

# Index

- $\lambda$ -reduction, **565**
- $\sqcup$  unification operator, 498
- \***, **373**
- ., 283, 688, 823
- .wav format, 233
- #**, **61**
- ^**, **140**
- De Meulder, F., 13, 746, 775
- ISO8601., 758, 759
- Schulte im Walde, S., 691
- Tjong Kim Sang, E. F., 458
- ^**, **140**
- 10-fold crossvalidation, **155**
- () (optional constituents), **400**
- $\rightarrow$  (derives), **392**
- $\Sigma$  to mean finite alphabet of symbols, 28
- $\delta$  as transition function between states in FSA, 28
- $\epsilon$ -transitions in finite-state automata, **32**
- #**, **61**
- ^**, **61**
- \*** (RE Kleene \*), 20
- +** (RE Kleene +), 20
- .** (RE any character), **20**
- ,** **\** (RE or symbol)21
- \$** (RE end-of-line), 21
- (** (RE precedence symbol), 21
- [** (RE character disjunction), 19
- \b** (RE non word-boundary), 21
- \b** (RE word-boundary), 21
- ]** (RE character disjunction), 19
- ^** (RE start-of-line), 21
- [ ^ ]** (single-char negation), 19
- $\exists$  (there exists), 563
- $\forall$  (for all), 564
- $\Rightarrow$  (implies), 567
- $\lambda$ -expressions, 565
- $\lambda$ -reduction, 565
- $\wedge$  (and), 563
- $\neg$  (not), 563
- $\vee$  (or), 567
- 2001: A Space Odyssey, 1
- 3sg, 403
- 4-tuple, 395
- , (short pause), 421
- . (long pause), 421
- [] (non-verbal events), 421
- Ait-Kaci, H., 527, 532
- A\* decoder, **339**
- compared to Viterbi, 344
- for information ordering in summarization, 823
- pseudocode, 345
- A\* evaluation function, **346**
- A\* search, **363**
- A\* search
- in MT, 907
- A-D conversion, **232, 297**
- Aarseth, P., 690
- AB test, **283**
- abbreviation expansion, 253
- Abduction, **569, 725**
- Abe, N., 688
- Abeillé, A., 490, 491
- Abella, A., 851, 852
- Abelson, R. P., 764
- Aberdeen, J., 724, 852
- Abney, S., 456
- Abney, S. P., 264, 462, 483, 484, 491, 532, 550, 688
- ABox, **579**
- ABSITY, 688
- absolute
- temporal expression, 755
- absolute discounting, **109**
- abstract
- in text summarization, **802**
- acceleration feature in ASR, **304**
- Accent ratio, **266**
- accented syllables, **224**
- accepting
- by finite-state automaton, 27
- accepting state, 27
- Accessibility scale, **712**
- accessing a referent, **708**
- Accomplishment expressions, **577**
- accomplishment expressions, 575
- accuracy
- as metric for POS tagging, 155
- in WSD, 656
- Accusative, **404**
- accusative, **404**
- Acero, A., 247, 285, 306, 326, 335
- Achievement expressions, **577**
- achievement expressions, 575, **577**
- ACL, 15
- Acoustic model, **292**
- acoustic model, **292**
- frame based, **364**
- segment based, **364**
- acoustic modeling, 292
- gender dependent, **356**
- Acquaah-Mensah, G. K., 769, 773
- acronyms
- in TTS, 253
- Action schema, **865**
- action schema, **865**
- active list (in beam search), **326**
- active voice sentence, **407**
- activity expressions, 575, **576**
- Ad hoc retrieval, **782**
- ad hoc retrieval, **782, 784**
- Adafre, S. F., 760
- Adams, J., 285
- Adamson, M. J., 284
- adaptation
- language model, **111**
- speaker in ASR, **356**
- Add-one, **98**
- add-one smoothing, **98**
- Additive noise, **355**
- adequacy
- in MT, 909
- Ades, A. E., 428
- adjacency pair, **828**
- Adjective phrase, **400**
- adjective phrase, **400, 426**
- adjectives, **125, 400**
- Adjukiewicz, K., 420
- adjunction
- in TAG, 427
- Adriaans, P., 492
- adverb, **126, 400**
- days of the week coded as noun instead of, 126
- degree, **126**
- directional, **126**
- locative, **126**
- manner, **126**
- temporal, **126**
- affix, 47
- affricate sound, **221**
- African-American Vernacular English, **230**
- agent, as thematic role, **633**
- agglomerative clustering, **687**
- agglutinative
- language, **877**
- morphology, 48
- Agichtein, E., 753, 799
- Agirre, E., 688, 689
- agreement, **52, 402, 403**
- determiner-nominal, 502, 503
- number, 404
- explosion of grammar size to deal with, 404
- subject-verb, 426, 502, 503
- Ahlsén, E., 854
- Ahn, D., 734, 760

- Aho, A. V., 434, 461, 462  
Aist, G., 871  
Albelson, R. P., 11, 629  
Albert, S., 772  
Albright, A., 387  
Alderete, J., 387  
Algoet, P. H., 115  
ALGOL, 427  
algorithm  
  A\* search, 345  
  beam search stack decoder for MT, 908  
  bottom-up parsing, 434  
  Brill tagger, 154  
  centroid for summarization, **823**  
  chart parsing, 452  
  CKY, 440  
  Corpus Lesk, 658, 659  
  decision list classifier, 655  
  Earley, **447**  
  Extended Gloss Overlap, 668  
  Extended Lesk, 668  
  forward, 183, 321  
  forward-backward, 192  
  Good-Turing smoothing, 101  
  Hobbs tree search algorithm for pronoun resolution, 716  
  IBM Model 1, 895  
  IBM Model 3, 914  
  Jiang-Conrath word similarity, 668  
  Kneser-Ney discounting, **109**  
  Lesk, 658  
  Lin association measure, 674  
  Lin word similarity, 668  
  log-linear models for MT, 918  
  minimum edit distance, 76  
  N-gram tiling for question answering, 800  
  naive Bayes classifier, 653  
  named entity recognition, 745  
  path-length based similarity, 665  
  pointwise mutual information, 674  
  probabilistic CKY, 469  
  Resnik word similarity, 667  
  semantic role labeling, 684  
  Simplified Lesk, 658  
  Soundex, 82  
  stack decoding, 345  
  stack decoding for MT, 905  
  SumBasic, **823**  
  t-test for word similarity, 675  
  TBL, 154  
  TextTiling, 697  
  top-down parsing, 433  
  unsupervised word sense disambiguation, 687  
  Viterbi, 184, 186, 323  
  Yarowsky, **662**  
Algorithm schema, **454**  
alignment, **74**  
  forced, in TTS, 275  
  HMM, for MT, 897  
  in ASR, **330**  
  in letter-to-sound rules, 260  
  minimum cost, **75**  
  of accents to syllables, **271**  
  of sentences in extractive summarization, 809  
  phrase, in MT, 892  
  sentence in MT, 899  
  string, **74**  
  via minimum edit distance, **75**  
  word, in MT, 893  
alignment templates, 920  
all-words task in WSD, **651**  
Allaria, G., 691  
Allen, J., 12, 16, 260, 284, 285, 364, 387, 587, 764, 853, 854, 859, 864, 869–871  
allophone, **226**  
allophonic rules, 246  
Allwood, J., 854  
alphabet  
  in defining a formal language, **30**  
Alshawi, H., 600, 606, 608, 611, 619, 733  
alternation, subcategorization, **425**  
Alternations, **510**  
Althaus, E., 813, 823  
Altun, Y., 266  
alveolar sound, **220**  
Alwan, A., 285  
Ambiguity, **4**  
ambiguity, **4**, 461  
  amount of part-of-speech in Brown corpus, 137  
  attachment, 436  
  coordination, 436, 461, 474  
  dialogue act, 856  
  duck as example, 4  
  homograph, 256  
  in semantic analysis, 590  
  in meaning representations, 554  
  in morphology, **67**  
  in named entity recognition, 741  
  lexical, 650  
  local, **437**  
  noun phrase bracketing, 461  
  of referring expressions, 709  
  part-of-speech, **133**  
  PCFG in , 467  
  PP-attachment, 461  
  prepositional phrase attachment, 473  
  resolution, 4  
  resolution of tag, **133**  
  specific/non-specific, 710  
  speech act, 5, 856  
  syntactic, 4  
  tests to distinguish from vagueness, 555  
  word sense, 650  
Ambiguous, **4**  
American Structuralism, 427  
amplitude  
  of a signal, **231**  
  RMS, 234  
Amsler, R. A., 645  
analogical reasoning and metaphor, 646  
Anaphora, **708**  
anaphora resolution, **709**  
anchors in regular expressions, **21**, 41  
Anderson, A. H., 854, 856, 871  
Anderson, M. J., 271  
Anderson, S. R., 531  
Andreou, A., 364  
Andrew, G., 81  
Ang, J., 364  
Answer type, **794**  
Answer type recognition, **795**  
Answer type taxonomy, **795**  
Antecedent, **708**  
Antonym, **628**  
Anttila, A., 135, 137, 139, 418, 419, 428  
Antworth, E. L., 55, 64, 65, 67, 81, 368  
Aone, C., 733  
AP, 400  
Appelt, D. E., 726, 733, 766–768, 774  
Apple  
  AIFF, 233  
appropriateness conditions for types (see also typed feature structures), 527  
approximant sound, **221**  
Arabic, 215  
  Egyptian, 241  
Araki, M., 852, 871  
Aramaic, 215  
Archangeli, D., 369, 371, 373  
arcs in directed graphs, **27**  
Arens, R., 771  
Arens, Y., 620  
Argamon, S., 462, 775  
argmax, **140**, 291  
Ariel, M., 712, 733

- Aristotle, 123, 575  
 Arity, **570**  
 Arons, B., 832, 850  
 ARPA, 334  
 ARPA format for N-grams, 108  
 ARPabet, 245  
 article (part-of-speech), **127**  
 articulatory phonetics, 218, **218**  
 Articulatory phonology, **244**  
 articulatory phonology, **244**  
 articulatory synthesis, 284  
 artificial intelligence  
   founding of field, 10  
 Ariles, J., 775  
 Aseltine, J., 775  
 Ashburner, M., 772  
 Asher, N., 734  
 Asher, R. E., 170  
 Aslin, R. N., 365, 382  
 Aspect, **575**  
 aspect  
   computational models of, 585  
 aspirated sound, **226**  
 ASR, **287**  
   confidence, 342, 860  
   decoding  
     recombining hypotheses, 340  
     neural nets for, **354**  
 ASR architecture  
   HMM-MLP hybrid, **354**  
 assertive, 830  
 Assimilation, **226**  
 assimilation, **226**  
 association, **672**  
   measure of for word  
     similarity, **672**  
   probability as measure of, 673  
 Atal, B. S., 333, 355  
 ATIS, 335, **390**, 393, 396, 421,  
   437, 502  
 Atkins, S., 651  
 Atkinson, M., 828  
 ATN, *see* augmented transition  
   network  
 atomic types, **527**  
 ATRANS, 642  
 Attachment ambiguity, **436**  
 Attar, R., 792  
 attribute grammars, 619  
 attribute-value matrix, 494  
 Atwell, E., 113  
 Aubert, X., 340, 341  
 augmentative communication,  
   12, **84**  
 augmented transition network,  
   531  
 Austin, J. L., 829, 870  
 Austin, S., 339, 343  
 Authorship attribution, **121**  
 authorship attribution, **121**  
 automatic speech understanding,  
   287  
 Automaton, **27**  
 automaton, finite state, *see*  
   finite-state automaton  
 auxiliary verb, **125**, 126, **128**,  
   **406**  
 Ayers, G. M., 268  
 Aylett, M. P., 118, 119  
 Baayen, R. H., 79, 81, 126, 158,  
   241, 487  
 Babbage, C., 249  
 Babko-Malaya, O., 688  
 Babyonyshev, M., 547  
 Bacchiani, M., 111  
 Bach, E., 428, 591, 619, 734  
 Bach-Peters sentences, 734  
 Bachenko, J., 264  
 backchannel, **831**  
 backoff in N-gram smoothing,  
   **104**  
 backtrace  
   in minimum edit distance, **77**  
 backup as solution to  
   non-determinism, 33  
 Backus, J. W., 10, 391, 427  
 Backus-Naur Form, **391**  
 Backward chaining, **568**  
 Backward looking center, **718**  
 Backward probability, **188**  
 Bacon, F., 431  
 Badulescu, A., 682, 690  
 Baeza-Yates, R., 779, 786, 822  
 Bag generation, **121**  
 bag-of-words, **653**  
   features in WSD, **652**  
   in IR, **781**  
 Bagley, W. C., 362  
 Bahl, L. R., 119, 139, 334, 339,  
   347, 354  
 Bailey, D., 585  
 Bailey, T., 387  
 Bake-off, **335**  
 bake-off  
   speech recognition  
     competition, **335**  
 Baker, C. F., 508, 509, 629, 637,  
   646  
 Baker, J. K., 119, 334, 471, 490  
 Bakis network, **178**, 295  
 Bakis, R., 285  
 Balashek, S., 10, 333  
 Baldrige, J., 428, 734  
 Baldwin, T., 491, 532  
 Ball, C. A., 772  
 Balogh, J., 832, 840, 841, 846,  
   850, 860, 871  
 Banerjee, S., 668, 670, 921  
 Bangalore, S., 491  
 Banko, M., 170, 797, 800, 824  
 Bar-Hillel, Y., 417, 420, 549,  
   919  
 Barge-in, **850**, **851**  
 barge-in, **851**  
 Barker, D. T., 846  
 Barney, H. L., 247  
 Baron, A., 359  
 Baroni, M., 387  
 Barrett, L., 477, 490  
 Bartels, C., 364  
 Barto, A. G., 863, 871  
 Barton, Jr., 81, 550  
 Barzilay, R., 733, 809, 812–815,  
   823, 824  
 baseline  
   for part-of-speech tagging  
     evaluation, 155  
   most frequent sense, 657  
   take the first sense, 657  
 Baseline frequency, **271**  
 Bates, R., 84, 857, 858, 871  
 Batliner, A., 871  
 Bauer, F. L., 10, 427  
 Bauer, L., 56  
 Baum, L. E., 104, 187, 212, 334  
 Baum, L. F., 850  
 Baum-Welch algorithm, *see*  
   forward backward  
   algorithm  
 Baumgartner, Jr., 769, 771  
 Bayer, S., 646  
 Bayes, T., 139, 290  
 Bayesian, **290**  
 Bayesian inference, **139**, **164**,  
   **290**  
 Bazell, C. E., 427  
 BBN, 334  
 BDI, **864**  
 beam search, **325**, 326  
   in MT, 907  
 Beam width, **326**  
 beam width, **326**  
 Bean, D., 733  
 bear pitch accent, 224, 264  
 Bear, J., 766, 767, 774  
 Beaver, D., 266  
 Becker, 620  
 Becket, R., 837  
 Beckman, M. E., 267, 268, 285  
 Beebe-Center, J. G., 908, 910,  
   920  
 Beeferman, D., 699, 700  
 Beesley, K. R., 70, 81, 387  
 belief  
   logic of, 865  
 Bell, A., 119, 229, 358  
 Bell, T. C., 120, 482  
 Bellegarda, J. R., 113, 284, 362  
 Bellman equation, **862**

- Bellman, R., 74, 863  
 Bender, E. M., 428  
 Bennett, C., 285  
 Bennett, S. W., 733  
 Benoît, C., 282  
 Benson, D. B., 619  
 Berck, P., 170  
 Berger, A., 13, 112, 207, 213, 699, 700  
 Bergler, S., 645  
 Bergsma, S., 718, 733  
 Berkeley Restaurant Project, 89  
 Bernstein, J., 334  
 Bertram, R., 79  
 Berwick, R. C., 81, 550  
 Bestgen, Y., 733  
 Beutnagel, M., 284  
 Bever, T. G., 424, 487  
 Bhatt, R., 387  
 Bhembé, D., 827  
 Biber, D., 428  
 Bickel, B., 878  
 Biddulph, R., 10, 333  
 Bies, A., 408, 410  
 Bigham, J., 679  
 Bigram, **88**  
 bigram, **88**  
 Bikel, D. M., 482, 491, 775  
 bilabial, **220**  
 Bilmes, J., 213, 364  
 Bimbot, F., 284  
 Binary branching, **416**  
 binary branching, **416**  
 binary tree, 416  
 binary vectors  
   for word similarity, 676  
 Binding theory, **714**  
 binyan, 370  
 Birch, A., 920  
 Bird, S., 16, 387  
 Birnbaum, L., 620  
 Bisani, M., 284  
 Bitext, **899**  
 bits  
   for measuring entropy, 114  
 Black, A. W., 254, 255, 258, 261, 264, 271, 272, 279–281, 283, 285  
 Black, E., 483, 484, 490, 688  
 Black, H. A., 80  
 Blair, C. R., 170  
 Blair-Goldensohn, S., 818, 823  
 Blake, J. A., 772  
 Blaschke, C., 772, 774  
 Bledsoe, W. W., 11, 290  
 bleeding, **367**  
 Blei, D. M., 13  
 BLEU  
   compared to ROUGE, 819, 820  
 Bleu, 909  
 Blevins, J., 377  
 Blitzler, J., 773  
 Blizzard Challenge, **285**  
 blocks world, 11  
 Bloedorn, E., 809, 823  
 Bloomfield, L., 413, 426  
 Blunsom, P., 491, 532  
 BNC, 94, *see* British National Corpus, 157  
 BNF (Backus-Naur Form), **391**  
 Bobrow, D. G., 587, 825, 836  
 Bobrow, R. J., 532, 837, 839  
 Bock, K., 424  
 Bod, R., 490  
 Body of STRIPS plan, 865  
 Boersma, P., 236, 244, 247, 375, 385  
 Boguraev, B., 645, 733  
 Bohus, D., 846  
 Bojar, O., 920  
 Boland, J., 852  
 Bolinger, D., 224, 265  
 Bonaparte, N., 249  
 BOOK-FLIGHT, 866  
 bootstrapping  
   in IE, 751  
 Booth, T. L., 464, 465, 490  
 bootstrapping, **662**  
   generating seeds, 662  
 bootstrapping for WSD, **662**  
 Borst, J. M., 284  
 Bos, E., 733  
 Boser, B. E., 13  
 Botstein, D., 772  
 Böttner, M., 827  
 bottom-up, 431, 433, 460  
 bottom-up parsing, **434**  
 Bouayad-Agha, N., 733  
 Bouillon, P., 836  
 Bound, **711**  
 bound  
   pronoun, **711**  
 Boundary tone, **267**  
 Bourassa, D., 377  
 Bourlard, H., 351, 354  
 Bouwman, G., 860  
 Boves, L., 860  
 Bowey, J., 377  
 Boye, J., 837  
 Brachman, R. J., 532, 587  
 Bracketed notation, **395**  
 bracketed notation, **395**  
 Branco, A., 733  
 Brants, T., 120, 149–151, 159–162, 168, 733  
 Brasoveanu, A., 387  
 Bratt, E., 852  
 Bratt, H., 871  
 Break index, **268**  
 breath  
   transcription of, 422  
 Breiman, L., 213  
 Brenier, J. M., 266  
 Brennan, S. E., 718, 731, 733, 735  
 Brent, M. R., 383, 387  
 Bresch, E., 285  
 Bresnan, J., 11, 427, 527, 529, 530, 532  
 Brew, C., 813  
 bridging inference, **712**  
 Brill tagger, **135**  
 Brill, E., 135, 151, 154, 155, 159, 163, 168, 170, 171, 343, 490, 797, 800, 824  
 Brin, S., 680, 799  
 Briscoe, T., 484, 491, 645  
 British National Corpus  
   use of C5 tagset to tag, 130  
 broadcast news  
   discourse segmentation, 698  
   speech recognition of, 334  
 Brockett, C., 810, 823, 824  
 Bromberger, S., 388  
 Brookes, D. M., 276  
 Broschart, J., 125  
 Browman, C. P., 244, 284  
 Brown corpus, 85, 86, 94, 130, 135, 152, 155, 169  
   POS tagset for, 130  
 Brown, J. S., 619  
 Brown, M. K., 844, 871  
 Brown, P. F., 13, 86, 111, 118, 354, 690, 894, 895, 904, 915–918, 920  
 Browning, I., 11, 290  
 Bruce, R., 651, 689  
 Bryant, M., 94  
 Buckley, C., 791  
 Budanitsky, A., 171, 665, 669  
 Budzikowska, M., 823  
 Bulut, M., 285  
 Bulyko, I., 112, 282, 359, 858  
 Bunesco, R. C., 750  
 Bunker, R. T., 652  
 Bunt, H., 854  
 Burger, J. D., 724, 733  
 Burges, C., 364  
 Burnett, D., 850  
 Burton, R. R., 619  
 Butler, H., 772  
 Butzberger, J. W., 339, 340  
 Bybee, J. L., 229  
 Byrd, D., 229  
 Byrd, R. J., 284  
 Byrne, W., 243, 343, 351, 920  
 cache language model, **112**  
 Cafarella, M., 690, 754, 755

