

# Index

- A**
- Accents
    - Cockney, 76
    - inferred personality traits, 75–76
    - inferred physical traits, 77
    - persona design, 75–76
    - Southern, 76
  - Accuracy
    - challenges, 27–28
    - maximizing, 219–221
    - measuring, 218–219
    - optimizing by N-best list processing, 220–221
    - optimizing in grammar, 220
  - Acoustic information, limits on, 27–28
  - Acoustic model adaptation, 30
  - Acoustic model files, 30
  - Acoustic models, 20
  - Active grammar, 21–22
  - Actually*, 144–145
  - Agents, 150
  - Alphabetization, 285
  - Ambiguity, 27
  - Ambiguity testing, 269
  - Antithetic discourse markers, 140
  - Applications
    - analyzing, 201
    - developing, 246, 297–298
    - persona, 99
    - persona design, 81
    - requirements definition, 53–55
    - testing, 247–248
  - Apposition discourse markers, 139
  - Architecture, spoken language systems, 16–26
  - Attention, 129–131
  - Attitudinal discourse markers, 140
  - Audio production, 247, 300–301
  - Audio recordings, and prosody, 201–202
  - Auditions, voice actors, 290–291
  - Auditory environment, 7, 58–60
  - Auditory interfaces. *See also* grammars; speech technology; spoken language interfaces; VUI.
    - definition, 6
    - description, 6–7
    - detailed design, 110
    - earcons, 6
    - icons. *See* earcons.
    - multimodal interfaces, 6–7
    - nonpersistent messages, 6
    - nonspeech output, 6
    - sound and feel, 7
    - transient messages, 6
- B**
- Background music, 59
  - Barge-in, 24, 213–215
  - Basic sounds. *See* phonemes.
  - Best-matching path, 23
  - Bigram, 263
  - Biographical sketches, 82
  - Bootstrapping, 264
  - Brand, and persona, 80–81, 98–100
  - Branding sounds, 59
  - Budgets, 42
  - Business model, 46
  - Business needs, 35, 46–48
  - By the way*, 142

## C

- Call flow, 105–106, 229–235
- Call log analysis, 252–253
- Call routing, 10, 65, 210–211
- Callbacks, tuning, 251–252
- Callers. *See* users.
- Card-sorting technique, 116
- Challenges to accuracy, 27–28
- Checklists. *See* guidelines.
- Clarification, universals, 122
- Clarity. *See also* efficiency.
  - barge-in, 213–215
  - definition, 205–206
  - delayed help, 213
  - description, 209
  - disfluency, 211
  - vs.* efficiency, 212–215
  - fast-tracking, 212–213
  - hesitations, 211
  - in individual prompts, 212–213
  - landmarking, 211–212
  - mental models, 210–211
  - natural language understanding, 210–211
  - navigational, 211–212
  - tapering prompts, 213
  - word fragments, 211
- Clauses, intonation contours, 182–183
- Closed-class items, 149
- Coachability, voice actors, 290
- Coaching voice actors, 293–294
- Cockney accents, 76
- Cognition, 119
- Cognitive load. *See also* universals.
  - attention, 129–131
  - challenges, 119–120
  - conceptual complexity, 120–125
  - consistency, 120, 123–124
  - constancy, 120, 121–123
  - context setting, 121, 124–125
  - definition, 119
  - demos, 127
  - human memory, 126
  - instructions, 126–129
  - just-in-time instruction, 127–129
  - memory load, 125–129
  - mental toolbars, 121
  - menu size, 126
  - recency effect, 126
  - tutorials, 127, 128
- Coherence, conversations, 137
- Cohesion, conversations, 137–147
- Colons, 287
- Commands, universal, 72
- Commas, 286
- “Commenting on truth” discourse markers, 140
- Communication, high-level design, 60
- Competitive environment, requirements definition, 46, 47–48, 54
- Computer Phonetic Alphabet (CPA)
  - definition, 20
  - list of, 313–314
- Concatenating phone numbers, 189–192
- Concatenation splices, 192–196
- Concatenative synthesis, 25
- Conceptual complexity, 120–125
- Concessive discourse markers, 140
- Confidence measures, 23–24
- Configuration files, 31
- Confirmation, 97, 221–224. *See also* error handling.
- Consistency
  - conversations, 161–164
  - universals, 120, 123–124
- Consistent character. *See* metaphors; NVA (nonverbal audio); persona.
- Constancy, universals, 120, 121–123
- Content, persona design, 81
- Context
  - audio cues. *See* landmarking.
  - conversational, 108–110
  - corporate, requirements definition, 46, 87–89
  - definition, 12
  - methodological principle, 37
  - setting, cognitive load, 121, 124–125
- Contextualization, in scripts, 280–284
- Contractions, 155–156
- Contrastive stress, 176, 284
- Conversation management, 141
- Conversational context, 108–110
- Conversational design principle, 36
- Conversations. *See also* dialogs; prompts.
  - actually*, 144–145
  - agents, 150
  - assumptions made by users, 166–168
  - closed-class items, 149
  - coherence, 137
  - cohesion, 137–147
  - consistency, 161–164
  - contractions, 155–156
  - Cooperative Principle, 166–168
  - definition, 133–134
  - detailed design, 108–110
  - as discourse, 135–137
  - discourse markers, 139–147
  - discourse sensitivity, 140

