

# Spis treści

## **Wstęp**

1. Przedmiot i zakres pracy oraz podstawy metodologiczne badań i opisu .....	9
2. Stan badań grafematycznych za granicą i w Polsce .....	12
3. Struktura pracy .....	16
4. Konwencje zapisu .....	17

## **Część I. W kierunku teorii pisma**

Wprowadzenie .....	21
1.1. Język – mowa a/i pismo w myśli lingwistycznej XX i XXI w. ....	22
1.2. Przedmiot i zadania badawcze grafematyki .....	24
1.3. Język pisany a język mówiony .....	27
1.4. Pismo a mowa jako substancja .....	32
1.5. Od systemu pisma do kroju pisma – ustalenia terminologiczne .....	41
1.6. Najmniejsze jednostki systemu pisma .....	46
1.7. Specyfika grafemów w systemach alfabetycznych .....	55
1.8. Psycholingwistyczna interpretacja pojęcia grafemu a proces czytania i pisania .....	58
Podsumowanie .....	60

## **Część II. Klasyfikacja typologiczna współczesnych systemów pisma**

Wprowadzenie .....	61
2.1. Przegląd ważniejszych ustaleń badawczych dotyczących klasyfikacji systemów pisma .....	62
2.2. Podstawowe typy współczesnych systemów pisma .....	68
2.2.1. Systemy semazjograficzne .....	68
2.2.2. Systemy glottograficzne .....	72
2.2.2.1. Pismo frazograficzne .....	72
2.2.2.2. Pismo logograficzne .....	73
2.2.2.3. Pismo morfemograficzne (morfemiczne) .....	76
2.2.2.4. Pismo sylabograficzne (sylabiczne) .....	82
2.2.2.4.1. Sylabariusz (arbitralny) .....	82
2.2.2.4.2. Abugida .....	88

2.2.2.5. Pismo fonemograficzne (fonemiczne).....	91
2.2.2.6. Pisma mieszane (niejednorodne).....	98
2.2.2.6.1. Pismo fonemo-morfemo-sylabograficzne .....	98
2.2.2.6.2. Pismo morfemo-fonemograficzne .....	102
Podsumowanie .....	103
<b>Część III. Analiza systemu grafematycznego współczesnej polszczyzny</b>	
Wprowadzenie .....	105
3.1. Różnice ilościowe i jakościowe w opisie inwentarza fonemów języka polskiego....	106
3.2. Analiza polskiego systemu grafematycznego .....	116
3.2.1. Grafemy fonemograficzne.....	116
3.2.1.1. Uwagi wstępne.....	116
3.2.1.2. Podstawa materiałowa analizy.....	122
3.2.1.3. Inwentarz polskich grafemów fonemograficznych .....	125
3.2.1.3.1. Alografy małe i wielkie .....	125
3.2.1.3.2. Monografem ⟨a, A⟩ .....	126
3.2.1.3.3. Monografem ⟨ą⟩ .....	126
3.2.1.3.4. Monografem ⟨b, B⟩ .....	127
3.2.1.3.5. Monografem ⟨c, C⟩ .....	127
3.2.1.3.6. Digrafem ⟨ch, Ch⟩ .....	128
3.2.1.3.7. Digrafem ⟨ci, Ci⟩ .....	128
3.2.1.3.8. Digrafem ⟨cz, Cz⟩ .....	129
3.2.1.3.9. Monografem ⟨ć, Ć⟩ .....	129
3.2.1.3.10. Monografem ⟨d, D⟩ .....	130
3.2.1.3.11. Digrafem ⟨dz, Dz⟩ .....	130
3.2.1.3.12. Trigrafem ⟨dzi, Dzi⟩ .....	131
3.2.1.3.13. Digrafem ⟨dź, Dź⟩ .....	131
3.2.1.3.14. Digrafem ⟨dż, Dż⟩ .....	132
3.2.1.3.15. Monografem ⟨e, E⟩.....	132
3.2.1.3.16. Monografem ⟨ę⟩ .....	133
3.2.1.3.17. Monografem ⟨f, F⟩ .....	133
3.2.1.3.18. Monografem ⟨g, G⟩ .....	134
3.2.1.3.19. Monografem ⟨h, H⟩ .....	134
3.2.1.3.20. Monografem ⟨i, I⟩ .....	135
3.2.1.3.21. Monografem ⟨j, J⟩ .....	136
3.2.1.3.22. Monografem ⟨k, K⟩ .....	136
3.2.1.3.23. Monografem ⟨l, L⟩ .....	137
3.2.1.3.24. Monografem ⟨ł, Ł⟩ .....	137
3.2.1.3.25. Monografem ⟨m, M⟩ .....	137
3.2.1.3.26. Monografem ⟨n, N⟩ .....	138
3.2.1.3.27. Digrafem ⟨ni, Ni⟩ .....	139
3.2.1.3.28. Monografem ⟨ń⟩ .....	140

