

# Spis treści

Uwagi redakcyjne . . . . .	9
Skróty i symbole . . . . .	13
Wstęp: obszar oraz dziedzina badania . . . . .	17
<b>I. Podstawy BL: jednostka badawcza oraz Surowiec R . . . . .</b>	<b>31</b>
I.1. Dobór jednostki badawczej: poziom morfoskładni KWN . . . . .	39
I.2. Frazem vs. nie-frazem: metody, dyskusje, wnioski. . . . .	43
I.2.1. Dyskusja I: frazem vs. nie-frazem, reprodukt vs. produkt . . . . .	44
I.2.2. Dyskusja II: rola metaforyczności, odtwarzalności oraz autonomiczności w KWN-PF . . . . .	60
I.2.3. Dyskusja III: klasa swobodna na skali frazematycznej w drodze do kontekstowości . . . . .	62
I.2.4. Dyskusja IV: autonomiczność vs. nieautonomiczność . . . . .	66
I.2.5. Dyskusja V: ukontekstowanie vs. frazematyczność . . . . .	69
I.2.6. Dyskusja VI: filary odtwarzalności: OS, OM, OK, OL, OT . . . . .	72
I.2.7. Dyskusja VII: parametr legislacyjności oraz parametr lektowości PF . . . . .	75
I.2.8. Metoda pozycjonowania KWN-PF na skali frazematycznej (w obrębie OM) . . . . .	79
I.2.9. Stan teoretyczny: założenia i hipotezy dotyczące BL przed adaptacją przez BI . . . . .	87
I.2.10. Zastosowanie praktyczne: ograniczenia formalizmu BL – byty oraz błędy . . . . .	91
I.2.11. Klucz frazematyczny: kwalifikatory decyzyjne. . . . .	96
I.3. Zasoby tekstowe: charakterystyka . . . . .	98
I.3.1. Korpus <sub>1</sub> (monoteksty RU): encyklopedia historii MŚ w piłce nożnej . . . . .	98
I.3.2. Korpus <sub>2</sub> (monoteksty RU): „Новая газета” (archiwum z roku 2013) . . . . .	99
I.3.3. Korpus <sub>3</sub> (dwuteksty RU-PL): literatura piękna . . . . .	99
I.3.4. Korpus <sub>4</sub> (monoteksty RU): KRPLOM . . . . .	101
I.3.5. Korpus <sub>5</sub> (monoteksty RU): Common Crawl . . . . .	102

<b>II. Metodologia preadaptacyjna SARWKF: fazy uzyskania Surowca R</b>	104
II.1. Metoda ekscerpcyjna „od góry”: zasób Common Crawl	106
II.2. Metoda ekscerpcyjna „od dołu”. Ekscerpcja kompleksowa w kierunku klasy swobodnej.	109
II.3. Etapy ekscerpcji i selekcji – koncepcja Listy R.	114
II.3.1. Etap I. Metody i techniki ekscerpcyjne	115
II.3.2. Etap I.1. Ograniczenie materiału badawczego.	121
II.3.3. Etap II.1. Pierwsza faza selekcji.	124
II.3.4. Etap II.2. Druga faza selekcji	127
II.3.5. Etap II.3. Trzecia faza selekcji.	135
II.3.6. Etap II.4. Czwarta faza selekcji	145
II.4. Ścieżka ku weryfikacji frazematycznej – koncepcja Tabeli R	148
II.4.1. Operacje weryfikacyjne na próbce Korpus <sub>2</sub>	153
II.4.2. Weryfikacja uzyskanych danych.	158
II.4.3. Randomizowana kwalifikacyjno-weryfikacyjna Tabela R.	160
II.4.3.1. Pierwotne kwalifikatory w Tabeli R.	161
II.4.3.2. Tytuły kolumn weryfikacyjnej Tabeli R	162
II.4.3.3. Kolumna KWN	162
II.4.3.4. Kolumny Q <sub>1</sub> oraz Q <sub>2</sub>	163
II.4.3.5. Kolumna MFR (megalista faktu rejestracji).	163
II.4.3.6. Kolumna Korpus <sub>2</sub>	166
II.4.3.7. Kolumna YANDEX	170
II.4.3.8. Decyzja kwalifikacyjna w ramach Q <sub>1</sub> oraz Q <sub>2</sub> .	171
II.4.4. Rodzaje konfiguracji decyzji kwalifikacyjnych: dylematy i komentarze	173
II.4.5. Priorytety: kwalifikatory w ramach decyzji Q <sub>1</sub> → Q <sub>2</sub> → Q <sub>3</sub>	188
II.4.6. Priorytety: kwalifikatory w ramach decyzji Q <sub>3</sub>	190
II.4.7. Statystyczny rozkład ZG KWN-PF w ramach decyzji Q <sub>3</sub> .	192
II.4.8. Randomizacja Tabeli R	196
<b>III. Metodologia adaptacyjna SARWKF</b>	203
III.1. Ograniczenia: klasyfikacja uproszczona (F oraz N-F)	207
III.2. Pierwsze próby adaptacji	208
III.3. Finalny algorytm – SARWKF	210
III.4. Wprowadzanie nowych danych do SARWKF	216
<b>IV. Metodologia postadaptacyjna SARWKF: powrót do formalizmu lingwistycznego</b>	218
IV.1. Prognozy wyników: KWN-PF <sub>output-bingo</sub> jako typy danych	220
IV.1.1. Pierwsza klasyfikacja możliwych typów bingo: zebranie danych	221
IV.1.2. Druga klasyfikacja możliwych typów bingo: uproszczenie.	224
IV.2. Uzyskane wyniki: tabela plików.	227
IV.2.1. Fragmenty plików: dane wyjściowe pośrednie	229
IV.2.2. Fragmenty plików: dane wyjściowe końcowe	234
IV.3. Uzyskane wyniki: definiowanie uzyskanych typów KWN-PF <sub>output-bingo</sub> (klasyfikacja trzecia, ostateczna)	235
IV.3.1. Przykłady KWN-PF <sub>output-bingo</sub> X3	241

<b>Zakończenie</b> . . . . .	247
<b>Bibliografia</b> . . . . .	259
<b>Megalista faktu rejestracji (MFR)</b> . . . . .	272
<b>Terminy oraz opisy techniczne</b> . . . . .	275
<b>Spis tabel i ilustracji</b> . . . . .	283
Tabele . . . . .	283
Ryciny . . . . .	285
Wykresy . . . . .	286
<b>Załączniki. Płyta CD-ROM</b> . . . . .	287
<b>Adapting lexicographic tools to automated phraseme recognition in a machine learning system. On the example of Russian verb-noun phrases (Summary)</b> . . . . .	289