- End Focus Principle, 148
- field, 162
- focus, 148
- formal verbs *vs.* colloquial, 161
- going to*, 158–160
- information structure, 147–152
- jargon, 164–166
- level of formality, 161–164
- Maxim of Quantity, 167–168
- may*, 156–158
- mode, 162
- must*, 156–158
- now*, 142
- oh*, 142–144
- okay*, 146–147
- open-class items, 149
- otherwise*, 145–146
- patients, 150
- placing “new” *vs.* “old” information, 147–152
- pointer words, 153–155
- pronouns, 138–139
- register, 161–164
- sorry*, 147
- speech *vs.* language, 134
- spoken *vs.* written, 152–161
- synonyms, 158–160
- tenor, 162
- time adverbs, 138–139
- topics, 149
- by the way*, 142
- will*, 158–160
- Cooperative Principle, 166–168
- Coordination of phases, 41–42
- Core phrases, 259
- Corporate context, 87–89
- Corporate image, 46
- Coverage testing, 267
- CPA (Computer Phonetic Alphabet)
  - definition, 20
  - list of, 313–314
- Creating a VUI. *See* detailed design; high-level design; methodology; requirements definition.
- Customer touchpoints, 47–48
- D**
- Data analysis, usability testing, 115
- Debriefing, usability testing, 114
- Declarative sentences, intonation contours, 181–182
- Definition files, 29–31
- Delayed help, 213
- Delayed login, 72–73
- Demonstration tapes, voice actors, 290–291
- Demos, 127
- Design criteria, 55, 95
- Design decisions, and speech technology, 26–31
- Design goals, and voice actors, 276
- Design process. *See* detailed design; high-level design; methodology; requirements definition.
- Designers as usability testers, 34, 110–111
- Detailed design. *See also* high-level design; methodology; requirements definition.
  - auditory, 110
  - call flow, 105–106, 229–235
  - conversations, 108–110
  - in development methodology, 39
  - dialog states, 104–105
  - principles of, 116–118
  - prompts, 107–110
  - sample application, 229–241
  - subdialogs, 231–235
  - usability testing, 110–116
- Development phase, 39–40, 245–247
- Development process. *See* detailed design; high-level design; methodology; requirements definition.
- Dialog design, tuning, 251–253
- Dialog elements
  - confirmation, 97
  - delayed login, 72–73
  - error handling, 67–72
  - error thresholds, 71–72
  - escalating detail, 68–69, 70
  - first-level error prompt, 69
  - global error counts, 71–72
  - hang up rates, 71–72
  - high-level design, 56, 96–97
  - login, 72–73, 97
  - no-speech timeout, 68, 71
  - progressive prompting, 68–69
  - rapid reprompt, 69–71
  - rejects, 68
  - sample application, 96–97
  - state-specific error counts, 71–72
  - universal commands, 72
- Dialog logic, 5
- Dialog manager, 18
- Dialog states, 104–105
- Dialog strategy, 56, 95–96
- Dialog structure, 78–79
- Dialogs. *See also* conversations; prompts.
  - directed, 63–64
  - emulating a live person, 67
  - form-filling, 64
  - maximizing accuracy, 219–221
  - menu hierarchies, 64

Dialogs (*continued*)  
 mixed initiative, 64–67  
 nonrecognition states, 106  
 prompting for missing information, 64–65  
 subdialogs, 107  
 system personality. *See* persona.  
 traversal, testing, 247–248  
 tuning, 303–306  
 Dictionaries, 20–21, 255  
 Dictionary files, 30  
 Digit-by-digit phone numbers, 190–191  
 Directed dialogs, 63–64  
 Direction notes, in scripts, 278–280  
 Discourse, 139–147. *See also* conversations; dialogs;  
 prompts.  
 Disfluency, 211  
 Distortion, 28  
 Dog toy, voice activated, 15, 311–312  
 Dropout analysis, tuning, 252

