

# INDEX

- “#” and “.” Keys, 202
- Acoustic database, 154
- Acoustic immittance, 452
- Acoustic model, 4, 6
  - adaptation schemes, 305
  - of speech recognizers, 6
- Acoustic modeling, 143
- Acoustic phonetic perception, 456
- Acoustic Research, current topics in, 229
- Acoustics for French consonants and vowels, 398
- Acoustics to speech signal, 410
- Adaptive modeling process, 317
- Advanced recognition technologies, 140
- Agent-caller dialog, 20–22, 28, 31
- Algorithm optimization, 149
- Alliance for telecommunications industry solutions, 435
- Alternative ID, 27
- Alt ID Entry, 25–26
- Ambient intelligence, 139
- American Heritage Dictionary, 331
- Americans with Disabilities Act (ADA), 429, 433
- Analysis of Variance (ANOVA), 366, 371, 448
- Animated speech, 159
- Annotation Guide Development, 47
- Anthropomorphic versions, 55, 59, 79
- Anthropomorphism, 4, 7–8, 12, 16, 55–56, 60, 79
- Application code, 49
- Application programming interfaces, 143
- Architecture simplification, 148
- Artificial Neural Networks (ANNs), 144
- Artificial Passenger, 292, 295, 308–310, 322
  - interaction, 309
- Aspiration noise, 411
- ASR, quality of recognition and decoding by the, 311
- Assistive devices, 425
- Assistive technology, 429
- AT&T, 5, 36, 54, 57, 62–64, 97, 100, 209
  - mail, 223
  - operators, 54
- Audio-codec communication, 157
- Auditory Short Term Memory (STM), 233
- Augmentative and Alternative Communication (AAC), 417, 429
- Automated
  - banking services, 50
  - billing information, 27
  - customer service, 64–65, 276
  - Customer Service Rep, 269, 275–278, 287–288
  - data cleaning, 123
  - directory assistance, 117
  - directory services, 165
  - fulfillment, 27
  - pre-processing of the data, 122
- Automatic call reception, 186
- Automatic diacritic reconstruction, 172
- Automatic Number Identification, 29
- Automatic segmentation models, 363
- Automatic speech recognition, 83, 94, 137, 219, 232, 234, 246–247, 301, 304, 306, 331, 364, 424, 436
  - accuracy, 237, 243
  - algorithm, 85

- applications employing in, 424
  - error, 5
  - grammars, 85–87, 94–95, 106
  - IVR systems, 217
  - speech recognition performance, 234
  - transcript, 222
  - and Voice Mail, 215
- Automatic Teller Machines (ATMs), 166
- Automatic translation, 173
- Babel technologies, 141, 144
- Barged-in (BI0), 241–242, 245, 247–248
- Batch-enrolled speakers, 428
- Bayesian classifier, 172
- Binary codification table, 154
- Blacklist, 190
- Blogs, 420–421
- Budapest Binet diagnostic test, 20, 458
- Buffers, 337
- Bulletin boards, 194
- Business listing
  - automation, 127
  - disambiguation, 130
- Business security, 13
- “Busy notification” notion of, 77
- Calendar Manager, 259–260, 287
- Call answering, 194, 203
  - greetings, 206
- Call delivery, 194
- Caller-path diagrams, 19, 23–5, 27, 36, 38
- Caller population, characteristics of the, 51
- Caller tolerance of system delay, 77
- Call flow, 23
  - development and annotation, 48
- Call profiles, 31
- Call-reason distributions, 23–24, 27, 38
- Call-redirection options, 204
- Call resolution, 24
- Call routing, 27, 36–37, 44–45, 54, 60, 66, 73
- Car media module, 312
- Car navigation devices, 136, 140, 310
- Car system messages, navigation, 166
- Cellular phone embedded speech
  - recognition, 145
- Child-directed speech, 109
- Children’s speech perception, 457
- Children with learning disabilities, 458
- Children with specific needs, 457
- Chi-square test, 369
- Chunks, 9, 235, 237–238
- Circumlocution, 97
- Classification errors, 70
- ‘Cleaning rules’ for the directory, 123
- Client-server
  - architecture, 173
  - communication, 180
  - configuration, 139, 149
- Closed-domains synthesis, 152
- Code division multiple access mobile (CDMA), 147
- Cognitive
  - aids, 423
  - disabilities, 458
- Command-line interfaces, 428
- Common sense reasoning, 420
- Communication, form of, 109
- Complexity of the conversion for English and French, 329
- Complex text transformation, 171
- Computer-based dialog
  - management, 231
- Computer-generated interfaces, 316
- Concatenating words, 385
- Consumer electronic devices, 140
- Conversational
  - maxims, 105
  - planner, 311
- Corel Draw, 278
- Corpus-based unit selection
  - technology, 150, 170
- Customized integrated circuits, 143
- Data arithmetic logic unit (DALU), 156
- Database Inadequacies, 122
- Decoding, 144
- Developmental order in language
  - acquisition, 109
- Dialog Design, 102, 133
- Dialog Manager (DM), 233, 294–295
  - prompts and responses in, 233