- Cahn, J. E., 285  
 Calder, J., 532  
 Caley, R., 272  
 Calhoun, S., 266  
 Callison-Burch, C., 912, 920  
 Campbell, I., 822  
 Campbell, N., 285  
 Canadian raising, **388**  
 Candide, **920**  
 Canon, S., 491, 532  
 Canonical form, **556**  
 Cantonese, 877  
 Cao, X., 873  
 Caporaso, J. G., 771  
 Caraballo, S. A., 682  
 Caramazza, A., 730  
 Carberry, S., 151, 869  
 Carbonell, J., 646, 811, 919, 920  
 Cardie, C., 460, 620, 717, 733, 774, 775  
 cardinal number, **399**  
 Carletta, J., 713, 854, 856, 871  
 Carlson, L., 13, 703, 706, 734  
 Carpenter, B., 527–530, 532, 826  
 Carpuat, M., 689  
 carrier phrase in TTS, 275, **282**  
 Carroll, G., 492  
 Carroll, J., 484, 491, 524, 532, 670, 688, 689  
 Carroll, J. B., 919, 921  
 Carstensen, F. V., 139  
 Carter, D., 837  
 Carter, D. M., 363  
 Casari, G., 772  
 Cascade, **65**  
 cascade, 26, 367  
   transducer, 10, 65, **65**  
 case  
   as feature in IE, 742  
 Case frame, **634**  
 case sensitive  
   regular expression search as, 19  
 case, nominal, **404**  
 Castano, J., 756, 759  
 Castao, J., 762  
 Cataphora, **711**  
 categorial grammar, 420, **420**, 428, 619  
 Cavedon, L., 871  
 cd, *see* conceptual dependency  
 CD phone, **348**  
 Cedergren, H. J., 230  
 ceiling  
   for part-of-speech tagging evaluation, 155  
   human, 155  
 CELEX, 81, 126, 241  
 cell on tape of finite-state automaton, 27  
 center-embedded, **542**  
 center-embedding, 417, 536, **542**, 546, 549  
 Centering  
   algorithm for information ordering in summarization, 813  
 Centering algorithm for pronoun resolution, 720  
   psychological tests of, 731  
 Centering theory, **718**  
   algorithm for pronoun resolution, 718  
   model of entity-based coherence, 718  
 centrality  
   of sentences in content selection for summarization, **805**  
 Centroid, **823**  
 cepstral coefficients  
   uncorrelated, 303  
 cepstral distance  
   in TTS, 280  
 cepstral mean normalization, **355**  
 Cepstrum, **302**  
 cepstrum, 297, 333  
   and covariance matrix, 303  
   delta feature, 304  
   formal definition, 303  
   in TTS, 280  
 CFG, 390, **391**  
   multiplying probabilities, 489  
 Chafe, W. L., 733  
 Chan, Y. S., 689  
 Chandiox, J., 919  
 Chandler, S., 646  
 Chang, E., 357  
 Chang, J. S., 688  
 Chang, N., 81, 585  
 Chang, P., 81  
 channels in stored waveforms, **232**  
 Chanod, J., 70  
 Charles, W. G., 669  
 Charniak parser, **477**  
 Charniak, E., 16, 118, 119, 156, 364, 413, 477, 486, 490–492, 683, 688, 726, 733, 823, 824  
 Chart, **447**  
   chart  
     in parsing, **447**  
 Chart parsing, **452**  
 chart parsing  
   fundamental rule, **453**  
 Chawla, N. V., 364  
 Chelba, C., 468, 487  
 Chen, F., 733, 823  
 Chen, J., 690  
 Chen, J. N., 688  
 Chen, S. F., 89, 112, 120, 213, 254, 255, 284  
 Chen, X., 243  
 Cherry, C., 914  
 Chi, Z., 491, 532, 550  
 Chiang, D., 491, 689, 913, 914  
 Chiba, S., 213, 333  
 Chierchia, G., 587  
 Chinchor, N., 774  
 Chinese, 215, 873, 878, 880  
   POS tagging, 160  
   word segmentation, **70**  
 Chklovski, T., 688–690  
 Chodorow, M. S., 284, 665, 731  
 Choi, F. Y. Y., 698, 733  
 Chomsky  
   adjunction, 416  
 Chomsky hierarchy, **536**, 548  
 Chomsky normal form, **416**, 426, 469  
 Chomsky, C., 586, 823  
 Chomsky, N., 9, 10, 12, 63, 83, 119, 228, 285, 369, 389, 391, 416, 417, 427, 531, 536, 538, 544, 546, 547, 549, 714  
 Chomsky-adjunction, **416**  
 Chou, W., 351, 919  
 Chow, Y.-L., 339, 348, 491  
 Christiansen, M. H., 387  
 Chronological ordering, **812**  
 Chu-Carroll, J., 826, 844, 869, 871  
 Chung, G., 850  
 Chunking, **455**  
 chunking, 455  
 Church, A., 565  
 Church, K. W., 12, 81, 86, 100, 102, 103, 113, 119, 120, 139, 155, 164–167, 170–172, 257, 259, 380, 437, 462, 549, 657, 663, 674, 920, 931  
 CI phone, **348**  
 Ciaramita, M., 682, 688  
 Cieri, C., 334  
 circumfix, 47  
 CIRCUS, 620, 774  
 Citation form, **623**  
 Citizen Kane, 693  
 CKY algorithm, 431, 461, 463  
   probabilistic, 489  
 clarification subdialogue, **860**  
 Clarissa, **871**  
 Clark, A., 170, 387, 492  
 Clark, E. V., 424  
 Clark, H. H., 85, 229, 424, 713, 730, 830, 832, 833, 854, 871, 872

- Clark, J., 247  
 Clark, S., 491  
 Clarke, J., 823, 824  
 Clarkson, P. R., 109  
 clash (prosodic), 266  
 Class-based N-gram, **111**  
 class-based N-gram, 111, **111**  
 Classen, W., 733  
 Classification, **199**  
 classification  
   logistic regression for, 199  
 classification versus regression, 194  
 Clause, **398**  
 Cleft, **713**  
 click due to bad joins in TTS, 276  
 click studies, 424  
 Clifton, C., 362  
 Clinton, W., 122  
 Clitic, **48**  
 clitic, **48**, 51  
   origin of term, 123  
 Cliticization, **48**  
 cloning of states in acoustic modeling, **350**  
 closed class, **124**  
 Closed vocabulary, **94**  
 closed vocabulary, 97  
 closure, 40  
   stop, **221**  
 Clow, J., 287  
 cloze task  
   for evaluating translations, 909  
 Cluster N-gram, **111**  
 clustering, **306**  
   in word sense disambiguation, 689  
   K-means, **306**  
 Clustering By Committee, 690  
 CMLE, *see* conditional maximum likelihood estimation, **353**  
 CMU., 225, 242, 258, 316  
 CNF, **416**  
 co-occurrence  
   first order, 686  
   second order, 686  
 Coarticulation, **273**  
 coarticulation, **227**  
   as motivation for triphone acoustic model, **348**  
 Coccaro, N., 113, 362, 857, 858, 871  
 Cochlea, **239**  
 cochlea, **239**  
 Cocke, J., 13, 904, 920  
 Cocke-Kasami-Younger algorithm  
   probabilistic, 489  
 coda  
   syllable, **223**  
 codebook  
   in VQ, **305**  
 codeword  
   in VQ, **305**  
 Cohen, B. K., 773  
 Cohen, D. N., 213  
 Cohen, J., 364  
 Cohen, K. B., 769–771, 773  
 Cohen, M. H., 358, 832, 840, 841, 846, 850, 860, 871  
 Cohen, M. M., 362  
 Cohen, P. R., 12, 156, 287, 825, 850, 864, 867, 870, 871  
 Cohen, W. W., 813  
 Coherence, **695**  
 coherence  
   different from cohesion, 697  
   relations, **695**, **701**  
   compendium of relations, 734  
 coherence conditions in LFG, **527**  
 cohesion  
   different from coherence, 697  
   in discourse, 697  
   lexical, 697  
 Cohesion chain, **697**  
 Coker, C., 284  
 Colby, K. M., 870  
 cold languages, 878  
 Cole, J. S., 369, 371  
 Cole, R. A., 9, 362, 363, 850  
 Coleman, J., 247, 379, 387  
 COLING, 15  
 collaborative completion, **831**  
 collection  
   in IR, **782**  
 Collier, N., 773  
 Collins parser, **477**  
 Collins, M., 264, 413–415, 419, 474, 475, 477–479, 481, 482, 484, 490–492, 683  
 collocation, **652**  
 collocational features, **652**  
   in WSD, **652**  
 collocations, **674**  
 Colmerauer, A., 11, 531  
 Colombe, J. B., 771  
 Colosimo, M. E., 771  
 Combinatory categorial grammar, **421**  
 combinatory categorial grammar, **538**  
 COMLEX, 405  
 commissive, 830  
 common ground, **830**  
 common noun, **125**  
 complement, 405, **405**, 426  
 complementizers, **128**  
   table of, 128  
 Complete, **569**  
 completeness  
   in FOL, 569  
 complex onsets, 372  
 Complex question, **781**  
 Complex types, **527**  
 complex types, **527**  
 Complex-term, **599**  
 complexity, **535**, 536, 546, 548  
   caused by memory limitations, 549  
   center-embedding, 536, 546, 549  
   generative power, 548  
   Swiss German, 544  
 componential analysis, 641  
 composition  
   transducer, **58**, 79, 367  
 compound nominals, 608  
   semantic attachments, 608  
 Compounding, **48**  
 compression, 231  
 Computational Grammar Coder (CGC), 169  
 computational lexical semantics, 649  
 computational phonology, 365, **365**  
 computer-aided human translation, **876**  
 Comrie, B., 877  
 concatenation, 41  
 concatenation as basic FSA operation, 40  
 concatenative morphology, **52**  
 Concept accuracy, **851**  
 conceptual dependency, 552, 642  
 concordance  
   semantic, **651**  
 conditional independence, 489  
 Conditional maximum likelihood estimation, **200**  
 conditional maximum likelihood estimation, 352  
 Conditional Random Field, **214**, 266  
 Condon, S., 854  
 confidence  
   ASR, **342**, 860  
   in relation extraction, 754  
 confusion matrix  
   in noisy channel spelling, **166**  
 confusion matrix for error analysis, **156**  
 confusion network, **343**  
 conjoined phrase, 407  
 conjugate gradient, 201  
 conjunctions, 126, **128**, **407**



- table of, 128
- Conkie, A., 275, 282, 284
- Connective, **705**
- Connine, C. M., 362, 377
- CoNNL, 15
- Connolly, D., 724, 733
- Conrad, S., 428
- Conrath, D. W., 668
- Conroy, J. M., 810, 813, 823
- Consistent, **465**
- consonant, **219**
- constant entropy rate hypothesis, **118**
- Constantin, A., 920
- constants
  - in FOL, **562**
- constituency, **389**, 417
  - evidence for, 391
- constituent, 405, 425
  - book titles which are not, 389
- Constituent Likelihood
  - Automatic Word-tagging System, *see* CLAWS and tagset
- constraint
  - unification features as, 493
- Constraint Demotion, **385**
- Constraint Grammar, 135, 419, 428
- constraint-based formalism, **493**
- Construction Grammar, 427, 620
- content planner, 839
- Content selection, **804**
- context-free, **537**
- context-free grammar, **395**, 425
  - see* CFG, **391**
  - and Chomsky hierarchy, 537
  - Chomsky normal form, 416
  - finite-state approximations, 426, 428
  - invention of, 427
  - non-terminal symbol, 392
  - productions, 391
  - rules, 391
  - semantic attachments, 591
  - terminal symbol, 392
  - weak and strong equivalence, 416
- context-free language, 425
- context-sensitive grammar, **536**, 537
- context-sensitive spell correction, **167**
- Conti, J., 850
- contingency table for error analysis, **156**
- Continuation rise, **267**
- Continuer, **831**
- continuer, **831**
- Continuous speech, **288**
- continuous speech recognition, **288**
- contrastive knowledge
  - in MT, 884
- Contribution, **830**
- contribution, **830**
- Control, **509**
- conventional metaphor, **644**
- conversation, **825**
- Conversation Analysis, **828**, 871
- conversational act, **854**
- Conversational agent, **1**
- conversational agent, 8, 825
  - acknowledgement in, 832
- conversational implicature, 834
- conversational move, **854**
- Conversational speech, **288**
- Convex optimization, **201**
- Convolutional noise, **355**
- Cooley, J. W., 333
- Cooper, F. S., 247, 284
- coordinate noun phrase, **407**
- Coordination ambiguity, **436**
- coordination ambiguity, 474
- Copestake, A., 529, 532, 645
- copula, **128**
- Córdoba, R., 842, 860
- Core Language Engine, 606, 619
- Core, M., 854
- Corefer, **708**
- Coreference chain, **708**
- Coreference resolution, **708**
- coronal sound, **220**
- Corpora, **85**
- corpora, **85**
- CORPORATION AS PERSON
  - metaphor, 644
- Corpus, **85**
- corpus, **85**
  - ATIS, 335, 393
  - BNC, 94, 157
  - Broadcast news, 334
  - Brown, 85, 86, 94, 130, 135, 152, 155, 169
  - CASS phonetic of Mandarin, 243
  - fisher, 334
  - Hansards, **899**
  - Hong Kong Hansards, 899
  - Kiel of German, 243
  - LOB, 169
  - MULTEXT-East, 162
  - regular expression searching inside, **18**
  - Resource Management, 334
  - Switchboard, 85, 86, 225, 229, 231, 232, 242, 243, 297, 318, 335, 423
  - Switchboard phonetic, 243
  - TimeBank, 762
  - TIMIT, 243
  - Wall Street Journal, 97, 334
  - Ziff-Davis, 807, 809
- Corpus Lesk, **659**
- correction act detection, **857**
- Corston-Oliver, S. H., 734
- Cosine, **783**
- cosine, 822
  - as a similarity metric, **783**
- cosine similarity metric between vectors, 677
- Cote, S., 718, 733
- Cotton, S., 13, 651, 652
- Cottrell, G. W., 688
- count noun, **125**
- counters, 41
- counts
  - treating low as zero, 160
- covariance matrix, 311
  - diagonal, **311**
- Cover, T. M., 113–115, 118
- Cowan, B., 920
- Cowie, J., 689
- Cowper, E. A., 547
- Cox, S., 332
- Crammer, K., 491
- Crawley, R. A., 730
- Crestani, F., 822
- CRF, **214**
- Croft, W., 285, 877
- Croft, W. B., 790
- Cronin, B., 646
- cross-brackets, 484
- Cross-entropy, **116**
- cross-entropy, **116**
- cross-serial dependencies, 535, **544**
- crossvalidation, **155**
  - 10-fold, **155**
- Crouch, C. J., 792
- Crouch, R., 484, 491, 532, 823
- Cruse, D. A., 645
- Crystal, D., 285
- Csaba Pléh, 547
- Cuayáhuitl, H., 871
- Cue phrase, **705**
- cue phrase
  - sentential use, **705**
- cue word
  - as feature for discourse segmentation, **699**
- Culicover, P. W., 427
- Cullingford, R. E., 11, 919
- Culotta, A., 750
- Culy, C., 544, 734
- Cumming, S., 285
- Cunningham, H., 766
- Curran, J. R., 491, 672, 674, 675, 677, 679, 689
- Currying, **566**

- Cutler, A., 124, 363  
Cutting, D., 150, 687  
CV skeleton, 370  
cycles in a wave, 231  
Cycles per second, **231**  
cycles per second, 231  
Czech, 162
- d**  
parameter in IBM Model 3, **914**  
Daelemans, W., 170, 284, 460, 491, 532  
Dagan, I., 462, 674, 677, 678, 689  
Dahlbäck, N., 854  
Dale, R., 734, 840, 859  
Daly, N. A., 871  
Damerau, F. J., 73, 165, 168, 170  
Damper, R. I., 284  
DAMSL, 854  
Dang, H. T., 13, 635, 651, 652, 656, 657, 688  
Danieli, M., 845, 851  
Darroch, J. N., 201  
data-directed search, 433  
date  
fully qualified, **759**  
date normalization, **837**  
Dative alternation, **635**  
Daumé III, H., 809, 823, 824  
David, Jr., 333  
Davidson, D., 572, 587  
Davidsonian, **572**  
Davis, A. P., 772  
Davis, E., 587  
Davis, K. H., 10, 333  
Davis, S., 302  
Day, D., 13, 762  
Day, D. S., 733  
de Courtenay, B., 247  
decision list, 171, 653, 655  
learning, 655  
decision tree  
use in WSD, 688  
declaration, 830  
declarative, 396, 426  
declarative constraint, 367  
declarative sentence structure, **396**  
Declination, **271**  
Decoder, **184**  
decoder, 184  
Decoding, **184, 317**  
decoding, **145, 184, 290, 292, 317**  
A\*, *see* A\* decoder  
as search, 290  
beam search, 325
- dynamic programming, *see*  
Viterbi algorithm  
MT, using IBM Model 1, 897  
multiple-pass, 339  
N-best, 339  
pruning in, 325  
stack, *see* A\* decoder  
Viterbi, *see* Viterbi algorithm, **145, 184**  
word lattice, 340
- deduction  
as type of inference, **725**  
deep role, **633**  
Deerwester, S., 690  
default pair, **61**  
Defaults, **529**  
defeasible inferences, **729**  
Definite Clause Grammars, **531**  
definite reference, 710  
degree adverb, **126**  
Dejean, H., 13  
DeJong, G. F., 646, 774, 823  
Delattre, P. C., 247  
deleted interpolation, **150**  
deletion, **227**  
Delfs, L., 13, 651, 652  
Delgado, R. L.-C., 852, 871  
Deline, S., 284  
Della Pietra, S. A., 13, 86, 118, 201, 207, 213, 894, 895, 904, 915–918, 920  
Della Pietra, V. J., 13, 86, 111, 118, 201, 207, 213, 690, 894, 895, 904, 915–918, 920  
delta feature, 304, **304**  
Demberg, V., 256  
Demetriou, G., 113  
Demner-Fushman, D., 824  
Dempster, A. P., 104, 167, 187  
Denes, P., 333  
Deng, L., 289, 348  
Deng, Y., 920  
Denotation, **558**  
dental sound, **220**  
dependency, 417  
grammar, 418  
lexical, 474  
relations, 390  
dependency grammar, **418**  
dependent-marking languages, **877**  
depth-first search, **37, 41**  
Dereferencing, **513**  
Derivation, **48**  
derivation  
compared to inflection, 51  
productivity of in English, 51  
syntactic, 392, **392, 395, 396, 425**  
direct, **395**  
derivational, 79  
Dermatas, E., 158  
DeRose, S. J., 135, 137, 139, 150  
Det, 392  
determiner, 126, **127, 392, 399, 426**  
deterministic algorithm, **28**  
deterministic finite-state  
automaton, 32  
development test set, **92, 153**  
devset, **92**  
devtest, **153**  
De Jong, N. H., 79  
de Lacalle, O. L., 688  
de Marcken, C., 383  
De Mori, R., 112, 362, 871  
de Rijke, M., 760, 775  
de Souza, P. V., 111, 347, 354, 690  
de Villiers, J. H., 850  
DFSA, **32, 41**  
DFT, **300**  
Diab, M., 688  
Diagnostic Rhyme Test, **282**  
diagonal covariance matrix, **311**  
dialect variation, **229**  
Dialogic pair, **828**  
Dialogue, **825**  
dialogue, **825**  
compared to monologue, **693**  
information-state model of, **853**  
dialogue act, 829, **854**  
acknowledgement, 831  
backchannel, 831  
continuer, 831  
correction, 857  
dialogue acts  
question detection, 857  
dialogue manager, 841  
design, 850  
finite-state automaton, 841  
frame-based, 843  
dialogue structure  
intentional, 868  
Dialogue system, **1**  
dialogue systems  
design, 850  
evaluation, 850  
Diathesis alternation, **635**  
Dice similarity metric, **677**  
Dienes, P., 161, 162, 887  
Dietterich, T. G., 156  
diff program, 77  
Digalakis, V. V., 339, 340  
Digilakis, V., 364  
digit pronunciation  
paired method, 254  
digitization, **232, 297**



