

# Contents

<b>1</b>	<b>Word forms and their frequencies</b>	<b>1</b>
1.1	Structure of entries and frequency data . . . . .	2
1.2	Data in electronic form . . . . .	3
1.3	Conception of a dictionary of word forms . . . . .	4
1.4	On the definition of a word . . . . .	5
1.5	Multi-word units . . . . .	6
<b>2</b>	<b>Corpus</b>	<b>9</b>
2.1	Data basis and extent of data collection . . . . .	9
2.2	Data pre-processing . . . . .	10
2.3	American and British English . . . . .	11
<b>3</b>	<b>Statistical data on the word list</b>	<b>13</b>
3.1	Character statistics . . . . .	14
3.2	Vowels and consonants . . . . .	14
3.3	Word length . . . . .	16
3.4	Word structure . . . . .	17
3.5	Zipf's law . . . . .	19
3.6	Text coverage by the top-N most frequent words . . . . .	20
3.7	Growth in word length . . . . .	21
3.8	Number of syllables . . . . .	23
3.9	Number of letter bigrams and trigrams . . . . .	25
3.10	Summary of statistical data . . . . .	26
<b>4</b>	<b>References</b>	<b>27</b>
<b>5</b>	<b>Most frequent words ordered by rank</b>	<b>29</b>
<b>6</b>	<b>Alphabetical frequency list</b>	<b>41</b>