- Digital signal processor (DSP), 142
- Diphone concatenation method, 168
- Directory assistance, 134
- Discount usability engineering, 2
- Distributed speech recognition (DSR), 149, 299
- Distribution of calls, 50
- Dragon PDSay, 183
- Driver profile module, 312
- Driver safety manager, 295–296
- DTMF data, 194–198, 213, 243
  - detector, 21
- Dundee University, Scotland, 430
  
- E-commerce, 436
- E-mail reading, 165
- Embedded
  - interfaces, 140
    - future of, 158
    - software implementation, 142
    - speech applications, 138–139
    - speech recognition, 141, 143, 297
- Emotional categories, 315
- Endpoint detection, 6
- End-to-end calls, 38
  - data capturing, 20
  - recording, 21–22
- Error consequences prediction system (ECPS), 300–302
- Error recovery technique, 17
- Ethnographic study, 430
- European Telecommunications Standards Institute (ETSI), 149
- Exception dictionary, 385
- Explicit confirmation, 69–71, 126
- External
  - code, 335
  - memory (memory card), 181
  
- F0 Patterns, 372–374
- Fax messaging, 220
- Field-programmable gate array (FPGA), 143
- Filtered speech samples, 366
- Finite vocabulary and grammar, 4–5
- Floating-point values, tables of, 155
- Fonix iSpeak, 183
- Formal modeling techniques, 430
  
- Frame relay, 13
- French consonants, 398
- French-speaking people, 330, 333
- Frequently requested listings (FRL), 120
- Fricative consonants, 400
- FRL approach, 121, 127–128, 130
- Front-end telephone recognition, 419
  
- Gaussian Mixture Model (GMM), 145, 298
- Gesture-based interface, 320, 322
- Gesture recognition system, 321
- Global commands, 201
- GOH test, 457–458
- GOMS family of task analysis methods, 210
- Grammar design for business listing automation, 127
- Grapheme-to-Grapheme, 382
- Grapheme-to-Phoneme conversion, 327–328, 353, 381–382, 384
- Graphical user interface (GUI), 222, 264, 279, 317, 428
- Great vowel shift, 329
- Greeting
  - experiment, 63
  - management, 240
- GSM
  - cellular networks, 186
  - character-set, 171
  
- Hamming weighting function, 154, 157
- Handheld computing devices, 140
- Hands-free and eyes-free communication, 158
- Hands-free mode, 219
- Hearing screening procedure, 450
- Heritage dictionary, 331
- Heuristic evaluations, 12, 18, 430
- Hidden Markov models (HMMs), 83, 144, 298
- HMI design, 291–292, 322
- Holophrases, 110
- Homonyms, 132
- “How May I Help You” (HMIHY) System, capabilities of the, 63–65, 69, 72–73, 75, 77
- Human cognition, 3–4, 9, 12