3.2.1.3.29. Monografem ⟨o, O⟩	140
3.2.1.3.30. Monografem ⟨ó, Ó⟩	140
3.2.1.3.31. Monografem ⟨p, P⟩	141
3.2.1.3.32. Monografem ⟨q, Q⟩	141
3.2.1.3.33. Monografem ⟨r, R⟩	141
3.2.1.3.34. Digrafem ⟨rz, Rz⟩	142
3.2.1.3.35. Monografem ⟨s, S⟩	143
3.2.1.3.36. Digrafem ⟨si, Si⟩	144
3.2.1.3.37. Digrafem ⟨sz, Sz⟩	144
3.2.1.3.38. Monografem ⟨ś, Ś⟩	144
3.2.1.3.39. Monografem ⟨t, T⟩	145
3.2.1.3.40. Monografem ⟨u, U⟩	145
3.2.1.3.41. Monografem ⟨v, V⟩	146
3.2.1.3.42. Monografem ⟨w, W⟩	146
3.2.1.3.43. Monografem ⟨x, X⟩	146
3.2.1.3.44. Monografem ⟨y⟩	146
3.2.1.3.45. Monografem ⟨z, Z⟩	147
3.2.1.3.46. Digrafem ⟨zi, Zi⟩	147
3.2.1.3.47. Monografem ⟨ź, Ź⟩	148
3.2.1.3.48. Monografem ⟨ż, Ź⟩	148
3.2.2. Grafemy niefonemograficzne	149
3.2.2.1. Grafem zerowy (spacja międzywyrazowa) ⟨ $\square$ ⟩	149
3.2.2.2. Grafemy sylabograficzne	150
3.2.2.3. Grafemy logograficzne i frazograficzne	150
3.2.2.4. Syngrafemy	152
3.2.3. Właściwości polskiego systemu grafematycznego na tle innych systemów	153
3.2.3.1. Rozkład liczebności reprezentacji grafematycznej	154
3.2.3.2. Ekonomiczność systemu grafematycznego	156
3.2.3.3. Ortograficzna niepewność fonemu	157
3.2.3.4. Fonologiczna niepewność grafemu	162
3.2.3.5. Rozmiar grafemu	162
3.2.3.6. Obecność liter w grafemach	165
3.2.3.7. Użyteczność liter w grafemach	166
3.2.4. Modelowanie zależności między wybranymi właściwościami systemów pisma	168
3.2.4.1. Hipoteza pierwsza	168
3.2.4.2. Hipoteza druga	170
3.2.4.3. Hipoteza trzecia	171
3.2.4.4. Hipoteza czwarta	172
Podsumowanie	173
<b>Wnioski z badań własnych i uwagi końcowe</b>	<b>175</b>

---

Bibliografia .....	187
Summary. The graphemic system of contemporary Polish language as compared with other writing systems .....	202
Spis rycin .....	207
Spis tabel .....	209
Spis wykresów .....	210

*Wyobrażenia jednostek fonetycznych możemy nazwać dla krótkości fonemami, wyobrażenia zaś jednostek graficznych – grafemami.*

(J. Baudouin de Courtenay, 1904)

## WSTĘP

Rozumienie zależności między językiem a mową i pismem nie jest jednoznaczne nie tylko dla zwykłych użytkowników języka, lecz także dla lingwistów. Nie wdając się w szczegóły zbyteczne na tym etapie wywodu, można stwierdzić, iż od czasów A. Schleichera, który w połowie XIX stulecia zaczął konsekwentnie oddzielać głoskę od litery, lingwiści skierowali swoją uwagę przede wszystkim ku stronie dźwiękowej języka. Znaki języka naturalnego były dla wielu badaczy wyłącznie ciągami dźwięków wytwarzanymi przez narządy mowne nadawcy i percypowanymi przez narząd słuchu odbiorcy. Mowa stanowiła użycie tak rozumianego systemu języka, a pismo – narzędzie służące do utrwalania mowy. W niniejszej monografii przedstawione zostanie odmienne stanowisko w tym względzie. Zgodnie z nim język stanowi pojęcie nadrzędne wobec dwóch swoich głównych subkodów – mowy i pisma, a grafem jest podstawową jednostką grafematycznego podsystemu języka.

### **1. PRZEDMIOT I ZAKRES PRACY ORAZ PODSTAWY METODOLOGICZNE BADAŃ I OPISU**

Głównym celem niniejszej monografii jest ustalenie inwentarza grafemów i alografów współczesnego języka polskiego, a następnie dokonanie funkcjonalnego opisu jednostek podsystemów grafematycznego i fonologicznego polszczyzny (por. część trzecią opracowania).

Analiza systemu grafematycznego pisma alfabetycznego, jakim jest polszczyzna, nie jest możliwa bez odwołań do jednostek systemu fonologicznego, ponieważ między oboma systemami zachodzą bezpośrednie, dwukierunkowe relacje. Wśród językoznawców opisujących kompleksowo system fonologiczny współczesnej polszczyzny nie ma zgody ani co do liczby fonemów, ani ich jakości (barwy). W zależności od przyjętych priorytetów i celów opisu badacze języka polskiego wyróżniają od 36 do 57 fonemów (por. podrozdz. 3.1). Dlatego też autorka monografii podjęła się zadania stworzenia – na podstawie przeanalizowanych systemów

fonologicznych – własnego inwentarza jednostek, który będzie najlepiej służył ukazaniu złożonych relacji fonem–grafem i grafem–fonem we współczesnej polszczyźnie (por. podrozdz. 3.2.1.1). Wzorem dla takiego trybu postępowania były działania K. Krakowiak, która – na potrzeby tworzonego przez siebie systemu fonogestów – zaadaptowała listę fonemów ustaloną na gruncie fonologii kazańsko-praskiej [por. Krakowiak 1995, s. 89].

