Subject index

absolute ill-formedness, 27
accented expressions, 272–7
acyclic data flow graphs, 17
adjacencies, 203–5
adverbs: and event sentences, 137–9; "fact" versus "manner" use of, 137–9; and focus, 274–5; relationships in sentences, 135–44
ambiguity: in current natural language systems, 156–7; and direct inference, 167; in discourse segments, 218, 243–5; and disjunction, 166–7; equality reasoning, 168–9; general reasoning approaches, 157–60; and intonation, 234, 243–5; in knowledge representation, 147–72; lexical aspects, 147; in machine-readable dictionaries, 126; requirements for representation of, 166–72; in syntax and structure, 147–8
anaphoric expressions: in discourse, 219–20; interpretation problem, 14; and pronouns, 180–1, 219–20
annotations, 15, 30
antecedent/pronoun sequence, 212–15, 222–5
argument-taking properties, 80, 82, 89–90
articulation, 260
attentional state, 179–225; centering model, 192–220; context effects, discourse, 208–23; and discourse structure, 179–225; and grammar, 203; and pronoun demonstrativity, 179–225; and surface form of referent, 201–3
augmented transmission network, 5–6
automatic methods, 100–2, 284; *see also* computational lexicography

backward-looking center, 179, 182, 192, 236*n*6
bilingual dictionaries, 94
bootstrapping, 11
BORIS system, 156
bottom-up classification, 269
boundary tones, 235, 242–3; and pragmatic meaning, 270–2
"bridges," 14

CASES system, 16, 22
CASREP, 31
center continuation, 195–7
center rule, 194–7
centering model (*see also* local center); definition, 194–5; discourse study, 187–220; pronoun interpretation, 179–225; review of, 192–200
"coarticulation," 242–3, 258
Cobuild Corpus, 40*n*2, 43, 95, 126*n*6, 127
cohesive chain data, 212–13, 222
collective/distributive distinction, 169
Collins Cobuild English Language Dictionary: ambiguity in communication component, 43–50; comparative word meanings, 39–73; homonymy versus polysemy, 39–43
Collins English Dictionary: ambiguity in communication component, 43–50; comparative word meanings, 37–73; homonymy versus polysemy, 39–43; verbs of sound, 93
Collins-Robert English-French Dictionary, 94, 124
combinatory grammars, 228–41
Common Facts Data Base, 13
compaction schemes, 109
complete inference strategy, 153
complete sentences, 24–9
compositional approaches: and logical form, 164–6; semantic interpretation, 151–2
computational lexicography, 99–130; and distributed lexical knowledge, 113–17; future of, 283–8; 'one-off' programs, 103–4; path specifications, 114–17; semantics in, 117–30; theory, 120–30
Consortium for Lexical Research, 286
constraint satisfaction, 172
context: and attentional state, discourse, 179–225; intonation effect, 233–6, 263–9; modeling of, 7; in speech segment pronunciation, 263–9
contour, 275–6
conversational data, *see* discourse
Core Language Engine, 155

Published online by Cambridge University Press

cross-references, 109–11; path specifications, 114–17
cross-system assessment, 31–2

DARPA program, 287
Data Collection Initiative, 286
"Davidsonian position," 135–44
DEC synthesizer, 259
declarative sentences, 156
declarative specifications, 6
decomposition schemes, 109
default inference strategies, 153
deictic pronouns, 180–1, 215–20
demonstrative pronouns, 179–225; antecedent effects, 215–20, 222; discourse study, 179–225; indexical signs, 206–8; local center establishment, 192–208; pragmatic functions, 208–20; use in transitions, 209
demonstratum, 215–20
dialects, 260
dialogue phenomena (*see also* discourse; spoken language), 285
diathesis alternations, 80, 81, 121
dictionaries (*see also* machine-readable dictionaries), 37–73; ambiguity in communication component, 43–50; computational lexicography approach to, 123–8; conversion to lexical knowledge base, 92–4, 123–8; extracting information from, 11–14, 30, 77, 99–102, 123–8; flaws in, 38, 77, 100–2; homonymy versus polysemy, 39–43; prototypical entries, 88; reliability, 37–9, 100–2; systematizing of polysemy, 50–2; verb entries, 87–90
Dictionary Entry Parser, 104–9
differentiae, 88–9, 93–4
direct arguments, 90
disambiguation, 157–60, 167–72
discourse: ambiguity, 218; and attentional state, 179–225; centering model, 187–208; context effects, 208–23, 263–9; deixic expression in, 215–20; and focus, 272–5; Grosz and Sidner's model, 182–4; indexical signs in, 206–8; and intonation, 228–50; meaning in, 270–8; pronoun choice, 183–4, 187–225; research opportunities, 31; statistical study, 187–225
discourse deixis, 215
disjunction, and ambiguity, 166–7
disjunctive normal form, 21–2
distributed lexical knowledge, 113–17, 121–30
domains, 8