- Human-computer
  - dialogs, 60–62, 231, 269
  - interaction, 101, 430–431
- Human-human
  - communication, 288
  - dialogue, 67, 76–77, 233
  - speech interactions, 240
- Human-machine dialogue, 54–55, 62, 66–67, 76, 79, 138, 291
- Hybrid dynamic time warping, 147
- IBM, 292–293, 309, 334, 351, 422, 426
  - Haifa research lab, 299
  - human language technologies, 293
  - speech filing system, 214
  - speech recognition system, 297
  - voice technologies, 292
- IconSpeak, 183
- ID entry, 25
- Idiosyncrasies, 281
- Impedance audiometry, 450
- Implicit confirmation, 69–70
- Indirect indicators, 318
- Initial greeting, 60, 63
- Initial teaching alphabet, 334
- Integrated development environments (IDEs), 181
- Intelligent agents, 421
- Intelligent spoken dialogue systems, 138, 158
- Interactive voice response (IVR), 1–2, 19–21, 23–24, 27, 30–31, 33–34, 51, 82, 95, 118, 194, 256, 292, 304–305, 434, 436
  - applications, 2, 98–99, 112, 195
  - automation, 31, 33–34
    - analysis, 32–33, 35, 38
    - quantifying, 38
    - rates, 32
  - benefit, 33
  - billing, 28
  - caller dialog, 20–21, 31
  - caller interaction, 33
  - code, 21
  - communication protocol, 435
  - cost-effectiveness, 19, 29–30, 34
  - design, best practices for, 10, 18
    - fine print of, 3
    - guidelines for 12;
  - forum, 435
  - modified touch-tone, 36
  - reports, flaws of standard, 20
  - speech-enabled, 2, 6, 37
  - touch-tone based, 2, 36, 435–436
  - usability problems, 26
- Interface flexibility and intelligence, 423
- International Electrotechnical Commission (IEC), 433
- International phonetic alphabet, 446
- International Standards Organization (ISO), 197, 201–202, 204, 211–212, 433, 434, 438
  - /ANSI voice messaging user interface, 202
  - technical committee, 434
- Interruptibility, 198
- Intonation (pitch), 327, 385–386
- ISDN, 13
- JAVA MIDP (Mobile Information Device Profile), 180
- “Just Say No” strategy, 67–69
- Klatt model, 413
- Labeled “traffic”, 32
- Language
  - identification, 174, 177, 187, 304, 306
  - advantages of, 175
  - model, 4, 296
  - translators, 140
- Last stressed word of the phrase, 396
- Learning
  - disabilities, 458
  - objects, 422
  - transformation system (LT), 295
- Letter-to-phone converter, 154
- Letter-to-sound, 328, 333, 382, 414
  - rules for English and French, 327
  - transcription, 382
- Lexical items, 91
- Liaison problems, 350
- Linguistic
  - analysis, 366
  - expression, 102

- pragmatics methodology, 99
- and psycholinguistics, 113
- theory of discourse, 103
- Liquids, 404
- Listening to e-mail, 223
- Logographic, 383
- Lombard effect, 6
- Long automated prompt, 65
- Long Term Memory (LTM), 235
- MacDraw, 278
- MailTALK, 223
- Markov model, 177
- Mel-frequency cepstral coefficients (MFCCs), 149, 297
- Message management, 240
  - and greetings management, 240
  - menu, 251
  - and message transfer, 240
- Message reading, 189
- Message retrieval and management, 194
- Microsoft Smartphones, 180
- Microsoft Speech API (SAPI), 178, 181, 431
- Mobile information device profile, 180
- Mobile phone
  - application, 186
  - as a speaking aid, 183
- Modeling process, 320
- Monotonous voice, 389
- Morph decomposition, 343–344
- Morpheme boundaries, 348
- Morphology, 84, 86–88, 96, 400
  - violations of, 363
- Motorola, 156
- Multiband resynthesis pitch
  - synchronous overlap-add, 151
- Multilevel-block architecture, 153
- Multilingual speech processing, 138
- Multimodal drawing, 255, 278, 280, 284, 286–287
  - editor, 284
- Multimodality, 138
- Nasal consonants, 403
- Natural dialogue design, issues for, 60
- Natural language application, 43, 46, 97
  - flexibility, pros and cons of, 45
  - steps to building, 46
- Natural language call routing, 36
- Natural language processing (NLP), 311, 422
- Natural language systems, 45–47, 49–73
- Natural language understanding (NLU), 85, 94, 97, 293
  - files, 296
  - system, 294
- Natural speech, analyzing, 392
- Network-based telecommunication services, 139
- N-gram
  - frequency statistics, 176
  - usefulness values, 177
- Noise, 6
- Nokia, 165, 181
- Non-homophone homographs, 385
- Non-human speech systems, 7
- Non-lexical categories, 88
- Non-natural dialogues, 79
- Non-speech
  - based approach, 321
  - interfaces, 419
  - technologies, 420
- Objective functions, 318–319
- Objective fundamental frequency, 373
- Office Monitor software, 264, 267, 272, 287–289
- On-board speech solutions, 136
- One-touch dialing, 186
- One-word utterances, 113
- Open-domain synthesis, 150
- Open-ended prompting style, 97
- Oto-Rhino-Laryngology, 450
- Out-of-grammar (OOG), 96, 106
- Out-of-vocabulary words, 5
- Overgenerative grammars, 128–129
- Pager user profile, 189
- Partially observable Markov processes, 317
- Parts of speech (POS), 88
- Pauses, 369, 376
  - comma, 398
  - duration of, 369
  - perception, 370