## E

Earcons  
 branding sounds, 59  
 definition, 6  
 high-level design, 59, 60  
 as landmarks, 59  
 Efficiency. *See also* clarity.  
 barge-in, 213–215  
 caller modeling, 208–209  
 vs. clarity, 212–215  
 definition, 205  
 delayed help, 213  
 description, 206  
 losing work, 207  
 shortcuts, 207–208  
 tapering prompts, 213  
 task frequency, 207  
 Einstein, Albert, on teaching, 85  
 Either/or questions, intonation contours, 187–188  
 Emulating a live person, 67  
 End Focus Principle, 148  
 End-of-speech timeout, tuning, 254  
 Endpointing, 17  
 Enumerative discourse markers, 139  
 Environmental sounds, 59  
 Equative discourse markers, 139  
 Error handling  
 confirmation, 97, 221–224  
 error thresholds, 71–72  
 escalating detail, 68–69, 70  
 false accepts, 221–224  
 first-level error prompt, 69

global error counts, 71–72  
 graceful recovery, 117  
 guidelines, 227–228  
 hang up rates, 71–72  
 no-speech timeouts, 29, 68, 71  
 progressive prompting, 68–69  
 rapid reprompt, 69–71  
 recognition failure, 28–29  
 rejects, 28–29, 68, 224–228  
 repeat errors, 224  
 sample application, 96–97  
 state-specific error counts, 71–72  
 three-tiered-confidence, 222  
 timeouts, 224–228  
 Error messages. *See* error handling.  
 Escalating detail, 68–69, 70  
 Eset, 219  
 ETSI (European Telecommunications Standards  
 Institute), 122  
 European Telecommunications Standards Institute  
 (ETSI), 122  
 Evaluative usability testing, 40, 302–303  
 Examples, listening to, xxiii  
 Experience, voice actors, 289–290

## F

False accepts, 221–224  
 Fast-tracking, 212–213  
 Feature extraction, 17  
 Feature vectors, 17  
 Field, 162  
 Filler phrases, 259  
 First-level error prompt, 69  
 Focus, conversations, 148  
 Focus groups, 52, 90–92  
 Form-filling dialogs, 64  
 Formal verbs vs. colloquial, 161  
 Formality, level of, 161–164  
 Formats, TTS prosody, 203  
 Frequency of system use, 80, 98–100

## G

Global error counts, 71–72  
*Going to*, 158–160  
 Grammar files, 29–30  
 Grammars. *See also* auditory interfaces; speech  
 technology; spoken language interfaces; VUI.  
 active, 21–22  
 definition, 5  
 high-level design, 56, 95–96  
 probabilities, tuning, 255–256  
 prosody, 172–173

