ł>!NC.!NC.!NC./r/./z/.!NC;

Użycie jej sprawia, że zostają zarejestrowane następujące zmiany:

Segment przed zmianą	*r*	*z*
Indeks przed zmianą	3	4
Segment po zmianie	/r/	/z/
Indeks po zmianie	3	4

Rejestracja tych zmian powoduje blokadę segmentów o indeksach 3 oraz 4:

Segment	#	*m*	*a*	*r*	*z*	*ł*	*e*	*m*	#
Indeks	0	1	2	3^	4^	5	6	7	8

Ze względu na długość powyższa reguła ma wyższy priorytet niż podana reguła podstawowa. W odniesieniu do segmentów niezablokowanych mogą być użyte następujące reguły jednoelementowe:

m>/m/;
a>/a/;
ł>/w/;
e>/e/;
m>/m/;

Ich obecność w WTP powoduje rejestrację kolejnych zmian oraz zablokowanie pozostałych segmentów:

Segment	#	*m*	*a*	*r*	*z*	*ł*	*e*	*m*	#
Indeks	0	1^	2^	3^	4^	5^	6^	7^	8

Następnie wszystkie zarejestrowane zmiany zostają wykonane, a wyraz przybiera następującą postać:

Segment	#	/m/	/a/	/r/	/z/	/w/	/e/	/m/	#
Indeks	0	1	2	3	4	5	6	7	8

Wyraz o tej strukturze jest przetwarzany kolejno w WSD oraz w WM. Warstwy te jednak nie zawierają żadnych reguł, które w tym wypadku mogłyby być użyte. Ostateczna transkrypcja fonologiczna razem z projekcją na zapis ortograficzny (przy uwzględnieniu indeksów początkowo przypisanych do poszczególnych liter) wygląda więc następująco:

Segment	#	/m/	/a/	/r/	/z/	/w/	/e/	/m/	#
Indeks	0	1	2	3	4	5	6	7	8
Zapis ort.		*m*	*a*	*r*	*z*	*ł*	*e*	*m*	

4.5.7. Przetwarzanie wyrazu: *roztwór*

Wyraz *roztwór* jest przekazywany do WTP w takiej postaci:

Segment	#	*r*	*o*	*z*	*t*	*w*	*ó*	*r*	#
Indeks	0	1	2	3	4	5	6	7	8

Następujące reguły jednoelementowe są odpowiednie dla przetwarzania tego wyrazu (w WTP nie występują reguły dłuższe, które mogłyby być użyte):

r>/r/;
o>/o/;
z>/z/;
t>/t/;
w>/v/;
ó>/u/;
r>/r/;

Po zarejestrowaniu wszystkich zmian oczekujących wynikających z wymienionych reguł stan wyrazu jest następujący:

Segment	#	*r*	*o*	*z*	*t*	*w*	*ó*	*r*	#
Indeks	0	1^	2^	3^	4^	5^	6^	7^	8

Zmiany te zostają wprowadzone, a także następuje usunięcie blokad. W takiej postaci wyraz jest przekazany do WSD:

Segment	#	/r/	/o/	/z/	/t/	/v/	/u/	/r/	#
Indeks	0	1	2	3	4	5	6	7	8

WSD zawiera dwie reguły właściwe dla przetwarzanego wyrazu (segment /t/ drugiego stopnia należy do zbioru WB2):

/z/.[WB2]>/s/.!NC;
[WB2]./v/>!NC./f/;

W związku z tym rejestrowane są dwie zmiany oczekujące:

Segment przed zmianą	/z/
Indeks przed zmianą	3
Segment po zmianie	/s/
Indeks po zmianie	3

Segment przed zmianą	/v/
Indeks przed zmianą	5
Segment po zmianie	/f/
Indeks po zmianie	5

Następuje również blokada segmentów, do których są przypisane indeksy: 3 oraz 5:

Segment	#	/r/	/o/	/z/	/t/	/v/	/u/	/r/	#
Indeks	0	1	2	3^	4	5^	6	7	8

Ponieważ WSD nie zawiera żadnych innych reguł odpowiednich dla przetwarzanego wyrazu, zmiany oczekujące zostają wykonane, a segmenty odblokowane. Wyraz zostaje przekazany do WM w następującej postaci:

Segment	#	/r/	/o/	/s/	/t/	/f/	/u/	/r/	#
Indeks	0	1	2	3	4	5	6	7	8

Również WM nie zawiera żadnej reguły dla przetwarzanego wyrazu, dlatego powyższy zapis jest transkrypcją końcową ortograficznego wyrazu *roztwór*. Projekcja tej transkrypcji na pierwotny wyraz ortograficzny jest następująca:

Segment	#	/r/	/o/	/s/	/t/	/f/	/u/	/r/	#
Indeks	0	1	2	3	4	5	6	7	8
Zapis ort.		*r*	*o*	*z*	*t*	*w*	*ó*	*r*	

4.5.8. Przetwarzanie wyrazu: *szczwacz*

Wyraz *szczwacz* ma wyjątkową konstrukcję, ponieważ zawiera aż 3 digrafemy oznaczające spółgłoski właściwe bezdźwięczne. Poza tym w tym wyrazie nie występują segmenty udźwięczniające. Jest on przekazywany do WTP w następującej postaci:

Segment	#	*s*	*z*	*c*	*z*	*w*	*a*	*c*	*z*	#
Indeks	0	1	2	3	4	5	6	7	8	9

W WTP zdefiniowano dwie reguły dwuelementowe, które mają większy priorytet niż jednoelementowe reguły dotyczące liter: *s, z* oraz *c*. Mają one następującą strukturę:

s.z>/S/.!ML;
c.z>/tS/.!ML;

Na podstawie tych reguł zostają zarejestrowane następujące modyfikacje:

Segment przed zmianą	*s*	*z*
Indeks przed zmianą	1	2
Segment po zmianie	/S/	–
Indeks po zmianie	1,2	–

Segment przed zmianą	*c*	*z*
Indeks przed zmianą	3	4
Segment po zmianie	/tS/	–
Indeks po zmianie	3,4	–

Segment przed zmianą	*c*	*z*
Indeks przed zmianą	7	8
Segment po zmianie	/tS/	–
Indeks po zmianie	7,8	–

Segmenty związane z tymi modyfikacjami zostają zablokowane:

Segment	#	*s*	*z*	*c*	*z*	*w*	*a*	*c*	*z*	#
Indeks	0	1^	2^	3^	4^	5	6	7^	8^	9

Następujące reguły jednoelementowe mogą być użyte dla modyfikacji segmentów, które nie zostały zablokowane:

w>/v/;
a>/a/;

Zostają zarejestrowane dwie zmiany wynikające z podanych reguł oraz następuje blokada kolejnych segmentów:

Segment	#	*s*	*z*	*c*	*z*	*w*	*a*	*c*	*z*	#
Indeks	0	1^	2^	3^	4^	5^	6^	7^	8^	9

Następnie wszystkie zarejestrowane modyfikacje zostają wykonane, a wyraz przybiera następujący kształt:

Segment	#	/S/	/tS/	/v/	/a/	/tS/	#
Indeks	0	1,2	3,4	5	6	7,8	9

Wyraz zostaje przekazany do WSD, do której jest włączona reguła dotycząca modyfikacji statusu dźwięczności segmentu /v/ drugiego stopnia (ma to związek ze zjawiskiem perseweracji artykulacyjnej):

[WB2]./v/>!NC./f/;

Jej obecność sprawia, że zostaje zarejestrowana odpowiednia modyfikacja i następuje blokada segmentu powiązanego z indeksem 5:

Segment	#	/S/	/tS/	/v/	/a/	/tS/	#
Indeks	0	1,2	3,4	5^	6	7,8	9

Ponieważ WSD nie zawiera żadnych innych reguł odpowiednich dla tego wyrazu, to jedyna zarejestrowana modyfikacja zostaje wykonana. W rezultacie przetwarzany wyraz przybiera następujący kształt:

Segment	#	/S/	/tS/	/f/	/a/	/tS/	#
Indeks	0	1,2	3,4	5	6	7,8	9

Taka struktura jest przekazywana do WM. Nie zawiera ona żadnych reguł, które mogłyby być użyte w tym wypadku; jest to zatem transkrypcja ostateczna. Projekcja tej transkrypcji na pierwotny zapis ortograficzny wygląda następująco:

Segment	#	/S/	/tS/	/f/	/a/	/tS/	#
Indeks	0	1,2	3,4	5	6	7,8	9
Zapis ort.		*s,z*	*c,z*	*w*	*a*	*c,z*	

4.5.9. Przetwarzanie wyrazu: *jeźdźca*

Wyraz *jeźdźca* jest przekazywany do WTP w następującej formie:

Segment	#	*j*	*e*	*ź*	*d*	*ź*	*c*	*a*	#
Indeks	0	1	2	3	4	5	6	7	8

W WTP zdefiniowano jedną regułę dwuelementową, która może być użyta w czasie przetwarzania tego wyrazu:

d.ź>/dz'/.!ML;

Dlatego rejestrowana jest zmiana związana z przesunięciem indeksu 5 w lewo:

Segment przed zmianą	*d*	*ź*
Indeks przed zmianą	4	5
Segment po zmianie	/dz'/	–
Indeks po zmianie	4,5	–

Jednocześnie segmenty, do których są przypisane indeksy 4 i 5, zostają zablokowane dla dalszego przetwarzania w ramach WTP:

Segment	#	*j*	*e*	*ź*	*d*	*ź*	*c*	*a*	#
Indeks	0	1	2	3	4^	5^	6	7	8

Poza tym w WTP zdefiniowano reguły jednoelementowe, które są odpowiednie dla tego wyrazu:

j>/j/;
e>/e/;
ź>/z'/;
c>/ts/;
a>/a/;

Ich użycie wiąże się z rejestracją odpowiednich zmian oraz z zablokowaniem kolejnych segmentów:

Segment	#	*j*	*e*	*ź*	*d*	*ź*	*c*	*a*	#
Indeks	0	1^	2^	3^	4^	5^	6^	7^	8

Następnie wszystkie zarejestrowane zmiany zostają wykonane:

Segment	#	/j/	/e/	/z'/	/dz'/	/ts/	/a/	#
Indeks	0	1	2	3	4,5	6	7	8

Wyraz w takiej postaci jest przekazywany do WSD, która zawiera jedną odpowiednią regułę związaną ze zjawiskiem desonoryzacji:

[WD2].[WD2].[WB2]>!LVL.!LVL.!NC;

Z podrozdziału 4.3 wynika, że ta uogólniona reguła została zaproponowana w miejsce reguły:

/z'/./dz'/./ts/>/s'/./ts'/.!NC;

W pierwszej z dwóch podanych wersji został użyty mechanizm list odwzorowań. Zastosowanie tej reguły powoduje rejestrację następującej zmiany:

Segment przed zmianą	/z'/	/dz'/
Indeks przed zmianą	3	4,5
Segment po zmianie	/s'/	/ts'/
Indeks po zmianie	3	4,5

