

Contents

1	Word forms and their frequencies	1
1.1	Previous works on Afrikaans word frequency lists	1
1.2	Structure of entries and frequency data	3
1.3	Data in electronic form	5
1.4	Conception of a dictionary of word forms	6
1.5	On the definition of a word	7
2	Corpus	9
2.1	Data basis and extent of data collection	9
2.2	Corpus timeline	10
2.3	Data pre-processing	10
3	Statistical data on the word list	13
3.1	Comparison criteria	13
3.2	Character statistics	14
3.3	Vowels and consonants	15
3.4	Word length	16
3.5	Word structure	17
3.6	Zipf's law	22
3.7	Text coverage of the top-N most frequent words	23
3.8	Growth in word length	23
3.9	Number of syllables and syllable length	27
3.10	Number of letter bigrams and trigrams	28
3.11	Summary of statistical data	29
4	References	31
5	Most frequent words ordered by rank	33
6	Alphabetical frequency list	45