EFL representation, 152, 155
Electronic Dictionary Research Project, 286
embedded labels, 106
encyclopedias, 11
enumerate-and-filter model, 149, 157, 171–2
equality reasoning, 168–9
errorful language: current technology, 25–9; probabilistic models, 28–9
event sentences, 125–44; adverbs in, 135–44; situational theory, 139–44
evidential reasoning, 30
expressive power problem: in knowledge representation, 151–2, 163–4; and logical form, 163–4
external arguments, 90

field structure, 114–17
first-order logic, 7
first-order predicates, 151, 153
focus, 272–5
forward-looking centers, 194, 236*n*6
fragments: in computational lexicography, 123; partial understanding of, 24–9
framed-base representations, 7
FRESH system, 16, 22
fundamental frequency, 261–2; in word recognition, 277–8
fuzzy sense boundaries, 49–50

gapping, 232–3
generality encoding, 171
generalizations, 258–9, 263
generative lexicon, 121
genus words, 88, 93–4, 101
goodness of fit test, 187
grammar: centering rule, 201–3; research opportunities, 30, 284; role in attentional state, referents, 201–3; rules in domains, 8

/h/ pronunciation, 265–9
HARC system, 23
harmonic ratio, 267–8
Hear and Respond to Continuous Speech, 23
Hidden Markov Models, 258
high tone, 261–2, 272–3, 275–7
homonymy: errors in, 27; versus polysemy, lexicography, 39–43
horn-clause formalism, 149

imperative sentences, 156
implicit cross-references, 109–11, 114–17
incomplete words: current technology, 25–9; probabilistic models, 29
"incredulity" contour, 275–6
indexical signs: in discourse, 206–8; local center basis, 206–8
indirect arguments, 90
Integrated Database, 23

Published online by Cambridge University Press

intentional logic, 17, 19
internal arguments, 90
interrogative sentences, 156
intonation, 228–50, 257–78; combinatory grammar of, 229–41; and discourse model, 245–50; effect on vowel spectra, 265–6; and meaning, 270–8; non-local effects, 263–9; notation, 234–5; and parsing, 233–6, 269–70; representation of, current theory, 258–70; sentence prosody relationship, 270–8; and sound structure, 261–3; and speech segment pronunciation, 263–9; in spoken language analysis, 241–50, 257–78; and syntax, discourse meaning, 228–50
intonational phrasing, 261–3; pragmatic meaning element, 270–2
intransitive verbs, 78–80
it: antecedent effects, 215–20; centering model, 187–208; cohesive chain data, 212–13; linguistic properties, 184–7

Janus language: in multiple underlying systems, 17, 23; semantic interpretation in, 154–5

knowledge acquisition, 8–15
knowledge representation, 120, 146–72; ambiguity in, 147–9, 152–72; expressive power issue, 151–2; general reasoning problem, 154–7; levels of, 149–50; logical form in, 154–6, 160–3
KRYPTON system, 153

laryngeal articulation, 261
learner's dictionaries, 77, 92
lexical ambiguity: interpretive techniques, 170–1; in knowledge representation, 147, 156
lexical database: building of, 53, 102–9, 99–103; definition, 103; framework for development of, 99–130; partial load in, 104; path specifications, 107–8, 114–17
Lexical Function Grammar, 90
lexical knowledge base: building of, 53, 99–130; mapping from machine-readable dictionaries, 54–5; on-line resources contribution, 90–2; semi-automatic generation, 55–60, 77
lexical seed, 125–8
lexical subordination, 86
Limited Domain Assumption, 9
local center, 179, 182, 192–208; definition, 200–1; and demonstrative pronoun functions, 208–20; discourse study, 192–208; establishment of, 203–5, 212; indexical basis, 206–8; pure index difference, 208
local processing, 24, 101
localized enumeration, 171–2
logical form, 154–5, 241; in compositional semantic interpretation, 164–6; and expressive power issue, 163–4; weaknesses, 160–6
longitudinal assessment, 31–2
Longman Dictionary of Contemporary English, 126*n*6; ambiguity in communication component, 43–50; comparative word meanings, 39–73; homonymy versus polysemy, 39–43; parsing, 105; verb transitivity entries, 118; verbs of sound, 93
low tone, 261–2, 272–3, 275–7
lumping versus splitting, 39–43