Stworzony na potrzeby niniejszego opracowania inwentarz 35 fonemów spełnia następujące warunki:

- stanowi sieć opozycji fonologicznych służących do rozróżniania znaczenia wyrazów;
- pozwala na maksymalną redukcję inwentarza fonemów, a w konsekwencji także grafemów;
- jest zgodny z faktami fonetycznymi ustalonymi w toku badań instrumentalnych.

Zaproponowany inwentarz został ponadto poddany obiektywnej ocenie w świetle teoriiinformacyjnego pojęcia entropii. Entropia maksymalna systemu wyniosła ok. 5,12 bita/fonem. Jest to najmniejsza wartość entropii maksymalnej spośród piętnastu porównywanych w podrozdziale 3.1 systemów fonologicznych współczesnej polszczyzny stworzonych od 1966 do 2007 r.

Ustalenie inwentarza grafemów i alografów współczesnego języka polskiego i ich relacji do jednostek fonologicznych umożliwiło w dalszym etapie analizy przeprowadzenie badań kwantytatywnych systemu grafematycznego polszczyzny oraz porównanie otrzymanych wyników z rezultatami podobnych analiz odnoszących się do pięciu–siedmiu (w zależności od dostępności określonych danych) innych języków z rodziny słowiańskiej, germańskiej i romańskiej.

Badania kwantytatywne objęły takie właściwości systemu grafematycznego polszczyzny jak:

- rozkład liczebności reprezentacji grafematycznej (ang. *distribution of graphemic representations*),
- ekonomiczność systemu grafematycznego (ang. *economy of the script system*),
- ortograficzna niepewność fonemu (ang. *orthographic uncertainty of the phonem*),
- fonologiczna niepewność grafemu (ang. *phonemic uncertainty of the grapheme*),
- rozmiar grafemu (ang. *grapheme length, grapheme sizes*),
- obecność liter łańcuchowych w grafemach (ang. *graphemic load of letters, L-graphemic load*),
- użyteczność liter w grafemach (ang. *graphemic usefulness of a letter, L-utility*) [por. Best i Altmann 2005, s. 29–39].

W kolejnym etapie analizy poddano weryfikacji cztery najpopularniejsze hipotezy zapoczątkowujące empiryczną analizę kontrastywną różnych systemów pisma. Hipotezy te zakładają, że:

- związek pomiędzy inwentarzem grafemów oraz liter łańcuchowych w fonemograficznych systemach pisma wpływa na średni rozmiar grafemów w tych

systemach (ang. *relationship between G-inventory and L-inventory influences the mean size of graphemes*),

- im większy średni rozmiar grafemów, tym większa średnia obecność liter łacińskich w grafemie (ang. *the greater the size of graphemes, the greater the mean graphemic load of letters*),
- im większa średnia obecność liter łacińskich w grafemie, tym większa średnia ortograficzna niepewność całego systemu pisma (ang. *the greater the mean G-load of letters, the greater the orthographic uncertainty*),
- im większy średni rozmiar grafemów, tym większa średnia użyteczność liter w grafemach (ang. *the greater the size of graphemes, the greater the L-utility*) [por. Altmann 2008, s. 149–164].

Podstawę materiałową wszelkich analiz zawartych w części trzeciej niniejszego opracowania stanowią formy wyrazów ortograficznych, a więc ciągi grafemów ograniczone z obu stron spacjami (w izolacji) bądź też spacjami oraz znakami interpunkcyjnymi (w tekście). Analiza dotyczy zarówno wyrazów pospolitych, jak i nazw własnych, głównie osobowych i geograficznych. Przedmiot badań obejmuje – w kategoriach opisu diachronicznego – wyrazy rodzime oraz spolszczone wyrazy obce. Analiza nie uwzględnia jednak tzw. wyrazów-cytatów, które zachowują obcą postać grafematyczną (często również fonologiczną). Bazę materiałową analizy oraz źródło reprezentatywnych przykładów stanowią wyrazy pochodzące z wydań elektronicznych *Uniwersalnego słownika języka polskiego PWN* [Dubisz i Sobol (red.) 2008] oraz *Wielkiego słownika ortograficznego PWN* [Polański (red.) 2016].

Dodatковым celem niniejszego opracowania była metaanaliza współczesnych systemów pisma (por. część drugą opracowania). Metaanaliza ta zaowocowała stworzeniem własnej klasyfikacji typologicznej współcześnie funkcjonujących systemów pisma. Autorka, korzystając z ustaleń poczynionych przez poprzedników, podzieliła pisma wielostopniowo. Zaproponowała też własną nomenklaturę, która czyni terminologię typologiczną jednoznaczną pod względem formy i treści.

Bez przeprowadzenia metaanalizy współczesnych systemów pisma niemożliwy byłby również kompleksowy opis poszczególnych typów grafemów i ich funkcji we współczesnej polszczyźnie. Jak się bowiem okazało, polski system grafematyczny koduje określone elementy języka nie tylko fonemograficznie, lecz także – choć oczywiście w dużo mniejszym zakresie – w sposób sylabo-, logo- oraz frazograficzny, co zasadniczo uchodziło dotąd uwadze badaczy zajmujących się polską grafią.