Zastosowanie drugiej reguły dałoby identyczny efekt. Rejestracja tej zmiany powoduje zablokowanie segmentów powiązanych z indeksami: 3, 4 i 5:

Segment	#	/j/	/e/	/z'/	/dz'/	/ts/	/a/	#
Indeks	0	1	2	3^	4^,5^	6	7	8

Do WSD nie włączono innych reguł, które mogłyby być użyte w wypadku tego wyrazu, dlatego zarejestrowane zmiany zostają zrealizowane. Po tej operacji przetwarzany wyraz wygląda następująco:

Segment	#	/j/	/e/	/s'/	/ts'/	/ts/	/a/	#
Indeks	0	1	2	3	4,5	6	7	8

WM nie zawiera odpowiednich reguł dla tego wyrazu; można jednak rozważyć utworzenie reguły dotyczącej redukcji segmentu /ts'/. Ostateczna transkrypcja wyrazu z projekcją na zapis ortograficzny (przy zachowaniu indeksacji pierwotnego zapisu ortograficznego) wygląda następująco:

Segment	#	/j/	/e/	/s'/	/ts'/	/ts/	/a/	#
Indeks	0	1	2	3	4,5	6	7	8
Zapis ort.		*j*	*e*	*ś*	*d,ś*	*c*	*a*	

4.5.10. Przetwarzanie wyrazu: *działo*

Wzorem omówionych już przykładów najpierw prezentowany jest wyraz w postaci, w jakiej jest on przekazywany do WTP:

Segment	#	*d*	*z*	*i*	*a*	*ł*	*o*	#
Indeks	0	1	2	3	4	5	6	7

Dla triady *dzi* umiejscowionej przed samogłoską ustalono jedną podstawową regułę:

d.z.i.[SA1]–i>/dz'/.!ML.!ML.!NC;

Do WTP włączono reguły uzupełniające, które dotyczą alternatywnej transkrypcji triady *dzi* (zob. podrozdz. 4.2.2). Żadna z tych reguł nie jest odpowiednia dla wyrazu *działo*. Użycie przytoczonej reguły podstawowej wiąże się z rejestracją następującej zmiany:

Segment przed zmianą	*d*	*z*	*i*
Indeks przed zmianą	1	2	3
Segment po zmianie	/dz'/	–	–
Indeks po zmianie	1,2,3	–	–

Towarzyszy temu blokada segmentów, do których są przypisane indeksy: 1, 2 oraz 3:

Segment	#	*d*	*z*	*i*	*a*	*ł*	*o*	#
Indeks	0	1^	2^	3^	4	5	6	7

Dla pozostałych segmentów (które nie zostały zablokowane) stosowane są następujące reguły jednoelementowe:

a>/a/;
ł>/w/;
o>/o/;

Po zarejestrowaniu zmian związanych z tymi regułami następuje blokada kolejnych segmentów:

Segment	#	*d*	*z*	*i*	*a*	*ł*	*o*	#
Indeks	0	1^	2^	3^	4^	5^	6^	7

Następnie zmiany te zostają wykonane i wyraz zostaje przekazany do WSD w następującej postaci:

Segment	#	/dz'/	/a/	/w/	/o/	#
Indeks	0	1,2,3	4	5	6	7

WSD i WM nie zawierają reguł odpowiednich dla przetwarzanego wyrazu, zatem podana transkrypcja jest ostateczna. Oto projekcja otrzymanej transkrypcji na pierwotny zapis ortograficzny:

Segment	#	/dz'/	/a/	/w/	/o/	#
Indeks	0	1,2,3	4	5	6	7
Zapis ort.		*d,z,i*	*a*	*ł*	*o*	

4.5.11. Przetwarzanie wyrazu: *nadziemny*

Początkowe przypisanie indeksów do liter w wyrazie *nadziemny* jest następujące:

Segment	#	*n*	*a*	*d*	*z*	*i*	*e*	*m*	*n*	*y*	#
Indeks	0	1	2	3	4	5	6	7	8	9	10

W tym wyrazie triada *dzi* jest umiejscowiona przed samogłoską, ale nie stanowi ona trigrafemu ze względu na obecność junktury morfologicznej. WTP zawiera odpowiednią regułę ośmioelementową, która umożliwia zastąpienie diady *zi* segmentem /z'/:

#.n.a.d.z.i.e.m>!NC.!NC.!NC.!NC./z'/.!ML.!NC.!NC;

Jest to reguła uzupełniająca dla reguły podstawowej, która dotyczy konwersji triady *dzi* w kontekście samogłoskowym (zob. podrozdz. 4.2.2):

d.z.i.[SA1]–i>/dz'/.!ML.!ML.!NC;

Reguła nie może być użyta, ponieważ obecność dłuższej reguły uzupełniającej sprawia, że rejestrowana zostaje następująca zmiana:

Segment przed zmianą	*z*	*i*
Indeks przed zmianą	4	5
Segment po zmianie	/z'/	–
Indeks po zmianie	4,5	–

Ta rejestracja powoduje blokadę segmentów powiązanych z indeksami: 4 i 5. Wyklucza to możliwość użycia przytoczonej reguły podstawowej:

Segment	#	*n*	*a*	*d*	*z*	*i*	*e*	*m*	*n*	*y*	#
Indeks	0	1	2	3	4^	5^	6	7	8	9	10

W odniesieniu do segmentów niezablokowanych mogą być użyte następujące reguły jednoelementowe:

n>/n/;
a>/a/;
d>/d/;
e>/e/;
m>/m/;
n>/n/;
y>/y/;

Po rejestracji odpowiednich zmian z nich wynikających stan przetwarzanego wyrazu jest następujący:

Segment	#	*n*	*a*	*d*	*z*	*i*	*e*	*m*	*n*	*y*	#
Indeks	0	1^	2^	3^	4^	5^	6^	7^	8^	9^	10

Następnie zarejestrowane zmiany zostają wykonane:

Segment	#	/n/	/a/	/d/	/z'/	/e/	/m/	/n/	/y/	#
Indeks	0	1	2	3	4,5	6	7	8	9	10

W WSD oraz w WM nie ma żadnych reguł, które mogłyby być użyte, więc powyższa transkrypcja jest ostateczna. Wraz z projekcją na zapis ortograficzny wygląda ona następująco:

Segment	#	/n/	/a/	/d/	/z'/	/e/	/m/	/n/	/y/	#
Indeks	0	1	2	3	4,5	6	7	8	9	10
Zapis ort.		*n*	*a*	*d*	*z,i*	*e*	*m*	*n*	*y*	

4.5.12. Przetwarzanie wyrazu: *rzęsy*

W momencie przekazywania wyrazu *rzęsy* do WTP przypisanie indeksów do liter jest następujące:

Segment	#	*r*	*z*	*ę*	*s*	*y*	#
Indeks	0	1	2	3	4	5	6

WMT zawiera odpowiednią regułę dwuelementową:

r.z>/rz/.!ML;

Dlatego rejestrowana jest następująca zmiana:

Segment przed zmianą	*r*	*z*
Indeks przed zmianą	1	2
Segment po zmianie	/rz/	–
Indeks po zmianie	1,2	–

Oprócz tego segmenty powiązane z indeksem 1 oraz z indeksem 2 zostają zablokowane:

Segment	#	*r*	*z*	*ę*	*s*	*y*	#
Indeks	0	1^	2^	3	4	5	6

Do WTP włączono następujące reguły jednoelementowe, które są odpowiednie dla transkrypcji tego wyrazu:

ę>/ę/;
s>/s/;
y>/y/;

Pierwsza reguła umożliwia zamianę litery *ę* na notację pośrednią (zob. podrozdz. 4.2.6). Po zarejestrowaniu stosownych zmian zostają zablokowane kolejne segmenty:

Segment	#	*r*	*z*	*ę*	*s*	*y*	#
Indeks	0	1^	2^	3^	4^	5^	6

Zmiany zostają wykonane i następuje zwolnienie wszystkich blokad. Wyraz w takiej postaci zostaje przekazany do kolejnej warstwy:

Segment	#	/rz/	/ę/	/s/	/y/	#
Indeks	0	1,2	3	4	5	6

WSD nie zawiera reguł odpowiednich dla przetwarzania takiej struktury. Do WM włączono jednak regułę, która umożliwia przekształcenie tymczasowego zapisu litery ę na transkrypcję docelową:

/ę/./s'/+/z'/+/f/+/v/+/s/+/z/+/S/+/Z/+/rz/+/x/>/e/&/w~/.!NC;

Dlatego rejestrowana jest kolejna zmiana:

Segment przed zmianą	/ę/
Indeks przed zmianą	3
Segment po zmianie	/e/&/w~/
Indeks po zmianie	3

Jednocześnie segment /ę/ zostaje zablokowany:

Segment	#	/rz/	/ę/	/s/	/y/	#
Indeks	0	1,2	3^	4	5	6

Jest to jedyna reguła w WM, którą można zastosować dla przetwarzanego wyrazu, zatem zostaje ona wykonana:

Segment	#	/rz/	/e/&/w~/	/s/	/y/	#
Indeks	0	1,2	3	4	5	6

Otrzymana transkrypcja jest ostateczna i jej projekcja na zapis ortograficzny wygląda następująco:

Segment	#	/rz/	/e/&/w~/	/s/	/y/	#
Indeks	0	1,2	3	4	5	6
Zapis ort.	#	*r;z*	*ę*	*s*	*y*	#

4.5.13. Przetwarzanie wyrazu: *jabłko*

W momencie przekazania do WTP przypisanie indeksów do segmentów w wyrazie *jabłko* jest następujące:

Segment	#	*j*	*a*	*b*	*ł*	*k*	*o*	#
Indeks	0	1	2	3	4	5	6	7

WTP zawiera reguły jednoelementowe, które są odpowiednie dla transkrypcji tego wyrazu:

j>/j/;
a>/a/;
b>/b/;
ł>/w/;
k>/k/;
o>/o/;

Dlatego zostają zarejestrowane odpowiednie zmiany – wiąże się to z zablokowaniem odpowiednich segmentów:

Segment	#	*j*	*a*	*b*	*ł*	*k*	*o*	#
Indeks	0	1^	2^	3^	4^	5^	6^	7

Następnie zarejestrowane zmiany zostają wykonane i następuje usunięcie blokad:

Segment	#	/j/	/a/	/b/	/w/	/k/	/o/	#
Indeks	0	1	2	3	4	5	6	7

W WSD została zdefiniowana jedna reguła (związana ze zjawiskiem desonoryzacji), która jest odpowiednia dla tego wyrazu:

/b/.[SO2].[WB2]>/p/.!NC.!NC;

Rejestracja zmiany wynikającej z tej reguły powoduje zablokowanie segmentu /b/:

Segment	#	/j/	/a/	/b/	/w/	/k/	/o/	#
Indeks	0	1	2	3^	4	5	6	7

Zarejestrowana zmiana zostaje zrealizowana, a przetwarzany wyraz przybiera następującą postać:

Segment	#	/j/	/a/	/p/	/w/	/k/	/o/	#
Indeks	0	1	2	3	4	5	6	7

Następnie wyraz jest przekazywany do WM, która zawiera regułę dotyczącą uproszczenia w grupie spółgłoskowej:

/b/./w/./k/>!NC.!RM.!NC;

W jej konstrukcji zostało użyte polecenie !RM, dlatego rejestrowana jest następująca zmiana:

Segment przed zmianą	/w/
Indeks przed zmianą	4
Segment po zmianie	EMP
Indeks po zmianie	4

Rejestracja tej zmiany wiąże się z zablokowaniem segmentu /w/:

Segment	#	/j/	/a/	/p/	/w/	/k/	/o/	#
Indeks	0	1	2	3	4^	5	6	7

WM nie zawiera żadnych innych reguł, które są odpowiednie dla tego wyrazu. Jedyna zarejestrowana zmiana zostaje wykonana:

Segment	#	/j/	/a/	/p/	EMP	/k/	/o/	#
Indeks	0	1	2	3	4	5	6	7

Jest to ostateczna transkrypcja przetwarzanego wyrazu, przy czym projekcja tej transkrypcji na zapis ortograficzny wynika z pierwotnego przypisania indeksów do liter:

Segment	#	/j/	/a/	/p/	EMP	/k/	/o/	#
Indeks	0	1	2	3	4	5	6	7
Zapis ort.		*j*	*a*	*b*	*ł*	*k*	*o*	

4.5.14. Przetwarzanie wyrazu: *barszcz* (w kontekście udźwięczniającym)

W czasie przetwarzania danego wyrazu może być uwzględniona struktura nagłosu następnego wyrazu. Jest on przetwarzany w pierwszej kolejności i w wyniku tego procesu może być wygenerowana określona etykieta. Jeżeli te dwa wyrazy są rozdzielone spacją, to etykieta zostaje dołączona do aktualnie przetwarzanego wyrazu (w ramach konkretnej warstwy) jako dodatkowy segment. Etykiety mogą zawierać informację między innymi o międzywyrazowym kontekście udźwięczniającym (tutaj przyjęto skrót UDZ) lub o międzywyrazowym kontekście ubezdźwięczniającym (skrót UBDZ). Zagadnienie to zostało szczegółowo omówione w podrozdziale 5.6.

Przykład dotyczy przetwarzania grupy umiejscowionej w wygłosie wyrazu i złożonej z dwóch spółgłosek właściwych bezdźwięcznych. Grupa taka może ulec sonoryzacji. Struktura wyrazu *barszcz* przekazanego do WTP jest następująca:

Segment	#	*b*	*a*	*r*	*s*	*z*	*c*	*z*	#
Indeks	0	1	2	3	4	5	6	7	8

WTP zawiera dwie reguły dwuelementowe właściwe dla tego wyrazu:

s.z>/S/.!ML;
c.z>/tS/.!ML;

Ich obecność sprawia, że zostają zarejestrowane zmiany:

Segment przed zmianą	*s*	*z*
Indeks przed zmianą	4	5
Segment po zmianie	/S/	–
Indeks po zmianie	4,5	–

Segment przed zmianą	*c*	*z*
Indeks przed zmianą	6	7
Segment po zmianie	/tS/	–
Indeks po zmianie	6,7	–

Jednocześnie zostają zablokowane segmenty:

Segment	#	*b*	*a*	*r*	*s*	*z*	*c*	*z*	#
Indeks	0	1	2	3	4^	5^	6^	7^	8

Poza tym WTP zawiera reguły jednoelementowe właściwe dla tego wyrazu:

b>/b/;
a>/a/;
r>/r/;

Następuje rejestracja zmian związanych z tymi regułami oraz blokowane są kolejne segmenty:

Segment	#	*b*	*a*	*r*	*s*	*z*	*c*	*z*	#
Indeks	0	1^	2^	3^	4^	5^	6^	7^	8

Zarejestrowane zmiany zostają wykonane. Przed przekazaniem wyrazu do WSD tymczasowo zostaje do niego dołączony segment zawierający etykietę UDZ. Zakłada się, że etykieta została uzyskana w wyniku przetworzenia następnego wyrazu:

Segment	#	/b/	/a/	/r/	/S/	/tS/	#	UDZ
Indeks	0	1	2	3	4,5	6,7	8	9

WSD zawiera regułę umożliwiającą modyfikację dźwięczności w wygłosowej grupie spółgłoskowej:

[SO2].[WB2].[WB2].#.UDZ>!NC.!LVO.!LVO.!NC.!NC;

W konstrukcji tej reguły został użyty mechanizm list odwzorowań dla zmiany statusu dźwięczności sekwencji segmentów drugiego stopnia, dlatego rejestrowane są zmiany:

Segment przed zmianą	/S/	/tS/
Indeks przed zmianą	4,5	6,7
Segment po zmianie	/Z/	/dZ/
Indeks po zmianie	4,5	6,7

Powoduje to blokadę segmentów:

Segment	#	/b/	/a/	/r/	/S/	/tS/	#	UDZ
Indeks	0	1	2	3	4^,5^	6^,7^	8	9

Następnie zarejestrowane zmiany są wykonywane. Przed przekazaniem wyrazu do WM następuje zwolnienie blokad oraz usunięcie tymczasowego segmentu UDZ:

Segment	#	/b/	/a/	/r/	/Z/	/dZ/	#
Indeks	0	1	2	3	4,5	6,7	8

Do WM nie włączono żadnych reguł, które mogłyby być użyte w przetwarzanym wyrazie. Podana transkrypcja jest więc ostateczna, a razem z projekcją na zapis ortograficzny prezentuje się następująco:

Segment	#	/b/	/a/	/r/	/Z/	/dZ/	#
Indeks	0	1	2	3	4,5	6,7	8
Zapis ort.		*b*	*a*	*r*	*s,z*	*c,z*	

4.5.15. Przetwarzanie wyrazu: *mierzw* (w kontekście ubezdźwięczniającym)

Wyraz *mierzw* zawiera w wygłosie strukturę *rzw*, która domyślnie jest dźwięczna (zob. podrozdz. 1.4). Przykład dotyczy międzywyrazowego kontekstu ubezdźwięczniającego. W przekazywanym do WTP wyrazie przypisanie indeksów do znaków ortograficznych jest następujące:

Segment	#	*m*	*i*	*e*	*r*	*z*	*w*	#
Indeks	0	1	2	3	4	5	6	7

WTP zawiera dwie reguły dwuelementowe właściwe dla tego wyrazu:

m.i.[SA1]–*i*>!NC./j/.!NC;
r.z>/rz/.!ML;

Na ich podstawie rejestrowane są zmiany:

Segment przed zmianą	*i*
Indeks przed zmianą	2
Segment po zmianie	/j/
Indeks po zmianie	2

Segment przed zmianą	*r*	*z*
Indeks przed zmianą	4	5
Segment po zmianie	/rz/	–
Indeks po zmianie	4,5	–

Rejestracja tych zmian powoduje blokadę trzech segmentów:

Segment	#	*m*	*i*	*e*	*r*	*z*	*w*	#
Indeks	0	1	2^	3	4^	5^	6	7

Poza tym WTP zawiera reguły jednoelementowe, które są odpowiednie dla transkrypcji niezablokowanych segmentów:

m>/m/;
e>/e/;
w>/v/;

Zmiany zostają zarejestrowane i następuje blokada pozostałych segmentów:

Segment	#	*m*	*i*	*e*	*r*	*z*	*w*	#
Indeks	0	1^	2^	3^	4^	5^	6^	7

Powyższe zmiany zostają wykonane i następuje usunięcie blokad. Przed przekazaniem do WSD do wyrazu zostaje dołączony segment zawierający etykietę UBDZ (zakłada się, że etykieta została wygenerowana w czasie przetwarzania następnego wyrazu):

Segment	#	/m/	/j/	/e/	/rz/	/v/	#	UBDZ
Indeks	0	1	2	3	4,5	6	7	8

WSD zawiera jedną regułę, którą można użyć w przetwarzanym wyrazie:

/rz/./v/.#.UBDZ>/S/./f/.!NC.!NC;

Na jej podstawie zostają zarejestrowane następujące modyfikacje:

Segment przed zmianą	/rz/	/v/
Indeks przed zmianą	4,5	6
Segment po zmianie	/S/	/f/
Indeks po zmianie	4,5	6

Powoduje to zablokowanie dwóch segmentów:

Segment	#	/m/	/j/	/e/	/rz/	/v/	#	UBDZ
Indeks	0	1	2	3	4^,5^	6^	7	8

Ponieważ WSD nie zawiera żadnych innych reguł, które są odpowiednie dla przetwarzanego wyrazu, zarejestrowane zmiany zostają wykonane, następuje zwolnienie blokad segmentów oraz usunięcie segmentu tymczasowego UBDZ:

Segment	#	/m/	/j/	/e/	/S/	/f/	#
Indeks	0	1	2	3	4,5	6	7

WM nie zawiera reguł odpowiednich dla przetwarzania takiej struktury, dlatego podana transkrypcja jest końcowa. Jej projekcja na zapis ortograficzny (przy zachowaniu początkowych wartości indeksów przypisanych do poszczególnych liter) wygląda następująco:

Segment	#	/m/	/j/	/e/	/S/	/f/	#
Indeks	0	1	2	3	4,5	6	7
Zapis ort.		*m*	*i*	*e*	*r,z*	*w*	

4.5.16. Przetwarzanie wyrazu: *liczb* (dwie możliwości kontekstów)

Ostatni omawiany przykład dotyczy wyrazu *liczb*. Zawiera on diadę ortograficzną kodującą w przyjętej transkrypcji podstawowej spółgłoskę właściwą bezdźwięczną oraz literę oznaczającą spółgłoskę właściwą dźwięczną (zob. podrozdz. 1.3 i 1.4). Poniżej podano początkowe przyporządkowanie indeksów w tym wyrazie:

Segment	#	*l*	*i*	*c*	*z*	*b*	#
Indeks	0	1	2	3	4	5	6

WTP zawiera jedną odpowiednią regułę dwuelementową:

c.z>/tS/.!ML;

Na podstawie tej reguły jest rejestrowana zmiana:

Segment przed zmianą	*c*	*z*
Indeks przed zmianą	3	4
Segment po zmianie	/tS/	–
Indeks po zmianie	3,4	–

Powoduje to zablokowanie dwóch segmentów:

Segment	#	*l*	*i*	*c*	*z*	*b*	#
Indeks	0	1	2	3^	4^	5	6

Do WTP włączono też reguły jednoelementowe, które są odpowiednie dla tego wyrazu:

l>/l/;
i>/i/;
b>/b/;

Dlatego kolejne segmenty zostają zablokowane:

Segment	#	*l*	*i*	*c*	*z*	*b*	#
Indeks	0	1^	2^	3^	4^	5^	6

Zarejestrowane zmiany zostają wykonane. Przed przekazaniem wyrazu do WSD następuje zniesienie blokad segmentów. Dalszy opis obejmuje dwa warianty kontekstu prawostronnego. Wiąże się to z dodaniem segmentu zawierającego etykietę UDZ lub etykietę UBDZ:

Segment	#	/l/	/i/	/tS/	/b/	#	UDZ
Indeks	0	1	2	3,4	5	6	7

Segment	#	/l/	/i/	/tS/	/b/	#	UBDZ
Indeks	0	1	2	3,4	5	6	7

WSD zawiera dwie odpowiednie reguły. Każda z nich odnosi się do jednego z omawianych tutaj wariantów. Reguła dotycząca zjawiska sonoryzacji międzywyrazowej ma następującą postać:

/tS/./b/.#.UDZ>/dZ/./b/.!NC.!NC;

Na podstawie tej reguły zostaje zarejestrowana zmiana:

Segment przed zmianą	/tS/
Indeks przed zmianą	3,4
Segment po zmianie	/dZ/
Indeks po zmianie	3,4

Sprawia to, że segment /tS/ zostaje zablokowany:

Segment	#	/l/	/i/	/tS/	/b/	#	UDZ
Indeks	0	1	2	3^,4^	5	6	7

Po wykonaniu tej modyfikacji następuje zniesienie blokady oraz usunięcie segmentu tymczasowego UDZ:

Segment	#	/l/	/i/	/dZ/	/b/	#
Indeks	0	1	2	3,4	5	6

WSD nie zawiera żadnych innych reguł właściwych dla przetwarzanego wyrazu, dlatego zostaje on przekazany do WM. Również ta warstwa nie zawiera żadnych odpowiednich reguł, dlatego podana transkrypcja jest ostateczna. Projekcja tej transkrypcji na początkowy zapis ortograficzny jest następująca:

Segment	#	/l/	/i/	/dZ/	/b/	#
Indeks	0	1	2	3,4	5	6
Zapis ort.		*l*	*i*	*c,z*	*b*	

Drugi wariant modyfikacji wyrazu w WSD (w międzywyrazowym kontekście ubezdźwięczniającym) polega na desonoryzacji segmentu /b/ drugiego stopnia:

/tS/./b/.#.UBDZ>!NC./p/.!NC.!NC;

Obecnosć tej reguły sprawia, że zostaje zarejestrowana zmiana:

Segment przed zmianą	/b/
Indeks przed zmianą	5
Segment po zmianie	/p/
Indeks po zmianie	5

Powoduje to zablokowanie segmentu /b/:

Segment	#	/l/	/i/	/tS/	/b/	#	UBDZ
Indeks	0	1	2	3,4	5^	6	7

Po wykonaniu zarejestrowanej modyfikacji następuje zniesienie blokady oraz usunięcie segmentu UBDZ:

Segment	#	/l/	/i/	/tS/	/p/	#
Indeks	0	1	2	3,4	5	6

Podana transkrypcja jest ostateczna. Projekcja tej transkrypcji na początkowy zapis ortograficzny wygląda następująco:

Segment	#	/l/	/i/	/tS/	/p/	#
Indeks	0	1	2	3,4	5	6
Zapis ort.		*l*	*i*	*c,z*	*b*	

5. Zaawansowane zagadnienia związane z WMT

5.1. Wprowadzenie

Ten rozdział jest poświęcony zaawansowanym zagadnieniom związanym z wielowarstwowym modelem transkrypcji. Opisano w nim między innymi mechanizm list odwzorowań, który był wielokrotnie używany w omówionym projekcie systemu transkrypcji (zob. rozdz. 4). Dzięki temu mechanizmowi można znacznie zredukować liczbę reguł. W rozdziale przedstawiono też koncepcję systemów probabilistycznych. W ich ramach istnieje możliwość tworzenia reguł opartych na podejściu statystycznym. Inna zaprezentowana koncepcja obejmuje systemy generatywne. Główne założenie dla takich systemów dotyczy możliwości definiowania alternatywnych względem siebie zbiorów reguł. Informacje te mogą posłużyć do rozbudowy przedstawionego systemu transkrypcji. W tym rozdziale omówiono również metodę generowania etykiet dla zależności międzywyrazowych. Przedstawiono również opcję załączania dodatkowych informacji do segmentów. Ostatnia część niniejszego rozdziału dotyczy zapisywania projektu opartego na WMT w pliku tekstowym – przedstawiono propozycję struktury takich plików. Ta struktura nie jest częścią omawianego modelu (podobnie jak program komputerowy, który umożliwia realizację założeń związanych z WMT).

5.2. Listy odwzorowań

Mechanizm list odwzorowań może być używany przy konstruowaniu reguł. Umożliwia on definiowanie zbioru przekształceń dla różnych wariantów segmentu, przy czym z góry nie wiadomo, jaki segment będzie częścią konkretnego przetwarzanego wyrazu. Opis tego mechanizmu można poprzeć przykładem. W podrozdziale 4.3.22 (w tabeli 4.5) zawarto podsumowanie odnoszące się między innymi do struktur obejmujących dwa segmenty drugiego stopnia, które są spółgłoskami właściwymi bezdźwięcznymi i są umiejscowione w kontekście udźwięczniającym. Dotyczy to następujących grup spółgłoskowych: /f.t.b/, /k.S.Z/, /k.s'.Z/, /k.tS.Z/, /l.f.s.b/, /p.s.b/, /p.S.Z/, /p.tS.Z/, /p.ts'.Z/, /s'.ts'.Z/, /s.k.dZ/, /S.tS.b/, /s'.ts'.dz'/, /S.tS.Z/, /t.tS.Z/, /ts.tS.Z/. Zamiast tworzenia oddzielnej reguły dla każdej wymienionej grupy w tabeli 4.5 zawarto jedną regułę alternatywną:

[WB2].[WB2].[WD2]>!LVO.!LVO.!NC;

Żaden element w pierwszym członie tej reguły nie zawiera etykiety konkretnego segmentu – wszystkie są nazwami zdefiniowanych zbiorów: WB2 lub WD2. W drugim członie reguły dwukrotnie występuje polecenie !LVO – powoduje ono wywołanie listy odwzorowań o nazwie VO w odniesieniu do analogicznego elementu umiejscowionego w pierwszym członie reguły. Struktura wyrażona przy użyciu nazw zbiorów odnosi się do segmentów należących do konkretnych wyrazów. Segmenty te są przekształcane zgodnie z utworzoną listą VO. Struktura reguły nie określa z góry etykiety segmentu, który będzie modyfikowany – określa ona jedynie przynależność tego segmentu do danego zbioru. Definiując listę odwzorowań, trzeba określić wszystkie przekształcenia, które są możliwe w ramach używanego zbioru (muszą być uwzględnione wszystkie elementy tego zbioru). Przykładowa lista VO obejmuje przekształcenia bezdźwięcznych segmentów drugiego stopnia na ich dźwięczne odpowiedniki:

/p/>/b/;
/t/>/d/;
/k/>/g/;
/tS/>/dZ/;
/ts/>/dz/;
/ts'/>/dz'/;
/f/>/v/;
/x/>/x/;
/s/>/z/;
/S/>/Z/;
/s'/>/z'/;

Odwrotne przekształcenia są zawarte w definicji listy VL. Umożliwia ona zamianę dźwięcznych segmentów drugiego stopnia na ich bezdźwięczne odpowiedniki:

```
/b/>/p/
/d/>/t/;
/g/>/k/;
/dZ/>/tS/;
/dz/>/ts/;
/dz'/>/ts'/;
/v/>/f/;
/z/>/s/;
/Z/>/S/;
/z'/>/s'/;
```

Sposób zapisywania list odwzorowań w pliku przechowującym definicję systemu transkrypcji został omówiony w części 5.7. W poniższym podsumowaniu wymieniono najważniejsze informacje dotyczące omawianego mechanizmu:

– każda lista odwzorowań ma swoją nazwę, w której nie mogą być użyte znaki specjalne oraz znaki diakrytyczne;
– odwołanie do określonej listy odwzorowań jest zawarte w drugim członie reguły i odnosi się ono do analogicznego elementu umiejscowionego w pierwszym członie reguły, przy czym ten element musi być zdefiniowanym zbiorem segmentów;
– listy odwzorowań są wywoływane poprzez użycie symbolu !L połączonego z nazwą zdefiniowanej listy;
– definicja listy odwzorowań zawiera zbiór par etykiet segmentów. Pierwszy element każdej pary dotyczy segmentu przed zamianą, natomiast drugi element określa etykietę tego segmentu po modyfikacji;
– definicja listy odwzorowań użytej w drugim członie danej reguły musi zawierać przekształcenia wszystkich elementów zbioru użytego na analogicznej pozycji w pierwszym członie reguły.

5.3. Systemy probabilistyczne

W podrozdziale 3.3 omówiono strukturę reguł, które odnoszą się do wyrazów zapisanych w sposób właściwy dla danej warstwy. Reguły mają statyczny charakter. Oznacza to, że po spełnieniu określonych warunków są zawsze stosowane. W WMT przewidziano również możliwość tworzenia reguł opartych na podejściu statystycznym. Trzeba podkreślić, że w czasie pisania tej książki nie istniał konkretny system probabilistyczny; informacje przedstawione w tej części mają zatem charakter koncepcyjny. Ustalenie zasad i składni tworzenia reguł daje możliwość projektowania takich systemów w przyszłości.

Istnieją dwa warianty budowy reguł opartych na rachunku prawdopodobieństwa. Pierwszy polega na określeniu prawdopodobieństwa (mniejszego niż 1) dla konkretnej reguły (niezależnie od prawdopodobieństwa ustalonego dla innych reguł). Przed średnikiem, który zamyka zapis danej reguły, powinna być podana w nawiasach liczba całkowita z zakresu od 1 do 100. Wartość ta oznacza setne części – na przykład wartość 50 oznacza, że prawdopodobieństwo użycia reguły wynosi 0.5. Takie rozwiązanie jest właściwe dla modelu wielowarstwowego i można je stosować między innymi w odniesieniu do fakultatywnych zmian uwzględnionych w warstwie modyfikacyjnej omówionego projektu. Na przykład do warstwy transkrypcji podstawowej może zostać włączona reguła, która umożliwia uzyskanie pełnej transkrypcji grupy spółgłoskowej /b.w.k/ (jak w wyrazie *jabłko*). W warstwie modyfikacyjnej natomiast może zostać ustalona reguła dotycząca uproszczenia zapisu tej grupy (poprzez użycie polecenia !RM w odniesieniu do segmentu /w/). Jeżeli prawdopodobieństwo wykonania tej reguły zostałoby ustalone na 0.95 (wartość 95 podana w nawiasach), to prawdopodobieństwo zachowania transkrypcji niezmienionej wyniosłoby 0.05. Konstrukcja tej reguły byłaby więc następująca:

```
/b/./w/./k/>!NC.!RM.!NC(95);
```

Druga możliwość budowy reguł opartych na rachunku prawdopodobieństwa polega na ustaleniu prawdopodobieństwa dla kilku (co najmniej dwóch) wariantów konstrukcji drugiego członu w ramach jednej reguły. Takie rozwiązanie jest właściwe, jeżeli w ramach jednej warstwy istnieje kilka możliwości określonego przekształcenia, ale tylko jedna z nich może być zrealizowana. Po utworzeniu pierwszego członu reguły trzeba utworzyć

co najmniej dwie wersje drugiego członu i podać prawdopodobieństwo dla każdej możliwości transkrypcji (liczba od 1 do 100). Suma tych wartości (wszystkich wariantów) nie może przekroczyć 100. Może być jednak mniejsza niż 100. Jeżeli suma wyniosłaby przykładowo 80, to prawdopodobieństwo tego, że żaden wariant reguły nie zostanie użyty, byłoby równe 0.2. Poniżej podano przykład reguły omawianego typu. Zastosowanie takiej konstrukcji dałoby identyczny rezultat jak użycie reguły już omówionej w tej części:

/b/./w/./k/>!NC.!RM.!NC(95),!NC.!NC.!NC(5);

Trzeba zaznaczyć, że jeżeli prawdopodobieństwo dla danej reguły nie zostało określone (jak we wszystkich przykładach omówionych w rozdziale trzecim oraz w rozdziale czwartym), to przyjmuje ono domyślną wartość 1. Reguła taka na pewno zostanie użyta, jeśli będą spełnione podstawowe warunki dotyczące stosowania reguł.