machine-readable dictionaries, 37–73; ambiguity problem, 126; assignment of lexicographical sense, 52; computational lexicography approach to, 123–8; conversion to lexical databases, 90–2, 99–102, 99–130; disadvantages, 100–1, 113; future needs, 287; hierarchically structured entries, 55–60; mapping to lexical knowledge base, 54–5, 90–2, 99–102; 'one-off' programs, 103–4; parsing into lexical databases, 104–9; reliability, 37–9, 100–1; semi-automatic extraction from, 55–60, 99–101, 103–4; structure and analysis, 102–13; systematization of polysemy, 50–2; template entries, 91–2
mapping dictionary entries, 53
melody, 275–7
misreferences, 27
MITalk, 245
mixed-depth text understanding, 167
modulation phenomenon, 50–1
Montague-style semantics, 164
multiple underlying systems: execution problem, 19, 22–3; formulation problem, 19–20, 22; interfacing, 15–25; representation problem, 18–22

named roles, 113
narrative understanding systems, 156
negation errors, 27
'noise,' 103
nominal compounds, 14
nominals: and expressive power, 163–4; knowledge representation issue, 151–2; qualia theory, 121–2
non-local center retention, 209, 214–15
novel language: partial understanding of, 24–9; probabilistic model, 28–9

Published online by Cambridge University Press

omitting words, 27
on-line resources (*see also* machine-readable dictionaries): lexical knowledge base contribution, 90–2; template entries, 91–2
'one-off' program, 103
open propositions, 229–30, 239, 248
OSGP system, 16
Oxford Advanced Learner's Dictionary, 126*n*6; ambiguity in communication component, 43–50; comparative word meanings, 39–73; homonymy versus polysemy, 39–43
Oxford English Dictionary, 103–4

PAM system, 156
parenthetical expressions, 123–4
Parlance system, 4
Parlance User Manual, 16
parsing: constraints in, 149–50; in current representational systems, 154–7; dictionaries into lexical database, 104–9; handling of nonlocal dependencies, 269; and intonation, 233–6; in large-scale lexicon building, 53; path concept in, 107–8; in speech recognition, 269–70; tagging differences, 112
part-whole relations, 14
partial lexical databases, 104
partial load, 104
partial solutions, 26
path concept: in lexical databases, 114–17; and parsing, 107–8
phoneme sequences, 260
phonetics, and sound structure, 260–3
pitch accent: pragmatic meaning element, 270–5; in prosodic phrase, 235, 242–3; tune meanings, 276–7
polysemy: dictionary variations, 45–7; versus homonymy, lexicography, 39–43; systematizing approaches, 50–2
pragmatics, 7
predicate-modifying operators, 151
prepositional phrase, 24, 162
probabilistic language models, 28–9
PROLOG, 149
pronominal case errors, 28
pronominal expression: attentional shifts in, 179–225; centering model, 192–208; context effects, 208–23; and pragmatics, 7, 179; structural constraints, 148
pronouns: and attentional state, 182–4, 192–223; centering model, 192–208; context effects, 208–23; demonstrativity property, 179–223; in pragmatics, 7, 179
pronunciation errors, 27
property-sharing constraint, 196–7
propositional logic, 7
propositions, and situations, 139–41
prosodic phrase boundaries, 235, 242–3, 262
prosodic structure, 262–3
prosody, 257–78; combinatory grammar, 236–41; discourse structure relationship, 31, 257–78; intonation relationship, 270–8; and meaning, 270–8; non-local effects, 263–9; notation, 234–5; and parsing, 269–70; representation of, 258–70; and sound structure, 260–3; and speech segment pronunciation, 263–9; and speech technology, 257–78; training procedure, 258–70
punctuation errors, 27
pure index, 207–8

qualia theory, 121–30
quantifier scoping, 159, 165, 169–70
quasi-logical form, 155–6

RAINFORM, 31
redundant representations, 158
referring function, 215
relational database, 16
"relational invariance," 264
resolved quasi-logical form, 155–6
resumptive pronouns/noun phrases, 27
"rheme," 236–40, 246–50
run-together sentences, 27