- recognizer, 21–22
- rule-based, 21–22, 66
- SLM (statistical language model), 22, 66
- stub, 298
- Grammars, developing
  - bigram, 263
  - bootstrapping, 264
  - N-gram, 263
  - natural language, 264–266
  - operands, 260
  - order, 263
  - rule-based grammars, 259–263
  - sample application, 298–300
  - SLM (statistical language model), 263–264
  - statistical natural language, 266
  - training, 263
  - training set, 263
  - trigram, 263
  - unigram, 263
- Grammars, testing
  - ambiguity, 269
  - coverage, 267
  - natural language, 268, 270
  - overcoverage, 267–268
  - perplexity, 270
  - pronunciation, 269
  - regression tests, 272
  - rule-based, 267–269
  - SLM (statistical language model), 270
  - spelling, 269
  - statistical natural language, 270
- Grammars, tuning
  - natural language, 272–273
  - rule-based, 271–272
  - sample application, 307–308
  - statistical language models, 272–273
- Grouping related items in scripts, 280–284
- Groups-of-digits phone numbers, 191–192
- Guidelines
  - error handling, 227–228
  - persona design checklist, 79–81
  - prosody, 200–203
  - scripts, 288–289
  - TTS (text-to-speech), 199–203
  - usability, 200–203
- H*
- Hang up rates, 71–72
- Help, delayed, 213
- Hesitations, 211
- High-level design. *See also* detailed design; methodology; requirements definition.
  - background music, 59
  - branding sounds, 59
  - communication, 60
  - in development methodology, 39
  - dialog elements, 67–73, 96–97
  - dialog strategy, 56, 63–67, 95–96
  - earcons, 59, 60
  - environmental sounds, 59
  - grammar type, 56, 63–67, 95–96
  - keeping it simple, 60
  - key design criteria, 55, 95
  - landmarking, 59
  - look and feel, 59
  - metaphor, 57, 97–98
  - NVA (nonverbal audio), 58–60
  - pervasive dialog elements, 56
  - recurring terminology, 56–57, 97
  - sample application, 95–100
  - system personality. *See* persona.
  - usability, 59
- Hotspot analysis, 252
- Human-centered design, 4. *See also* usability testing; user-centered design.
- Human memory, clustering, 126
- The Humane Interface*, 4
- I*
- Icons, auditory. *See* earcons.
- Image, 81, 98–100
- Inferential discourse markers, 140
- Information questions, intonation contours, 185–187
- Information structure, 147–152
- Instructions, 126–129
- Interface design. *See* detailed design; high-level design; methodology; requirements definition; testing; usability testing.
- Interrupting a prompt, 24, 213–215
- Interviewing, 51–53
- Intonation contours. *See also* prosody.
  - basic contours, 180–183
  - clauses, 182–183
  - in context, 183–188
  - declarative sentences, 181–182
  - either/or questions, 187–188
  - information questions, 185–187
  - lists, 183–185
  - rising, 182
  - rising-falling, final, 181–182
  - rising-falling, nonfinal, 182–183
  - wh- questions, 185–187
  - yes/no questions, 182, 185

## J

Jargon, 164–166  
Just-in-time instruction, 127–129

## L

Labs, usability testing, 112–113  
Landmarking, 59, 211–212  
Language experience, 81–82  
Language vs. speech, 134  
Lexington Brokerage, 86. *See also* sample application.  
Likert scales, 114  
Lists, intonation contours, 183–185  
Live person, emulating, 67  
Load test, 248  
Login, 72–73, 97  
Longitudinal studies, 253  
Look and feel, 59  
Losing work, 207

## M

Marking up text, 203  
Markup language, 246  
Maxim of Quantity, 167–168  
Maximizing accuracy, 219–221  
*May*, 156–158  
Measuring. *See also* testing; usability testing.  
    accuracy, 218–219  
    confidence levels, 23–24  
    success, 89  
Meetings with company personnel, 48, 50–51, 54–55, 87–89  
Memory load, 125–129  
Mental models, 12, 210–211  
Mental toolbars, 121  
Menu hierarchies, and dialogs, 64  
Menu size, 126  
Messages  
    error. *See* error handling.  
    nonpersistent, 6  
    system. *See* prompts.  
    transient, 6  
Metaphors  
    benefits of, 125  
    high-level design, 57, 97–98  
    persona design, 80–81  
Methodology. *See also* detailed design; high-level design; requirements definition.  
    application development, 246, 297–298  
    audio production, 247, 300–301  
    grammar development, 246, 298–300  
Methodology, ideal process  
    detailed design, 39

development, 39–40  
high-level design, 39  
testing, 39–40  
tuning, 40–41

Methodology, principles of  
    business needs vs. user needs, 35  
    context, 37  
    conversational design, 36  
    thorough planning, 35–36  
    user input, 34–35  
Methodology, real-world process  
    budgets, 42  
    coordination of phases, 41–42  
    goals, 41–42  
    time constraints, 42  
Mixed initiative dialogs, 64–67  
Mode, 162  
Multimodal interfaces, 6–7  
*Must*, 156–158

## N

N-best processing, 24, 220–221  
N-gram, 263  
Natural language  
    call routing, 10  
    grammar development, 264–266  
    testing, 268, 270  
    tuning, 272–273  
    understanding, 17–18, 210–211  
Naturalness, 203  
Navigation, universals, 122  
Navigational clarity, 211–212  
“New” vs. “old” information, placing, 147–152  
No-speech timeouts, 29, 68, 71  
Noise, 28  
Nonpersistent messages, 6  
Nonrecognition states, 106  
Nonspeech output. *See* NVA (nonverbal audio).  
Nonverbal audio (NVA). *See* NVA (nonverbal audio).  
Nonverbal meaning. *See* prosody.  
*Now*, 142  
Nowlin, Rebecca, 86  
Nuance V-World, 86  
NVA (nonverbal audio)  
    definition, 6  
    high-level design, 58–60  
    sample application, 100  
    testing over the telephone, 60