5.4. Systemy generatywne

W tej części została omówiona koncepcja systemów generatywnych. Główne założenie związane z tymi systemami dotyczy możliwości generowania konkretnych systemów konwersji tekstu na transkrypcję, które spełniają określone założenia. Rozwiązanie polega na tworzeniu alternatywnych względem siebie grup reguł. Do jednego systemu generatywnego można włączyć dowolną liczbę zbiorów alternatywnych grup reguł (co najmniej jeden). W czasie pisania tej książki nie istniał gotowy system generatywny oparty na WMT. Uściślenie i sformalizowanie założeń dotyczących funkcji oraz sposobu zapisywania takich systemów umożliwi w przyszłości tworzenie ujednoliconych projektów.

Zakłada się, że w czasie użytkowania systemu generatywnego z każdego zdefiniowanego zbioru grup musi być wybrana dokładnie jedna grupa reguł. W ten sposób można utworzyć standardowy system transkrypcji oparty na WMT. Warto zauważyć, że suma wszystkich (różnych) systemów transkrypcji, które można utworzyć przy użyciu systemu generatywnego, jest równa iloczynowi mocy wszystkich zbiorów grup reguł.

Przykładowy schemat ujęty w tabeli 5.1 dotyczy systemu generatywnego, w którym zdefiniowano 3 zbiory grup reguł. Pierwszy zbiór zawiera 3 grupy reguły, drugi zbiór obejmuje 5 grup, natomiast w zbiorze trzecim są zdefiniowane 4 grupy reguł. Przy użyciu takiego systemu można wygenerować 60 różnych systemów transkrypcji. Trzeba zaznaczyć, że dany system generatywny może zawierać dowolną liczbę reguł nienależących do zdefiniowanych zbiorów. Reguły takie są dołączane do każdego systemu transkrypcji, który zostanie wygenerowany.

Tabela 5.1 – Schemat struktury przykładowego systemu generatywnego

Zbiory grup reguł	Grupy reguł	Konkretne reguły w ramach danej grupy reguł
zbiór grup reguł 1	grupa reguł 1	reguła 1, reguła 2, ..., reguła n
	grupa reguł 2	reguła 1, reguła 2, ..., reguła n
	grupa reguł 3	reguła 1, reguła 2, ..., reguła n
zbiór grup reguł 2	grupa reguł 1	reguła 1, reguła 2, ..., reguła n
	grupa reguł 2	reguła 1, reguła 2, ..., reguła n
	grupa reguł 3	reguła 1, reguła 2, ..., reguła n
	grupa reguł 4	reguła 1, reguła 2, ..., reguła n
	grupa reguł 5	reguła 1, reguła 2, ..., reguła n
zbiór grup reguł 3	grupa reguł 1	reguła 1, reguła 2, ..., reguła n
	grupa reguł 2	reguła 1, reguła 2, ..., reguła n
	grupa reguł 3	reguła 1, reguła 2, ..., reguła n
	grupa reguł 4	reguła 1, reguła 2, ..., reguła n

W części 5.7 przedstawiono propozycję dotyczącą sposobu zapisywania w pliku definicji systemu generatywnego. W dalszym ciągu tej części omówiono przykłady zbiorów grup reguł, które mogłyby zostać włączone do systemu generatywnego. W opisie odwoływano się do konstrukcji reguł w systemie transkrypcji przedsta-

wionym w rozdziale czwartym. Informacje, które znajdują się w dalszym ciągu niniejszej części, można uznać za opcje rozbudowy tego projektu.

Przed omówieniem przykładów trzeba wspomnieć o istotnych założeniach. Każdy zbiór grup reguł ma swoją indywidualną nazwę (identyfikator). Poza tym każda grupa reguł zdefiniowana w ramach danego zbioru również ma swój identyfikator. Funkcjonalność systemu generatywnego jest ograniczona do generowania standardowych zbiorów reguł stanowiących konkretne systemy transkrypcji (nie wykonuje się transkrypcji tekstu przy użyciu systemu generatywnego). Identyfikatory zbiorów grup reguł oraz identyfikatory grup w ramach tych zbiorów pełnią kluczową rolę w czasie użytkowania systemu generatywnego. Użytkowanie polega na wyborze jednej grupy reguł z każdego zdefiniowanego zbioru. Istotny jest również fakt, że dany zbiór grup reguł zasadniczo jest definiowany w ramach konkretnej warstwy, jednak nie jest to formalny wymóg – istnieje możliwość zdefiniowania grupy zawierającej reguły należące do różnych warstw.

Pierwszy prezentowany przykład dotyczy zbioru grup reguł przeznaczonych dla przetwarzania segmentu /ą/ trzeciego stopnia. Grupy w tym zbiorze zawierają reguły włączone do warstwy modyfikacyjnej. Zgodnie z informacjami podanymi w podrozdziale 2.2.3 istnieje kilka możliwości transkrypcji, a w wypadku możliwości podstawowej nie jest uwzględniony wpływ kontekstu poprzedzającego. W omawianym przykładzie odniesiono się do informacji zawartych w ósmym wierszu tabeli 2.2 oraz w tabelach 2.3 i 2.4. Pierwsza grupa reguł (w ramach zbioru grup reguł) ma strukturę opartą na informacji zapisanej w wierszu ósmym w tabeli 2.2. Jest to zatem grupa jednoelementowa i zawiera ona następującą regułę:

```
/ą/./s'/+/z'/+/f/+/v/+/s/+/z/+/S/+/Z/+/x/=/o/&/w~/.!NC;
```

Druga grupa reguł należąca do omawianego zbioru zawiera reguły utworzone na podstawie tabeli 2.3 (po stosownych modyfikacjach związanych z notacją trzeciego stopnia). W takim wariancie kontekst lewostronny również nie jest uwzględniony; zakłada się natomiast, że rezonans nosowy miękki występuje tylko przed fonemem /s'/ oraz przed fonemem /z'/:

```
/ą/./s'/+/z'/>/o/&/n'/.!NC;
/ą/./f/+/v/+/S/+/Z/+/s/+/z/+/x/=/o/&/w~/.!NC;
```

Ostatnia grupa dotyczy wariantu, przy którym zakłada się, że rezonans nosowy miękki występuje tylko po spółgłoskach palatalnych oraz przed fonemami: /s'/ i /z'/. Możliwość ta została ujęta w tabeli 2.4. Reguły zawarte w niej wymagały znacznych modyfikacji (w celu dostosowania ich struktury do transkrypcji trzeciego stopnia):

```
/i/./ą/./s'/+/z'/>!NC./o/&/n'/.!NC;
/ą/./s'/+/z'/+/f/+/v/+/S/+/Z/+/s/+/z/+/x/=/o/&/w~/.!NC;
```

Pierwsza reguła jest trójelementowa zatem ma wyższy priorytet niż reguła druga (obie reguły obejmują prawostronny kontekst: /s'/ oraz /z'/).

Systemy generatywne dają nieograniczone możliwości tworzenia reguł. Poniżej wymieniono kilka propozycji ich zastosowania. W projekcie systemu generatywnego można uwzględnić przekształcenia związane ze statusem dźwięczności – oddzielne grupy reguł mogłyby dotyczyć wymowy krakowskopoznańskiej oraz wymowy warszawskiej. Inna możliwość polega na zawarciu w poszczególnych grupach reguł statystycznych związanych ze zmianami statusu dźwięczności właściwymi dla różnych obszarów geograficznych. W projekcie systemu generatywnego można też uwzględnić różnice dotyczące tempa wymowy. Do poszczególnych grup reguł mogłyby być włączone reguły związane z określonym tempem wymowy – dzięki temu zostałyby uwzględnione różnice redukcji w grupach spółgłoskowych. Z informacji zawartych w niniejszym akapicie wynika, że funkcjonalność systemów generatywnych może być łączona z funkcjonalnością systemów probabilistycznych. W celu efektywnego wykorzystania tych możliwości niezbędne jest prowadzenie dodatkowych badań.

5.5. Załączanie dodatkowych informacji o segmentach

W czasie przetwarzania wyrazów do poszczególnych segmentów mogą być dołączane dodatkowe informacje w nawiasach klamrowych. Możliwość ta nie obejmuje segmentów, do których przypisano polecenie !ML. Informacje można natomiast bez ograniczeń dodawać do segmentów przepisywanych, do segmentów modyfikowanych

lub do segmentów usuwanych przy użyciu polecenia !RM (w takim wypadku dodatkowa informacja zostanie dołączona do segmentu pustego EMP). Treści umieszczone w nawiasach klamrowych przy danym segmencie nie uczestniczą w przetwarzaniu wyrazu w kolejnych warstwach. Poza tym, jeżeli w czasie przetwarzania w ramach określonej warstwy do danego segmentu zostanie dodana informacja i w czasie późniejszego przetwarzania w jednej z następnych warstw do tego segmentu zostanie dodana kolejna informacja, to informacja dołączona później zastąpi informację wcześniejszą. Jeżeli w czasie przetwarzania w ramach określonej warstwy do danego segmentu zostanie dołączona informacja w nawiasach klamrowych i później w odniesieniu do tego segmentu zostanie użyte polecenie !ML, to ta informacja zostanie usunięta łącznie z segmentem.