Niniejsza monografia wpisuje się zasadniczo w nurt współczesnego językoznawstwa (post)strukturalnego, którego centrum zainteresowań są relacje jednostek. Całość ujęcia ma charakter synchroniczny, choć niemal we wszystkich analizach i opisach odwoływano się do danych historycznojęzykowych oraz faktów historycznych. W opinii autorki nie można tego uznać za błąd metodologiczny, coraz częściej bowiem w pracach synchronicznych wskazuje się na wagę procesów

historycznych. Metaforycznie często mówi się o „pamięci języka”, tzn. o obecności w elementach i strukturach językowych danej epoki skumulowanej wiedzy o fazach rozwojowych poprzednich epok. W odniesieniu do badań nad pismem, które ze swej natury ma konserwatywny charakter, podejście to jest szczególnie uzasadnione. Czasami po prostu wyjątkowo trudno jest oddzielić opis synchroniczny od opisu historycznego. Odwołania do danych historycznych oraz innych o pozajęzykowej naturze służą także klarowności wywodu.

Badania i opisy w wybranym przez autorkę paradygmacie metodologicznym nie oznaczają wszakże, iż nie korzysta ona z ustaleń poczynionych przez uczonych prowadzących badania wedle odmiennych założeń metodologicznych. Rzeczelny opis zjawisk językowych ma bowiem wartość niezależnie od przyjętych paradygmatów. Autorka zresztą sama czyni wyjątek od przyjętych dla prowadzonych analiz założeń metodologicznych, kiedy w części pierwszej rozprawy formułuje psycholingwistyczną definicję grafemu, który w takiej interpretacji staje się kluczowym pojęciem dla obu form komunikacji językowej, jakimi są czytanie i pisanie.

## 2. STAN BADAŃ GRAFEMATYCZNYCH ZA GRANICĄ I W POLSCE

Współczesna teoria grafematyczna uprawiana jest przede wszystkim na gruncie języków angielskiego, francuskiego, niemieckiego oraz rosyjskiego [por. np. Catach 1973, 1979, 1982, 1983, 1984ab, 1986; Catach (ed.) 1980; Amirova (Амирова) 1977, 1985; Zalizniak (Зализняк) 1979; Columas 1982, 1984, 2003; Gallmann 1985, 1986; Kohrt 1985, 1986; August 1986; Günther i Ludwig (Hrsg.) 1994–1996]. Poszczególni badacze proponują rozmaite definicje pojęć związanych z pismem, m.in. takich jak alfabet (ang. *alphabet*), grafem (ang. *grapheme*), system grafematyczny (ang. *graphematic system*), ortografia (ang. *orthography*), pismo (ang. *script*), pismo alfabetyczne (ang. *alphabetic script*), pisownia (ang. *spelling*), system pisma (ang. *writing system*) czy znak (ang. *character*).

Od 1901 r., kiedy to J. Baudouin de Courtenay ukuł termin „grafem” jako jednostkę paralelną do fonemu, brak jest ogólnej zgody co do definicji tego pojęcia. Jest ono rozumiane na co najmniej kilka sposobów. Po pierwsze, grafem bywa utożsamiany z jednostką alfabetu danego pisma niezależnie od języka, w którym jest użyta, np. litera [p, P] w alfabecie łacińskim niezależnie od tego, czy występuje w angielszczyźnie, francuszczyźnie, polszczyźnie czy jakimkolwiek innym języku posługującym się pismem łacińskim. Po drugie, pojęcie grafemu jest odnoszone do litery jako jednostki alfabetu danego języka niezależnie od jej konkretnej realizacji graficznej. Po trzecie, w definicjach z pogranicza językoznawstwa i nauk poznawczych grafemem nazywa się psychiczne wyobrażenie litery. Po czwarte w końcu, w odniesieniu do pism alfabetycznych grafem rozumiany jest jako reprezentant fonemu, czyli jako znak literowy bądź połączenie literowe



odnoszące się do jednego fonemu [por. np. Coltheart 1984; Henderson 1984; Kaczmarek 2003; Polański 2003a]. Rozbieżności dotyczą również traktowania jako grafemów znaków diakrytycznych, znaków interpunkcyjnych itp.

Zagraniczne badania empiryczne systemów pisma prowadzono od początku lat sześćdziesiątych XX w. i miały przeważnie charakter ilościowy [por. np. Bohn 1998, 2002; Gibson i in. 1963; Koch 1971; Prün 1994; Yu 2001]. Na początku pierwszej dekady XXI stulecia dwóch badaczy skupionych wokół czasopisma naukowego „Glottometrics” – G. Altmann z Uniwersytetu Ruhry w Bochum oraz K.-H. Best z Uniwersytetu J. Augusta w Getyndze – poddało ten dorobek analizie krytycznej. Naukowcy ci uznali, iż publikowane wyniki badań, które były prowadzone na podstawie różnych podejść statystycznych, są właściwie nieporównywalne. Nie da się, opierając na nich, wykrywać nowych faktów oraz związków występujących pomiędzy badanymi zjawiskami, zwłaszcza w duchu heurystycznym z wykorzystaniem hipotez.

W 2005 r. K.-H. Best i G. Altmann w artykule *Some properties of graphemic systems*, opublikowanym na łamach „Glottometrics”, sformułowali kilka założeń teoretycznych odnośnie do właściwości alfabetycznych systemów pisma. Następnie właściwości te zostały zapisane w formie wzorów matematycznych. Swoje podejście kwantytatywne do zagadnień związanych z systemami pisma badacze uzasadniali w następujący sposób: „A description of the properties strives for their numerical expression, i.e. quantification and measurement, because processing of this kind is more appropriate for the search for regularities, and the pertinent hypotheses can be more easily tested (corroborated or rejected) if one analyzes different languages and scripts. For the present analysis we do not need to know the frequencies of letters or graphemes since we are concerned here merely with the graphic system” [Best i Altmann 2005, s. 29].