## O

Observational studies, 51  
*Oh*, 142–144

- Okay*, 146–147  
 “Old” vs. “new” information, placing, 147–152  
 Open-class items, 149  
 Operands, grammar, 260  
 Optimizing accuracy  
   in grammar, 220  
   by N-best list processing, 220–221  
 Order, grammar, 263  
*Otherwise*, 145–146  
 Overcoverage testing, 267–268
- P**  
 Parameter sweeps, 306  
 Parentheses, 287–288  
 Patients, 150  
 Pauses, 196–199  
 Performance challenges, 27–28  
 Periods, 287  
 Perplexity, 270  
 Persona  
   applications, 81, 99  
   biographical sketches, 82  
   brand, 81, 98–100  
   content, 81  
   defining, 81–83, 99  
   definition, 77  
   design checklist, 79–81  
   dialog structure, 78–79  
   differentiating one from another, 78–79  
   fitting voice actors to, 290  
   frequency of system use, 80, 98–100  
   high-level design, 57–58  
   image, 81, 98–100  
   inferred personality traits, 75–76  
   inferred physical traits, 77  
   language experience, 81–82  
   metaphor, 80–81  
   origins of, 78–79  
   prosody, 79  
   role, 80–81  
   sample application, 98–100  
   speaker’s accent, 75–76  
   target audience, 80, 98–100  
   task-related issues, 81  
   users, 80, 98–100  
   vocal attributes, 82  
   vocal register, 82–83  
   voice-activated dialer, 81  
   wording, 78–79  
 Persona definition document, 82  
 Personality of the system. *See* persona.  
 Personality traits, inferred from accent, 75–76  
 Phone numbers, concatenating, 189–192  
 Phonemes, 20  
 Phonetic alphabet. *See* CPA (Computer Phonetic Alphabet).  
 Phonological words, 193  
 Physical traits, inferred from accent, 77  
 Pilot phase, tuning, 250  
 Planning, 35–36. *See also* detailed design; high-level design; methodology; requirements definition.  
 Pointer words, 153–155  
 Principles  
   Cooperative Principle, 166–168  
   detailed design, 116–118  
   development methodology, 34–37  
   End Focus Principle, 148  
   of VUI. *See* methodology, principles of.  
 Problem solving. *See* error handling.  
 Professionalism, voice actors, 289–290  
 Progressive prompting, 68–69  
 Prompts. *See also* conversations; dialogs.  
   clarity, 212–213  
   detailed design, 107–110  
   interrupting. *See* barge-in.  
   for missing information, 64–65  
   tapering, 213  
 Prompts, designing  
   *actually*, 144–145  
   agents, 150  
   assumptions made by users, 166–168  
   closed-class items, 149  
   coherence, 137  
   cohesion, 137–147  
   consistency, 161–164  
   contractions, 155–156  
   Cooperative Principle, 166–168  
   definition, 133–134  
   as discourse, 135–137  
   discourse markers, 139–147  
   discourse sensitivity, 140  
   End Focus Principle, 148  
   field, 162  
   focus, 148  
   formal verbs vs. colloquial, 161  
   *going to*, 158–160  
   information structure, 147–152  
   jargon, 164–166  
   level of formality, 161–164  
   Maxim of Quantity, 167–168  
   *may*, 156–158  
   mode, 162  
   *must*, 156–158  
   *now*, 142

Prompts, designing (*continued*)

*oh*, 142–144  
*okay*, 146–147  
 open-class items, 149  
*otherwise*, 145–146  
 patients, 150  
 placing “new” vs. “old” information, 147–152  
 pointer words, 153–155  
 pronouns, 138–139  
 register, 161–164  
 sample application, 235–237  
*sorry*, 147  
 speech vs. language, 134  
 spoken vs. written, 152–161  
 synonyms, 158–160  
 tenor, 162  
 time adverbs, 138–139  
 topics, 149  
*by the way*, 142  
*will*, 158–160

Pronouns, 138–139

Pronunciation, 20–21, 269

## Prosody

application analysis, 201  
 audio recordings, 201–202  
 choosing a voice, 201  
 concatenating phone numbers, 189–192  
 concatenation splices, 192–196  
 contrastive stress, 176  
 definition, 172–173  
 formats, 203  
 functions of, 173–175  
 grammar, 172–173  
 keeping it simple, 202–203  
 loudness. *See* stress.  
 marking up text, 203  
 naturalness, 203  
 nonverbal meaning, 172  
 pauses, 196–199  
 persona design, 79  
 phonological words, 193  
 pitch. *See* intonation contours.  
 rhythm, 197–198  
 stress, 175–179  
 stress shift, 179  
 TTS (text-to-speech) guidelines, 199–203  
 usability guidelines, 200–203  
 verbal meaning, 172