Opcja dołączania informacji do poszczególnych segmentów znacznie zwiększa możliwości WMT. Zagadnienie to wykracza poza definicję WMT, ponieważ model nie precyzuje, co konkretnie może być zapisane w nawiasach klamrowych. Nie jest też z góry określona składnia takiego zapisu. Jest jednak z niego wykluczony znak: > oraz nawiasy klamrowe (za wyjątkiem nawiasu otwierającego i nawiasu zamykającego dany zapis). Można wymienić kilka przykładowych zastosowań związanych z omawianą funkcją. W niektórych wypadkach funkcjonalność systemów probabilistycznych może okazać się niewystarczająca, ponieważ działanie poszczególnych reguł w tych systemach polega na wyborze najwyżej jednego wariantu przekształcenia (zgodnie z ustalonym prawdopodobieństwem). Jeżeli wraz zapisem segmentu miałaby być przekazana informacja o prawdopodobieństwie różnych transkrypcji, to da się to osiągnąć tylko poprzez zapisanie informacji w nawiasach klamrowych. Istnieje kilka możliwości takiego rozwiązania. Pierwsza polega na zapisaniu wariantu najbardziej prawdopodobnego w treści segmentu przy jednoczesnym podaniu w nawiasie klamrowym jego prawdopodobieństwa oraz prawdopodobieństw innych wariantów. Druga polega na zapisaniu w treści segmentu umownej etykiety, która informowałaby o jego wariantywności. Jednocześnie w nawiasach klamrowych należałoby zapisać informacje o prawdopodobieństwie poszczególnych wariantów transkrypcji. Istnieje znacznie więcej możliwości użycia omawianej funkcjonalności. Na przykład w nawiasach klamrowych mogą być umieszczane komentarze lub wskaźniki do danych zewnętrznych. Można też zaprojektować oddzielną warstwę przeznaczoną tylko i wyłącznie do dodawania komentarzy i innych informacji. Poniżej podano przykładową regułę. Pozwala ona na dodanie trzech komentarzy:

/s/./j/./i/.#>!NC{komentarz_1}.!RM{komentarz_2}.!NC{komentarz_3}.!NC;

5.6. Tworzenie reguł dla zależności międzywyrazowych

W tej części omówiono zasady dotyczące tworzenia reguł związanych z procesami fonetyki międzywyrazowej. Trzeba podkreślić, że procesy te nie obejmują tylko modyfikowania statusu dźwięczności segmentów (jak w przykładach omówionych w rozdziale czwartym). Mogą występować również inne rodzaje przekształceń, szczególnie na poziomie alofonicznym. Poza tym przyjmuje się, że w języku polskim oddziaływanie międzywyrazowe jest realizowane jednokierunkowo i ma związek ze zjawiskiem antycypacji artykulacyjnej (tutaj antycypacji międzywyrazowej). Dlatego wyrazy są przetwarzane w kolejności od strony prawej do strony lewej. Wynika to z faktu, że struktura nagłosu może mieć wpływ na zmiany w wygłosie wyrazu poprzedzającego. Nawiązując do ogólnego schematu budowy reguły zamieszczonego w tabeli 3.2 (w podrozdziale 3.3), można przedstawić składnię reguły, która pozwala na przekazywanie określonej etykiety do konkretnej warstwy wyrazu poprzedzającego.

Tabela 5.2 – Schemat budowy reguły umożliwiającej przekazywanie etykiety do wyrazu poprzedzającego dany wyraz

Pierwszy człon reguły			Drugi człon reguły
element 1	element 2	element n	!SEND:etykieta,nazwa_warstwy

Z tabeli 5.2 wynika, że konstrukcja pierwszego członu reguły niczym nie różni się od konstrukcji standardowych reguł WMT. Funkcja takiego zapisu jest jednak inna – służy on wyłącznie do identyfikowania odpowiednich wyrazów. W drugim członie reguły nie są definiowane żadne modyfikacje. Jego budowa jest całkowicie

inna niż budowa drugiego członu w standardowych regułach WMT. W skład tej konstrukcji wchodzi jedno polecenie !SEND. Po tym poleceniu (po dwukropku) podawane są dwie informacje rozdzielone przecinkiem: etykieta, która ma być przekazana, oraz nazwa warstwy, w której ta etykieta ma być dołączana do poprzedzającego wyrazu (w czasie jego przetwarzania). Istotne jest to, że:

– przekazywana etykieta jest dołączana do wyrazu poprzedzającego dany wyraz jako segment; zostaje zatem do niego przypisany indeks o 1 większy niż indeks przypisany do znacznika #;
– do reguły zawierającej polecenie !SEND może być dołączona lista ciągów ortograficznych umożliwiających identyfikację wyjątków (polecenie !EXC) lub włączeń (polecenie !INC);
– określona etykieta jest przekazywana do konkretnej warstwy w czasie przetwarzania wyrazu poprzedzającego dany wyraz. Nie może więc być obecna w czasie przetwarzania tego wyrazu w innych warstwach;
– dana reguła zawierająca polecenie !SEND należy do reguł związanych z określoną warstwą; w pierwszym członie tej reguły mogą więc być używane tylko segmenty i zbiory segmentów zdefiniowane w tej warstwie;
– do konkretnej warstwy związanej z przetwarzaniem wyrazu poprzedzającego dany wyraz może być przekazana tylko jedna etykieta. Ewentualne przekazanie kolejnej etykiety powoduje usunięcie etykiety przekazanej wcześniej;
– można tworzyć dowolne etykiety z zastrzeżeniem, że nie mogą one zawierać spacji oraz znaków specjalnych;
– zdefiniowana etykieta jest przekazywana do wyrazu poprzedzającego dany wyraz tylko jeżeli między tymi wyrazami znajduje się pusty odstęp (spacja). W innych wypadkach (jeżeli jest to: przecinek, kropka, średnik itp.) pozycja poprzedniego wyrazu jest traktowana jako wygłos absolutny.

Na podstawie informacji zawartych w tabeli 5.8 w *Asymilacji…* oraz na podstawie konstrukcji zbiorów segmentów trzeciego stopnia ustalonych w warstwie modyfikacyjnej omówionego przykładowego projektu (zob. podrozdz. 4.4) zostały utworzone reguły zawierające polecenie !SEND:

```
#.[WD3]>!SEND:UDZ,WSD;
#.[WB3]>!SEND:UBDZ,WSD;
#.[SO3]>!SEND:UBDZ,WSD;
#.[SA3]>!SEND:UBDZ,WSD;
```

Są one właściwe dla wymowy warszawskiej i mogą zostać włączone do zbioru reguł w warstwie modyfikacyjnej omówionego projektu. Reguły umożliwiają generowanie etykiet: UDZ lub UBDZ. Warunkiem użycia danej reguły jest jej zgodność ze strukturą nagłosu aktualnie przetwarzanego wyrazu (w warstwie modyfikacyjnej). Więcej informacji dotyczących zjawisk sonoryzacji oraz desonoryzacji wygłosu wyrazów zawarto w rozdziale czwartym.

5.7. Zapisywanie projektu opartego na WMT

Definicja konkretnego projektu systemu transkrypcji opartego na WMT jest przechowywana w pliku tekstowym. Trzeba zaznaczyć, że zaproponowane tutaj rozwiązanie dotyczące sposobu zapisywania danych nie jest częścią omawianego modelu – podobnie jak program komputerowy, który jest technicznym narzędziem umożliwiającym praktyczną realizację założeń WMT [por. Pluciński 1996]. Przyjęta struktura pliku musi zapewniać możliwość przechowywania wszystkich informacji, które są istotne dla WMT. Można zatem utworzyć program spełniający założenia WMT, z którym powinien być powiązany określony sposób zapisywania projektów systemów konwersji tekstu na transkrypcję.

Autor przyjął rozwiązanie polegające na sekwencyjnym zapisie danych w pliku tekstowym zgodnym z kodowaniem UTF-8. W rozwiązaniu tym etykiety oraz dane powiązane z nimi są umieszczane w kolejnych liniach. W przyszłości można również wprowadzić rozwiązanie oparte na języku XML. Istotny jest fakt, że dla transkrybowania konkretnych tekstów ortograficznych potrzebny jest program komputerowy, który spełnia założenia WMT oraz plik przechowujący kompletny zestaw danych dotyczących zdefiniowanych warstw i reguł.

Na początku pliku tekstowego zamieszczone są informacje ogólne dotyczące konkretnego projektu opartego na WMT. Dane są oznaczone odpowiednimi etykietami, przy czym przed każdą etykietą stoi wykrzyknik, a po każdej etykiecie jest zapisywany dwukropek. Każdy zapis jest zakończony średnikiem:

```
!TYPE:static;
!ENCODING:UTF8;
```

!AUTHOR:*dane autora;*
!NAME:*autorska nazwa systemu transkrypcji;*
!VERSION:*dane dotycz*ące wersji opracowanego systemu *transkrypcji;*
!DESCRIPTION:*opis systemu transkrypcji;*

Informacje dotyczące autora oraz systemu transkrypcji (nazwa, opis, informacja o wersji) pełnią rolę wyłącznie informacyjną i nie mają wpływu na działanie konkretnego rozwiązania opartego na WMT. Znacznik !ENCODING zawiera informację o standardzie kodowaniu pliku tekstowego. Dodatkowego wyjaśnienia wymaga znacznik !TYPE. Dopuszczalne słowa, które mogą znajdować się za tym znacznikiem, to: *static* lub *generative*. Oznaczają one, że dany projekt stanowi system statyczny lub system generatywny.

W pliku przeznaczonym dla przechowywania danych o konkretnym projekcie opartym na WMT muszą znaleźć się definicje wszystkich warstw. Definicja każdej warstwy rozpoczyna się od znacznika !LAYER. Za nim musi zostać umieszczona nazwa warstwy. Jest ona używana do identyfikowania poszczególnych reguł zapisanych w tym samym pliku. Nie może ona zawierać spacji i innych znaków specjalnych ani znaków diakrytycznych – zatem dopuszczalne są tylko litery łacińskie i cyfry. Kolejny znacznik: !DES ma charakter informacyjny – można po nim umieścić opis danej warstwy. Za znacznikiem !SET jest zapisana definicja zbioru segmentów używanego w regułach w ramach danej warstwy. Nazwa każdego zbioru jest ujęta w nawiasach kwadratowych. Następnie w nawiasach klamrowych (po znaku równości) są wyliczone wszystkie elementy danego zbioru, przy czym ich etykiety są rozdzielone przecinkami (bez użycia spacji). Definicję warstwy zamyka znacznik !END (oraz średnik). Oto przykładowa definicja warstwy, która jest zgodna z powyższym opisem:

!LAYER:WTP;
!DES:opis_warstwy;
!SET:[X1]={*a,e,i,o,y,u,ó,ę,ą,ł,j,l,r,m,n,ń,f,w,s,z,ż,ś,ź,h,p,b,t,d,k,g,c,ć*};
!SET:[SA1]={*a,e,i,o,y,u,ó,ę,ą*};
!SET:[SP1]={*ł,j,l,r,m,n,ń,f,w,s,z,ż,ś,ź,h,p,b,t,d,k,g,c,ć*};
!END;

Każda definicja warstwy może również zawierać dowolną liczbę definicji list odwzorowań. Poniżej zamieszczono zapis przykładowej listy VO, która dotyczy przekształceń segmentów bezdźwięcznych drugiego stopnia na ich dźwięczne odpowiedniki:

!LIST:VO={/p/>/b/,/t/>/d/,/k/>/g/,/tS/>/dZ/,/ts/>/dz/,/ts'/>/dz'/,/f/>/v/,/x/>/x/,/s/>/z/,/S/>/Z/,/s'/>/z'/};

Po zapisaniu informacji dotyczących warstw, zbiorów segmentów oraz list odwzorowań można przystąpić do definiowania reguł. Każdą regułę trzeba poprzedzić etykietą !R połączoną z nazwą warstwy, do której jest ona przypisana oraz ze znakiem dwukropka, na przykład:

!RWTP:*s.i.*[SA1]–i>/s'/.!ML.!NC;
!RWTP:*c.i.*[SA1]–i>/ts'/.!ML.!NC;
!RWTP:*z.i.*[SA1]–i>/z'/.!ML.!NC;

Nieistotna jest kolejność zapisywania reguł (reguły należące do różnych warstw mogą być przemieszane). Istotne jest tylko to, żeby definicje wszystkich warstw były umiejscowione przed definicjami reguł.