- digits task  
   ASR, **288**  
 Dimitrova, L., 162  
 Dingare, S., 713  
 Diphone, **273**  
 diphone synthesis  
   compared to unit-selection, 278  
 Diphthong, **222**  
 diphthong  
   origin of term, 123  
 Direct translation, **881**  
 directed graph as representation  
   of FSA, 27  
 directional adverb, **126**  
 directive, 830  
 Directive prompt, **844**  
 directive prompt, **850**  
 directly derive, **395**  
 disambiguation, **67, 135**  
   homograph, 256  
   PCFGs for, 465  
   period, 251  
   role of probabilistic parsing, 463  
   syntactic, **437**  
   via PCFG, 467  
   word sense, 649  
 Discount, **99**  
 discount, **99, 100**  
 discounted reward, **862**  
 Discounting, **99**  
 discounting, **99**  
 Discourse, **693**  
 discourse  
   context  
   role in reference resolution, **708**  
   model, **708**  
   purpose, **868**  
   segment, **702**  
   purpose, **868**  
 Discourse marker, **705**  
 discourse marker  
   as feature for discourse segmentation, **699**  
 Discourse parsing, **704**  
 Discourse purpose, **868**  
 Discourse segment purpose, **868**  
 Discourse segmentation, **696**  
 discourse segmentation  
   broadcast news, 698  
   discovery procedure, 427  
 Discrete Fourier Transform, **300**  
 discriminative training  
   in ASR, **351**  
 discriminative versus generative  
   models, 209  
 disfluency, **85, 360, 422, 422**  
 disfluent, 426  
 Disjunction, **530**  
 disjunction, 41  
   in feature structures, **530**  
   pipe in regular expressions as, **21**  
   square braces in regular expression as, 19  
 dispreferred response, **829, 872**  
 Distal, **711**  
 Distance, **481**  
 distance  
   edit, **74**  
   Euclidean, 307, 676  
   Mahalanobis, **307**  
   Manhattan, 675  
   or distortion, in MT, 892  
 Distance metric, **306**  
 distance metric, **306**  
 distinctive feature, **227**  
 distortion metric, **306**  
 distortion probability  
   in IBM Model 3, **915**  
   in MT, 891, 892  
 distributional similarity, 427  
 ditransitive, 425  
 ditransitive alternation, 425  
 Divay, M., 256  
 divergences between languages  
   in MT, **876**  
 Dixon, N., 283  
 Di Eugenio, B., 733  
 document  
   in IR, **782**  
 Document Understanding  
   Conference, 824  
 Doddington, G., 335, 390, 825, 921  
 Dogil, G., 271  
 Doherty-Sneddon, G., 854, 856, 871  
 Dolan, W. B., 688, 824  
 Dolbey, A., 771  
 Dolinski, K., 772  
 Doolittle, Doctor, 27  
 Domain, **558**  
 dominance  
   discourse relation, 868  
 dominates, **393**  
 domination in syntax, **392**  
 Donovan, R. E., 281, 284  
 Dorr, B., 646, 810, 823, 876, 921  
 Dostert, L., 919  
 Dot product, **196, 784**  
 dot product, **676**  
 Dotted rule, **447**  
 dotted rule, **447**  
 double delta feature, **304**  
 Doumpiotis, V., 343, 351  
 Dowding, J., 871  
 Downey, D., 690, 754, 755  
 Downing, P., 620  
 downstepped high in ToBI, 272  
 Dowty, D. R., 428, 575, 578, 587, 642  
 DP, 868  
 Dragon Systems, 334  
 Dream of the Red Chamber, 873  
 Drew, P., 828  
 DRT, **282**  
 DSP, 868  
 Dubner, S. J., 194  
 DUC, 824  
 duck  
   as example of ambiguity, 4  
 Duda, R. O., 213, 306, 687  
 Dudík, M., 213  
 Dumais, S. T., 679, 690, 797, 800, 824  
 Dunn, H. K., 10, 247  
 Dunning, T., 805  
 duration  
   in TTS, 277  
   temporal expression, **755**  
 Durbin, R., 213  
 Dutoit, T., 285  
 Du Bois, J. W., 285  
 Dwight, S. S., 772  
 Dyer, C., 920  
 dynamic programming, **74, 146, 213, 431, 439, 461**  
   forward algorithm as, 181  
   invariant, 338  
   Viterbi as, 184  
 Dynamic time warping, **333**  
 Džeroski, S., 162  
 Di Eugenio, B., 733  
 E-step (expectation step) in EM, **192**  
 E. E., 333  
 EACL, 15  
 Eagon, J. A., 212, 334  
 Earley algorithm  
   integrating with semantic analysis, 615  
 Earley algorithm, 431, 461, 518  
 Earley, J., 431, 447, 449, 462  
 Earnest  
   Importance of Being, 72  
 Eberlein, J., 773  
 Echihiabi, A., 706, 707, 799  
 Eckert, W., 861–863  
 Eddy, S., 213  
 edit cost  
   of post-editing MT, 909  
 edit distance, **74**  
   minimum, **74**  
 edit term, **423**  
 Edmonds, P., 689  
 Edmunson, H., 823

- Effects of STRIPS plan, 865
- EKG (electroglottograph), **276**
- EGYPT, **920**
- Ehlen, P., 871
- Eide, E. M., 281, 285, 364
- Eisner, J., 177, 387, 491, 492, 914
- Ejerhed, E. I., 462
- Elaboration (as coherence relation), **701**
- Electroglottograph, **276**
- Elhadad, M., 532, 809
- Elhadad, N., 812
- ELIZA, **7**
  - implementation, 26
  - sample conversation, 25
- Ellis, D., 225, 243
- Ellison, T. M., 374, 379, 387
- Ellsworth, M., 637
- Elman, J. L., 363
- Elson, D., 810
- Elvgren III, G., 380
- EM, 104, 167
  - E-step, 192
  - in training MEMM, 211
  - M-step, 192
- EM algorithm, 187
- Embedded training, **327**
- embedded training, **327**, 355
- embedded verb, 401
- emission probabilities, **145**, **177**, **293**, **296**, **318**, **327**
- EMNLP, 15
- Empiricist, **433**
- empiricist, **433**
- empty category, 398
- enclitic, **51**
- energy in frame, 304
- EngCG, 135, **137**
- Engel, R., 871
- English, 878
  - amount of utterance overlap in dialogue, 828
  - argument by Jim McCawley that infixes exist in, 47
  - derivational morphology of, 50
  - entropy of, 117
  - grammar of spoken, 421
  - habitual present form, 49
  - inflectional morphology of, 48
  - lexical differences from French, 881
  - number of irregular verbs in, 49
  - simplified grammar rules, 393
- Entity grid, **813**
- entropy, **114**
  - and branching factor in letter prediction for boundaries, 382
  - and MDL, 382
  - and perplexity, **114**
  - cross-entropy, 116
  - of English, 117
  - of morphological paradigm, 79
  - relation to thermodynamics, 10
  - relative, 661, 678
- Entropy rate, **115**
- entropy rate, **115**
  - constancy hypothesis, 118
- epoch (pitch pulse), 276
- epoch detection, 276
- Eppig, J. T., 772
- Ergodic HMM, **178**
- Erjavec, T., 162
- Erkan, G., 806, 809
- errors
  - analyzing, 156
- Espy-Wilson, C., 364
- Essen, U., 113
- Estonian, 162
- Etzioni, O., 690, 754, 755
- et al, 364, 919, 921
- Euclidean distance, **306**, 676
- Eugene Oegin, 119, 173, 212
- Euler, 300
- Euler's formula, **300**
- EUROSPEECH, 16
- EVAL in Optimality Theory, 372
- EVALB, **484**
- evaluating parsers, 483
- evaluation, **91**
  - 10-fold crossvalidation, 155
  - adequacy in MT, 909
  - baseline, 155
  - most frequent class, 156
  - baselines in summarization, 821
  - Bleu in MT, 909
  - ceiling, 155
  - cloze task, 909
  - comparing models, 97
  - coreference resolution, 724
  - crossvalidation, 155
  - development test set, 92, 153
  - devset, 92, **153**
  - dialogue systems, 850
  - discourse segmentation using  $P_k$ , 700
  - discourse segmentation using WindowDiff, 700
  - distributional word similarity, 679
  - error analysis, 156
  - extrinsic, **95**, **656**
  - fidelity in MT, 909
  - fluency in MT, 909
  - human ceiling, 155, 657
  - in vitro, **656**
  - in vivo, **95**
  - Kendall's  $\tau$ , 819
  - Matched-Pair Sentence Segment Word Error (MAPSSWE), 156, 331
  - mean opinion score, 283
  - METEOR for MT, 910
  - model-theoretic coreference scoring scheme, 724
  - most frequent class baseline, 156
  - MT, 908
  - MUC coreference, 724
  - named entity recognition, 744
  - of N-gram, 91
  - of N-grams via perplexity, **95**
  - of summarization, 819
  - precision and recall for MT, 910
  - pseudoword, 657
  - Pyramid Method in summarization, 820
  - relation detection, 754
  - relevance feedback, 791
  - ROUGE for summarization, 819
  - test set, 91
  - training on the test set, 92
  - training set, 91
  - choosing, 153
  - TTS, 282
  - unsupervised WSD, 687
  - word similarity, 669
  - WSD systems, 656
- evaluation
  - TER for MT, 910
- Evans, N., 125, 126
- Evans, R., 532
- Event detection and classification, **738**, **761**
- Event variable, **572**
- events
  - representation of, 569
- Evermann, G., 316, 343, 358
- Evidence, **703**
- evoking a referent, **708**
- existential there, **129**
- expansion, 393, 397
- Expectation step, **471**
- expectation step in EM, 192
- Expectation-Maximization, 187
- Explanation (as coherence relation), **701**
- Explicit confirmation, **845**
- exponential classifiers, 193
- expressive, 830
- expressiveness

- of a meaning representation, 557
- Extended Gloss Overlap, **668**
- Extended Lesk, **668**
- extract
  - in text summarization, **802**
- extraposition, **713**
- Extrinsic evaluation, **95**
- extrinsic evaluation, **656**
- eye-tracker, **487**
- F (for F-measure), 459, 483, 745
- $f^*$ , 346
- $F_0$ , **233**
- F-measure, **459, 483**
  - in NER, 744
- F-M., 387
- Fackrell, J., 259
- factoid question, **780**
- FAHQT, 875, **875**
- fail state, **30**
- Fail type, **527**
- fail type, **527**
- FAITHC, 372
- faithfulness in MT, 889
- faithfulness in phonology, 373, 380
- FAITHV, 372
- False negative, **23**
- False positive, **23**
- Fano, R. M., 674
- Fanshel, D., 856
- Fant, C. G. M., 284
- Fant, G., 247
- Fant, G. M., 247, 284, 361
- Fanty, M., 850
- Farrar, K., 285
- Farsi, 878
- Fass, D., 620, 646
- Fast Fourier Transform, **301, 333**
- fast match in ASR decoding, **344**
- fast match, 345
- FASTUS, **766**
- Fauconnier, G., 587
- feasible pair, **61**
- feature
  - distinctive, 227
  - extraction of spectral, 292
  - morphological, **53**
  - phonetic, 227
- feature structure
  - AGREEMENT, 495
  - reentrant, 496
- feature co-occurrence
  - restrictions, **527**
- Feature path, **495**
- feature structure, 404
  - NUMBER, 494
  - CAT, 495
  - as DAGs, 511
- constraining values of
  - features, 527
- DAG content field, 512
- DAG pointer field, 512
- description of versus instance
  - of, 530
- disjunction in, **530**
- need for typing in, 526
- negation in, **530**
- path inequalities in, **530**
- reentrant, 498
- set-valued features, **530**
- Feature structures, **494**
- Feature vector, **297**
- feature vector, **652**
  - for distributional word similarity, 670
  - in IR, **782**
  - in WSD, **652**
- features
  - articulatory, 244
  - in linear regression, 195
  - phonetic, 244
- Federalist papers, 11, 121
- Federico, M., 920
- Fedorenko, E., 359
- Feldman, J., 585
- Feldman, L. B., 79
- Fellbaum, C., 13, 629, 645, 651, 652, 688
- Feng, F., 766
- Fensel, D., 584
- Ferguson, G., 853, 859, 864
- Ferguson, J., 179
- Ferguson, M., 408, 410
- Ferreiros, J., 842, 860
- Ferrer, L., 335, 871
- Ferro, L., 13, 756, 759, 762
- fertility
  - in MT, **914**
- FFT, **301, 333**
- fidelity
  - in MT, 909
- FIFO, 37, 41
- Fikes, R. E., 865, 870
- file format
  - .wav, 233
- Filippova, K., 733
- filled pause, 85, **423, 426**
- Filler, **511**
- filler, **85, 423**
- Fillmore, C. J., 12, 427, 508, 509, 532, 587, 620, 629, 633, 637, 645, 646
- Final fall, **267**
- final fall, 267
- Final lowering, **856**
- final state, 27
- Finegan, E., 428
- Finin, T., 83, 620
- finite automaton, 80
- finite-state approximations to CFGs, 428
- Finite-state automaton, **26**
- finite-state automaton, 17, **26, 27, 27, 40, 41**
  - $\epsilon$ -transitions in, 32
  - accepting by, 27
  - accepting state, 27
  - agenda in non-deterministic recognition, 33
  - concatenation, 40
  - converting non-deterministic to deterministic, 38
  - deterministic, 32, 41
  - deterministic recognition, 29
  - deterministic recognition of, **28**
  - dialogue manager, 841
  - failure in recognition of, 29
  - final state, **27**
  - for English money phrases, 31
  - for English nominal inflection, 54
  - intersection, **67**
  - Kleene closure as basic operation for combining, 40
  - Moore vs. Mealy machines, 80
  - non-deterministic, 32, 41
  - non-deterministic recognition, 33, 35
  - recognition by, 27
  - rejection by, **28**
  - search and non-deterministic recognition, 36
  - sink state, 30
  - start state, **27**
  - state-transition table for, **28**
  - states in, **27**
  - union as basic operation for combining, 40
  - weighted, 144, 174, 378
- finite-state morphology, 60
- finite-state transducer, 46, 57, **57, 79, 386**
  - and OT, 374
  - as generator, 57
  - as recognizer, 57
  - as set relater, 57
  - as translator, 57
  - cascade of, 10, 65
  - closure properties, 58
  - composition, 58
  - epsilon in, 58
  - machine learning of, 380
  - non-determinism, 67
  - other, 64
  - regular relations, 58

- two-level morphology, 65
- union, 58
- Fink, G. A., 228
- First In First Out, 37
- first order co-occurrence, 686
- First Order Logic, *see* FOL
- first-order Markov chain, 175
- Firth, J. R., 670, 689
- Fischer, M. J., 74, 170, 213
- fish
  - plural of, 45
- Fisher, D., 620, 766, 774, 775
- Fisher, W., 334, 378
- Fitt, S., 258
- Fitzpatrick, E., 264
- Flammia, G., 854
- Flanagan, J. L., 284
- flap (phonetic), **222**, 226
- flapping, 228
- Flat start, **328**
- flat start, **328**
- Flemming, E., 229
- Flickinger, D., 483, 484, 491, 532
- fluency
  - in MT, 909
- fluency in MT, 889
- Fodor, J. A., 424, 587, 646, 660
- FOL, 552, 561
  - $\exists$  (there exists), 563
  - $\forall$  (for all), 564
  - $\Rightarrow$  (implies), 567
  - $\wedge$  (and), 563, 567
  - $\neg$  (not), 563, 567
  - $\vee$  (or), 567
  - and verifiability, 561
  - constants, 562
  - expressiveness of, 561
  - expressiveness of, 558
  - functions, 562
  - inference in, 561
  - terms, 562
  - variables, 562
- fold (in crossvalidation), 155
- Fong, C., 264
- food in NLP
  - ice cream, 177
- Forced alignment, **329**
- forced alignment
  - for TTS, 275
- Ford, C., 285
- Ford, M., 547
- Form-based, **843**
- formal language, **30**, 41, 395, 536
  - for soda machine, 31
- Formant, **239**
- formant, **239**
- formant synthesis, **284**
- Forney Jr., 213
- Forward algorithm, **181**
- forward algorithm, 74, 183
  - use in stack decoding, 344
- Forward chaining, **568**
- Forward looking center, **718**
- forward trellis, 181
- forward-backward algorithm, 187, 213, 310, 343, 471
  - backward probability in, 188
  - for GMMs, 314
- forward-backward search, 343
- FORWARD-BACKWARD ALGORITHM, 192
- FORWARD ALGORITHM, 183, 321
- Fosler, E., 90, *see* Fosler-Lussier, E., 487
- Fosler-Lussier, E., 229, 358
- Foster, D. W., 121, 122
- Foster, G., 919
- Fox, B., 285, 423
- Fox, H., 839
- Fox, L., 769
- Fox Tree, J. E., 85, 229
- Fraenkel, A. S., 792
- fragment
  - transcription of, 421
  - word, 422, **423**
- fragment of word, **85**
- Frakes, W. B., 786, 822
- Frame, **299**
- frame
  - semantic, 629, **637**
- Frame element, **637**
- Frame shift, **299**
- Frame size, **299**
- frame-and-slot semantics, 836
- Frame-based, **843**
- frame-based acoustic model, **364**
- FrameNet, 629, **637**
- Frames, **578**
- frames, 836
- Francis, H. S., 472
- Francis, W. N., 11, 85, 94, 123, 130, 134, 135, 169
- Frank, E., 213
- Frank, R., 374
- Frankel, J., 364
- Franz, A., 120, 156, 168, 170, 491
- Fraser, A., 920
- Fraser, N. M., 532, 850, 852
- Free word order, **419**
- Freitag, D., 159, 213, 765, 775
- French, 877
  - homographs, 256
- frequency
  - of a signal, **231**
- fricative sound, **221**
- Friedl, J. E. F., 42
- Friedman, J. H., 213
- Friedman, M. W., 718, 733, 735
- Frisch, S. A., 387
- Fromer, J. C., 845
- Fromkin, V., 78
- fronting of [uw], 216
- Frump, 774
- Fry, D. B., 247, 333
- FSA, **26**, *see* finite-state automaton
- FST, **57**, *see* finite-state transducer
  - as sequence classifier, 173
  - projection of, **59**
- Fujisaki, **285**
- Fujisaki, H., 285
- Fukui, Y., 871
- full listing hypothesis, **78**
- Fully qualified date expressions, **759**
- fully-connected HMM, **178**
- function word, **125**, 168
- functional grammar, 428
- functions
  - in FOL, **562**
- Fundamental frequency, **233**
- fundamental frequency, **233**
- Fundamental rule, **453**
- fundamental rule
  - of chart parsing, **453**
- Funeral Elegy, 122
- Fung, P., 243, 689, 914
- Furnas, G. W., 690
- Furui, S., 824
- fusion language, **877**
- g, 346
- G. D., 213
- G. E., 81, 550
- Gabrilovich, E., 690
- Gaizauskas, R., 13, 762, 766, 775
- Gale, W. A., 81, 86, 100, 102, 103, 119, 120, 155, 164–167, 170, 171, 657, 663, 920, 931
- Gales, M., 316, 343, 364
- Galescu, L., 284
- Galley, M., 823, 914
- Galliers, J. R., 95
- Gao, J., 113, 120
- Gap list, **511**
- gap-threading
  - and semantic analysis, 619
- García, P., 380
- García, P., 380
- Garden-path, **487**
- garden-path sentences, **487**, 490
- Garofolo, J., 852
- Garrett, M. F., 78

