

# Index

- algorithms 67, 72, 73
- ambiguity, semantic 71, 74
- American Heritage Word Frequency Book, The 30
- authorship attribution 43, 51
- automatic annotation 157
  
- BNC *see* British National Corpus
- bottom up approach 9
- British National Corpus 22, 35, 53, 60, 61, 63, 64, 65
- Brown Corpus 30, 35
  
- Cambridge Advanced Learners' Dictionary 31
- Cambridge Grammar of the English Language 31
- CASE *see* Corpus of American Spoken English
- CLAWS *see* Constituent Likelihood Automatic Word-tagging System
- clustering 77
- COBUILD 31
- cognitive metaphor theory 12, 141, 157
- colligates/colligation 2
- colligation(s) 10
- Collins English Dictionary 27
- collocate(s) 2, 23, 62, 68
- collocation(al) 2, 10, 64
- concordance/concordancer 2, 4, 10, 72
- Constituent Likelihood Automatic Word-tagging System 11
- content word(s) 36, 40
- corpus
  - annotated 72
  - historical 78
  - historical, ambiguity 74
  - historical, spelling 73
  - normative 2
  - 'off the shelf', *see* BNC, Brown, FLOB, ICE, LLC, LOB
  - reference 2, 8, 9, 91
  - regional 78
  - size 9, 59
  - Corpus del Español* 53, 54, 56, 57, 58
  - Corpus of American Spoken English
    - see* Corpus of American Spoken English
  - Corpus of Dramatic Texts in Scots 5, 17, 26
  - Corpus of Historical English 6
  
- data mining 77
- Delta 6, 43, 51
- Delta-Lz 46
- Delta-Oz 46
- Delta Prime 51
- dialogue 38, 39
- discourse(s) 10
- discourse marker(s) 31
- discourse prosody/ies 128
- discourse role(s) 9
  
- editing 40, 41
  - manual 38, 39, 40
  
- frequency data 62
- function word(s) 3, 35, 40
  
- granularity 59
- Gutenberg 42
  
- hapax legomena 35
- Historical Thesaurus of English (HTE) 69, 77
- historical variation 58
- homonym(s) 74
- homophone(s) 73
- HTE *see* Historical Thesaurus of English
  
- ICE 19, 33
- ICE-GB 19, 22
- ICE-Ireland 19, 22, 23, 24, 25, 31
- International Corpus of English. *see* ICE

- key domain(s) 11, 14  
 key keyword(s) 10  
 keywords/keywords analysis 9  
 KFN-gram 55
- lemmatizing/lemmatization 38, 51, 56  
 LICHEN. *see* Linguistic and Cultural Heritage Electronic Network  
 Linguistic and Cultural Heritage Electronic Network 69  
 Literature Online 43  
 LLC *see* London-Lund Corpus  
 LOB *see* Lancaster-Oslo-Bergen Corpus  
 London-Lund Corpus 31  
 Longman Mexican of Contemporary English 11, 133
- Macmillan English Dictionary For Advanced Learners 31  
 manual analysis  
   sorting/sifting 14, 73  
 meaning, multiple 70  
 MEMT *see* Middle English Medical Texts  
 MI 101, 141, 153, 156  
 Middle English Medical Texts 72  
 MindNet 74  
 MWC *see* Mary Whitehouse Corpus
- n-gram/n-gram architecture 6, 53, 54, 56, 64, 65  
 narration 38, 39  
 Natural Language Processing 74  
 NECTE *see* Newcastle Electronic Corpus of Tyneside English  
 NICTS *see* Northern Ireland Transcribed Corpus of Speech  
 NLP *see* Natural Language Processing  
 Northern Ireland Transcribed Corpus of Speech 26, 33
- noun(s)  
   common 37, 38, 41, 56  
   pronoun(s) 38, 39, 40  
   proper 37, 41
- part of speech tagging *see* tagging  
 Pathfinder 74  
 personal pronouns. *see* noun(s)  
 Phrases in English 55, 56, 59
- polysemy 74  
 POS tagging. *see* tagging  
 pronoun(s). *see* noun(s)  
 proper noun(s) 40
- register 58, 67  
 register variation 68  
 relational database 2, 54  
 relational databases 6
- SARA 53  
 SCOTS *see* Scottish Corpus of Text and Speech  
 Scottish Corpus of Text and Speech 71  
 semantic prosody 23  
 spelling variation 71, 73  
 SQL query 56, 58, 62  
 statistical stylistics 51  
 synonym(s) 68  
 synsets 63, 64
- Tagged Corpus of Scottish Correspondence 72
- tagger  
   semantic 11
- tagging  
   manual 38  
   manual:correction 38  
   part of speech (POS) 38, 41, 51, 56  
   pragmatic 31  
   prosodic 31  
   tagger 38
- Thesaurus of Old English 71  
 TOE *see* Thesaurus of Old English
- token(s) 9, 35, 82  
 top down approach 12  
 type(s) 82
- UCREL Semantic Analysis Systemm 11, 75, 133, 138  
 Unstructured Information Management Architecture 179  
 USAS *see* UCREL Semantic Analysis System
- variable spelling *see* spelling  
   variation
- variant spelling detector 13, 77

VIEW 60, 63

VIEW interface 54, 63, 65

wildcard searches 72

word frequency 1, 5, 42, 51, 68

list 2, 45

list, modified 41

list, problems with 39

WordNet 54, 63, 64, 74

WordSmith 8, 55

XAIRA 53

z scores 42, 61