- temporal characteristics of, 369
- of the speech sample, 368
- Pearson correlation, 366
- People with special needs, 431, 433
- Perceptual disorders, 457
- Perceptual linear predictive (PLP), 145
- Personal digital assistants (PDAs), 136, 164
- Personal speech assistant for palm PDAs, 158
- Pervasive computing, 139
- Phoneme
  - duration, 293, 327, 382, 385, 391
  - formant motions, 381
  - transitions, 405, 409
- Phonetic
  - converter algorithm, 156
  - values, 154
- Phonological rules, 327
- Pidgin, 109
- Pitch variation, range of, 389
- Play speech, 157
- Plosives, 401
- Pocket palms, 164
- Polysemy, 94
- Portable media centers, 140
- Port speech synthesis, 178
- Pre-automation methodology, 124
- Primacy versus recency, 9
- ProfiVox text-to-speech engine, 184
- Projective world model, 319
- Pronunciation
  - defining, 171
  - of names, 123
  - of words, 330
- Prosodic phrase, 362
- Prosody, 385
- Punctuation mark detection, 154
- Pure-tone audiometry, 450–451, 457
  
- Question inflection strategy, 67
  
- Recognition accuracy, 75, 306
- Recognition error, 4–5, 10, 59, 70, 268
  - identifying, 66
  - limitation of, 11
  - repairing, 72
- Recognition of environmental situations, 315
- Recording natural speech, 391
- Reduction–elongation coefficient curves, 397
- Rehabilitation Act, 433
- Residential listings, 131
- Resolving internet addresses, 174
- Resonators, 411
- Response time of recognizer, 267
- Retrieving and manipulating messages, 206
- Return on investment (ROI) for the call center, 20
- RISC processors, 182
- Robotic, 169
- Robustness of speech, 159
- Role of prosody and segmentation, 375
- Roman invasion, 329
  
- Safety driver manager components, 295–296
- SayIt, 184–186
- SCANMail, 222–223
- ScanSoft, 141, 151, 183
- Second language learners, 427
- Self-corrections, 281
- Semantic, 86, 93, 96
  - and prosodic cues, 376
  - situational, 105
  - specificity, 111
  - taxonomy, 98, 100–102
- Semivowels, 384, 405
- Short Message Service (SMS), 163, 173, 175, 219, 415
  - Mobile Phone, 186
  - reading, 165, 166, 177, 186, 190
  - application, 166
- Short-term memory, 9
- Signal synthesizer, 415
- Singing, 414
- SLU model, 100
- Smartphones, 140, 164, 180, 183–186
- Smiley faces, 223
- SMSrappier, 186–189
- Social computing, 420–421
- Social network analysis (SNA), 420