- Garside, R., 94, 123, 130, 136, 139, 169, 170  
 Garvey, C., 730  
 Gaston, J., 850  
 Gates, B., 823  
 Gaussian, **308**  
   estimators for acoustic likelihood, **308**  
   mixtures, **313**  
   prior on weights, 205  
 Gaussian Mixture Model, **308**, **313**  
 Gaussier, E., 384  
 Gaustad, T., 657  
 Gazdar, G., 16, 408, 527, 532, 542, 546, 549, 619  
 gazetteer, **742**  
 Gdaniec, C., 483, 484  
 Ge, N., 733  
 Gee, J. P., 264  
 Geman, S., 491, 532  
 GEN in Optimality Theory, 372  
 gender  
   marking on words, **52**  
 Gender agreement, **404**  
 gender agreement, **404**  
 gender dependent acoustic modeling, **356**  
 Gene normalization, **771**  
 Generalized semantic role, **635**  
 generate, 393  
 generation  
   template-based, 840  
 Generative grammar, **395**  
 generative grammar, **395**, 426  
   relation to formal language, **31**  
 Generative Lexicon, **645**  
 generative power, **536**, 548  
 generative syntax, 428  
 generator, 392  
 Generic summary, **802**  
 generics, 713  
 genitive NP, **429**  
 Gentner, D., 646  
 genus  
   in definition questions, **817**  
 Genzel, D., 118, 119  
 Georgila, K., 871  
 Gerber, L., 756, 759  
 Gerbino, E., 845, 851  
 Gerdemann, D., 387, 532  
 German, 241, 877  
 Germann, U., 920  
 Gershman, A. V., 620, 919  
 Gerund, **50**  
 gerund, **50**, 401  
 gerundive, 426  
 gerundive postmodifier, **401**  
 Gestural score, **244**  
 gestural score, **244**  
 Gesture, **244**  
 gesture  
   articulatory, **244**  
 Giangola, J. P., 832, 840, 841, 846, 850, 860, 871  
 Gibbon, D., 285, 871  
 Gibson, E., 359, 546–548, 734  
 Gil, D., 125  
 Gilbert and Sullivan, 72, 287  
 Gilbert, G. N., 850  
 Gildea, D., 13, 112, 229, 358, 380, 381, 491, 585, 683, 684, 690, 914  
 Gillick, L., 332  
 Gilliom, L. A., 731  
 Gillis, S., 170  
 Girand, C., 229, 358  
 Girju, R., 682, 690  
 Gish, H., 364  
 Givenness hierarchy, **712**  
 Givón, T., 472  
 GIZA++, **920**  
 Glance, N., 775  
 Glass, J., 245, 358, 364, 851  
 Glennie, A., 461  
 Gloss, **630**  
 Glottal, **221**  
 glottal, **221**  
 Glottis, **218**  
 GMM, **308**  
   covariance matrix, 303  
 GMM (Gaussian Mixture Model), **313**  
 goal-directed search, 433  
 Godfrey, J., 85, 225, 334, 335, 390, 825  
 Godzilla, speaker as, 639  
 Gold Standard test set, **155**  
 Gold, B., 247, 335  
 Goldberg, J., 858  
 Goldfinger, S. D., 363  
 Golding, A. R., 171  
 Goldman, R., 726  
 Goldsmith, J., 370, 383  
 Goldstein, J., 810, 811, 813, 823  
 Goldstein, L., 244  
 Goldwater, S., 170, 377–379, 387  
 The Gondoliers, 72  
 Gonzalo, J., 775  
 Good, I. J., 101  
 Good, M. D., 850  
 Good-Turing smoothing, **101**  
 Goodenough, J. B., 669  
 Goodine, D., 851  
 Goodman, J., 89, 112, 120, 213, 491, 532  
 Goodman, K., 920  
 Goodwin, C., 856, 871  
 Gopalakrishnan, P. S., 347  
 Gordon, D., 870  
 Gordon, P. C., 731  
 Gordon, S. A., 871  
 Gorin, A. L., 826  
 Götz, T., 532  
 Gould, J. D., 850  
 Gould, S. J., 649  
 Government and Binding, 427  
 Grabe, E., 285  
 gradient ascent, 201  
 Gradual Learning Algorithm, **385**  
 Graff, D., 334  
 Grammar  
   Constraint, 419  
   Construction, 427  
   Dependency, 418  
   Government and Binding, 427  
   Head-Driven Phrase Structure (HPSG), 413, 427  
   Lexical-Functional (LFG), 427  
   Link, 419  
   Tree Adjoining, 427  
   probabilistic, 491  
 grammar  
   binary branching, **416**  
   categorical, 420, **420**  
   CCG, 421, **538**  
   checking, 431  
   combinatory categorical, 421, **538**  
   equivalence, 415  
   generative, **395**  
   Head, **538**  
   hierarchy, 536  
   inversion transduction, 912  
   Minimalist, **538**  
   regular, 537  
   strong equivalence, 416  
   systemic, **407**  
   TAG, **538**  
   Tree-Adjoining, 538  
   weak equivalence, 416  
 Grammar Rock, 123  
 grammars  
   mildly context-sensitive, **538**  
 grammatical relations, **390**  
 grammatical sentences, **395**  
 Granacki, J., 620  
 Granoien, N., 359  
 Grapheme-to-phoneme, **259**  
 Graphone, **284**  
 Gravano, L., 753, 799  
 Gray, R. M., 305  
 greedy RE patterns, **22**  
 Greek, 215  
 Green, B. F., 586, 823  
 Green, J., 10, 427  
 Greenbaum, S., 49, 127, 428

- Greenberg, S., 225, 243  
 Greene, B. B., 137, 169  
 greeting, **129**  
 Grefenstette, G., 70, 71, 81, 677, 679, 690, 823  
 Gregory, M. L., 119, 229, 266, 358, 472  
 Greiff, W., 688  
 Grenager, T., 691  
 grep, 17, **18**, 42  
 Grice, H. P., 834, 866  
 Grice, M., 282  
 Gricean maxims, 834  
 Gries, S. T., 646  
 Griffiths, T. L., 170, 387  
 Grishman, R., 359, 364, 405, 483, 484, 508, 510, 766  
 Grober, E., 730  
 Groliers, 680  
 Grosjean, F., 264, 362  
 Grosjean, L., 264  
 Grosz, B. J., 12, 16, 695, 718, 731–733, 868–871  
 grounding  
   five kinds of, 831  
 grounding contributions, **830**  
 Grover, C., 743, 809  
 Gruber, J. S., 633, 645  
 Grudin, J. T., 166  
 Gu, L., 357  
 Gulikers, L., 81, 126, 241  
 Gundel, J. K., 712, 733  
 Gupta, V., 120, 347  
 Gusfield, D., 77, 81  
 Gustafson, K., 284  
 Guthrie, J. A., 689  
 Guthrie, L. M., 688, 689  
 Guy, G. R., 229  
 Guyon, I. M., 13  
  
 h\*, 346  
 H\* pitch accent, 268  
 Habash, N., 81  
 Hachey, B., 809  
 Hacioglu, K., 690  
 Haeb-Umbach, R., 347  
 Hafer, M. A., 382  
 Haghighi, A., 492  
 Hahn, U., 387, 733, 823  
 Häikiö, T., 79  
 Hain, T., 316, 343, 358, 364  
 Hajič, J., 13, 161–163, 419, 484, 491, 492  
 Hakkani-Tür, D., 360, 364  
 Hakkani-Tür, D., 161, 162, 359, 364  
 Hakwing, S., 249  
 HAL 9000, 1  
 Hale, J., 487, 733  
 Hall, J., 491  
 Hall, K., 486  
 Hall, R., 78  
 Halle, M., 63, 228, 285, 369, 388, 536  
 Halliday, M. A. K., 407, 697, 733, 734  
 Hamada, H., 283  
 Hamilton, A., 121  
 Hamming, **299**  
 Hammond, M., 374  
 Hamza, W., 285  
 Hanauer, S., 333  
 Hankamer, J., 80  
 Hanks, P., 13, 674, 762  
 Hansards corpus, **899**  
 Hansen, B., 850  
 hapax legomena, 101, **158**  
 hapax legomenon, 101, **158**  
 Hapsburg, M. T., 249  
 Harabagiu, S., 690, 795, 823  
 hard decisions  
   avoiding them, 197  
 Harman, D. K., 790  
 harmonic, **241**  
 Harmonic Grammar, **387**  
   and log-linear models, **387**  
 harmonic mean, **459, 483**  
 harmony, **369, 387**  
   in Harmonic Grammar, **387**  
 Harper, M. P., 150, 159, 364  
 Harris, C. M., 283  
 Harris, R. A., 828, 850, 871  
 Harris, Z. S., 10, 137, 156, 169, 382, 387, 462, 672, 689  
 Hart, P. E., 213, 306, 687  
 Hart, T., 113, 120  
 Harvey, T., 151  
 Hasan, R., 697, 733, 734  
 Hasegawa, M., 700, 734  
 Hasegawa-Johnson, M., 364  
 Hastie, T., 213  
 Haviland, S. E., 713  
 Hayes, B., 375, 385, 387  
 Hayes, E., 646  
 Hays, D., 418  
 Hazan, V., 282  
 Hazen, T., 351  
 head, 413, 477  
   tag, **477**  
 Head features, **505**  
 Head Grammar, **538**  
 head nouns, 426  
 Head tag, **477**  
 Head-Driven Phrase Structure Grammar (HPSG), 413, 427  
 head-marking languages, **877**  
 Hearst, M. A., 112, 488, 657, 662, 680, 688, 697, 699, 700, 733, 771, 772  
 Hebrew, 215  
 Hecker, M. H. L., 282  
 HECTOR corpus, 651  
 Hedberg, N., 712, 733  
 Heeman, P. A., 120, 364  
 Hehmsoth, C., 282  
 Heikkilä, J., 137, 138  
 Heikkilä, J., 135, 139, 418, 419, 428  
 Heim, I., 566, 587  
 Heinz, J. M., 247  
 Held-out, **104**  
 held-out data, **92**  
 Helmreich, S., 646  
 Hemphill, C. T., 335, 390, 825  
 Henderson, J., 491, 871  
 Hendler, J. A., 584  
 Hendrickson, C., 156  
 Henrix, J., 225  
 Hepple, M., 762, 775  
 Herbst, E., 920  
 Hermansky, H., 361  
 Hermjakob, U., 795, 799  
 Hernon, W. P., 78  
 Herold, K., 358  
 Hertz as unit of measure, **231**  
 Herzog, M., 359  
 Hickey, M., 84  
 hidden events  
   in HMM, 144  
 Hidden Markov Model, *see* HMM, **177**  
 Hidden Understanding Model, **839**  
 Hieronymus, J. L., 871  
 Higgins, H., 247  
 Higuchi, N., 285  
 Hilf, F. D., 870  
 Hill, D. P., 772  
 Hill, E. A., 387  
 Hillard, D., 359, 364  
 Hindi, 877, 878  
 Hindle, D., 423, 483, 484, 490, 672, 674, 675, 689  
 Hinkelman, E. A., 620, 854, 855, 871  
 Hintikka, J., 587, 865  
 Hirschberg, J., 264, 265, 267, 268, 285, 364, 705, 734, 810, 854, 857–860, 869  
 Hirschman, L., 724, 771, 774, 850–852  
 Hirst, G., 171, 660, 665, 669, 688, 689, 733  
 Hitzeman, J., 733  
 Hjelmslev, L., 646  
 Hladká, B., 162  
 HMM, **144, 177, 289**  
   accepting states, 178  
   alignment, for MT, 897



- cloning, **350**  
 deleted interpolation, **150**  
 ergodic, **178**  
 for alignment in  
   summarization, 809  
 for speech synthesis, 285  
 formal definition of, 144, 177, 293  
 fully-connected, **178**  
 history in speech recognition, 334  
 illustrative example of, 294  
 initial distribution, 178  
 multiple observations for a  
   single hidden state, 839  
 observation likelihood, 145, 177, 293, 296, 318, 327  
 observations, 145, 177, 293  
 part-of-speech tagger, **135**  
 phone models, 296  
 semi-HMM, **838, 839**  
 states, 145, 177, 293, 296, 318, 327  
 tied states, **348**  
 transition probabilities, 145, 177, 293, 296, 318, 327  
 trigram POS tagging, 149  
 HMM part-of-speech tagging  
   simplifying assumptions, 140  
 HMM state, **296**  
 HMM synthesis, **285**  
 HMM tagger, **135**  
 HMM-MLP hybrid, **354**  
 Hoang, H., 920  
 Hobbs algorithm, **716**  
 Hobbs tree search algorithm for  
   pronoun resolution, 716, 717  
 Hobbs, J. R., 12, 619, 646, 701, 716, 726, 733–735, 766, 767, 774  
 Hockenmaier, J., 491  
 Hockey, B. A., 836, 839, 871  
 Hofmann, T., 112  
 Hofstadter, D. R., 173  
 Hollenback, J., 225, 243  
 Holliman, E., 85, 225, 334  
 Holmes, D. I., 122  
 Holonym, **629**  
 Homograph, **626**  
 homograph, **256**  
 homonym, **624**  
 Homonymy, **624**  
 homonymy  
   and IR, 790  
 Homophone, **626**  
 Hon, H.-W., 247, 285, 306, 326, 335  
 Honal, M., 364  
 Hong Kong Hansards corpus, **899**  
 Hopcroft, J. E., 38, 39, 42, 81, 395, 538, 539, 541, 549  
 Hopely, P., 10, 169  
 Hopkins, M., 914  
 Hori, C., 824  
 Horning, J. J., 492  
 hot languages, **878**  
 hourglass metaphor for TTS, **250**  
 House, A. S., 247, 282  
 Householder, F. W., 170  
 Hovanyecz, T., 850  
 Hovy, E., 795, 805  
 Hovy, E. H., 688, 734, 799, 809, 819  
 Howes, D., 362  
 Huang, C., 357  
 Huang, E.-F., 339  
 Huang, L., 491  
 Huang, X., 247, 285, 289, 306, 326, 335  
 Huang, Z., 359, 364  
 Huddleston, R., 396, 428  
 Hudson, R. A., 428, 532  
 Huffman, D. A., 80  
 Huffman, S., 775  
 Hukari, T. E., 527  
 Huls, C., 733  
 Hulteen, E. A., 832, 850  
 HUM, **839**  
 human  
   sentence processing, 487  
 human parsing, 487  
 Human sentence processing, **487**  
 human sentence processing, 487  
 Humphreys, K., 766  
 Hundreds digits, **256**  
 Hungarian, 162, 877  
   part-of-speech tagging, 160  
 Hunnicke-Smith, K., 357  
 Hunnicut, M. S., 260, 284, 285  
 Hunt, A. J., 271, 279–281, 283  
 Hunter, J., 734  
 Hunter, L., 769–773  
 Hutchins, J., 919, 921  
 Hutchins, W. J., 880, 881, 883, 887, 919, 921  
 Huybregts, R., 544  
 Hyperarticulation, **858**  
 Hypernym, **628**  
 hypernym  
   and information content, 667  
   and WordNet supersenses, 682  
   as lexical cohesion, 697  
   bootstrapping, 681  
   extending WordNet, 681  
   in Extended Lesk, 669  
   induction of relation, 680  
   lexico-syntactic patterns for, 680  
 hyperplane  
   logistic regression as, 200  
 Hypoarticulation, **226**  
 hypoarticulation, **226**  
 Hyponym, **628**  
 hyponym  
   and WordNet supersenses, 682  
   induction of relation, 680  
 hypnym  
   extending WordNet, 681  
 Hz as unit of measure, **231**  
 IBM, 119  
 IBM clustering, **111**  
 IBM Model 1, 895  
 IBM Model 3, 914  
 IBM Thomas J. Watson  
   Research Center, 11, 12, 119, 334  
 ICASSP, 16  
 ice cream, 177  
 ichi Tsujii, J., 491  
 ICSLP, 16  
 Ide, N. M., 162, 652, 689  
 IDF, **659**, 785, **785**, 822  
 IDF term weighting, **785**  
 idioms, 553, 617  
   and compositional semantics, 617  
 if then reasoning  
   in FOL, 567  
 Iida, M., 718, 733  
 III, J. T. M., 484, 491, 532  
 Ilhan, H. T., 899  
 illocutionary act, 829  
 Illocutionary force, **829**  
 illocutionary force, **829**  
 Imai, S., 285  
 immediately dominates, **393**  
 imperative, 396, 426  
 imperative sentence structure, **396**  
 imperatives  
   semantic attachments, 606  
 Implicature, **834**  
 implicature, **834**  
 Implicit confirmation, **845**  
 in vitro, **656**  
 in vitro evaluation, **656**  
 in vivo, **656**  
 in vivo evaluation, **95, 656**  
 indefinite article, 399  
 indefinite reference, 710  
 Indicator function, **202**  
 indirect object, 425  
 indirect speech act, **856**  
 Inference, **199, 557**  
 inference

- bridging, **712**
- in FOL, 567
- inferable referent and bridging inferences, **712**
- infinite employment of finite means, 30
- infinitives, 390, 405, 426
- infix, 47
- Inflection, **48**
- inflectional, 79
- INFORM, 866
- information content word similarity, 666
- Information extraction, **737**
- information extraction, 455, 461
- Information retrieval, **781**
- information retrieval, 79
- Information status, **712**
- Information structure, **712**
- information theory, 113
- information-state architecture, **853**
- informativeness
  - in MT, 909
- Ingria, R., 483, 484, 756, 759, 762, 837, 839
- Initiative, **842**
- initiative, 871
  - mixed, 843
  - single, 842
  - system, 842
- inner ear, 239
- inner product, **676**
- input form, 45
- Inside-outside, **471**
- inside-outside algorithm, 490
- Instance checking, **583**
- intelligibility
  - for TTS evaluation, **282**
- intensity of sound, **234**
- Intentional structure, **868**
- intentional structure, **868**
- interjection, **129**
- Interlingua, **887**
- interlingua, **887**
- intermediate phrase, **263**
  - in ToBI, 268
- intermediate semantic representations, 555
- intermediate tape, 62
- internal rule
  - in a CFG parse, **478**
- International Phonetic Alphabet, **216**, 245
- International Space Station, 8, 871
- Interpolated Kneser-Ney, **110**
- Interpolated Kneser-Ney discounting, **109**
- Interpolated precision, **788**
- interpolation in N-gram smoothing, **104**
- Interpretation, **559**
- Interruption point, **361**
- interruption point, **423**
- Intersection, **903**
- intersection
  - transducer, 79
- Intersective semantics, **610**
- Intonation phrase, **263**
- Intonation unit, **285**
- intransitive, 426
- intransitive verbs, **405**
- Intrinsic evaluation, **95**
- Inverse document frequency, **785**
- inverse document frequency, **659**, *see* IDF
- inversion
  - of transducer, **58**
- Inversion transduction grammar, **912**
- IOB tagging, **457**
- Iolanthe, 72
- IPA, **216**, 245
- IR, *see* information retrieval and WSD, 790
  - feature vector, 782
- IDF term weighting, 785
- probabilistic model, **822**
- term, 782
- term frequency, 785
- term weighting, 782, 785, 822
- vector space model, 782
- Irons, E. T., 461
- irregular verb, **49**
- Isard, A., 854, 856, 871
- Isard, S., 275, 854, 856, 871
- Ishizaka, K., 284
- Ishizaki, M., 854
- ISO 8601, 758
- Isolated word, **288**
- isolated-word recognition, **288**
- isolating language, **877**
- Israel, D., 766–768, 774
- Issar, S., 837
- Issel-Tarver, L., 772
- Itakura, F., 333
- iterative scaling, 201
- ITG, *see* inversion transduction grammar
- Iverson, E., 646
- Iwahashi, N., 283
- Iyer, R. M., 111
- Jaccard similarity metric, **677**
- Jaccard, P., 677
- Jackendoff, R., 427, 529, 575, 646
- Jackson, P., 746
- Jacobs, P., 532
- Jacobs, P. S., 774, 823
- Jacobson, N., 156
- Jacquemin, C., 384
- Jahr, M., 920
- Jakimik, J., 363
- Jakobson, R., 531
- Janssen, T. M. V., 619
- Japanese, 215, 241, 877, 878, 880
- Jarosz, G., 387
- Järvinen, T., 418, 419
- Jasperson, R., 423
- Jay, J., 121
- Jefferson, G., 828, 832, 856, 870
- Jeffreys, H., 98, 100, 119
- Jekat, S., 855
- Jelinek, F., 13, 104, 119, 120, 150, 170, 213, 334, 335, 339, 345, 346, 364, 468, 483, 484, 487, 490, 904, 920
- Jenson-Shannon divergence, **678**
- Ji, H., 359, 364
- Jiang, J. J., 668
- Jiang-Conrath distance, **668**
- Jianping, Z., 358
- Jilka, M., 271
- Jiménez, V. M., 491
- Jing, H., 809, 810, 823
- Jitter, **361**
- Johansson, S., 428
- Johnson, C., 508–510
- Johnson, C. D., 80, 365, 386, 539
- Johnson, C. R., 637
- Johnson, H. L., 771
- Johnson, K., 247
- Johnson, M., 364, 377–380, 387, 428, 475, 484, 486, 491, 530, 532, 550, 575, 644, 646, 682, 688
- Johnson, S. C., 619
- Johnson, W. E., 98, 100, 119
- Johnson-Laird, P. N., 629
- Johnston, M., 287, 871
- Join cost, **279**
- joint intentions, 871
- joint probability, 466
- Jones, D. A., 359
- Jones, M. A., 646
- Jones, M. P., 171
- Jones, R., 753
- Jones, S. J., 850
- Joos, M., 689
- Jordan, M. I., 13
- Jordan, P. W., 827
- Joshi, A. K., 10, 13, 169, 427, 490, 491, 538, 695, 718, 733, 734
- Joslyn, C., 773