/s/ phoneme, 264–5
SAM system, 156
SCISOR system, 154
scope ambiguity, 159, 165
seed technique, 125–8
segments: centering model, 194; in discourse structure, 182–3, 215–20; in sound structure, 260–3
selectional restrictions, 158
semantic fields, 114–17
semantics: compositionality, 151; in domain model, 8–9; fields in, 114–17; in large-scale lexicon building, 53, 120–30; logical form weaknesses, 160–6; research opportunities, 30; state of the art, 6–7; systematizing approaches, 50–2; theory in computational lexicography, 120–30; transitivity alternations in, 117–20; unification strategies, 120–30; verb interpretations, 79–81, 117–20
semi-automated procedures, 100–2, 284; *see also* computational lexicography
"Sense Unit Condition," 229–30
sentences, 24–9
serial processing strategy, 160
server module, 20
"situational semantics position," 135
situational theory, 135–44
'small caps' notation, 115

Published online by Cambridge University Press

smoothing algorithms, 243*n*13
speech act interpretation, 148–50, 156
speech recognition, 257–78; allophonic effects, problem, 258–70; coarticulation in, 242; intonation role, 270–8; parsing in, 269–70; probabilistic language models, 29; prosody role, 270–8; surface structure and intonation in, 228–50; training procedures, 258–70
speech segments: non-local effects in pronunciation, 263–9; prosody and intonation in recognition, 269–70; in sound structure, 260–3
speech synthesis, 257–78; allophonic effects, problems, 258–70; dialect differences, 260; training procedures, 258–70
spelling errors, 27
spoken language (*see also* discourse): errors in, 27–8; intonation in the analysis of, 241–50; limitations in understanding, 156; partial understanding, 25–9; probabilistic models, 28–9; research opportunities, 30–2; surface structure and intonation in, 228–50
Spoken Language System, 23
spoonerisms, 27
sprouting techniques, 124–8
"spurious" ambiguity, 234
statistics: conversational data, 187–225; errorful language interpretation, 28–9; use in text analysis, 14–15
stimulus emission verbs, 87
stressed syllables: and fundamental frequency, 277–8; and pragmatic meaning, 270–2; in tunes, 275–7
strings, in dictionary entries, 109–11
structural ambiguity, 147–8, 157
subject-verb disagreement, 27
suprasegmentals, 260–1
surface form (*see also* syntax): and attentional state, referents, 201–3; and centering rule, 201–3; discourse meaning effect, 228–50; and intonation, 228–50
syllable parser, 270
syntax: and ambiguity, 147–8; and intonation, discourse meaning, 228–50; as knowledge type, 8; logical form use, weaknesses, 165–6; research opportunities, 31; state of the art, 5–6

tagging, parsing differences, 112
TEAM system, 154
telephone number recognition, 263
template entries, 91–3
text corpora: contribution to lexical knowledge base, 90–2, 94–5; richness of, 128
text-to-speech systems, 259, 261
that: antecedent effects, 215–20; centering model, 187–208; coherence functions, 208–20; in cohesive chains, 212–13, 222; as demonstrative pronoun, 186; linguistic properties, 184–7; and non-local center retention, 214–15; pragmatic functions, 208–20
"theme," 236–40, 246–50
top-down processing, 24
trailing prepositions, 127
tunes, 275–7
type-coercing verbs, 127
type constraints, 158
typographical errors, 27

"uncertainty" contour, 275–6
unification grammars, 6
unification strategies, 30–1, 121–30

verb ambiguity, 170–1
verbs: ambiguity, 170–1; argument-making properties, 80, 82, 89–90; computation lexicography, 122–30; contribution of linguistics, 76–95; diathesis alternations, 80–1, 121; dictionary entries, 87–90; in lexical knowledge bases, 77–92, 117–20; lexical subordination, 86; means or manner relations, 82; stimulus emission class of, 87; transitivity alternations, 117–20; type coercion, 122–30
verbs of sound, 81–92
virtual adjacency, 200
vocatives, 275–7
voiced stops, 269
vowel spectra, 265–6

Webster's New World Dictionary: ambiguity in communication component, 43–50; comparative word meanings, 39–73; homonymy versus polysemy, 39–43
Webster's Ninth New Collegiate Dictionary, 54, 69, 93
weighted control strategies, 30–1
wh-questions, 23, 229–30
word order errors, 28
word order switching, 27
World Model Language, 19
written language: errors in, 26–8; partial understanding, 25–9; probabilistic models, 28–9; research opportunities, 30–2

/z/ phoneme, 264–5

Published online by Cambridge University Press