Prototypes, usability testing, 112

Pruning thresholds, 254

Punctuation, 284, 286–289

## Q

Questionnaires, usability testing, 114, 239

## Questions, intonation contours

either/or, 187–188  
 information, 185–187  
 wh-, 185–187  
 yes/no, 182, 185

## R

Radio Rex, voice activated toy, 15, 311

Rapid reprompt, 69–71

Raskin, Jef, 4

Recency effect, 126

## Recognition

challenges, 53–54  
 failure, 28–29  
 hypothesis, 23–24  
 model, 19–20  
 search, 22–23  
 test, 248–249  
 tuning, 253–256, 306

## Recognizer

acoustic models, 20  
 barge-in, 24  
 best-matching path, 23  
 confidence measures, 23–24  
 dictionaries, 20–21  
 grammar, 21–22  
 N-best processing, 24  
 phonemes, 20  
 pronunciation, 20–21  
 recognition hypothesis, 23–24  
 recognition model, 19–20  
 recognition search, 22–23  
 in spoken language architecture, 17  
 training process, 20

Recommendations. *See* guidelines.

Recording sessions, 291–294

Recording whole-calls, 251

Recordings, and prosody, 201–202

Recurring terminology, 56–57, 97

Reformulatory discourse markers, 140

Register, 82–83, 161–164

Regression tests, 272

Reinforcing discourse markers, 139

Rejection thresholds, 254

Rejects, 28–29, 68, 224–228

Replacive discourse markers, 140

Requirements definition. *See also* detailed design;  
 high-level design; methodology.

corporate context, 46, 87–89



- goals of, 38–39, 45–46
  - ideal process, 38–39
  - rollout plans, 46
  - schedules, 46
  - success metrics, 89
  - system feature list, sample, 93
  - Requirements definition, application needs
    - application environment, 54
    - competitive environment, 54
    - meetings with company personnel, 54–55
    - recognition challenges, 53–54
    - sample application, 92–95
    - task and subtask analysis, 53
    - user interface technologies, 54
  - Requirements definition, business needs
    - business model, 46
    - competitive environment, 46, 47–48
    - corporate context, 46
    - corporate image, 46
    - customer touchpoints, 47–48
    - meetings with company personnel, 48, 87–89
    - rollout plans, 46
    - sample application, 87–89
    - schedules, 46
    - understanding, 46–48, 87–89
  - Requirements definition, user needs
    - caller profiles, 49, 89–92
    - focus groups, 52, 90–92
    - interviewing customer service representatives, 51–52
    - meetings with company personnel, 50–51
    - observational studies, 51
    - sample application, 86–95
    - surveys, 53
    - usage profiles, 50
    - user interviews, 52–53
  - Result discourse markers, 139
  - Rex, voice activated toy, 15, 311
  - Rhythm, 197–198. *See also* prosody.
  - Rising-falling, final intonation contours, 181–182
  - Rising-falling, nonfinal intonation contours, 182–183
  - Rising intonation contours, 182
  - Robust natural language grammars, 264–266
  - Robust parsing, 264–266
  - Role, persona design, 80–81
  - Rollout plans, 46
  - Romans perspire..., 160–161
  - Rule-based grammars
    - definition, 21–22
    - developing, 259–263
    - testing, 267–269
    - tuning, 271–272
    - uses for, 66
  - Rules. *See* guidelines.
- S
- Sample application
    - application development, 297–298
    - audio production, 300–301
    - detailed design, 229–241
    - dialog elements, 96–97
    - dialogs, tuning, 303–306
    - error handling, 96–97
    - evaluative usability testing, 302–303
    - grammar development, 298–300
    - grammars, tuning, 307–308
    - high-level design, 95–100
    - Lexington Brokerage, 86
    - NVA (nonverbal audio), 100
    - persona, 98–100
    - prompts, designing, 235–237
    - recognition tuning, 306
    - requirements definition, application needs, 92–95
    - requirements definition, business needs, 87–89
    - requirements definition, user needs, 86–95
    - scripting, 276–278
    - testing, 301–303
    - tuning, 303–310
    - usability testing, 237–241, 302–303
    - user surveys, 308–310
  - Schedules, 42, 46
  - Schwab application, 3–4, 9
  - Scripting
    - alphabetization, 285
    - case study, 276–278
    - colons, 287
    - commas, 286
    - contextualization, 280–284
    - contrastive stress, 284
    - direction notes, 278–280
    - grouping related items, 280–284
    - guidelines, 288–289
    - parentheses, 287–288
    - periods, 287
    - punctuation, 284, 286–289
    - thought groups, 286
  - Sheeder, Tony, 86
  - Shortcuts, 207–208
  - SLM (statistical language model)
    - definition, 22
    - developing, 263–264