- Joyce, J., 84  
 Juang, B. H., 11, 213, 335, 351, 919  
 jump width  
   in MT, **898**  
 Jun, S.-A., 268  
 Juneja, A., 364  
 Junqua, J. C., 355  
 Juola, P., 550  
 Jurafsky, D., 81, 90, 113, 119, 160, 229, 266, 357, 358, 362, 380–384, 487, 491, 525, 681–684, 688, 690, 857, 858, 871  
 Just, M. A., 547
- K-means clustering, **306**  
 Kaalep, H. J., 162  
 Kager, R., 388  
 Kahn, D., 378  
 Kaiki, N., 283  
 Kajarekar, S., 335  
 Kameyama, M., 733, 766, 767, 774  
 Kamm, C. A., 850–852, 860  
 Kamm, T., 243, 364  
 Kan, M. Y., 733  
 Kanthak, S., 364  
 Kantrowitz, M., 811  
 Kaplan, R. M., 11, 12, 39, 67, 81, 365, 368, 431, 452, 484, 491, 529, 530, 532, 732, 823, 825, 836  
 Karamanis, N., 733, 813, 823  
 Karger, D. R., 687  
 Karlsson, F., 135, 137, 139, 418, 419, 428, 543, 547  
 Karttunen, L., 10, 70, 81, 169, 370, 374, 375, 386, 387, 524  
 Kasami, T., 431, 461  
 Kasarskis, A., 772  
 Kashyap, R. L., 170  
 Kasowski, S., 247, 284  
 Kasper, R. T., 501, 530  
 Kataja, L., 387  
 Katz backoff, **105**  
   details, 106  
 Katz, C., 10, 427  
 Katz, G., 756, 759, 762, 775  
 Katz, J. J., 587, 646, 660  
 Katz, K., 408, 410  
 Katz, S. M., 103, 105  
 Kawahara, T., 700, 734  
 Kawamoto, A. H., 688  
 Kay, M., 11, 12, 39, 67, 81, 365, 368, 371, 387, 431, 452, 461, 496, 524, 527, 530, 531, 825, 836, 920  
 Kay, P., 427, 532, 620
- Kazakov, D., 383  
 Kazemzadeh, A., 285  
 Keane, M. T., 646  
 Kearns, M., 844, 860, 863  
 Keating, P. A., 229  
 Kehler, A., 733, 734, 767  
 Keller, F., 112, 119  
 Kelley, L. G., 919  
 Kelly, E. F., 688, 689  
 Kendall's  $\tau$ , 819  
 Kennedy, C., 733  
 Kernighan, M. D., 164–167, 170  
 Kershaw, D., 316, 343  
 keyword in IR, **68**  
 Khudanpur, S., 357  
 Kibble, R., 709, 823  
 Kiefer, B., 524, 532  
 Kießling, A., 871  
 Kilgariff, A., 13, 651, 656, 658, 659  
 Kim, D., 364  
 Kim, G., 408, 410  
 Kim, J., 364  
 Kimball, O., 348, 364  
 King, J., 547  
 King, P., 532  
 King, R. C., 118  
 King, S., 364, 871  
 King, T. H., 484, 491, 532, 823  
 Kingsbury, P., 13  
 Kinoshita, S., 771  
 Kintsch, W., 11, 587  
 Kipp, M., 871  
 Kipper, K., 635  
 Kiraz, G. A., 81, 378, 379, 387  
 Kirchhoff, K., 228, 858  
 Kirsch, H., 772  
 Kisseberth, C. W., 369, 372  
 Kita, K., 871  
 Kitade, T., 700, 734  
 Kitamura, T., 285  
 Kittay, E., 646  
 KL divergence, **678**  
 KL-ONE, 587  
 Klatt formant synthesizer, 284  
 Klatt, D. H., 260, 270, 284, 285, 334  
 Klavans, J. L., 483, 484, 645, 733  
 Kleene \*, **20**  
   sneakiness of matching zero things, 20  
 Kleene +, **20**  
 Kleene, S. C., 9, 18, 42  
 Klein, A., 855  
 Klein, D., 156, 160, 170, 475–477, 490–492  
 Klein, E., 408, 527, 619  
 Klein, J., 122  
 Klein, S., 137, 169
- Kleinman, D., 730  
 Klesen, M., 871  
 Klimt, B., 250  
 Klovstad, J. W., 347  
 Kneser, R., 110, 113, 120  
 Kneser-Ney discounting, **109**  
 Knight, K., 532, 813, 823, 824, 883, 886, 901, 907, 912, 914, 916–918, 920–922  
 Knill, K., 335  
 Knott, A., 734, 823  
 Knowledge base, **554**  
 knowledge base, 555  
 Knuth, D. E., 82, 170, 619  
 Kobayashi, T., 285  
 Koehn, P., 264, 891, 903–906, 908, 912, 920  
 Koeling, R., 670, 689  
 Koenig, W., 10, 247  
 Koerner, E. F. K., 170  
 Kogan, Y., 773  
 Kogure, K., 524  
 Kokkinakis, G., 158  
 Koller, A., 813, 823  
 Koller, D., 491  
 Kompe, R., 871  
 Koo, T., 491  
 Koopmans-van Beinum, F. J., 118  
 Koppel, M., 775  
 Korean, 241  
 Kornai, A., 387  
 Koskenniemi, K., 61, 81, 365, 367, 387  
 Kostic, A., 487  
 Kothari, A., 266  
 Kowtko, J. C., 854, 856, 871  
 Kozima, H., 733  
 Krallinger, M., 772  
 Krasnwer, M., 348  
 Kratzer, A., 566, 587  
 Krauthammer, M., 773  
 Krbec, P., 163  
 Krieger, H.-U., 524, 529, 532  
 KRL, 587  
 Krogh, A., 213  
 Krovetz, R., 68, 664, 790  
 Kruskal, J. B., 102, 146, 212  
 Krymowski, Y., 462  
 Kryter, K. D., 282  
 Kučera, H., 11, 85, 86, 94, 123, 130, 134, 135, 169  
 Kudo, T., 491  
 Kuhl, P. K., 387  
 Kuhn, J., 532  
 Kuhn, R., 112, 362  
 Kuhn, S., 733  
 Kuhn, T., 871  
 Kukich, K., 72, 84, 167, 170  
 Kullback, S., 661, 678

- Kullback-Leibler divergence, **661**  
 Kumar, S., 254, 255, 343, 920  
 Kuno, S., 461  
 Kupiec, J., 150, 156, 170, 823  
 Květoň, P., 163
- L\* pitch accent, 268  
 L+H\* pitch accent, 268  
 L., 10, 247  
 L1 norm, **675**  
 L2 norm, **676**  
 labeled precision, 483  
 labeled recall, 483  
 labial place of articulation, **220**  
 labiodental consonants, **220**  
 Labov, W., 216, 227, 229, 230, 856  
 Lacy, 10, 247  
 Ladd, D. R., 263, 265, 285  
 Ladefoged, P., 222, 223, 242, 247, 377  
 Ladusaw, W. A., 247  
 Lafferty, J. D., 13, 170, 201, 213, 214, 490, 491, 699, 700, 775, 920  
 Laham, D., 679  
 Lai, J. C., 86, 111, 118, 690, 904, 920  
 Laird, N. M., 104, 167, 187  
 Lakoff, G., 370, 575, 585, 587, 642, 644, 646, 870  
 Lambda notation, **565**  
 Lambek, J., 428  
 Lambrecht, K., 733  
 Lancaster UCREL, 130  
 Landauer, T. K., 679, 690, 850  
 Landes, S., 652  
 Lane, H., 264  
 Langendoen, D. T., 285, 428  
 Langer, S., 84  
 Langlais, P., 659  
 Langsam, Y., 428  
 language  
   formal, 41  
   of certain sheep, 27  
   regular, 38  
 Language identification, **335**  
 language identification, **335**  
 Language model, **292**  
 language model, **83, 292**, 463  
   adaptation, **111**  
   for MT, 890  
   PCFG, 467  
 language model scaling factor, **317**  
 language universal, 877  
 Lapalme, G., 659  
 Lapata, M., 112, 705, 733, 813, 814, 819, 823, 824
- Laplace smoothing, **98**  
 Lappin, S., 733  
 lapse (prosodic), 266  
 Large, N. R., 387  
 Lari, K., 471, 492  
 Larsson, S., 853, 871  
 Laryngograph, **276**  
 larynx, **218**  
 Lascarides, A., 529, 707, 734  
 Last In First Out, 37  
 latent semantic analysis, 384, **690**  
   for N-gram smoothing, **113**  
   for spelling error correction, 171  
 latent semantic indexing, **690**  
   for N-gram smoothing, **113**  
 lateral sound, **222**  
 Lathrop, B., 871  
 lattice, **340**  
   pinched, **343**  
 lattice density, **341**  
 lattice error rate, **341**  
 Lau, R., 378, 379  
 Lauer, M., 620  
 Laughery, K., 586, 823  
 Lavie, A., 921  
 Lawrence, W., 284  
 Lazo, M., 13, 762  
 La Polla, R., 428  
 LCS, **666**  
 LDC, 471  
 LDC., 97, 242, 334  
 Le, A., 852  
 Leacock, C., 651, 652, 665  
 leading sentences baseline in summarization, **821**  
 leap year, 287  
 Lease, M., 364  
 Leass, H., 733  
 Lede, **696**  
 Lee, C.-H., 351, 837, 838, 919  
 Lee, H. B., 688  
 Lee, K.-F., 357  
 Lee, K.-S., 284  
 Lee, L., 364, 678, 689, 690, 813, 824  
 Lee, S., 285, 852  
 Leech, G., 49, 94, 123, 127, 130, 136, 170, 428  
 Lees, R., 620  
 left-linear, **537**  
 left-linear grammar, **537**  
 Legendre, G., 387  
 Leggetter, C. J., 341, 356  
 Lehiste, I., 247  
 Lehnert, W. G., 11, 620, 733, 766, 774, 775, 823  
 Lehrer, A., 646  
 Leibler, R. A., 661, 678
- Lemas, M., 822  
 Lemma, **623**  
 lemma, 85  
   versus wordform, 85  
 Lemmatization, **46, 623**  
 Lemon, O., 871  
 lenient composition, **374**  
 Lennig, M., 120, 347, 348  
 Lenzo, K., 258, 261  
 Leon, E. A., 772  
 Lerner, A. J., 47  
 Lesk  
   Corpus, **659**  
   Extended, **668**  
   Simplified, **658**  
 Lesk algorithm, **658**  
 Lesk, M. E., 619, 659, 688  
 letter-to-sound rules, **260**  
 Leung, H., 851  
 Levelt, W. J. M., 361, 422, 424  
 Levenshtein distance, 74, **675**  
 Levenshtein, V. I., 74  
 Levesque, H. J., 587, 871  
 Levi, J., 620  
 Levin, B., 425, 635, 645  
 Levin, E., 837, 838, 861–863  
 Levin, L., 854  
 Levine, R. D., 527  
 Levinson, S. C., 828  
 Levinson, S. E., 335  
 Levitt, S. D., 194  
 Levow, G.-A., 829, 832, 846, 850, 858  
 Levy, R., 487  
 Lewin, I., 837  
 Lewis, C., 850  
 Lewis, D. L., 774  
 Lewis, H., 38, 39, 42, 81, 538, 539  
 Lewis, S., 772  
 Lexeme, **623**  
 lexical  
   category, 392  
   cohesion, 697  
   dependency, 464, 474  
   gap, **880**  
   head, 490  
   rule, **529**  
   stress, **224**  
   tags, *see* part-of-speech  
   tape, **60**  
 Lexical access, **361**  
 lexical ambiguity resolution, 688  
 lexical database, 629  
 Lexical dependency, **474**  
 Lexical gap, **880**  
 lexical level in finite-state morphology, 60  
 lexical rule  
   in a CFG parse, **478**

- to capture lexical generalizations, **529**
- lexical sample, **651**
- lexical sample task in WSD, **651**
- lexical semantics, **623**
- Lexical stress, **224**
- Lexical tape, **60**
- lexical trigger
  - in IE, **756**
- Lexical-Functional Grammar, 550
- Lexical-Functional Grammar (LFG), 427
- Lexicalized grammar, **477**
- lexico-syntactic pattern, **680**
- lexicon, **391**
  - as list of lexemes, **623**
  - definition of, 54
  - tree-structured, **347**
  - use of in morphological parsing, **53**
- LFG, 11
- Li, A., 243
- Li, C. N., 885
- Li, H., 688
- Li, X., 795, 796
- Liaison, **262**
- Liberman, A. M., 247, 284
- Liberman, M. Y., 257, 259, 266, 271, 283, 483, 484, 775, 870
- Lidstone, G. J., 98, 100
- Lieber, R., 79
- Lieberman, H., 584
- LIFO, 37, 41
- Light, M., 688
- likelihood, **140**
- Lim, D. C. Y., 723
- Lin association measure, **674**
- Lin similarity, **668**
- Lin, C.-Y., 805, 809, 819
- Lin, D., 484, 666–668, 670, 672–674, 679, 689, 690, 718, 733, 824
- Lin, J., 794, 800, 810, 823, 824
- Lin, S., 920
- Lindblom, B. E. F., 118
- Lindemann, A., 771
- Lindsey, R., 586
- linear interpolation for N-grams, 104
- Linear Predictive Coding, *see* LPC
- linear regression, 194, 196
  - multiple, 195
  - setting weights, 196
- Linear segmentation, **696**
- Linguistic Data Consortium, 471
- Linguistic Discourse model, 734
- linguistic knowledge
  - why needed, 2
- linguists, firing, 83
- Link Grammar, **419**
- lipsmack
  - transcription of, 421
- Literal meaning, **553**
- literal meaning, 589
- Litman, D. J., 705, 734, 827, 844, 851, 852, 857–860, 863, 869
- Littman, J., 756, 759
- Littman, M., 679
- Liu, H.-M., 387
- Liu, Q., 920
- Liu, Y., 361, 364, 920
- Liu, Z., 871
- Livescu, K., 245, 358, 364
- LM, 83
- LMSF, **317**
- LOB corpus, 169
- Local ambiguity, **437**
- local ambiguity, **437**
- localization, **876**
- locative adverb, **126**
- Lochbaum, K. E., 869, 870
- locutionary act, 829
- Loebell, H., 424
- Loebner Prize, 7
- Lofting, H., 27
- log
  - compression of speech, 232
- Log likelihood ratio, **805**
- log probabilities, 107
- log-linear classifiers, 193
- Logical connectives, **563**
- Logical vocabulary, **558**
- Logistic function, **199**
- Logistic regression, **198**
- logistic regression, 193
  - conditional maximum likelihood estimation, 200
  - logit function in, 198
  - multinomial, 201
  - setting weights, 200
- Logit function, **198**
- Logprob, **336**
- logprob, 336
- logprob (log probability), **315**
- Loke, H. P., 276
- Lombard effect (Lombard speech), **355**
- Long-distance dependencies, **511**
- long-distance dependencies, 608
- long-distance dependency, 397, **408**
- Longacre, R. E., 734
- look-ahead as solution to non-determinism, 33
- Loper, E., 16
- loudness, **235**
- Lowe, J. B., 508, 509, 629, 637, 646
- lower tape, 61
- Lowerre, B. T., 325, 334
- lowest common subsumer, **666**
- LSA, *see* latent semantic analysis
- LSI, *see* latent semantic indexing
- LTS, *see* letter-to-sound
- Lu, Z., 771, 773
- Lua, K., 113
- Lucassen, J., 260, 284
- Luce, P. A., 363
- Luhn, H. P., 785, 804, 822, 823
- Luk, R. W. P., 284
- Lunar, 586, 619
- Luperfoy, S., 854
- LVCSR, **289**
- Lx (laryngograph), **276**
- Lyons, J., 587, 645
- Lyons, R. G., 247
- Lytel, D., 688
- M-step (maximization step) in EM, **192**
- Macdonald, J., 362
- MacEachern, M., 858
- machine, *see* finite-state automaton
- machine learning, 173
  - applied to document categorization, 824
- machine readable dictionaries in WSD, 688
- machine state as opposed to search state, 33
- Machine translation, **1**, **873**
- machine translation, *see* MT
- MacIntyre, R., 408, 410
- Macleod, C., 405, 508, 510
- Macon, M., 282
- MacWhinney, B., 547
- Madhu, S., 688
- Madison, J., 121
- Magerman, D. M., 170, 414, 490
- Mahalanobis distance, **307**
- Maier, E., 854, 855
- Main, M. G., 619
- Maiorano, S., 795
- Makatchev, M., 827
- Makhoul, J., 348, 359, 921
- Makkai, A., 620
- Maleck, I., 855
- Malouf, R., 201, 524, 532
- Mandarin, 241, 873, 877
- Mangu, L., 171, 343
- Manhattan distance, **675**

- Mani, I., 756, 759, 802, 803, 809, 823, 824  
Mann, W. C., 80, 703, 734  
manner adverb, **126**  
manner of articulation, **221**  
Manning, C. D., 16, 81, 156, 160, 170, 213, 471, 475–477, 490–492, 532, 675, 689, 691, 699, 734, 805, 822, 899  
Manurung, H. M., 823  
MAPSSWE, **156**  
Maratsos, M., 547  
Marcel, S., 772  
Marchand, Y., 284  
Marcinkiewicz, M. A., 13, 123, 130, 155, 408, 410, 471  
Marcu, D., 13, 703, 704, 706, 707, 734, 799, 807–809, 813, 823, 824, 891, 914, 920  
Marcus, M. P., 13, 123, 130, 155, 408, 410, 457, 462, 471, 483, 484, 490, 688  
Marcus, S., 674  
marker passing  
    for word sense  
        disambiguation, 688  
Markov, **88**  
    assumption, **88**  
    chain, 119, **144**  
    model, **88**  
    history, 119  
Markov assumption, 175  
Markov chain, **174**  
    accepting states, 176  
    first order, 175  
    formal definition of, 175  
    initial distribution, 176  
    N-gram as, 175  
    start and end states, 175  
    states, 175  
    transition probabilities, 175  
Markov decision process, **860**  
Markov model  
    formal definition of, 175  
Markov, A. A., 119, 173, 212  
Markovitch, S., 674, 690  
Marshall, C., 871  
Marshall, I., 139, 169  
Marslen-Wilson, W., 78, 362, 363  
Martin, D., 767  
Martin, J. H., 171, 620, 645, 646, 690, 775  
Martin, P., 726  
Martin, R., 857, 858, 871  
Martina, R., 871  
Marx, G., 436  
Marx, M., 829, 832, 846, 850  
Marzal, A., 491  
mass noun, **125**  
Massaro, D. W., 362  
Mast, M., 855, 871  
Masterman, M., 586, 688  
Masuko, T., 285  
Matiaszek, J., 387  
Matsumoto, Y., 491  
Matsuzaki, T., 491  
Matthews, A., 731  
Matui, E., 284  
max-margin methods, 214  
MaxEnt, 193, 201, **201**  
    Gaussian priors, 205  
    learning in, 204  
    normalization factor Z in, 201  
    regularization, 205  
    why called maximum entropy, 205  
Maxey, H., 283  
Maxim, **834**  
    of Manner, 834  
    of Quality, 834  
    of Quantity, 834  
    of Relevance, 834  
maxim, **834**  
Maximal Marginal Relevance, **811**  
Maximization step, **471**  
maximization step in EM, 192  
Maximum Entropy Markov Model, *see* MEMM  
Maximum Entropy modeling, 193  
maximum entropy modeling  
    intuition, 205  
maximum likelihood estimation, *see* MLE  
Maximum Likelihood Linear Regression, **356**  
maximum margin methods, 214  
Maximum matching, **70**  
maximum matching, **70**  
Maximum Mutual Information Estimation, 353  
Maximum Onset, **377**  
maxmatch, 70  
Maxwell, J. T., 484, 491, 532  
Mayan, 878  
Maybury, M. T., 802, 824  
Mayfield, J., 775  
Mays, E., 73, 168  
McAllaster, D., 364  
McAllester, D. A., 550  
McCallum, A., 159, 213, 214, 738, 765, 775  
McCarthy, D., 670, 688, 689  
McCarthy, J., 10, 427, 766  
McCarthy, J. F., 733  
McCarthy, J. J., 370  
McCawley, J. D., 47, 428, 587, 646, 647  
McClelland, J. L., 363  
McConnel, S. R., 80  
McConnell-Ginet, S., 587  
McCoy, K., 646  
McCulloch, W. S., 9, 42  
McCulloch-Pitts neuron, 42  
McDaniel, J., 85, 225, 334  
McDermott, D., 726  
McDermott, E., 351  
McDonald, D. B., 620  
McDonald, R., 491, 492, 773, 823  
McEnery, A., 123, 130, 136, 170  
McEnery, T., 733  
McGee, D., 287  
McGill, M. J., 822  
McGuiness, D. L., 584  
McGurk effect, **362**  
McGurk, H., 362  
McKeown, K. R., 265, 689, 733, 810, 812, 815, 818, 820, 823, 824  
McLuhan, M., 878  
McNamee, P., 775  
McNemar test, **156**, **331**  
McRoy, S., 688  
McTear, M. F., 871  
MDL, **382**  
MDP, **861**  
Mealy machine  
    and Moore machine, 80  
Mealy, G. H., 80  
Mean, **308**  
Mean average precision, **789**  
mean opinion score, 283  
Mean reciprocal rank, **801**  
meaning  
    as action, 585  
Meaning as action, **585**  
Meaning postulates, **571**  
Meaning representation  
    languages, **551**  
meaning representation  
    languages, 553  
Meaning representations, **551**  
meaning representations  
    as sets of symbols, 552  
    early uses, 586  
Mechanical Turk, 249  
mediated referent, **712**  
Medvedeva, O., 771  
Megyesi, B., 160  
Mel, **301**  
mel frequency cepstral  
    coefficients, **297**  
mel scale, **235**  
Melamed, I. D., 913, 914, 921  
Mel'čuk, I. A., 428