- Soft keys, 222
- Software development kits, 181
- Sony Ericsson, 181
- Speaker-independent recognition
  - tasks, 147
- Speaking aid, 183
- Speaking communication aid, 185
- Speaking telephone, 165
- Special working group on
  - accessibility, 433
- Spectrograms, 391–392, 398, 446
- SpeechActs, 236, 259, 267
  - calendar, 259, 263, 287
  - user study, 289
- Speech algorithms, 158
- Speech Application
  - accessibility standards, 434
  - group, 256
- Speech enabled services, codec for, 299
- Speech Engines, 295
- Speech interactions, 315
- Speech interface designers, 159
- Speech melodies, 363
- Speech recognition, 119, 125, 296, 304
  - early attempts to use, 119
  - limitations of, 3–4, 7–8
  - problems, 26
- Speech recognizer, 173
- Speech segmentation, 361, 364
- Speech synthesis, 123, 425
  - software for English or French, 354
- Speech-to-speech translation, 173
- Speech user interfaces (SUIs), 11
- Speech versus touch-tone, 218
- Speech voice messaging, advantage
  - to a, 218
- Spell-checker, 172
- Spoken language interface (SLI),
  - 135, 137
- Spoken language systems, 98
- Spoken natural language dialogue, 43
- Spontaneous speech
  - prediction of, 363
  - sample, 374
- Standardization in spelling, 329
- Statistical frequency measurements, 170
- Statistical language models (SLMs),
  - 45, 85
- Statistical modeling methods, 315
- Store and forward system, 118
- Stressed
  - and reduced vowels, 343
  - syllables, 354
- Strings and classes, 335, 338
- Subjective melody, 373
- Suprasegmental features, 365
- Symbian, 142, 164, 181–182, 187
- Synchronization of movements, 321
- Synonyms, 93, 96
- Syntactic
  - component of language, 109
  - specificity, 112
- Syntax based statistical language
  - models, 363
- Syntax of the formalism, 334
- Synthesis technologies, 141–142,
  - 158, 167
- Synthesize
  - material in kindergartens, 453
  - speech, 157
  - speech, 157, 392, 443, 446
- Synthesizers, 389
  - formant, 169
  - next generation of formant, 412
- Synthetic speech, applications of, 428
- Taxonomy of speech limitations, 11
- TCP/IP communication, 180–181
- Technical jargon, 29
- Telecommunication devices for the
  - deaf (TTYs), 435
- Telecommunications Act, 433
- Telematics, 291–292, 301, 308, 315,
  - 320, 322
  - Conversational Interactivity for,
    - 293, 309
- Telephone-based automation
  - systems, 426
- Terminal node, 91, 298
- Tester, 49, 265–266, 279
- Text normalization, 350
- Text-to-Speech (TTS), 63, 120, 137,
  - 140, 149–150, 168, 178, 223, 249,
    - 295, 328, 334
  - algorithm, 157
  - application, 152

- converter, 171
- engines, 142, 178–179, 184–185
- formant synthesis for French, 381
- processor, 223
- server component, 152
- Textual conventions to voice,
  - Transliteration of, 223
- 3G mobile networks, 13, 299
- Threshold curve, 446, 450–451, 453–454, 456, 458
- Tolerance of system delay, 77
- Topic-related text preprocessing, 170
- Touch gesture recognition, 321
- Touch-tone, 196
  - inputs, key mappings of, 435
  - IVRs, 2
  - transactions, elements of, 197
  - voice mail user interface, 196
- Transactional analysis, 238
- T-test, 366, 448
- Turn-taking model, 245, 248, 250–51
- Tympanometry, 450, 452–453
- Typographical errors, 121–122
  
- U.S. Access Board, 434
- Unified messaging, features of, 249
- Unique Identification (UID), 181
- Universal design vs. assistive technology, 418
- Unstressed word, 386, 392, 395
- Upper cases for lower cases,
  - substitution of, 154
- Usability analysis, 27
- Usability Professionals Association, 433
- Usability testing, 38, 217, 242, 289, 430
- User-centered design, 256, 430
- User-initiated model, 138
- User-system dialogue, 195
- Utterance, 61
  - corpus of, 85
  
- Verbal Communication, 239
- Very smart recognizer (VSR), 146
- ViaScribe, 294, 426–428, 439
  
- Video mail, 226
- Videotelephony, 226
- Virtual sentences, 364
  - identification of, 366–367
- Visual deficits, 425
- Viterbi algorithm, 144, 146
- Vocabulary concatenation, 167
- Voice activated services, 232, 251–252
- Voice based interactive
  - control, 320
- Voice coding scheme, 155
- Voice controlled
  - unified messaging, 249–250
  - voice mail, 232–234, 238–241, 243, 246, 249–252
- Voice dialing, 186
- Voice mail
  - application, 239
  - systems tape, 196
- Voice messaging user interface
  - standards, 211
- Voice onset time (VOT), 401
- VoiceSignal Technologies, 140
- Voice user interface, 1–2, 17–18, 117–118, 121–133, 234, 291
- Vowels, 398
  
- Wearable devices, 420
- Webservice, 180
- WIFI, 13
- Wikis, 420–421
- Wilcoxon signed ranks test, 368
- Winbond Electronics, 151
- Windows mobile, 142, 181
- WinSpeak, 183
- Wizard of Oz method, 54, 57, 59, 63–64, 73, 77, 257–258, 431
- Word recognition rate (WRR), 146
- Working memory, 9
- Workload manager (WM), 295
  
- XML, 180, 185
- Xoumphyvient, 131