- SLM (*continued*)  
 testing, 270  
 uses for, 66
- Slots, 18
- Sorry*, 147
- Sorting alphabetically, 285
- Sound and feel, 7
- Southern accents, 76
- Speaker verification, 25–26
- Speech recognition. *See* recognition; recognizer.
- Speech technology. *See also* auditory interfaces;  
 grammars; spoken language interfaces; VUI.  
 acoustic model files, 30  
 ambiguity, 27  
 configuration files, 31  
 definition files, 29–31  
 and design decisions, 26–31  
 dictionary files, 30  
 distortion, 28  
 grammar files, 29–30  
 limits on acoustic information, 27–28  
 markup language, 246  
 noise, 28  
 performance challenges, 27–28  
 problem solving, 28–29  
 speaker verification, 25–26  
 TTS (text-to-speech) synthesis, 25  
 VoiceXML, 246
- Speech vs. language, 134
- Spelling, testing, 269
- Spoken English vs. written, 152–161
- Spoken language interfaces. *See also* auditory  
 interfaces; grammars; speech technology; VUI;  
 VUI (voice user interfaces).  
 architecture, 16–26  
 description, 7–8  
 detecting start/stop points in speech. *See*  
 endpointing.  
 dialog manager, 18  
 elements of, 16–19. *See also* recognizer.  
 endpointing, 17  
 feature extraction, 17  
 feature vectors, 17  
 natural language understanding, 17–18  
 processing modules, 16–17  
 reasons for, 9–11  
 slots, 18  
 utterances, 16  
 waveforms, 17
- Standards bodies, 122
- Start-of-speech timeout, tuning, 254
- State-specific error counts, 71–72
- Statistical language model (SLM). *See* SLM (statistical  
 language model).
- Statistical language models  
 developing, 266  
 testing, 270  
 tuning, 272–273
- Stress, 175–179. *See also* prosody.
- Stress shift, 179
- Stub grammars, 298
- Subdialogs, 107, 231–235
- Success metrics, 89
- Summative discourse markers, 139
- Surveys, 53
- Synonyms, 158–160
- System messages. *See* prompts.
- System QA testing, 248
- T*
- Tapering prompts, 213
- Target audience, 80, 98–100
- Task and subtask analysis, 53
- Task completion analysis, 252
- Task design, usability testing, 113
- Task frequency, and efficiency, 207
- Task-related issues, persona design, 81
- Telephone-based usability testing, 112–113
- Telephone Speech Standards Committee  
 (TSSC), 122
- Temporal transition discourse markers, 140
- Tenor, 162
- Termination, universals, 123
- Terminology  
 high-level design, 56–57  
 jargon, 164–166  
 recurring, 56–57, 97
- Test sets, 270
- Testing. *See also* usability testing.  
 applications, 247–248  
 detailed design, 110–116  
 in development methodology, 39–40  
 dialog traversal, 247–248  
 load test, 248  
 recognition test, 248–249  
 sample application, 301–303  
 success metrics, 89  
 system QA, 248  
 unit test, 247
- Testing, grammars  
 ambiguity, 269  
 coverage, 267