- Mellish, C., 16, 823  
 Melz, E., 799  
 MEMM, 208  
   as discriminative model, 209  
   compared to HMM, 209  
   for unknown words in POS  
     tagging, 159  
   inference (decoding), 210  
   learning, 211  
   Viterbi decoding, 210  
 memory limitations, 549  
 Menezes, A., 914  
 Meng, H., 378, 379  
 Meng, Y., 871  
 Mensah, A. G., 771  
 Mercer, R. L., 13, 86, 104, 118,  
   119, 139, 150, 168, 170,  
   260, 284, 334, 339, 354,  
   490, 894, 895, 904,  
   915–918, 920  
 Merchant, N., 387  
 Merialdo, B., 150, 170  
 Merlo, P., 691  
 Mermelstein, P., 120, 302, 347,  
   348  
 Meronym, **629**  
 meronym  
   difficulty of extracting, 682  
 Meronymy, **629**  
 MERT  
   for training in MT, **919**  
 Mertins, I., 285  
 Merton, R. K., 14, 617  
 MeSH (Medical Subject  
   Headings), **650**  
 mesh, in ASR, **343**  
 Message Understanding  
   Conference, **766**  
 Metadata, **360**  
 metamorphosis grammar, 11  
 Metaphor, **643**  
 metaphor, 553  
   and analogical reasoning, 646  
   computational approaches,  
     646  
 metarule, **408**  
 Meter, M., 150, 159, 170, 857,  
   858, 871  
 METEOR, 910, 921  
 Metonymy, **625**  
 metonymy  
   in named entity recognition,  
     741  
 Meurers, W. D., 529, 532  
 Meyers, A., 405, 508, 510  
 MFCC, **297**  
 Micciulla, L., 921  
 Michaelis, L. A., 472  
 Michelizzi, J., 669, 691  
 Micro-Planner, 586  
 microgrammar, **856**  
 Microsoft  
   WAV format, 233  
 Microsoft Word, 171  
   regular expressions in, 17  
 Mihalcea, R., 688, 689  
 Mikheev, A., 81, 743  
 mildly context-sensitive  
   grammar, 537, **538**  
 Miller, C. A., 224  
 Miller, D., 334  
 Miller, G. A., 92, 119, 247, 538,  
   546, 547, 652, 669, 908,  
   910, 920  
 Miller, J. L., 362  
 Miller, R., 112  
 Miller, S., 482, 775, 837, 839  
 Miłtasakaki, E., 13, 734  
 Mimura, K., 283  
 min/max similarity metric, **677**  
 Minimalist Grammar, **538**  
 minimum description length, **382**  
 Minimum edit distance, **74**  
 minimum edit distance, 74, **74**,  
   184, 330  
   and Viterbi, 146  
   example of, 76  
   in summarization, 809  
 Minimum Error Rate Training,  
   **919**  
 minimum redundancy  
   hypothesis, **78**  
 MINIMUM EDIT DISTANCE, 76  
 Minnen, G., 529  
 Minsky, M., 10  
 Mishne, G., 775  
 Mishra, R., 871  
 MITalk, 285  
 Mitamura, T., 920  
 Mitchison, G., 213  
 Mitkov, R., 733  
 Mittal, V., 811  
 mixed initiative, **843**  
 Miyao, Y., 491, 532  
 MLE  
   alternatives to, **351**  
   for N-grams, **88**  
   for N-grams, intuition, **89**  
 MLLR, **356**  
 MMIE, **353**  
 MMR, **811**  
 Möbius, B., 274  
 Möbius, B., 378, 379  
 modal, **128**  
 modal logic, 587  
 modal operators, 587  
 modal verb, **128**, **406**  
 Model, **558**  
 model  
   to mean phone HMM, **296**  
 Modified N-gram precision, **910**  
 Modified Rhyme Test, **282**  
 modus ponens, 567, 725  
 Moens, M., 675, 679, 743, 807,  
   810  
 Mohammad, S., 689  
 Mohler, G., 271  
 Mohri, M., 57, 59, 60, 67, 81,  
   428, 542  
 Moldovan, D., 682, 688–690  
 Möller, S., 850, 852  
 Mondschein, L. F., 347  
 monologue  
   as type of discourse, **693**  
 Monotonic, **501**  
 Montague semantics, 587  
 Montague, R., 587, 619  
 Montero, J. M., 842, 860  
 Monz, C., 797  
 Mooney, R. J., 750  
 Moore machines and Mealy  
   machines, 80  
 Moore, E. F., 80  
 Moore, G., 316, 343  
 Moore, J., 698, 733  
 Moore, J. D., 854  
 Moore, R., 285  
 Moore, R. C., 168, 170, 587,  
   805, 920  
 Moran, C., 920  
 Morgan, A., 771  
 Morgan, A. A., 771  
 Morgan, N., 90, 229, 247, 335,  
   351, 354, 358, 487  
 Morimoto, T., 871  
 morpheme, **47**  
 Morpheme boundary, **61**  
 morphemes, 79  
   examples of in Turkish, 46  
   morphological classes, 123  
 Morphological family size, **79**  
 Morphological parsing, **45**  
 morphological parsing, 79  
   argument from English  
     productive affixes for  
       need for, 46  
   evidence from Turkish for  
     necessity of, 46  
   goals of, 53  
   requirements for, 53  
 morphological recognition, 56  
 morphological rules, 45  
 morphology, 79  
   agglutinative, 48  
   CELEX dictionary of, 81  
   cliticization, 48, 51  
   compounding, 48  
   concatenative, 52  
   derivation, 48  
   derivational, 50

- inflection, 48
- non-concatenative, 370
- root-and-pattern, 52
- templatic, 52, 370
- tier, **370**
- use in part-of-speech tagging
  - of unknown words, 158
- morphotactics, **53**, 79
- Morris, J., 733
- Morris, W., 626
- MOS (mean opinion score), **283**
- Moscato del Prado Martín, F., 79, 487
- MOSES, **920**
- Moshier, M. A., 501, 530
- most frequent sense, **657**
- Mosteller, F., 11, 121, 139, 290
- Moulinier, I., 746
- move, conversational, **854**
- MRR, **801**
- MRT, **282**
- MT, **873**
  - and dates, 879
  - and the web, 875
  - decoding
  - recombining hypotheses, 908
  - distortion probability, 891
  - divergences, **876**
  - faithfulness, 889
  - fluency, 889
  - interlingua, **887**
  - problems with, 888
  - jump width, 898
  - lexical differences and, 880
  - phrase-based, 891
  - post-editing, 876
  - sublanguage in, 876
  - transfer model of, **884**
  - transformation examples, 886
  - translation probability, 891
  - useful domains for, 875
- MUC, **766**, 774
- MUC evaluation for coreference, 724
- Müller, K., 378, 379
- Multi-document summarization, **810**
- Multinomial logistic regression, **201**
- multinomial logistic regression, 193
- multiple discovery, 13, 81
- Multiple document
  - summarization, **802**
- multiple linear regression, 195
- Munoz, M., 462
- Munteanu, C., 524, 532
- Murata, T., 585
- Murvet, H., 339, 340
- mutual information, **674**
- Myers, K., 767
- n
  - parameter in IBM Model 3, **914**
- N-best, **339**
- N-best list, **485**
- N-gram, **88**
  - absolute discounting, **109**
  - adaptation, **111**
  - add-one smoothing, **98**
  - applications of, 84
  - ARPA format, 108
  - as approximation, 88
  - as generators, 93
  - as Markov chain, 175
  - backoff, **105**
  - details, 106
  - cache language model, **112**
  - Cambridge-CMU toolkit, **109**
  - class-based, 111
  - equation for, 88
  - example of, 90
  - filtering in QA, **801**
  - for Shakespeare, 93
  - Good-Turing smoothing, **101**
  - history of, 119
  - IBM clustering, **111**
  - interpolation, **103**
  - Katz backoff, **105**
  - Kneser-Ney discounting, **109**
  - logprobs in, 107
  - mining in QA, **801**
  - normalizing, 90
  - of dialogue act sequences, 856
  - parameter estimation, 89
  - sensitivity to corpus, 92
  - skip, **113**, **820**
  - smoothing, 97
  - SRILM toolkit, **109**
  - test set, 91
  - tiling in QA, **801**
  - tokenization before
    - computing, **86**
  - topic-based, **112**
  - training set, 91
  - triggers, **113**
  - unknown words, **94**
  - variable length, **113**
- N-gram tiling, **800**
- Nádas, A., 118, 119, 354
- Nagao, M., 491
- Nagata, M., 871
- Nahamoo, D., 347
- naive Bayes classifier, **653**
  - training, 654
  - use in WSD, 653
- Nakajima, S., 283
- Nakatani, C., 364, 854, 869
- Nakov, P. I., 112, 657
- name
  - Soundex, 82
- Named entity, **739**
- named entity
  - list of types, 740
  - mentions, **738**
- named entity recognizer, 767
- named entity recognition, 737, **739**
- names
  - and gazetteers, 743
  - as unknown words, 258
  - Census lists, 743
  - detecting language of origin, 262
  - in TTS, 259
  - pronunciation in TTS, 259
  - rewriting in summarization, 814
- Nanjo, H., 700, 734
- Narayanan, S., 585, 646
- Narayanan, S. S., 285, 845, 852
- nasal sound, **219**, **221**
- nasal tract, **219**
- Nash-Webber, B. L., 732, 823
- Nass, C., 8
- natural languages contrasted
  - with formal languages, **31**
- Naur, P., 10, 427
- Navigli, R., 688
- Nederhof, M. J., 417, 428
- Needleman, S. B., 213
- Negation, **530**
- negative part-of-speech, **129**
- Neiser, J., 78
- Nenkova, A., 266, 810, 815, 820, 823, 824
- Neo-Davidsonian, **572**
- NER, **737**
- Nerbonne, J., 529
- Nespor, M., 285
- Nesson, R., 914
- nested, 536
- Neu, H., 229
- Newell, A., 84
- newline character, **24**
- Newman, E. B., 235, 301
- Newman, S., 369
- Newport, E. L., 365, 382
- Ney, H., 13, 110, 113, 120, 284, 340, 341, 347, 364, 469, 899, 904, 919, 920
- NFSA, **32**, 41
- Ng, A. Y., 13, 681, 682, 688, 690
- Ng, H. T., 656, 657, 688, 689, 723
- Ng, S.-K., 770
- Ng, V., 733
- Ngampatipatpong, N., 9

- Nguyen, L., 347, 359  
 Nicely, P. E., 247  
 Nichols, J., 877  
 Nicolov, N., 775  
 Nida, E. A., 646  
 Niekrasz, J., 871  
 Nielsen, J., 850  
 Niemann, H., 871  
 Niesler, T. R., 111, 113  
 Nilsson, J., 491  
 Nilsson, N. J., 346, 865, 870  
 Ninomiya, T., 532  
 Nirenburg, S., 920, 921  
 Nissim, M., 713  
 NIST  
   for MT evaluation, 910, 921  
 NIST., 13, 77, 243, 331  
 Nivre, J., 491, 492, 854  
 Niyogi, P., 364  
 Nock, H., 357  
 node as term for FSA state as  
   opposed to search state,  
   33  
 Noisy channel, **289**  
 noisy channel model, 291  
   for MT, 889  
   for speech, **289**  
   for spelling, **163**  
 Noisy-or, **754**  
 Nolan, F., 285  
 Nominal, 392  
 nominalization as example of a  
   morphological process,  
   **50**  
 Nominative, **404**  
 nominative, **404**  
 non-concatenative morphology,  
   52  
 non-deterministic FSA, **32**  
 non-finite postmodifier, **401**  
 Non-logical vocabulary, **558**  
 Non-projective dependencies,  
   **491**  
 non-standard words  
   in TTS, **253**  
 Non-stationary, **298**  
 non-terminal symbols, **392**, 393,  
   425  
 none in FUG, 527  
 Noordman, L. G. M., 734  
 Normal distribution, **308**  
 Normal form, **416**  
 normal form, 415, **416**  
 normalization  
   dates, 837  
   temporal, 758  
 normalization of probabilities,  
   **88**  
 normalized dot product, **677**  
 Norman, D. A., 11, 587, 825,  
   830, 836  
 Normandin, Y., 354  
 Norris, D., 363  
 Norvig, P., 16, 42, 84, 168, 172,  
   462, 562, 587, 646, 656,  
   691, 802, 863, 871  
 Nöth, E., 871  
 noun, **125**  
   abstract, 125, 399  
   common, 125  
   count, 125  
   days of the week coded as,  
   126  
   mass, **125**, 399  
   proper, **125**  
 noun class, **52**  
 Noun group, **418**  
 noun group, **418**  
 Noun phrase, **390**  
 noun phrase, 389, 392, 393, 426  
   genitive, 609  
 noun-noun compounds  
   stress in, 266  
 noun-noun sequences, 608  
 Novick, D. G., 850  
 Nowak, M. A., 119  
 NP, **392**, 393  
 NP attachment, **473**  
 NP-complete, 813  
 NP-completeness of LFG, 550  
 NP-completeness of natural  
   language parsing, 550  
 NP-completeness of two-level  
   morphology, 550  
 Nuclear accent, **265**  
 Nuclear scope, **594**  
 Nucleus, **703**  
 nucleus of syllable, **223**  
 number, 426, 493  
 numeral, 126  
 numerical expression, **740**  
 Nunes, J. H. T., 871  
 Nyberg, E. H., 920  
 Nyquist frequency, **232**, **297**  
 Ndellec, C., 460  
 O'Connor, M. C., 620  
 O'Donnell, M., 823  
 O'Shaughnessy, D., 247  
 Oberlander, J., 823  
 object, syntactic, 390  
   frequency of pronouns as, 472  
 obligatory rules, 367  
 observation likelihood, 148, 182,  
   185, 292, 321, 322  
 observations  
   in HMM, 144  
 Occasion (as coherence  
   relation), **702**  
 Och, F. J., 13, 891, 899, 904,  
   919, 920  
 OCR, **72**  
 Odden, D., 388  
 Odds, **198**  
 Odell, J. J., 341, 348–352  
 Odell, M. K., 82, 170  
 Oden, G. C., 362  
 Oehrle, R. T., 428  
 Oepen, S., 491, 532  
 Oerder, M., 347  
 Oettinger, A. G., 461  
 Oflazer, K., 81, 161, 162  
 Ogburn, W. F., 14  
 Ogren, P. V., 771, 773  
 Ohno, S., 285  
 Okurowski, M. E., 13, 703, 706,  
   734  
 OLA (overlap-and-add), **276**  
 old information, and word order,  
   472  
 Older, L., 78  
 Oliva, K., 163  
 Olive, J. P., 274, 283, 285  
 Ollason, D., 316, 343  
 Olshen, R. A., 213  
 Omura, H., 284  
 on-line handwriting recognition,  
   72  
 on-line nature of human speech  
   processing, **362**  
 on-line sentence-processing  
   experiments, 490  
 Oncina, J., 380  
 One Sense per Collocation, **662**  
 one sense per collocation  
   constraint, **662**  
 One Sense Per Discourse, **663**  
 onset  
   syllable, **223**  
 ontology, **579**, **629**  
 OntoNotes, **688**  
 Oommen, B. J., 170  
 OOV (out of vocabulary) words,  
   **95**  
 OOV rate, **95**  
 open class, **124**  
 Open prompt, **844**  
 Open quotient, **361**  
 open vocabulary system  
   unknown words in, **95**  
 operation list, 74  
 operator precedence, 22, **22**  
 Oppenheim, A. V., 333  
 optical character recognition, 72  
 Optimal coupling, **275**  
 Optimality Theory, **371**, 373  
   implementation via FST, 374,  
   375  
 optional rules, 367

- optionality
  - of determiners, 399
  - use of () in syntax, 400
  - use of ? in regular expressions for, 19
- Oracle accuracy, **486**
- oracle error rate, **341**
- oral tract, **219**
- Oravec, C., 161, 162
- ordinal number, **399**
- Orloff, J., 364
- orthographic rules, 45, 53, 63
- orthography
  - opaque, 217
  - transparent, 217
- Ortmanns, S., 341
- Ortony, A., 646
- Osborne, M., 912, 920
- Osgood, C. E., 689
- Ostendorf, M., 111, 112, 254, 255, 264, 267, 282, 359, 364, 858
- OSTIA, 380
- OT, **371**
- other, 64
- Ott, K., 871
- Overlap-and-add, **276**
- Oviatt, S. L., 287, 825, 850, 858
  
- p-subsequential transducer, 60, **60**
- p1
  - parameter in IBM Model 3, **914**
- Packard, D. W., 80
- Pagel, V., 258, 261
- paired method for digit pronunciation, 254
- paired t-tests, **156**
- paired tests, **156**
- Pakhomov, S., 773
- palatal sound, **220**
- palatalization, **227**
- Palate, **220**
- palato-alveolar sound, **220**
- Pallet, D., 334
- Palmer, D. D., 71, 81
- Palmer, M., 13, 83, 419, 635, 651, 652, 656, 657, 666, 688, 690
- Palmucci, J., 150, 159, 170
- Pan, S., 265, 860
- Pang, B., 824
- Pantel, P., 690, 824
- Pao, C., 851
- Paolino, D., 285
- Papadimitriou, C., 38, 39, 42, 81, 538, 539
- Papcun, G. J., 773
- Papineni, K., 852, 910, 921
  
- Pappuswamy, U., 827
- Paquette, J., 771
- PARADISE, 852
- Paragraph segmentation, **699**
- Parallel (as coherence relation), **701**
- Parallel corpus, **899**
- parallel text, 899
- parallelism as solution to non-determinism, 33
- Paraphrase detection, **824**
- Parent annotation, **475**
- Parker, A., 620
- PARRY, 870
- parse tree, **392**, 394
- parsed corpus, 490
- parsers
  - evaluation, 483
- parsing, **45**, 460
  - ambiguity, 436
  - as search, 432
  - bottom-up, 433, **434**
  - chart, 440, **452**
  - chunking, 455
  - CKY, 440, 469
  - complexity due to memory limitations, 547
  - CYK, 440
  - difficulty of self-embedding, 547
  - Earley, 440, **447**
  - history, 461
  - morphological, **45**
  - partial, 455
  - probabilistic CKY, 469
  - shallow, **455**
  - syntactic, 431
  - top-down, 433, **433**
  - top-down versus bottom-up, 435
  - well-formed substring table, 461
- parsing and grammar, 396
- Parsons, T., 572, 587
- part-of-speech, **123**, 392
  - adjective, **125**
  - adverb, **126**
  - closed class, **124**, 126
  - greeting, **129**
  - interjection, **129**
  - negative, **129**
  - noun, 125
  - open class, **124**
  - particle, **126**
  - possessive versus personal pronouns, 124
  - subclasses of verbs, 125
  - subtle distinction between verb and noun, 125
  - usefulness of, 123
  - verb, **125**
- part-of-speech tagger
  - CLAWS, 130
  - ENGTWOL, **137**
  - HMM, **135**
  - part-of-speech tagger PARTS, **170**
  - rule-based, **135**
  - stochastic, **135**
- part-of-speech tagger TAGGIT, 169
- TBL or Brill, 151
- part-of-speech tagging, 124, **133**
  - accuracy, 155
  - adverbial *that*, 138
  - ambiguity and, 133
  - amount of ambiguity in Brown corpus, 137
  - analysis of errors in, 157
  - and morphological analysis, 160
  - Brill or TBL
  - example of rule template from, 153
  - complementizer *that*, 139
  - contingency table or confusion matrix for error analysis of, 156
  - decision trees, 170
  - distinguishing preterites from participles, 157
  - early algorithms, 137
  - evaluation, 155
  - for phrases, 157
  - Gold Standard, **155**
  - history of, 169
  - human performance at, 155
  - Hungarian, 160
  - in agglutinative languages, 161
  - log-linear analysis, 170
  - maximum entropy, 170
  - percent correct as metric for, 155
  - SNOW, 170
  - TBL or Brill, rule learning in, 152
  - TBL or Brill, sample application, 152
  - TBL or Brill, sample transformations, 155
  - Turkish, 160
  - unigram baseline, 155
  - unknown word dealing with, 158
  - use of morphology, 158
  - use of subcategorization information, 137
- Part-whole, **629**