Zaproponowane teorie oraz metody kwantytatywne K.-H. Best i G. Altmann zweryfikowali empirycznie, badając i porównując język niemiecki oraz szwedzki. Wyniki swych badań podsumowali w następujący sposób: „The quantified and measured properties like graphemic uncertainty, grapheme size, graphemic load of characters, and character utility are merely the first possibilities towards the characterization of writing systems, that can be helpful for the observation of their development, comparison and effectivity, and at last for setting up of hypotheses about scripts. The next logical step would be the establishing of further properties, the evaluation of other script systems and the consideration of grapheme and letter frequencies and their relation to the ‘non-weighted’ properties” [Best i Altmann 2005, s. 39].

W kolejnych latach za namową G. Altmanna i niejednokrotnie z jego udziałem badaczom z kilku europejskich ośrodków akademickich udało się przeprowadzić badania większej lub mniejszej liczby wspomnianych powyżej właściwości dla języka chorwackiego i słoweńskiego [Kelih 2007, 2008, s. 61–74, 2013, s. 56–71], słowackiego [Nemcová i Altmann 2008, s. 79–87], ukraińskiego [Buk, Mačutek

i Rovenchak 2008, s. 63–79], włoskiego [Bernhard i Altmann 2008, s. 13–23] oraz angielskiego [Fengxiang i Altmann 2008, s. 25–59].

W 2008 r. G. Altmann, dysponując już częściowymi bądź też pełnymi wynikami badań właściwości pisma alfabetycznego kilku języków, sformułował w tekście pod znamienym tytułem *Towards a theory of script* kilkanaście praw rządzących systemami pisma alfabetycznego. Prawa te, ujęte w formę hipotez, odnosiły się zarówno do poziomu grafematycznego, jak i grafetycznego pisma [por. Altmann 2008, s. 149–164]. Empiryczna weryfikacja sformułowanych przez G. Altmanna hipotez nadal trwa.

Na gruncie polszczyzny autorce niniejszego opracowania znane są co najmniej trzy poprzedzone badaniami własnymi monografie naukowe wpisujące się w nurt analiz o charakterze grafematycznym. Omówmy je pokrótce w kolejności ukazywania się drukiem.

Książkę *Automatyzacja transkrypcji fonematycznej tekstów polskich* M. Steffen-Batogowej [1975] można określić mianem „pregrafematycznej”, ponieważ autorka nie posługuje się w treści publikacji pojęciem grafemu ani alografu, a reprezentantami fonemów czyni litery, przy czym litery nie są interpretowane jednoznacznie, tj. ani konsekwentnie jako najmniejsze znaki graficzne, z których buduje się teksty pisane, ani jako pojedynczy znak bądź wieloznak, który odnosi się do dźwięku. Wieloznakami są u autorki tylko tradycyjnie postrzegane dwuznaki [ch, cz, dz, dź, dż, rz, sz] oraz trójznak [dzi]. Konsekwencją takiego sposobu rozumienia litery i wieloznaku jest w automatycznej transkrypcji odmienne traktowanie wyrazu [cz-e-k] – /tʃ-ɛ-k/ (/č-e-k/), który składa się z trzech „liter” i trzech fonemów, oraz wyrazu [ci-o-c-i-a] – /tɕ-#-ɔ-tɕ-#-a/ (/č-#-o-ć-#-a/), który składa się z sześciu „liter” oraz czterech fonemów i dwóch zer fonologicznych odpowiadających literze [i]. Tymczasem – przy uwzględnieniu pojęcia grafemu – dekompozycja sześcioliterowego wyrazu [ci-o-c-i-a] przedstawia się następująco: ⟨ci-o-ci-a⟩ – /tɕ-ɔ-tɕ-a/ (/č-o-ć-a/), tzn. czterem grafemom odpowiadają cztery fonemy.

Pierwszą monografią grafematyczną – choć dotyczącą polszczyzny w wymiarze historycznym – jest praca *Grafia druków polskich z 1521 i 1522 roku. Problemy wariantywności i normalizacji* T. Lisowskiego [2001]. Autor definiuje grafem jako dyspozycję do graficznego substytuowania fonemu, która jest realizowana w tekście przez alografy. Wyróżnia alografy tekstowe i fakultatywne (prymarne i sekundarne). Alograf u T. Lisowskiego to przede wszystkim litera alfabetu, niekiedy wzbogacona o znak diakrytyczny (np. kreskę, kropkę, ogonek) bądź o literę-diakryt (tzw. dwuznaki [ʃʃ] lub trójznaki [ʃʃh]), ale też kombinacja litery (liter) z kontekstem graficznym (np. spacją) [por. Lisowski 2001, s. 17]. Alograf tekstowy prymarny, czyli tekstowy wariant główny grafemu, to najczęstszy tekstowy wariant grafemu tworzący opozycję na płaszczyźnie tekstu między danym grafemem a innymi grafemami. O tym, czy mamy do czynienia z alografem tekstowym prymarnym, czy z alografem tekstowym sekundarnym, decyduje frekwencja. Ten wariant grafemu, który ma większą frekwencję, należy uznać za