- natural language, 268, 270
- overcoverage, 267–268
- perplexity, 270
- pronunciation, 269
- regression tests, 272
- rule-based, 267–269
- SLM (statistical language model), 270
- spelling, 269
- statistical natural language, 270
- Text-to-speech (TTS) synthesis. *See* TTS (text-to-speech) synthesis.
- Thought groups, in scripts, 286
- Three-tiered-confidence, 222
- Time adverbs, 138–139
- Time constraints, 42
- Timeouts
  - end-of-speech, tuning, 254
  - error recovery, 224–228
  - no-speech, 29, 68, 71
  - start-of-speech, tuning, 254
- Tips. *See* guidelines.
- Topics, 149
- Training, 20
- Training set, 263
- Transient messages, 6
- Transitional discourse markers, 139
- Trigram, 263
- Troubleshooting. *See* error handling.
- TSSC (Telephone Speech Standards Committee), 122
- TTS (text-to-speech) synthesis
  - description, 25
  - prosody guidelines, 199–203
- Tuning
  - call log analysis, 252–253
  - callbacks, 251–252
  - in development methodology, 40–41
  - dialog design, 251–253
  - dictionary, 255
  - dropout analysis, 252
  - end-of-speech timeout, 254
  - hotspot analysis, 252
  - longitudinal studies, 253
  - pilot phase, 250
  - pruning thresholds, 254
  - recognition, 253–256
  - rejection thresholds, 254
  - sample application, 303–310
  - start-of-speech timeout, 254
  - task completion analysis, 252
  - user experience research, 253
  - whole-call recordings, 251
- Tuning, grammars
  - grammar probabilities, 255–256
  - natural language, 272–273
  - rule-based, 271–272
  - statistical language models, 272–273
- Tutorials, 127, 128
- U*
- Unigram, 263
- Unit test, 247
- Universal commands, 72
- Universals. *See also* cognitive load.
  - clarification, 122
  - consistency, 120, 123–124
  - constancy, 120, 121–123
  - mental toolbars, 121
  - navigation, 122
  - recommended list, 122–123
  - termination, 123
- Usability
  - high-level design, 59
  - prosody guidelines, 200–203
- Usability testing. *See also* human-centered design; testing; user-centered design.
  - basic approaches, 111–113
  - card-sorting technique, 116
  - data analysis, 115
  - debriefing, 114
  - designers as testers, 34, 110–111
  - in the development methodology, 40
  - evaluative tests, 40, 249
  - in labs, 112–113
  - Likert scales, 114
  - measurement, 113
  - participants, 113–114
  - prototypes, 112
  - questionnaires, 114, 239
  - rating user responses, 114–115
  - running the test, 114–115
  - sample application, 237–241, 302–303
  - Schwab application, 3–4, 9
  - task design, 113
  - telephone-based, 112–113
  - WOZ (Wizard of Oz) approach, 111–112, 237–241
- Usage profiles, 50
- User-centered design. *See also* human-centered design; usability testing.
  - definition, 4
  - gathering user input, 34–35
- User experience research, 253
- User interface technologies, requirements definition, 54

User tests. *See* usability testing.

#### Users

- caller modeling, 208–209
  - caller profiles, 49, 89–92
  - persona, 80, 98–100
  - in usability testing, 113–114
- Users, requirements definition
- caller profiles, 49
  - customer touchpoints, 47–48
  - focus groups, 52
  - input, gathering, 34–35
  - interviews, customer service representatives, 51–52
  - interviews, users, 52–53
  - meetings with company personnel, 48, 50–51
  - needs assessment, 48–53
  - observational studies, 51
  - surveys, 53, 308–310
  - usage profiles, 50

Utterances, 16

#### V

Verbal meaning, 172

Vocal attributes, persona design, 82

Vocal register, 82–83, 161–164

Voice, choosing, 201

Voice-activated dialing, 81

#### Voice actors

- auditions, 290–291
- choosing, 289–291
- coachability, 290
- coaching, 293–294
- demonstration tapes, 290–291
- and design goals, 276
- experience, 289–290
- fit with persona, 290
- professionalism, 289–290
- recording sessions, 291–294

#### Voice actors, scripting

- alphabetization, 285
- case study, 276–278
- colons, 287

commas, 286

contextualization, 280–284

contrastive stress, 284

direction notes, 278–280

grouping related items, 280–284

guidelines, 288–289

parentheses, 287–288

periods, 287

punctuation, 284, 286–289

thought groups, 286

Voice portals, 9

Voice user interfaces (VUI). *See* VUI (voice user interfaces).

VoiceXML, 246

VUI (voice user interfaces). *See also* auditory interfaces; grammars; speech technology; spoken language interfaces.

advantages, 11

creating. *See* detailed design; high-level design; methodology; requirements definition.

definition, 5

first successful application, 15, 311–312

motivation for, 10–11

Schwab application, 3–4, 9

#### W

Waveforms, 17

Wh- questions, intonation contours, 185–187

Whole-call recordings, 251

*Will*, 158–160

Wizard of Oz (WOZ) usability testing, 111–112, 237–241

Word fragments, 211

Wording, persona design, 78–79

WOZ (Wizard of Oz) usability testing, 111–112, 237–241

Written English vs. spoken, 152–161

#### Y

Yes/no questions, intonation contours, 182, 185