- Partee, B. H., 16, 541, 542, 549, 550, 587
- Partial parse, **455**
- partial parsing, 455, 460
- participle
- ing in English, 50
- particle, 126, **126**, 405
- table of, 127
- PARTS, **170**
- Pasca, M., 795, 797, 799, 800
- Passage retrieval, **797**
- passive auxiliary, **406**
- Passonneau, R., 734, 810, 820, 824, 852, 869
- past participial VP, 426
- Pastizzo, M., 79
- Pater, J., 387
- Path inequalities, **530**
- path-length based similarity, **665**
- Patil, R., 437
- PATR-II, 496
- pattern
- as target of regular expression search, 18
- Patwardhan, S., 669, 670, 691
- Paul, D. B., 345, 346
- PCFG, 463, **464**, 489
- for disambiguation, 465
  - lack of lexical sensitivity, 473
  - lexicalized, 490
  - parse probability, 466
  - poor independence
    - assumption, 472
  - rule probabilities, 464
  - use in language modeling, 467
- PCM (Pulse Code Modulation), **233**
- pdf, **308**
- Pearl, J., 13, 346
- Pedersen, J., 150, 687, 790, 823
- Pedersen, T., 668–670, 689, 691
- Peirce, C. S., 725
- Pelletier, F. J., 619
- Pellom, B., 9, 852
- Penn Treebank, 155, **408**, 471
- POS tags for phrases, 157
  - tagset, 130, **131**
- Penn, G., 524, 530, 532
- per-letter entropy, 118
- per-word entropy, 115
- percent correct
- use in evaluating
    - part-of-speech taggers, 155
- Percival, W. K., 426
- Pereira, F. C. N., 11, 16, 159, 213, 214, 378, 428, 491, 492, 501, 524, 530, 531, 550, 678, 689, 690, 773, 775
- perfect
- ed/-ed form in English, 50
- perfect auxiliary, **406**
- performance grammars, 619
- performative, **829**
- period disambiguation, 251
- period of a wave, **231**
- Perkowitz, M., 156
- Perl language, 17, 18
- Perles, M., 417, 549
- Perlis, A. J., 10, 427
- perlocutionary act, 829
- Perplexity, **95**, **117**
- perplexity, 92
- as weighted average branching factor, 96
  - defined via cross-entropy, **117**
- Perrault, C. R., 12, 864, 867, 870
- person, 495
- personal pronoun, **128**
- Peshkin, L., 765
- Peskin, B., 364
- Peters, I., 688
- Peters, S., 587, 734, 871
- Peters, W., 688
- Peterson, G. E., 247, 283
- Peterson, J. L., 73
- Petkevic, V., 162
- Petkevič, V., 163
- Petri net, 585
- as meaning representation, 585
- Petrie, T., 212, 334
- Petrov, S., 477, 490
- Petruck, M. R. L., 637, 646
- Pevzner, L., 700
- Pfeifer, A., 765
- PHARAOH, **920**
- Phillips, M., 851
- phone, **216**
- Phone model, **296**
- phone recognition, 292
- phone, context-dependent (CD), **348**
- phone, context-independent (CI), **348**
- phoneme, **226**, 246
- Phoneme restoration effect, **362**
- phones, 245
- phonetic feature, 227
- Phonetics, **216**
- phonetics
- articulatory, 218, **218**
  - field of, **216**
- phonological rule
- compiling into FST, 368
  - flapping, 228
  - ordering, 366
  - transducer for, 365
- phonological rules, 228
- Chomsky-Halle rewrite format, 536
- obligatory, 367
- optional, 367
- phonology
- articulatory, **244**
- Phonotactics, **223**
- PHRAN, 620
- phrasal alignment, 895
- phrasal verb, **126**, 405
- Phrase alignment, **892**
- phrase penalty
- in MT, **919**
- Phrase translation table, **893**, **904**
- phrase-based MT, 891, **891**
- phrase-structure grammar, **391**, 425, 427
- Picheny, M. A., 285, 347
- Piepenbrock, R., 81, 126, 241
- Pieraccini, R., 837, 838, 861–863
- Pierce, J. R., 919, 921
- Pierre, I., 620
- Pierrehumbert, J. B., 267, 268, 271, 285, 379, 387, 869, 870
- pinched lattice, **343**
- Pineau, J., 871
- pipe, **21**
- Pisoni, D. B., 362, 363, 387
- Pitch, **235**
- pitch, **235**
- changing in diphone TTS, 277
- Pitch accent, **224**, **264**
- pitch accent
- and sentence stress, 224
  - ToBI, 267
- Pitch extraction, **235**
- Pitch marking, **276**
- Pitch range, **271**
- Pitch track, **233**
- Pitch tracking, **276**
- pitch-synchronous joins in TTS, **276**
- Pito, R., 412
- Pitrelli, J. F., 267, 285
- Pitts, W., 9, 42
- place of articulation, **220**
- Placeway, P., 343
- planning
- and speech acts, 864
  - shared plans, 871
- Plato, 246
- Pleonastic, **713**
- plosive sound, **221**
- Plotkin, J. B., 119
- plural, 48, 399
- plural affix, **48**
- Ply, **433**
- ply of a search space, **433**

- PNAMBIC, 850  
 Poesio, M., 733, 823  
 pointwise mutual information, **674**  
 Polanyi, L., 734  
 Polifroni, J., 844, 850–853  
 politeness  
   by and to computers, 8  
 politeness marker, **129**  
 Pollard, C., 413, 414, 427, 529, 530, 532, 538, 718, 733, 735  
 Pollock, J. J., 823  
 Pols, L. C. W., 118  
 Polysemy, **625**  
 polysemy  
   and IR, 790  
 polysynthetic language, **877**  
 POMDP, 864, 871  
 Pon-Barry, H., 871  
 Ponzetto, S. P., 690  
 Popescu, A.-M., 690, 754, 755  
 Porter stemmer, 68, 79  
   use in IR, 786  
 Porter, M. F., 56, 68, 82, 786  
 POS, **123**  
 POS tagging, *see* part-of-speech tagging  
 possessive, 48  
 possessive NP, **429**  
 possessive pronoun, **128**  
 Post, B., 285  
 Post-editing, **876**  
 post-editing, **876**  
 post-editing MT output, 909  
 postdeterminer, 399  
 postmodifier, 400, 426  
 Postposed, **391**  
 postposed constructions, **391**  
 postposition, 877  
 Potamianos, A., 852  
 Potts, C., 387  
 Povey, D., 316, 343, 351, 354, 358  
 power, **535**  
 power of a signal, **234**  
 Power, R., 733, 823, 854  
 PP, 393  
 praat, 236, 237, 244  
 Prabhu, P., 852  
 Pradhan, S., 690  
 pragmatic grammars, 619  
 Prakash, S., 688  
 Prasad, R., 13, 734  
 Precedence, **21**  
 precedence, operator, **22**  
 Precision, **459**  
 precision, 459  
   for MT evaluation, 910, 921  
   in NER, 744  
   modified N-gram, in MT, 910  
 preconditions for STRIPS plan, 865  
 predeterminer, **402**  
 predicate, 405  
 predicate-argument relations, 405  
 preference semantics, 688  
 prefix, 47, 79  
 Preposed, **391**  
 preposed constructions, **391**  
 preposition, 126, **126**  
 prepositional phrases  
   semantic attachments, 613  
 prepositional dative, 425  
 prepositional phrase, 391, 393, **393**, 400  
   attachment, 473  
 prepositions  
   acquisition of semantics, 585  
   table of English, 127  
 Prescher, D., 379, 532  
 present tense, 402  
 Preterite, **49**  
 previous forward path  
   probability, 182, 321  
 previous Viterbi forward path  
   probability, 148, 185, 322  
 Price, P. J., 264, 267, 334, 850, 858  
 Primary Colors, 122  
 priming, 425  
   repetition, **78**  
 primitive decomposition, 641  
 Prince, A., 266, 371, 387  
 Prince, E., 712, 713, 733  
 principle of compositionality, 619  
 Principle of compositionality, **589**  
 principle of compositionality, 589  
 prior probability, **140**, 292  
 priority queue, **344**  
 Priority union, **529**  
 pro-drop languages, **878**  
 Probabilistic CKY, **469**  
 probabilistic CKY algorithm, 469  
 probabilistic context-free grammar, *see* PCFG  
 probabilistic model  
   in IR, **822**  
 probabilistic parsing  
   by humans, 487  
 Probability density function, **308**  
 probability density function, **308**  
 procedures  
   as meaning representations, 585  
 proclitic, **51**  
 production system, **568**  
 productions, **391**, 425  
 productive, 49  
 productive morphology  
   use in argument for not storing all possible English word forms, **46**  
 Progressive, **50**  
 progressive auxiliary, **406**  
 Progressive prompting, **846**  
 projection of an FST, **59**  
 Prolog, 568  
 Prominence, **264**  
 prominence, phonetic, **225**  
 Prompt, **840**  
 prompts  
   design of, 850  
 pronominal reference, 711  
 pronoun, 126, **128**  
   and old information, 472  
   bound, 711  
   demonstrative, 711  
   greater frequency in spoken English, 422  
   personal, **128**  
   possessive, **128**  
   table of, 129  
   wh-, **128**  
 pronunciation dictionary, **241**  
   CELEX, 241  
   CMU, **242**  
   PRONLEX, 242  
 PropBank, **636**  
 proper noun, **125**  
 Propositional meaning, **627**  
 Prosodic Phrasing, **263**  
 Prosody, **263**  
 prosody  
   accented syllables, **224**  
   reduced vowels, **224**  
 PROTO-AGENT, **635**  
 PROTO-PATIENT, **635**  
 prototype vector  
   in VQ, **305**  
 Proximal, **711**  
 Pruning, **325**  
 Prüst, H., 734  
 pseudoword, **657**  
 PTAG, 491  
 PTRANS, 642  
 Pullum, G. K., 247, 389, 396, 408, 428, 527, 542, 544, 546, 549, 619  
 pumping lemma, **539**, 540, 549  
   for context-free languages, 541, 549  
 punctuation  
   as feature in IE, 742  
 Punyakanok, V., 462



- Purple Haze, 225  
 Purver, M., 871  
 Pushkin, A., 173  
 Pustejovsky, J., 13, 645, 756, 759, 762, 775  
 Pyramid Method, **820**  
  
 q0 as start state, 28  
 Q-systems, 11  
 Qin, L., 690  
 Qu, Y., 775  
 Qualia structure, **645**  
 quality  
   for TTS evaluation, **282**  
 quantifier, **399**, 400, 426  
 Quantifiers, **563**  
 quantifiers  
   scoping, **598**  
 quantization, **232**, **297**  
 quantization of an analog signal, **232**  
 Quantz, J., 855  
 quasi-Newton optimization, 201  
 Query, **782**  
 query  
   expansion, **792**  
   improvement, 791  
   in IR, 782  
 Query expansion, **792**  
 query reformulation  
   in question answering, **794**  
 Query-based summarization, **781**  
 Query-focused summarization, **802**, **816**  
 question  
   automatic detection of, 857  
   factoid, **780**  
 Question answering, **1**  
 question answering  
   factoid questions, **780**  
   query reformulation in, **794**  
 question answering task, **730**  
 Question classification, **795**  
 Question ontology, **795**  
 Question rise, **267**  
 question rise, 267  
 questions, 397  
 queue  
   use of in breadth-first search, 37  
 Quillian, M. R., 11, 587, 688  
 Quinlan, J. R., 213  
 Quirk, C., 824, 914  
 Quirk, R., 49, 127, 428  
  
 R. D., 428  
 Rabin, M. O., 42  
 Rabiner, L. R., 11, 179, 190, 191, 213, 335  
 Radev, D., 13, 762, 806, 809, 823  
  
 Radford, A., 391, 428  
 Radio Rex, 333  
 Raghavan, P., 16, 822  
 Raghunathan, B., 871  
 Rambow, O., 81, 690, 840  
 Ramesh, P., 364  
 Ramshaw, L. A., 150, 159, 170, 359, 419, 457, 462, 484, 688, 839  
 Rand, D., 230  
 random sentences baseline in summarization, **821**  
 range  
   in regular expressions, **19**  
 Raphael, B., 586  
 Rapid reprompting, **846**  
 Rappaport Hovav, M., 635, 645  
 rarefaction, 231  
 Ratcliff, D., 201  
 rate of speech, **229**  
 Rationalist, **433**  
 rationalist tradition, **433**  
 Ratiu, F., 871  
 Ratnaparkhi, A., 155, 156, 159, 160, 170, 213, 482, 490, 491  
 Ratner, N. B., 78  
 Rau, L. F., 774, 823  
 Ravichandran, D., 690, 795, 799  
 Ravishankar, M. K., 347  
 Raymond, W. D., 119, 229, 358  
 Rayner, M., 836, 837, 839, 871  
 RE, **17**  
 read speech (as opposed to conversational speech), **288**  
 Reading time, **487**  
 reading time experiments, **730**  
 real estate ads  
   predictors of house prices, 194  
 real-word spelling errors, **73**  
 Rebholz-Schuhmann, D., 772  
 Recall, **459**  
 recall, 459  
   for MT evaluation, 910, 921  
   in NER, 744  
 recognition  
   by finite-state automaton, 27  
 recombining hypotheses  
   in MT, 908  
   in speech, 340  
 Rectangular, **299**  
 recursion, 391  
   center-embedding, 417  
 recursively enumerable, **536**  
 reduced vowel, **224**  
 reduction  
   phonetic, 224, **226**  
 reentrant structure, 496, 498  
 Reeves, B., 8  
  
 Reference, **707**  
 reference  
   bound pronouns, 711  
   cataphora, 711  
   definite, 710  
   generics, 713  
   indefinite, 710  
   pronominal, 711  
   resolution, 707  
   Centering algorithm, 718, 720  
   Hobbs tree search algorithm, 716, 717  
   psychological studies, 730  
 Reference line, **271**  
 Reference point, **574**  
 Reference resolution, **707**  
 Referent, **707**  
 referent  
   accessing of, **708**  
   evoking of, **708**  
   mediated, **712**  
 Referent naming task, **730**  
 Referential density, **878**  
 Referring expression, **707**  
 Reflexive, **714**  
 reformulation, **831**  
 Regier, T., 585  
 register  
   sociolinguistic, **230**  
 registers in RE pattern  
   substitution, **25**  
 Regression line, **195**  
 regression versus classification, 194  
 Regular expression, **17**  
 regular expression, 17, **17**, 18, 41  
   returning lines of documents, 18  
   substitutions, **25**  
 regular grammar, 537  
   and Chomsky hierarchy, 537  
 Regular language, **26**  
 regular language, **26**, **38**  
   proof that English isn't one, 543  
   pumping lemma, 540  
 Regular relation, **58**  
 regular sets, 38  
 regular verbs, **49**  
 regularization in MaxEnt learning, **204**  
 Rehder, B., 679  
 Reichenbach, H., 574  
 Reichert, T. A., 213  
 Reichl, W., 120  
 Reichman, R., 869  
 Reimer, U., 823  
 Reiter, E., 840, 859  
 Reiter, R., 529  
 Reithinger, N., 854, 871

- Rejection, **846**  
 rejection, 860  
   by finite-state automaton, **28**  
 Relation detection and classification, **738**  
 relative  
   temporal expression, 755  
 relative entropy, **661**  
 Relative frequency, **89**  
 relative frequency, **89**  
 relative pronoun, **401**  
 release  
   stop, **221**  
 Relevance feedback, **791**  
 relevance feedback, **791**, 822  
   evaluation of, 791  
   update formula, 791  
 Renals, S., 871  
 repair, 422  
   in conversation, **360**, 422, **423**  
 repair as disfluency, 426  
 Reparandum, **423**  
 reparandum, 423, **423**  
 repeated name penalty, **731**  
 repeated parsing of sub-trees, 461  
 repetition  
   of words as disfluency, 422  
 Repetition priming, **362**  
 REQUEST, 866  
 rescoring, **339**  
 Residual collection, **791**  
 residual collection, **791**  
 Resnik similarity, **667**  
 Resnik, P., 490, 491, 660, 661, 665–667, 688, 689  
 Resolution, **569**  
 resolve, **133**  
 Resource Management, 334  
 restart, 422  
   and disfluency, 422  
 restart as disfluency, 426  
 Restriction, **594**  
 Restrictive grammar, **836**  
 restrictive relative clause, **401**  
 Result (as coherence relation), **701**  
 Reverse translation model, **918**  
 rewrite, 392  
 Reynar, J. C., 491, 699, 733  
 Reynolds, D. A., 335, 359  
 rhetorical relations, 703  
 Rhetorical Structure Theory, *see* RST  
 Rhodes, F. L., iv  
 Rhodes, R. A., 229  
 rhyme  
   syllable, **223**  
 Riau Indonesian, 125  
 Ribarov, K., 491, 492  
 Ribeiro-Neto, B., 779, 822  
 Riccardi, G., 826  
 Rich transcription, **360**  
 Richards, C., 254, 255  
 Richardson, J. E., 772  
 Rieger, C., 688  
 Ries, K., 857, 858, 871  
 Riesbeck, C. K., 11, 620, 688  
 Riezler, S., 484, 491, 532, 823  
 Riggle, J., 387  
 right-linear, **537**  
 right-linear grammar, **537**  
 Rijsbergen, C. J. V., 822  
 Riley, M. D., 272, 378  
 Riloff, E., 620, 680, 733, 753, 774, 775  
 rime  
   syllable, **223**  
 Rindflesch, T. C., 772  
 Ringenber, M., 827  
 Ringwald, M., 772  
 Ristad, E. S., 81, 550  
 Rivest, R. L., 656  
 RMS amplitude, 234  
 Roark, B., 81, 111, 487, 491  
 Robins, R. H., 170, 246  
 Robinson, J. A., 531  
 Robinson, J. J., 619  
 Robinson, S. E., 822  
 Rocchio, J. J., 791, 822  
 Roche, E., 81, 153  
 Rochester, N., 10  
 Rohde, D. L. T., 412–414  
 Romanian, 162  
 Roossin, P. S., 13, 920  
 root-and-pattern morphology, 52  
 Rooth, M., 490  
 Roque, A., 827  
 Rosario, B., 772  
 Röscheisen, M., 920  
 Rosé, C., 827  
 Rose, R. C., 335, 364  
 Rosenberg, A. E., 361  
 Rosenberg, C. R., 284  
 Rosenfeld, R., 109, 112, 113, 213, 362  
 Rosenzweig, J., 651, 658, 659  
 Ross, K., 364  
 Roth, D., 170, 171, 462, 765, 795, 796  
 ROUGE, **819**  
   compared to BLEU, 819, 820  
 ROUGE-1, **819**  
 ROUGE-2, **819**  
 ROUGE-L, **820**  
 ROUGE-S, **820**  
 ROUGE-SU, **820**  
 Roukos, S., 170, 348, 483, 484, 490, 491, 910, 921  
 rounded vowels, **223**  
 Rounds, W. C., 501, 529, 530  
 Roussel, P., 531  
 Roy, N., 871  
 RST, **702**, 806  
   original formulation, 703  
   parsing, 706  
   TreeBank, 703  
 Rubenstein, H., 669  
 Rubin, D. B., 104, 167, 187  
 Rubin, G. M., 137, 169, 772  
 Rudnicky, A. I., 362, 826, 846, 852  
 Ruhi, U., 243  
 rule  
   lexical, **529**  
 rule operator, 367  
 rule-based, **135**  
 rule-to-rule hypothesis, 619  
 Rule-to-rule hypothesis, **591**  
 rules, 425  
   context-free, **391**  
   expansion, 392, 397  
   sample English, 393  
   use of () for optional constituents, 400  
   dotted, **447**  
   optional, 367  
   orthographic, 53, 63  
   phonological, 228  
   compiling into FSTs, 368  
   ordering, 366  
   two-level, 367  
   phonological and transducers, 365  
   spelling, 53, 63  
 Rumelhart, D. E., 11, 587  
 Ruppenhofer, J., 637  
 Rush, A., 914  
 Russell, G., 920  
 Russell, N. H., 364  
 Russell, R. C., 82, 170  
 Russell, S., 16, 42, 84, 562, 587, 656, 691, 863, 871  
 Russell, S. W., 646  
 Russian, 877, 878  
 Rutishauser, H., 10, 427  
 Ryder, M. E., 620  
  
 S, **393**  
 Sacks, H., 828, 856, 870  
 Sadek, D., 871  
 Sadek, M. D., 871  
 Sadock, J., 555  
 Saffran, J. R., 365, 382, 387  
 Sag, I. A., 408, 413, 414, 427, 428, 511, 527, 529, 530, 532, 619, 870  
 Sagerer, G., 228  
 Sagisaka, Y., 283, 285  
 Saint-Dizier, P., 645