prymarny [por. Lisowski 2001, s. 16–17]. Z kolei alograf fakultatywny prymarny to wariant stylistyczny grafemu, którego zasięg użycia może być regulowany arbitralnie (np. stosowanie wielkich i małych liter, spacji). Natomiast alograf fakultatywny sekundarny to wariant stylistyczny grafemu, którego zasięg użycia jest regulowany indywidualnymi upodobaniami estetycznymi (np. krój pisma drukarskiego lub indywidualny charakter pisma odręcznego) [por. Lisowski 2001, s. 17]. Na powyższą klasyfikację alografów nakłada autor dodatkowe kryterium związane z pozycją danego alografu w kontekście ortograficznym. Ze względu na frekwencję tekstową wyróżnia alografy tekstowe pozycyjne prymarne i sekundarne [por. Lisowski 2001, s. 17].

Wprowadzony przez T. Lisowskiego aparat pojęciowy wydaje się zbyt skomplikowany i niewygodny w użyciu. Wątpliwości budzą także definicje alografów fakultatywnych, które traktuje się równorzędnie z alografami tekstowymi jako reprezentacje grafemów, podczas gdy w rzeczywistości opisują one dodatkowe własności graficzne elementów tekstowych.

Pierwszą monografią, która w obszernej części poświęcona jest systemowi grafematycznemu współczesnej polszczyzny, jest książka *Lingwistyczne podstawy początkowej nauki czytania i pisania po polsku* E. Awramiuk [2006]. Autorka dzieli polskie grafemy na równorzędne i nierównorzędne. Grafemy równorzędne oznaczają takie pary jednostek grafematycznych, które odsyłają do wspólnej jednostki fonologicznej i mają tę samą podstawową wartość fonologiczną, np. ⟨u⟩ i ⟨ó⟩ oraz ⟨ś⟩ i ⟨si⟩, przy czym niektóre z grafemów równorzędnych są nimi niemal bezwyjątkowo (por. np. ⟨u⟩ i ⟨ó⟩), a inne mogą być uznane za grafemy równorzędne jedynie w określonym kontekście fonologicznym (por. np. ⟨ś⟩ i ⟨si⟩). Dlatego też autorka nazywa te pierwsze jednostki grafemami równorzędnymi stałymi, a drugie – grafemami równorzędnymi niestałymi. Grafemy nierównorzędne – w opinii autorki – mogą wystąpić w funkcji prymarnej lub sekundarnej. Jeśli ich wartość fonologiczna w danym kontekście zgadza się z podstawową wartością fonologiczną danego grafemu, to wówczas są to grafemy nierównorzędne w funkcji prymarnej, por. np. ⟨b⟩ = /b/ w wyrazie ⟨żaba⟩ – /zaba/ (/żaba/). Z kolei jeśli ich wartość fonologiczna w danym kontekście nie zgadza się z podstawową wartością fonologiczną danego grafemu, to wówczas są to grafemy nierównorzędne w funkcji sekundarnej, por. np. ⟨b⟩ = /p/ w wyrazie ⟨żabka⟩ – /zapka/ (/żapka/) [por. Awramiuk 103–104].

E. Awramiuk polskie językoznawstwo zawdzięcza nie tylko pierwszą klasyfikację współczesnych grafemów polszczyzny, lecz także pionierski – ujęty w formę tabelaryczną – opis relacji fonem–grafem oraz grafem–fonem [por. Awramiuk 107–108, 115–116]. Autorka nie ustrzegła się jednak pewnych nieścisłości i braków, które wynikły zapewne z niedostatecznej pod względem objętości bazy materiałowej służącej do wyodrębnienia wszystkich możliwych relacji grafemato-fonologicznych. Opis ten również w ograniczonym stopniu odnosi się do kontekstów, w jakich może wystąpić dany grafem, czyli reguł grafotaktycznych.

Nie umniejsza to oczywiście wartości samej publikacji, która poświęcona była przede wszystkim początkowej nauce czytania i pisania, a nie dogłębnej analizie systemu grafematycznego polszczyzny.

### 3. STRUKTURA PRACY

Na strukturę opracowania składają się trzy obszerne części – teoretyczna, meta-analityczna oraz empiryczna. Strukturę tę narzuciła niejako podjęta problematyka badawcza.

Część pierwsza, zatytułowana *W kierunku teorii pisma*, poświęcona jest omówieniu zagadnień podstawowych dla podjętych studiów. Stanowi krytyczny przegląd i systematykę wyników badań poświęconych zagadnieniom teoretycznym pisma w jego wielu aspektach. Określono w nim przedmiot i zadania badawcze grafematyki. Porównano język pisany i mówiony oraz pismo i mowę jako substancję. Przedyskutowano najważniejsze kwestie terminologiczne. Wydzielono najmniejsze jednostki pisma oraz scharakteryzowano specyfikę grafemów w systemach alfabetycznych. Część tę kończy własna definicja grafemu w ujęciu psycholingwistycznym.

Część druga, zatytułowana *Klasyfikacja typologiczna współczesnych systemów pisma*, zawiera szczegółową metaanalizę pism frazograficznych, logograficznych, morfemograficznych, sylabograficznych, fonemograficznych oraz mieszanych. Analiza ta pozwoliła odpowiedzieć na fundamentalne pytanie, czy istnieje związek między cechami systemowymi danego języka a jego pismem. W każdym języku mówionym da się bowiem wydzielić fonemy, sylaby i morfemy, ale – zapewne nie bez przyczyny – określony język uprzywilejowuje jedną z tych jednostek i przyjmuje ją za podstawę swojego systemu pisma.

