

# Contents

<b>1</b>	<b>Word forms and their frequencies</b>	<b>1</b>
1.1	Previous frequency dictionaries for Russian . . . . .	1
1.2	Present dictionary: Entries and frequency data . . . . .	2
1.3	Data in electronic form . . . . .	3
1.4	Word forms and lemmas . . . . .	4
1.5	On the definition of a word . . . . .	6
<b>2</b>	<b>Corpus</b>	<b>9</b>
2.1	Corpus compilation procedure . . . . .	9
2.2	Data pre-processing . . . . .	10
2.3	Linguistic processing . . . . .	10
2.4	Brief corpus description . . . . .	12
2.5	Corpus timeline . . . . .	13
<b>3</b>	<b>Statistical data on the word list</b>	<b>15</b>
3.1	Comparison criteria . . . . .	15
3.2	Character statistics . . . . .	16
3.3	Vowels and consonants . . . . .	16
3.4	Word length . . . . .	17
3.5	Word structure . . . . .	19
3.6	Zipf's law . . . . .	22
3.7	Text coverage of the top-N most frequent words . . . . .	22
3.8	Growth in word length . . . . .	23
3.9	Number of syllables . . . . .	26
3.10	Number of letter bigrams and trigrams . . . . .	27
3.11	Summary of statistical data . . . . .	28
<b>4</b>	<b>References</b>	<b>29</b>
<b>5</b>	<b>Most frequent words ordered by rank</b>	<b>31</b>
<b>6</b>	<b>Alphabetical frequency list</b>	<b>43</b>