

Table of contents

1	Introduction	11
1.1	Collocations	11
1.2	Aims and objects of study	15
1.2.1	The import of collocations in language	18
1.2.2	Collocational relations and constituent structures	21
1.3	Summary	25
2	State of the art in collocation research	27
2.1	Origins of the term collocation in linguistic theory	28
2.2	Establishment and development of the concept <i>collocation</i> in linguistic theory	30
2.2.1	J.R. Firth and contemporaries	30
2.2.2	Syntagmatic word field relations vs. paradigmatic lexical relations	33
2.2.3	Coseriu's "lexical solidarities"	35
2.2.4	Hausmann's conception of collocations	36
2.3	The status of collocations in mainstream linguistics	40
2.3.1	Transformational Generative Grammar and related theories	41
2.3.2	Dependency grammar: valency and collocational relations	44
2.3.3	Collocations in Head-Driven Phrase Structure Grammar (HPSG)	47
2.3.4	Cognitive grammar	49
2.4	Lexicographical treatments of collocations	51
2.4.1	The lexicographical approach: The BBI	52
2.4.1.1	Grammatical collocations in the BBI	53
2.4.1.2	Lexical collocations in the BBI	54
2.4.2	Lexical Functions (LFs) in Igor Mel'čuk's Meaning-Text Theory (MTT)	55
2.5	Criteria for collocations in the literature	58
2.5.1	Frequency of co-occurrence	59
2.5.2	Constrained lexical selection and pre-fabrication	60
2.5.3	Degree of semantic cohesion	63
2.5.4	Syntactic relations between constituents of collocations	64
3	Definition of collocation	65
3.1	Defining criteria	68
3.1.1	Quantitative criteria	68
3.1.2	Positional criteria	69
3.2	Qualitative criteria	70
3.2.1	Direct syntactic relation	70
3.2.2	Meaning and collocation: the moving target problem	71
3.2.3	Pragmatics: added or opaque meaning, communicative constraints	75
3.3	Working definition of collocations	76

3.4	Grammaticality, acceptability and naturalness	77
4	Statistical modelling of lexical co-occurrence	79
4.1	Statistical language modelling	81
4.1.1	The problem of elicitation of lexical co-selection	82
4.1.2	Statistical identification of potentially significant data	83
4.1.3	Statistics in modelling constraints on lexical co-occurrence	84
4.2	Some remarks on the probabilistic nature of language	85
4.2.1	Some basic examples	86
4.2.2	The cognitive reality behind probabilities in language	87
4.2.3	Probabilistic parsing	89
4.3	Statistical measures of co-occurrence significance	91
4.3.1	Employing raw frequencies	91
4.3.2	Positional variation as a feature in collocation discovery	96
4.3.3	The directionality of lexical co-selection	98
4.4	Gauging statistical significance: probabilities	99
4.4.1	The t-test in collocation discovery	100
4.4.2	The Chi-square-test (χ^2)	104
4.4.3	Information theoretic tests of association: Mutual Information ..	106
4.4.3.1	MI in collocation discovery	106
4.4.3.2	Applying the MI score	108
4.5	Lemmatisation	110
4.6	An integrated approach	111
4.7	Why statistics alone cannot be criterial to collocations	113
5	Corpus, tools and methodology	116
5.1	The corpus under study: the British National Corpus (BNC)	118
5.1.1	The composition of the BNC	119
5.1.2	Annotation of the BNC	120
5.2	Selection of nodes under study: an entry point to the corpus	123
5.2.1	Selection of items under study: Lexical database and thesaurus classification of verbal communication verbs	124
5.2.2	Items under study: Obligatorily premodified Adjectival Past Participles and Desubstantival Adjectives	127
5.3	Tools and procedures	128
5.3.1	WConcord: word-lists, concordances, collocate tables	128
5.3.2	Saving, maintaining and Accessing data	131
5.3.3	Creating statistical indices to a corpus: Excel spreadsheets	131
6	Lexical collocations: verbal communication verbs	133
6.1	Corpus-study of collocations of verbal communication verbs	136
6.2	Verbal Communication Verbs and related lexical items	141
6.2.1	Items under study	142
6.2.2	Distinctions between types of VCVs	143
6.2.3	Argument structure of verbal communication verbs	145
6.3	Collocations involving Verbal Communication Verbs	147
6.3.1	Phrase-level collocations	148
6.3.1.1	Adverb + Communication verb collocations	151

6.3.1.2	Semantic prosodies in collocations	156
6.3.1.3	Collocations across related words	158
6.3.2	Collocations above the phrase level	161
6.3.2.1	Speaker + Predicate Verb Collocations	163
6.3.2.2	Predicate Verb + Addressee Collocations	164
6.3.2.3	Predicate Verb – Message Collocations	164
6.3.2.3.1	Predicate Verb + Message-Label Collocations	164
6.3.2.3.2	Predicate Verb + Message-Content Collocations	166
6.3.2.4	Predicate Verb – Medium Collocations	168
6.4	Delexical verb collocations involving verbal communication nouns	169
6.5	Summary: Evaluation of the corpus evidence	171
7	Pragmatic constraints on lexical premodification	177
7.1	Obligatorily premodified adjectival past participles (APPs) and desubstantival adjectives (DesubstAdjs)	179
7.2	Corpus-analysis: observed data	182
7.3	Structural properties	188
7.3.1	Syntactic structure: pos-patterning	188
7.3.2	Semantic structure: relations between the constituents	190
7.3.2.1	Semantic properties of adjectival past participle collocations	190
7.3.2.2	Semantic properties of desubstantival adjective collocations	191
7.4	The nature of pragmatic constraints on lexical co-selection	192
7.4.1	Conversational maxims: avoiding redundant information	192
7.4.2	Extensions of meaning	194
8	Discussion and conclusion	196
	Appendix: THE BNC BASIC TAGSET	202
	Appendix: 6.3.1 Adverb (modifier) + VCV collocations	205
	Appendix: 6.3.1.1 Adverb + VCV collocations	216
	Appendix: 6.3.2 collocations above the phrase level	224
	Bibliography	231