Część trzecia, zatytułowana *Analiza systemu grafematycznego współczesnej polszczyzny*, rozpoczyna się od przeanalizowania różnic ilościowych i jakościowych w piętnastu opisach inwentarzy fonemów języka polskiego. W kolejnym kroku autorka proponuje własny inwentarz fonemów, który będzie służył ukazaniu relacji grafem–fonem (i na odwrót) we współczesnej polszczyźnie. Analiza wyrazów ortograficznych i fonologicznych języka polskiego prowadzi do wyznaczenia inwentarza grafemów fonemograficznych. Następnie charakteryzowane są grafemy niefonemograficzne: grafem zerowy, grafemy sylabo-, logo- i frazograficzne, a także syngrafemy. Całość wieńczy analiza kwantytatywna wybranych właściwości systemu grafematycznego polszczyzny oraz porównanie języka polskiego pod względem tych właściwości z innymi językami z rodziny słowiańskiej, germańskiej i romańskiej.

W podsumowaniu, zatytułowanym *Wnioski z badań własnych i uwagi końcowe*, zinterpretowano uzyskane w części empirycznej wyniki badań. Wnioski płynące z przeprowadzonych analiz mają charakter zarówno deskryptywny, jak i preskryptywny.

Rozprawę zamyka bibliografia oraz wykazy materiałów załącznikowych.

## 4. KONWENCJE ZAPISU

W pracy stosowane są różne konwencje zapisu elementów języka:

1) Zapis grafemacyjny prowadzony jest przy użyciu grafemów właściwych dla danego systemu pisma. Zapis ten umieszczany jest w nawiasach ostrokątnych (< >), por. np. <境>, <あ>, <ㄑ>, <ㄣ>, <ㄤ>, <ㄨ>, <ㄨ>. Dla odróżnienia litery i ich warianty (glify) – jako zjawiska z poziomu grafetycznego, a nie grafemacyjnego – są umieszczane w tzw. podwójnych nawiasach kwadratowych, np. [[ż, z]], [[R, R]], [[a, a, a]].

2) Zapis fonologiczny (tzw. transkrypcja szeroka) dokonywany jest przy użyciu międzynarodowego alfabetu fonetycznego (ang. *international phonetic alphabet* – IPA), opracowanego i przyjętego przez Międzynarodowe Towarzystwo Fonetyczne (fr. L'Association Phonétique Internationale, ang. International Phonetic Association – IPA). Zapis ten jest umieszczany w nawiasach w formie ukośnych kresek (/ /), por. np. /bædmitən/, /kɔt/. W pracy zastosowano standardową transkrypcję IPA. Afrykaty zapisywano za pomocą ligatur, a nie znaków połączonych łuczkiem, por. np. /dʒes/ zamiast: /dʒes/. Palatalizacja oznaczana jest znakiem joty w indeksie górnym (i), np. /bʲes/. W opisie systemu wokalicznego języka polskiego wyodrębniono sześć samogłosek ustnych, zgodnie ze zgłoszeniem W. Jassem do Międzynarodowego Towarzystwa Fonetycznego [por. Jassem 2003, s. 105–106]. Zastosowano znaki samogłoskowe w następującej formie: /a ε i ɔ u i/. W odniesieniu do samogłosek średnich W. Jassem zastosował znaki /e/ i /o/, gdyż zgodnie z zaleceniami IPA [1999] w transkrypcji szerokiej wykorzystuje się znaki, które powinny być jak najbardziej dostępne (np. z klawiatury komputera). W niniejszej monografii zastosowano symbole<sup>2</sup> bardziej szczegółowe w postaci samogłosek półotwartych (ang. *mid-open*): /ɛ/ i /ɔ/. Odpowiadają one najczęstszym realizacjom fonetycznym tych fonemów w polszczyźnie. Ponadto w zapisie fonologicznym zastosowano znak /#/ na oznaczenie fonemu zerowego (junktura)<sup>3</sup>, a /ø/ – na oznaczenie braku realizacji. Podział sylabowy oznaczono znakiem kropki, np. /mo.tʰa.nun.sa.la.mi/, /zɛt.xa.ɛn/, a połączenie międzywyrazowe znakiem łuczku, np. /ɔd\_rana/. W opisie elementów fonologicznych języka polskiego zastosowano dodatkowo – oprócz transkrypcji IPA – zapis przy użyciu sławistycznego alfabetu fonetycznego (SAF).

<sup>1</sup> Polska transkrypcja fonologiczna zaprezentowana w części wstępnej, pierwszej oraz drugiej niniejszej pracy odwołuje się do inwentarza fonemów oraz zasad zapisu przyjętych dla części trzeciej – analitycznej. Opracowany na potrzeby zawarty w pracy analiz inwentarz został przedstawiony w podrozdziale 3.2.1.1. Inny niż przyjęty inwentarz fonemów skutkowałby w niektórych wypadkach odmienną transkrypcją fonologiczną.

<sup>2</sup> W tekście niniejszego opracowania terminy „symbol” i „znak” są stosowane zamiennie, głównie ze względów stylistycznych oraz tradycji użycia, por. np. „znak interpunkcyjny”, „znak literowy” itd.