Można również przedstawić propozycję zapisywania reguł dotyczących koncepcyjnego systemu generatywnego (zob. podrozdz. 5.4). Definicja zbioru grup reguł rozpoczyna się od znacznika !SET, po którym podawana jest nazwa tego zbioru. Po znaczniku !SETDES można umieścić opis zbioru grup reguł. Definicja każdej grupy reguł rozpoczyna się od znacznika !GRP. Po znaczniku !GRPDES można umieścić opis grupy. W kolejnych wierszach wymieniane są wszystkie reguły należące do danej grupy reguł. Po zapisaniu wszystkich reguł (w ramach danej grupy reguł) można utworzyć kolejną definicję grupy reguł. Zapis: !END; kończy definicję zbioru grup reguł. Oto schemat zgodny z powyższym opisem:

!SET:nazwa_zbioru_grup_reguł;
!SETDES:opis zbioru grup;
!GRP:nazwa_grupy_reguł_1;
!GRPDES:opis_grupy_reguł_1;
reguła_1;

```
reguła_2;
reguła_3;
...
!GRP:nazwa_grupy_reguł_2;
GRPDES:opis_grupy_reguł_2;
reguła_1;
reguła_2;
reguła_3;
...
!END;
```

Podsumowanie

W niniejszej monografii została przedstawiona koncepcja wielowarstwowego modelu transkrypcji tekstu. Polega ona na stopniowym przetwarzaniu wyrazów ortograficznych. W kolejnych etapach działania algorytmu zapis tych wyrazów coraz bardziej upodabnia się do docelowej transkrypcji. W terminologii związanej z omówionym modelem etapy działania algorytmu są utożsamiane z kolejnymi warstwami. Każda warstwa obejmuje niezależny zbiór reguł. Model nie narzuca funkcji przypisanych do poszczególnych warstw, natomiast precyzuje on składnię, sposób działania oraz zasady tworzenia reguł. Na przykładzie projektu, który został omówiony w niniejszej publikacji, wykazano, że poszczególne warstwy można powiązać z określonymi problemami związanymi z transkrypcją tekstu (na przykład z asymilacją dźwięczności). Jest to główna zaleta omówionego rozwiązania.

W publikacji zostały przedstawione wyniki szczegółowych analiz związanych z fonologiczną transkrypcją wyrazów języka polskiego. Analizy zostały oparte na regułach pochodzących z monografii autorstwa Marii Steffen-Batogowej: *Automatyzacja transkrypcji fonematycznej tekstów polskich*. Poza tym reguły stanowiły dla autora główną inspirację zaprojektowania modelu wielowarstwowego. W projektach opartych na nowszej koncepcji poszczególne reguły mogą dotyczyć pojedynczych problemów, dlatego modyfikowanie zbiorów reguł stało się łatwiejsze.

Dzięki licznym przykładom omówionych w tej książce udało się wykazać skuteczność działania konkretnego projektu opartego na wielowarstwowym modelu transkrypcji. Ponadto w ostatnim rozdziale zostały przedstawione między innymi zagadnienia koncepcyjne. Pierwsze z nich dotyczy możliwości uwzględnienia w regułach prawdopodobieństwa różnych wariantów przekształceń. Drugie zagadnienie o charakterze koncepcyjnym obejmuje systemy generatywne. Główne założenie związane z tymi systemami dotyczy tworzenia alternatywnych względem siebie zbiorów reguł. W zbiorach tych można uwzględnić między innymi różne warianty wymowy, inwentarzy fonologicznych czy alfabetów transkrypcyjnych. Na podstawie jednego projektu systemu generatywnego da się utworzyć wiele standardowych systemów transkrypcji opartych na modelu wielowarstwowym. Systemy generatywne będą stanowiły kolejny etap rozwoju koncepcji uniwersalnych systemów transkrypcji tekstu.

Bibliografia

Bańczerowski J. 1982. *System fonemiczny.* [W:] Bańczerowski J., Pogonowski J., Zgółka T. *Wstęp do językoznawstwa. Skrypt dla studentów studiów uniwersyteckich*. Poznań, 176–182.

Bańko M. (red.) 2003. *Słownik wyrazów obcych PWN.* Warszawa.

Bargiełówna M. 1950. *Grupy fonemów spółgłoskowych współczesnej polszczyzny kulturalnej.* „Biuletyn Polskiego Towarzystwa Językoznawczego" 10, 1–25.

Bethin C.Y. 1992. *Polish Syllables. The Role of Prosody in Phonology and Morphology*. Columbus.

Biedrzycki L. 1963. *Fonologiczna interpretacja polskich głosek nosowych.* „Biuletyn Polskiego Towarzystwa Językoznawczego" 22, 25–45.

Biedrzycki L. 1978. *Fonologia angielskich i polskich rezonantów.* Warszawa.

Bralczyk J. (red.). 2005. *Słownik 100 tysięcy potrzebnych słów.* Warszawa. (Aktualizacja wersji online: Drabik L., sjp.pwn.pl/sjp).

Chomyszyn J. 1986. *A phonemic transcription program for Polish.* „International Journal of Man-Machine Studies" 25, 271–293.

Demenko G., Wypych M., Baranowska E. 2003. *Implementation of Grapheme-to-Phoneme Rules and Extended SAMPA Alphabet in Polish Text-to-Speech Synthesis.* „Speech and Language Technology" 7, 79–95.

Dobrogowska K. 1984. *Śródgłosowe grupy spółgłosek w polskich tekstach popularnonaukowych.* „Polonica" 10, 15–34.

Dobrogowska K. 1990. *Word internal consonant clusters in Polish artistic prose.* „Studia Phonetica Posnaniensia" 2, 43–67.

Dobrogowska K. 1992. *Word initial and word final consonant clusters in Polish popular science text and in artistic prose.* „Studia Phonetica Posnaniensia" 3, 47–121.

Doroszewski W., Wieczorkiewicz B. 1947. *Zasady poprawnej wymowy polskiej (ze słowniczkiem).* Warszawa.

Dukiewicz L. 1967. *Polskie głoski nosowe.* Warszawa.

Dukiewicz L. 1985. *Nagłosowe grupy spółgłosek w polskich tekstach popularnonaukowych i prasowych.* „Studia Gramatyczne" 6, 17–34.

Dukiewicz L., Sawicka I. 1995. *Fonetyka i fonologia*. Kraków.

Dunaj B. 1985. *Grupy spółgłoskowe współczesnej polszczyzny mówionej (w języku mieszkańców Krakowa).* Warszawa–Kraków.

Dunaj B. 1986. *Wygłosowe grupy spółgłoskowe współczesnej polszczyzny mówionej.* „ZN UJ: Prace Językoznawcze" 82, 103–117.

Dunaj B. 1991. *Dwa dyskusyjne problemy polskiej fonologii.* „Prace Językoznawcze UŚ" 19, 40–46.

Dunaj B. 1994. *Zmiany fonetyczne (i fonologiczne) w polszczyźnie po 1945 roku.* [W:] Gajda S., Adamiszyn Z. (red.) *Przemiany współczesnej polszczyzny*. Opole, 41–52.

Dunaj B. 2001a. *Rozwój systemu fonetyczno-fonologicznego polszczyzny w XX wieku.* [W:] Dubisz S., Gajda S. (red.) *Polszczyzna XX wieku. Ewolucja i perspektywy rozwoju*. Warszawa, 75−81.

Dunaj B. 2001b. *System fonetyczno-fonologiczny.* [W:] Gajda S. (red.) *Najnowsze dzieje języków słowiańskich. Język polski*. Opole, 65–75.

Dunaj B. 2006. *Zasady poprawnej wymowy polskiej.* „Język Polski" 86, 3, 161–172.

Dyszak A., Laskowska E., Żak-Święcicka M. 1997. *Fonetyczny i fonologiczny opis współczesnej polszczyzny*. Bydgoszcz.

Dziczek-Karlikowska H. 2012. *Error-Based Evidence for the Phonology of Glides and Nasals in Polish with Reference to English*. Frankfurt am Main.

Kamińska B. 2004. *Samogłoski nosowe w wymowie dzieci w wieku przedszkolnym*. Lublin.

Karaś M., Madejowa M. (red.) 1977. *Słownik wymowy polskiej PWN*. Warszawa−Kraków.

Klemensiewicz Z. 1930. *Prawidła poprawnej wymowy polskiej.* Kraków.

Lubaś W., Urbańczyk S. 1990. *Podręczny słownik poprawnej wymowy polskiej*. Warszawa.

Jassem W. 1960. *Wstępna analiza spektrograficzna głosek polskich.* „Rozprawy Elektrotechniczne" 6, 3, 333−361.

Jassem W. 1966. *The Distinctive Features and the Entropy of the Polish Phoneme System.* „Biuletyn Polskiego Towarzystwa Językoznawczego" 24, 87−108.

Kleśta J. 1998. *Lingwistyczne i paralingwistyczne zastosowania akustyczno-fonetycznej bazy danych samogłosek języka polskiego.* „Speech and Language Technology" 2, 47–63.

Gussmann E. 2007. *The Phonology of Polish*. Oxford.

Jassem K. 1996. *A phonemic transcription — syllable division rule engine*. „Proceedings of ONOMASTICA–COPERNICUS Research Colloquium Edinburgh", 106.

Jassem W. 1998. *An Acoustical Linear-Predictive and Statistical Discriminant Analysis of Polish Fricatives and Affricates*. „Speech and Language Technology" 2, 9–45.

Jassem W. 2003. *Illustrations of the IPA: Polish.* „Journal of the International Phonetic Association" 33, 1, 103–107.

Kaczmarek B.L.J. 2003. *Phonological Awareness and Ability to Read and Write*. [W:] Joshi R.M., Leong Ch.K., Kaczmarek B.L.J. (red.) *Literacy Acquisition. The Role of Phonology, Morphology and Orthography*, Amsterdam–Berlin–Oxford–Tokyo–Washington, 15–24.

