

# Contents

<b>1</b>	<b>Word forms and their frequencies</b>	<b>1</b>
1.1	Previous works on Danish word frequency lists . . . . .	1
1.2	Structure of entries and frequency data . . . . .	2
1.3	Data in electronic form . . . . .	4
1.4	Conception of a dictionary of word forms . . . . .	4
1.5	On the definition of a word . . . . .	6
<b>2</b>	<b>Corpus</b>	<b>9</b>
2.1	Data basis and extent of data collection . . . . .	9
2.2	Corpus timeline . . . . .	10
2.3	Data pre-processing . . . . .	10
<b>3</b>	<b>Statistical data on the word list</b>	<b>13</b>
3.1	Comparison criteria . . . . .	13
3.2	Character statistics . . . . .	14
3.3	Vowels and consonants . . . . .	14
3.4	Word length . . . . .	16
3.5	Word structure . . . . .	17
3.6	Zipf's law . . . . .	19
3.7	Text coverage of the top-N most frequent words . . . . .	20
3.8	Growth in word length . . . . .	20
3.9	Number of syllables and syllable length . . . . .	22
3.10	Number of letter bigrams and trigrams . . . . .	24
3.11	Summary of statistical data . . . . .	25
<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>