<sup>3</sup> Termin junktura jest używany w fonologii w odniesieniu do cech granic fonetycznych. Junktura może wytyczać granice morfemów, wyrazów lub fraz. Najbardziej oczywistą cechą oznaczaną za pomocą junktura jest pauza [por. Crystal 2008, s. 258; Sawicka 1995, s. 112].

Zapis ten jest umieszczany na drugim miejscu w nawiasie okrągłym, por. np. /w ǃ/ (/ǃ ǃ/), /awtentiǃni/ (/aǃtentyǃny/). W SAF fonemy palatalne oznacza się ukośną kreseczką nad znakiem, np. /ś/, a fonemy spalatalizowane – ukośną kreseczką za znakiem, np. /s'/. W niniejszym opracowaniu akcentu wyrazów wielosylabowych zasadniczo się nie oznacza.

3) Zapis fonetyczny (tzw. transkrypcja wąska) dokonywany jest przy użyciu znaków IPA, a w wypadku dźwięków języka polskiego dodatkowo przy użyciu SAF. Zapis ten jest umieszczany w nawiasach kwadratowych ([ ]), por. np. [j] ([j]). Cyfry w indeksie górnym oznaczają w wypadku języków tonalnych określony rodzaj tonu, por. np. [nǔ<sup>3</sup>] (w odniesieniu do języka chińskiego cyfra 3 oznacza ton łamany).

4) Zapis morfofonologiczny dokonywany jest przy użyciu znaków IPA. Umieszcza się go w nawiasach w formie podwojonych ukośnych kresek (// //), por. np. //z//, //mot-ha-nun-sa.lam-i//).

5) Zapis morfologiczny odpowiada zasadniczo ortografii danego języka. Umieszcza się go w nawiasach klamrowych ({ }), por. np. pol. {pocałować} + {-unek}, ang. {sign} + {ø}.

Ponadto w opisach i analizach – oprócz powszechnie przyjętych w pracach lingwistycznych skrótów (np. D. – dopełniacz, l.poj. – liczba pojedyncza) – stosowano następujące oznaczenia:

- C (od ang. *consonant*) – dowolna spółgłoska (głoska, fonem, grafem, litera);
- V (od ang. *vowel*) – dowolna samogłoska (głoska, fonem, grafem, litera);
- CV – sylaba o strukturze „spółgłoska + samogłoska”;
- T – ton w językach tonalnych;
- →, ← – przejście jednej formy w drugą, np. ang. ⟨discount⟩ – /diskaont/ → pol. ⟨dyskont⟩ – /diskont/ (/dyskont/);
- ↔ – relacja obustronna;
- = – jest równe, równa się;
- \* – formy niepoprawne, niezgodne z obowiązującymi obecnie normami, np. \*[auto] ([auto]), \*⟨kuzka⟩;
- ‘ ’ – znaczenie wyrazu.

\* \* \*

Książka ta nie powstałaby bez merytorycznego wsparcia Pani Profesor Piotry Łobacz. Pragnę w tym miejscu Jej serdecznie podziękować za chęć zmierzenia się z zagadnieniem grafemu, cierpliwą lekturę monografii, liczne sugestie, które znacząco wpłynęły na kształt rozprawy. Dziękuję Recenzentom monografii – Pani Profesor Piotrze Łobacz z Uniwersytetu im. A. Mickiewicza w Poznaniu i Panu Profesorowi Jerzemu Podrackiemu z Uniwersytetu Warszawskiego. Ich recenzencka wnikliwość umożliwiła mi samokrytyczne spojrzenie na własny tekst. Jestem głę-

boko wdzięczna Panu Profesorowi Jerzemu Podrackiemu za wsparcie i czas spędzony na wspólnych rozmowach o współczesnym języku polskim.

Chciałabym bardzo podziękować Profesorowi Stanisławowi Dubiszowi, długoletniemu dziekanowi Wydziału Polonistyki UW i dyrektorowi Instytutu Polonistyki Stosowanej, który od początku życzliwie wspierał moje poczynania badawcze. Słowa podziękowania kieruję również do Profesora Józefa Porayskiego-Pomsty za inspirację do szukania odpowiedzi na wiele pytań i wytyczenie przed laty mojej ścieżki naukowo-badawczej.

Dziękuję Pani Doktor habilitowanej Anicie Lorenc za przyjacielskie dyskusje i nieocenioną pomoc w realizacji zamierzeń autorki. Osobne podziękowania kieruję do dyrektora Instytutu Polonistyki Stosowanej Wydziału Polonistyki UW, Pana Profesora dr hab. Grzegorza Bąbiaka, za zaangażowanie w wydanie drukiem niniejszej publikacji. Dziękuję także Pani Doktor Izabeli Więcek-Poborczyk za konsultacje związane z przygotowaniem transkrypcji fonologicznej tekstu.

Pragnę również podziękować Pani Profesor Marii Frankowskiej za nieustanną motywację i niezachwianą wiarę w to, że książka ta powstanie, kiedy autorce wiary tej brakło.

Na koniec słowa szczerego podziękowania kieruję do moich Bliskich, szczególnie gorące do Męża – Adama, bez którego poświęcenia, ogromnego zaangażowania i wytrwałości Czytelnik z pewnością nie mógłby wziąć tej monografii do ręki.