Kłosowski P. 2016. *Algorithm and implementation of automatic phonemic transcription for polish*. „Proceedings of 20th IEEE International Conference Signal Processing Algorithms, Architectures, Arrangements and Applications", 298–303.
Koržinek D., Brocki Ł., Marasek K. 2016a. *Polish grapheme-to-phoneme tool and service*. CLARIN-PL digital repository: http://hdl.handle.net/11321/295.
Koržinek D., Marasek K., Brocki Ł., Wołk K. 2016b. *Polish Read Speech Corpus for Speech Tools and Services*. „Proceedings, CLARIN Annual Conference", 52–64.
Kuryłowicz J. 1952. *Uwagi o polskich grupach spółgłoskowych.* „Biuletyn Polskiego Towarzystwa Językoznawczego" 11, 54–69.
Klebanowska B. 1990. *Interpretacja fonologiczna zjawisk fonetycznych w języku polskim.* Warszawa.
Laskowski R. 1991. *System fonologiczny języka polskiego*. [W:] Urbańczyk S. (red.) *Encyklopedia języka polskiego*. Wrocław–Warszawa–Kraków, 344–347.
Lisowski T. 2001. *Grafia druków polskich z 1521 i 1522 roku. Problemy wariantywności i normalizacji.* Poznań.
Lorenc A., Ptaszkowska A. 2015. *Programowanie języka dziecka z uszkodzeniem słuchu z zastosowaniem metody audytywno-werbalnej. Studium przypadku*. „Logopedia Silesiana" 4, 2015, 229–247.
Lorenc A. 2016. *Wymowa normatywna polskich samogłosek nosowych i spółgłoski bocznej*. Warszawa.
Łobacz P. 1973. *Non-unique phonemic interpretation of the Polish speech sounds.* „Speech Analysis and Synthesis" 3, Warszawa, 53–74.
Łobacz P., Ciarkowski R. 1988. *Klasyfikacja fonemów polskich na podstawie skalowania wielowymiarowego*. Warszawa.
Łobacz P. 1999. *Problemy automatyzacji transkrypcji fonematycznej nazw własnych współczesnej polszczyzny.* [W:] Bańczerowski J., Zgółka T. (red.) *Linguam amicabilem facere. Ludovico Zabrocki in memoriam*. Poznań, 97−114.
Łobacz P. 2000. *The Polish Rhotic. A Preliminary Study in Acoustic Variability and Invariance.* „Speech and Language Technology" 4, 85−101.
Madejowa M. 1989. *Zasady wymowy polskiej.* „Biuletyn Audiofonologii" 1, 2–4, 69–83.
Madejowa M. 1990. *Modern Polish linguistic norm with special reference to the pronunciation of consonants.* „Studia Phonetica Posnaniensia" 2, 69–105.
Madejowa M. 1992. *Zasady współczesnej wymowy polskiej (w zakresie samogłosek nosowych i grup spółgłoskowych) oraz ich przydatność w praktyce szkolnej.* „Język Polski" 2–3, 187–198.
Madejowa M. 1993. *Normative rules of modern Polish pronunciation.* „Studia Phonetica Posnaniensia" 4, 19–30.
Maksymienko M., Bolc L. 1981. *Komputerowy system przetwarzania tekstów fonematycznych.* Warszawa.
Madelska L. 2005. *Słownik wariantywności fonetycznej współczesnej polszczyzny.* Kraków.
Mańczak W. 1988. *O nieregularnym rozwoju fonetycznym spowodowanym frekwencją.* „Biuletyn Polskiego Towarzystwa Językoznawczego" 41, 105−111.
Nawrocki G., Gonet W. 2004. *Realization of /x/ in intervocalic context in Southern Polish.* „Speech and Language Technology" 8, 87−106.
Nawrocki G. 2004. *Southern Polish /x/ − subphonemic variation. A study in female speakers.* „Speech and Language Technology" 8, 107−136.
Nowak I. 1995. *Transkrypcja fonematyczna tekstów polskich*. [W:] Pogonowski J. (red.) *Eufonia i logos*. Poznań.
Nowakowski P. 2002. *Fonetyka tekstu wtórnie mówionego (na przykładzie badań zróżnicowania współczesnej wymowy scenicznej).* „Poznańskie Spotkania Językoznawcze" 8, 91−116.
Nowakowski P., Wiatrowski P. 2013. *Informacje fonetyczno-ortofoniczne w podręczniku Tomasza Karpowicza Kultura języka polskiego. Wymowa, ortografia, interpunkcja.* „Slavia Occidentalis" 70/1, 87−100.
Nowakowski P., Wiatrowski P. 2020. *Kilka uwag o tak zwanych samogłoskach nosowych w języku polskim.* [W:] Strawińska A. B. (red.) *„Ojczysta święta mowo!... wiąże nas twoje słowo…". Polszczyzna w perspektywie diachronicznej*. Białystok, 371−382.
Osowicka-Kondratowicz M. 2005. *Assimilative palatalization within consonantal clusters in Polish.* „Studia Phonetica Posnaniensia" 7, 5−22.
Ostaszewska D., Tambor J. 2000. *Fonetyka i fonologia współczesnego języka polskiego.* Warszawa.
Owsianny M. 1998. *Zależność między barwą wokaliczną a częstotliwością podstawową F0 w percepcji samogłosek polskich.* „Speech and Language Technology" 2, 65–87.
Patryn R. 1988. *Phonetic-Acoustic Analysis of Polish Speech Sounds.* Warszawa.
Pluciński A. 1996. *Program transkrypcji fonematycznej polskich tekstów ortograficznych.* [W:] Basztura Cz., Dobrogowska K. (red.) *Podstawowe założenia fonetyczne i techniczne tłumaczenia różnojęzycznego dialogu w czasie rzeczywistym*, 121−149.
Pluciński A. 2002. *Rekonstrukcja reguł transkrypcji fonematycznej na podstawie próby uczącej.* Poznań.
Polański K. (red.). 2003. *Encyklopedia językoznawstwa ogólnego.* Wrocław–Warszawa–Kraków, 207.
Przybysz P., Kasprzak W. 2013. *The generation of letter-to-sound rules for grapheme-to-phoneme conversion*. „6th International Conference on Human System Interactions (HSI)".
Retz R.I. 1989. *Zanik korelacji palatalności we współczesnej polszczyźnie ogólnej.* „Poradnik Językowy" 1, 19−32; 2, 90−99; 3, 159−168.

Rocławski B. 1984. *Palatalność. Teoria i praktyka.* Gdańsk.
Rocławski B. 1986. *Zarys fonologii, fonetyki, fonotaktyki i fonostatystyki współczesnego języka polskiego.* Gdańsk.
Rocławski B. 1986. *Poradnik fonetyczny dla nauczycieli*. Warszawa.
Rocławski B. 2001. *Podstawy wiedzy o języku polskim dla glottodydaktyków, pedagogów, psychologów i logopedów.* Gdańsk.
Rubach J.J. 1984. *Cyclic and Lexical Phonology. The Structure of Polish.* Dordrecht.
Sawicka I. 1974. *Struktura grup spółgłoskowych w językach słowiańskich.* Wrocław.
Sawicka I. 1988. *Fonologia konfrontatywna polsko-serbsko-chorwacka*. Wrocław.
Sawicka I. 1995. *Fonologia*. [W:] H. Wróbel (red.) *Gramatyka współczesnego języka polskiego. Fonetyka i fonologia*, Kraków, 105–195.
Sawicka I., Grzybowski S. 1999. *Studia z palatalności w językach słowiańskich*. T. 1. Toruń.
Sawicka I. (red.) 2007. *Komparacja współczesnych języków słowiańskich 2, Fonetyka/Fonologia*. Opole.
Skulina T. 1964. *Rozwój grup spółgłosek zwarto-szczelinowych w języku polskim.* Poznań.
Skurzok D., Ziółko B., Ziółko M. 2015. *Ortfon2 – tool for orthographic to phonetic transcription*. „7th Language & Technology Conference". (Dostęp na stronie konferencji: http://ltc.amu.edu.pl/a2015/book/papers/SPEECH1-2.pdf).
Steffen-Batogowa M. 1973. *The problem of automatic phonemic transcription of written Polish*. „Biuletyn Fonograficzny" 14, 75−86.
Steffen-Batogowa M. 1975. *Automatyzacja transkrypcji fonematycznej tekstów polskich.* Warszawa.
Steffen-Batóg M., Nowakowski P. 1992. *An algorithm for phonetic transcription of orthographic texts in Polish.* „Studia Phonetica Posnaniensia" 3, 135−183. (Przedruk: Steffen-Batóg M. 1997. *Studies in Phonetic Algorithms*. Poznań, 93−142.)
Steffen-Batóg M. 1989−1990. *Rules for the mutual conversion of the phonemic and phonetic transcriptions of the Polish texts.* „Lingua Posnaniensis" 32−33, 211−223. (Przedruk: Steffen-Batóg M. 1997. *Studies in Phonetic Algorithms*. Poznań, 72−85.)
Stieber Z. 1966. *Historyczna i współczesna fonologia języka polskiego.* Warszawa.
Szczerbiński M. i in. 2012. *Online remediation of reading fluency problems: technology, pedagogy, psychology*. „6th European Conference on Game-Based Learning", Cork 4–5 October 2012 (poster).
Szpyra J. 1995. *Three tiers in Polish and English phonology*. Lublin.
Szpyra-Kozłowska J. 2007. *Inwentarze fonemów języka polskiego i ich konsekwencje.* „Logopedia" 31, 7–26.
Śledziński D. 2019. *Asymilacja dźwięczności w polskich grupach spółgłoskowych.* Poznań.
Śledziński D. 2022. I*dentyfikacja bifonematycznych wystąpień wybranych diad ortograficznych w języku polskim na potrzeby automatycznej transkrypcji tekstu*. „Biuletyn Polskiego Towarzystwa Językoznawczego", 78, 221−235. (w niniejszej monografii zostały użyte fragmenty tego artykułu w rozdziałach: 1, 2.5, 3)
Śledziński D. 2022. *A Multi-Layer Transcription Model – concept outline*. „Lingua posnaniensis" 64, 1, 49−71. (w niniejszej monografii zostały użyte fragmenty tego artykułu w rozdziałach: 1, 3, 4)
Warmus M. 1972. *Program na maszynę Odra-1204 dla automatycznej transkrypcji fonematycznej tekstów języka polskiego.* „Prace CO PAN" 66, 1−22.
Wells John C. 1997. *SAMPA computer readable phonetic alphabet.*[W:] Gibbon D., Moore R., Winski R. (red.) *Handbook of Standards and Resources for Spoken Language Systems.* Berlin and New York, 684.
Wierzchowska B. 1971. *Wymowa polska*. Wyd. 2 zmienione. Warszawa.
Wierzchowska B. 1980. *Fonetyka i fonologia języka polskiego.* Wrocław–Warszawa–Kraków–Gdańsk.
Wiśniewski M. 1998. *Zarys fonetyki i fonologii współczesnego języka polskiego. Skrypt dla studentów filologii polskiej*. Toruń.
Wolańska E. 2019. *System grafematyczny współczesnej polszczyzny na tle innych systemów pisma.* Warszawa.
Wypych Mikołaj. 1999. *Implementacja algorytmu transkrypcji fonematycznej*. „Speech and Language Technology" 3, 89−105.
Zagórska-Brooks M. 1968. *Nasal Vowels in Contemporary Standard Polish*. Paris.

Wydawnictwo Rys
ISBN 978-83-67287-28-9