- Sakoe, H., 213, 333  
 Salasoo, A., 362  
 salience  
   in discourse model, **711**  
 Salomaa, A., 80, 490  
 Salton, G., 782, 791, 822  
 Salvetti, F., 775  
 Samelson, K., 10, 427  
 Sampling, **297**  
 sampling, **297**  
   used in clustering, 687  
 sampling of an analog  
   waveform, **232**  
 Sampling rate, **297**  
 sampling rate, **232, 297**  
 Sampson, G., 102, 103, 120, 170  
 Samuel, A. G., 362  
 Samuel, K., 151  
 Samuelsson, C., 120, 159  
 San-Segundo, R., 842, 860  
 Sanders, G., 852  
 Sanders, T. J. M., 734  
 Sanderson, M., 790  
 Sanfilippo, A., 405, 484, 509,  
   529  
 Sang, E. F. T. K., 13, 746, 775  
 Sang, E. T. K., 460  
 Sankoff, D., 213, 230  
 Santorini, B., 13, 123, 130–132,  
   155, 408, 471, 483, 484  
 Saraclar, M., 111, 357  
 Sassano, M., 775  
 Satellite, **703**  
 satellite, **878**  
 satellite-framed language, 878  
 satisfaction-precedence  
   discourse relation, 868  
 Satta, G., 374  
 Saurí, R., 13  
 Sauri, R., 756, 759  
 Saur, R., 762  
 sausage, in ASR, **343**  
 Scaled likelihood, **355**  
 scaled likelihood, **355**  
 SCFG, *see* PCFG, **464**  
 Schönkfinkelization, 566  
 Scha, R., 734  
 Schabes, Y., 81, 153, 171, 490,  
   491, 914  
 Schachter, P., 126  
 Schaefer, E. F., 830, 854  
 Schafer, R. W., 333  
 Schäfer, U., 532  
 Schalkwyk, J., 850  
 Schank, R. C., 11, 587, 629, 764  
 Schapire, R. E., 213, 491, 813  
 Schasberger, B., 408, 410  
 Scheffczyk, J., 637  
 Schegloff, E. A., 828, 832, 856,  
   870  
 Scheideck, T., 871  
 Schilder, F., 762, 775  
 Schiller, A., 70  
 Schlaikjer, A. H., 818  
 Schlesinger, J. D., 810, 813, 823  
 Schmandt, C., 832, 850  
 Schmolze, J. G., 532, 587  
 Schone, P., 382–384  
 Schönkfinkel, M., 566  
 Schoolhouse Rock, 123  
 Schreiner, M. E., 679  
 Schreuder, R., 79  
 Schroder, M., 285  
 Schroeter, J., 284  
 Schubert, L. K., 619  
 Schuetze-Coburn, S., 285  
 Schultz, T., 357, 364  
 Schütze, H., 16, 170, 213, 471,  
   492, 657, 675, 686, 689,  
   690, 790, 805, 822  
 Schwa, **224**  
 schwa, **224**  
 Schwartz, A. S., 771  
 Schwartz, R., 150, 159, 170,  
   339, 343, 347, 348, 359,  
   482, 491, 775, 810, 823,  
   837, 839, 921  
 Schwartz, S., 9  
 SCISOR, 774  
 sclite, 331  
 sclite package, 77  
 Scott, D., 42, 733  
 Scott, M., 487  
 script, 629  
 Scripts, **764**  
 SCU, **820**  
 SDC, 334  
 SDRT (Segmented Discourse  
   Representation Theory),  
   **734**  
 search, 42, 460  
   A\*, **363**  
   as metaphor for  
     non-deterministic  
     recognition, 35  
   beam, 325  
   beam in MT, 907  
   breadth-first, **37, 41**  
   picture of, 37  
   pitfalls of, 38  
   data-directed, 433  
   depth-first, **36, 37**  
   pitfalls in, 37  
   FIFO, 37  
   First In First Out, 37  
   forward-backward, 343  
   goal-directed, 433  
   in ASR, 290  
   Last In First Out, 37  
   LIFO, 37  
   multiple-pass, 343  
   parsing as, 432  
   queue for breadth-first, 37  
   stack for depth-first, 36  
 search strategy, 41  
 search-state in non-deterministic  
   recognition by finite-state  
   automata, **33**  
 Searle, J. R., 7, 830, 870  
 secondary stress, **225**  
 See, A., 13, 762  
 seed pattern  
   in IE, **751**  
 Segal, J., 90, 487  
 segment based acoustic model,  
   **364**  
 segment or phone, **216**  
 segmentation  
   Chinese word, **70**  
   maximum matching, **70**  
   sentence, 70  
   word, **47, 69**  
 Seidenberg, M. S., 387  
 Seitz, F., 348  
 Sejnowski, T. J., 284  
 Sekine, S., 484, 775  
 selectional association, **661**  
 selectional preference strength,  
   **660**  
 Selectional restriction, **639**  
 selectional restriction violations  
   and WSD, 660  
 selectional restrictions, 640  
 self-embeddings, **547**  
 Selfridge, J. A., 92  
 Selfridge, M., 620  
 Selfridge, O. G., 333  
 Selkirk, E., 285  
 semantic analysis  
   syntax-driven, 590  
 semantic networks, 552  
 Semantic analysis, **552, 589**  
 semantic analysis  
   expectation-driven, 620  
   syntax-driven, 589  
   word-based, 620  
 Semantic analyzer, **590**  
 semantic analyzer, **590, 615**  
 Semantic attachments, **591**  
 semantic concordance, **651**  
 semantic distance, **664**  
 semantic drift  
   in IE, 753  
 semantic feature, **642**  
 Semantic field, **629**  
 semantic grammars, 619  
 semantic network  
   for word sense  
     disambiguation, 688  
 Semantic networks, **578**

- semantic networks
  - origins, 587
- semantic processing, 551
- semantic relations
  - list of for IE, 747
- semantic role labeling, **683**
- semantic transfer
  - in MT, 885
- semantically unpredictable
  - sentence, 282
- semantics, 551
  - frame-and-slot, 836
  - grounding in embodiment, 585
- semi-HMM, **838**, 839
- Semilattice, **501**
- semivowel, **220**
- Sen, Z., 358
- Seneff, S., 243, 378, 379, 837, 838, 841, 844, 850–853
- Senellart, J., 887
- Sengal, C. J., 730
- sense
  - word, **624**
- sense accuracy, **656**
- SENSEVAL
  - and WSD evaluation, 656
- SENSEVAL corpus, 651
- sentence, 426
- sentence centrality
  - in summarization, **805**
- Sentence completion task, **730**
- Sentence compression, **810**
- Sentence error rate, **331**
- sentence error rate, **331**
- Sentence fusion, **815**
- sentence segmentation, **70**
  - in MT, 899
- Sentence simplification, **810**
- Sentence tokenization, **251**
- sentential
  - use of cue phrase, **705**
- Sentential complements, **404**
- sentential complements, **404**
- sequence classifier, **173**
  - FST as, 173
- sequential transducer, **59**
- serial method for digit
  - pronunciation, **254**
- Set-valued features, **530**
- Sethi, R., 462
- Settles, B., 738
- Setzer, A., 756, 759, 762
- Seuss, D., 45
- Seymore, K., 112
- Shafraan, I., 357
- Shah, P. K., 773
- Shaked, T., 690, 754, 755
- Shakespeare
  - N-gram approximations to, 93
- author attribution, 122
- word count, 86
- Shallow parse, **455**
- Shamir, E., 417, 549
- Shanahan, J., 775
- Shannon, C. E., 9, 80, 92, 117, 119, 334
- shape
  - feature in IE, **742**
- shared plans, 871
- Sharpley, A., 282
- Shattuck-Hufnagel, S., 264
- Shaw, G. B., 247
- sheep language, 27
- Sheil, B. A., 462
- Shen, L., 81, 921
- Shen, W., 359, 920
- Sherlock, G., 772
- Shieber, S. M., 7, 16, 496, 501, 524, 530, 532–534, 544, 545, 619, 914
- Shih, C., 81, 274
- Shillcock, R., 487
- Shimodaira, H., 871
- Shimoni, A. R., 775
- Shimony, S. E., 726
- Shipley, K. L., 284
- Shitaoka, K., 700, 734
- Shnayder, V., 679
- Shockey, L., 247
- Shopen, T., 230
- Shoup, J. E., 216
- SHRDLU, 11, 586
- Shriberg, E., 85, 335, 360, 361, 364, 422, 850, 857, 858, 871
- sibilant sound, **221**
- Sibun, P., 150
- Siddharthan, A., 823
- Sidner, C. L., 12, 732, 868–871
- Sigleman, S., 824
- signal processing, 292
- Signature terms, **805**
- signatures, **383**
- significance tests
  - Kendall's  $\tau$ , 819
- Significant silence, **828**
- significant silence, **828**
- Siler, S., 827
- Silliman, S., 827
- Sills, D. L., 617
- Silverman, K., 267
- Simma, A., 733
- Simmons, R. F., 12, 137, 169, 586, 587, 620, 645, 683, 688, 690, 780, 823
- Simple Good-Turing smoothing, 102
- Simple types, **527**
- simple types, 527
- Simplified Lesk, **658**
- Singer, Y., 156, 160, 170, 491, 813
- Singh, S. P., 844, 863
- Single document summarization, **802**
- single initiative, **842**
- singleton
  - in authorship identification, 122
- singleton events, **101**
- singleton N-grams, **101**
- singular, 402
- singular affix, **48**
- singular value decomposition
  - in IR, 690
- sink state in finite-state
  - automaton, 30
- Situational context, **708**
- Sivertsen, E., 283
- skip bigrams, **820**
- Skip N-gram, **113**
- Skut, W., 259
- Slator, B. M., 688, 689
- Sleator, D., 419, 491
- slips of the tongue, 78
- Slobin, D. I., 878
- slot-filler representations, 578
- Slovene, 162
- Small, S. L., 688
- SMART, 822
- SMDP, 871
- Smedt, K. D., 532
- Smith, D. A., 492
- Smith, I., 287
- Smith, N. A., 492, 550
- Smith, R. W., 871
- Smith, V. L., 872
- Smolensky, P., 371, 385, 387
- Smoothing, **98**
- smoothing, 97, **98**
  - absolute discounting, **109**
  - add-one, **98**
  - backoff, **105**
  - details of Katz backoff, 106
  - discounting, **99**
  - for HMM POS tagging, 150
  - Good-Turing, **101**
  - interpolation, **103**
  - Katz backoff, **105**
  - Kneser-Ney discounting, **109**
  - Laplace, **98**
  - linear interpolation, 104
  - Simple Good-Turing, **103**
- Smrž, O., 81
- Smyth, R., 730
- Snippet, **816**
- snippet, **780**
- Snodgrass, A., 357

- Snover, M., 921  
 SNOW algorithm, 170  
 Snow, R., 681, 682, 688, 690  
 Snyder, L., 9  
 Social media, **775**  
 sociolinguistic factors in pronunciation, **229**  
 Soderland, S., 766, 775  
 Somers, H. L., 880, 881, 887, 921  
 Song, Z., 243  
 sonority, **377**  
 sonority hierarchy, **377**  
 Soon, W. M., 723  
 Soong, F. K., 339  
 Sorensen, J., 750  
 Soricut, R., 813, 823  
 Sound inference, **725**  
 Soundex, 82  
 source-filter model, **241**  
 Souter, C., 113  
 SOV language, **877**  
 space  
   as a regular character, 18  
 Spam detection, **824**  
 Spanish, 241  
 Sparck Jones, K., 16, 95, 688, 689, 785, 822, 824  
 Sparse data, **97**  
 sparse data, **97**  
 speaker adaptation, **356**  
 Speaker identification, **335**  
 speaker identification, **335**  
 Speaker independent, **289**  
 Speaker verification, **335**  
 speaker verification, **335**  
 speaker-independent, 289  
 species  
   in definition questions, **817**  
 spectral  
   features, 292  
 spectral subtraction, **355**  
 Spectral tilt, **298, 361**  
 spectrogram, **239**  
 Spectrum, **237**  
 spectrum, **237**  
 speech  
   telephone bandwidth, **232**  
   wideband, **232**  
 speech act, **829, 830**  
 speech error, 78  
 speech recognition  
   acoustic model training, 326, 327  
   architecture, 289  
   continuous, **288**  
   history of, 333  
   isolated-word, **288**  
   noisy channel model, 290  
   pruning, 325  
   speaker independent, 289  
   word segmentation, 318  
 Speech synthesis, **249**  
 speech synthesis, *see* TTS  
 spelling correction  
   use of N-grams in, 84  
 spelling errors  
   correction  
     context-dependent, 73  
     EM, 167  
     isolated-word, 73  
     noisy channel model, 165  
   detection  
     context-dependent, 73  
     morphology, 73  
     non-word, 72  
     real words via N-gram, 167  
   overview of detection and correction, 72  
   real word, 73  
   societal role, 72  
 spelling rule  
   English consonant doubling, 50  
 spelling rules, 53, **63**  
 Spiegel, M. F., 259  
 Split, **475**  
 Split and merge, **477**  
 spoken dialogue system, 825  
 spoken English, grammar of, 421  
 spoken language system, 825  
 Spooren, W. P. M., 734  
 Sporleder, C., 705, 707, 733  
 Sproat, R., 39, 55, 56, 81, 158, 254, 255, 266, 285, 357, 371, 378, 542  
 spurious word  
   in MT, 914  
 spurious words  
   in MT, 894  
 SRILM toolkit, **109**  
 Srinivas, B., 491  
 Srivastava, R., 827  
 Stabler, E., 538  
 stack  
   for depth-first search, 36  
 stack decoder, *see* A\* decoder  
 Stack decoding, **905**  
 Stallard, D., 359, 839, 852  
 Stalnaker, R. C., 830  
 Stanners, R. F., 78  
 Starr, R., 357  
 start state, **27**  
 Start symbol, **393**  
 start symbol, **393**  
 state  
   accepting, 27  
   final, **27**  
   in finite-state automaton, **27**  
 state-space search, **35**  
 state-transition table  
   finite-state automaton, **28**  
 states  
   representation of, 569  
 Stationary, **299**  
 stationary stochastic process, **116**  
 statistical paradigm  
   rise of, 11  
 statistical significance  
   MAPSSWE for ASR, 156, 331  
   McNemar test, 156, 331  
   paired t-tests, 156  
   Wilcoxon signed-rank test for sequence models, 156  
 statistical vs symbolic paradigms, 10  
 x expressions, **575**  
 stative expressions, 575  
 Steedman, M., 268, 421, 428, 491, 538, 713  
 Stefanowitsch, A., 646  
 stem, **47**  
 stemming, 79, **786**  
   and morphological parsing, **46**  
 stemming in IR, **68**  
 Stent, A., 853, 858, 859, 864  
 Stetina, J., 491  
 Stevens, K. N., 247, 284  
 Stevens, S. S., 235, 301  
 Stevenson, M., 688  
 Stevenson, R., 733  
 Stevenson, R. J., 730  
 Stevenson, S., 690, 691  
 Stickel, M. E., 726  
 Stifelman, L. J., 832, 850  
 Stochastic OT, **375**  
 Stockham, T. G. J., 333  
 Stoics, 123  
 Stolcke, A., 85, 90, 109, 112, 113, 120, 316, 335, 343, 360, 364, 487, 490, 857, 858, 871  
 Stolz, W. S., 139  
 Stone, C. J., 213  
 Stone, M., 734  
 Stone, P. J., 688, 689  
 stop (consonant), **221**  
 Stop list, **786**  
 stop list, **786**  
 stoplist, **672**  
 stopword, **672**  
 Stork, D. G., 213, 306  
 Story of the Stone, 873  
 Streeter, L., 285  
 stress  
   lexical, **224**  
   secondary, **225**  
 string, 395

- defined as sequence of symbols, **18**
- string distance, **74**
- Strom, V., 284
- strong equivalence
  - of grammars, 416
- Strube, M., 690, 733
- Structural ambiguity, **436**
- structurally ambiguous, 461
- Strzalkowski, T., 483, 484, 823
- Sturm, J., 860
- Stuttle, M., 871
- style
  - in phonetic variation, **230**
- Stylianou, Y., 284
- Su, J., 748
- Su, Y., 357
- subcategorization, 390, **405**, 417, 426, 464, 506, 510
  - alternation, **425**
- subcategorization frame, 405, **405**, 425
- Subcategorization frames, **508**
- subcategorize for, **405**
- subdialogue, **869**
  - knowledge precondition, 869
- subject, syntactic, 390, 397, 402, 426
  - frequency of pronouns as, 472
- subject-verb agreement
  - via unification, 495
- subject-verb agreement, 403
  - via unification, 498
- sublanguage, **876**
- subsequential transducer, **59**
- substitutability, 427
- substitution
  - in TAG, 427
- substitutions
  - regular expressions, **25**
- Subsumption, **580**
- subsumption, **582**
- subsumption in unification, **500**
- Subtype, **527**
- subtype, **527**
- Suci, G. J., 689
- suffix, 47, 79
- Sum-of-products, **270**
- Sum-squared error, **196**
- SumBasic, **823**
- summarization, 823
  - generic, **802**
  - multiple document, **802**
  - single document, **802**
  - via sentence centrality, **805**
- Summary Content Unit, **820**
- Sundheim, B., 13, 724, 756, 759, 762, 766, 774
- Superordinate, **628**
- supersenses, **682**
- Support Vector Machine, *see* SVM
- Suprasegmental, **263**
- Surdeanu, M., 690
- surface form, **45**
- surface level in finite-state morphology, **60**
- surface tape, 61
- SUS (semantically unpredictable sentence), **282**
- Sutton, C., 214
- Sutton, R. S., 863, 871
- Sutton, S., 850
- Suzuki, H., 810, 823
- Suzuki, T., 284
- Svartvik, J., 49, 127, 428
- SVM, 204
- SVO language, **877**
- Swedish, 878
- Sweet, H., 246
- Swerts, M., 857–860
- Swier, R., 690
- Swiss German
  - cross-serial constraints, 544
- Switchboard
  - Transcription Project, 243
- Switchboard Corpus, 85, 86, 225, 229, 231, 232, 242, 297, 318, 335
- Syllabification, **223**
- syllabification, 373, **376**
  - Maximum Onset Principle, **377**
- syllable, **223**
  - accented, **224**
  - coda, **223**
  - nucleus, **223**
  - onset, **223**
  - prominent, **224**
  - rhyme, **223**
  - rime, **223**
- syllables
  - complex codas, 372
  - complex onsets to, 372
- symbolic vs statistical paradigms, 10
- Symmetrizing, **903**
- Synchronous grammar, **912**
- Synonym, **627**
- synonymy
  - and IR, 790
- Synset, **630**
- Syntactic disambiguation, **437**
- Syntactic movement, **408**
- syntactic movement, **408**
- syntactic transformations, **885**
- Syntax, **389**
- syntax, **389**
  - origin of term, 123
- Syrdal, A. K., 282, 284, 285
- system initiative, **842**
- Systemic grammar, **407**
- systemic grammar, **407**
- t
  - parameter in IBM Model 3, **914**
- t-test
  - for word similarity, **675**
- T., 387
- tableau, **373**
- tableau in Optimality Theory, 373
- TAG, 427, 491
- tagger, *see* part-of-speech tagger
- CLAWS, 130
- tagging, *see* part-of-speech tagging, **133**
- tagging, **133**
- TAGGIT, 169
- tags, *see* tagsets or part-of-speech taggers/tagging
- tagset, **123**
  - Brown, 130
  - C5, 130
  - difference between C5 and Penn Treebank, 132
  - difference between Penn Treebank and Brown, 132
  - English, 130
  - history of Penn Treebank, 132
  - Penn Treebank, 130, **131**
  - table of Penn Treebank tags, 131
- Tajchman, G., 90, 487
- take the first sense, **657**
- Talbot, D., 920
- Talmy, L., 878
- Tamil, 878
- Tanabe, L., 772
- Tanenhaus, M., 688
- Tanimoto similarity metric, **677**
- Tannenbaum, P. H., 139, 689
- tap (phonetic), **222**, **226**
- Tapanainen, P., 418, 419
- tape in finite-state automaton, 27
  - picture of, 28
- Tapered prompt, **840**
- Target cost, **279**
- Target point, **271**
- Tarte Tatin
  - recipe for, 697
- Taskar, B., 491
- tau Yih, W., 765
- Taura, K., 532
- Taussig, K., 357
- taxonomy, **629**
- Taylor, L., 733



- Taylor, P., 250, 264, 268, 269, 272, 275, 283–285, 302, 303, 857, 858, 871
- Taylor, W. L., 909
- TBL, **151**, 171
- painting metaphor for, 151
- TBox, **579**
- TD-PSOLA, **276**, 277
- technai, 123
- Telephone-bandwidth, **297**
- telephone-bandwidth speech, **232**
- Telic eventualities, **578**
- Temperley, D., 419, 491
- template
- in IE, 764
- template, in TBL, **153**
- template-based generation, 840
- Template-filling, **739**
- templatic morphology, 52, 370
- temporal expressions
- metaphors for, 575
- temporal adverb, **126**
- Temporal analysis, **738**
- Temporal anchor, **760**
- temporal expression, **740**
- absolute, 755
- relative, 755
- Temporal expression
- recognition, **738**
- Temporal logic, **573**
- temporal normalization, **758**
- temporal reasoning, 587
- Tengi, R. I., 652
- Tense logic, **573**
- ter Meulen, A., 16, 541, 542, 549, 550, 587, 619
- Teranishi, R., 284
- term
- in IR, 782, **782**
- Term clustering, **792**
- term clustering, 688, 689
- term frequency, 785
- Term weight, **782**
- term weighting, 785, 822
- term weights
- in IR, **782**
- term-by-document matrix, **784**
- terminal symbol, **392**, 425
- Terminology
- in description logic, **579**
- terms
- in FOL, **562**
- Tesar, B., 385, 387
- Tesnière, L., 418
- Test set, **91**
- test set, 91
- development, **92**
- how to choose, 92
- Teufel, S., 807, 810, 824
- text analysis (stage of TTS), **250**
- Text categorization, **824**
- text normalization
- in TTS, 250
- text summarization, **802**, 823
- Text-to-speech, **249**
- text-to-speech synthesis, *see* TTS
- TextTiling, **697**
- TF\*IDF, **266**
- tf-idf, 785
- The Pirates of Penzance, 287
- Thede, S. M., 150, 159
- Thematic grid, **634**
- Thematic role, **633**
- thematic role, 632
- and diathesis alternation, 635
- examples of, 633
- in MT, 888
- labeling, 683
- problems, 635
- Theme, **633**
- theme, as thematic role, **633**
- there, **129**
- thesaurus, 687
- for query expansion in IR, **792**
- Thesaurus generation, **792**
- thesaurus generation, **792**
- thesaurus induction, **682**
- Thibaux, R., 477, 490
- Thiessen, E. D., 387
- Thione, G. L., 734
- third-person, 402
- Thomas, D. S., 14
- Thomas, J. A., 113–115, 118
- Thompson, H., 825, 836
- Thompson, K., 42
- Thompson, R. A., 465
- Thompson, S. A., 285, 703, 734, 885
- Thornton, J. H. S., 364
- Thrax, D., 123
- Thrun, S., 871
- Tibshirani, R., 213
- Tice, D. M., 13
- tied states in HMM, **348**
- tier
- of ToBI transcript, 268
- tier, morphological, **370**
- Tillmann, C., 419, 484, 920
- Tilt model, **268**
- tilt parameter in, 268
- tilt parameter in Tilt model, **268**
- time
- representation of, 572
- Time-aligned transcription, **243**
- time-synchronous beam search, 326
- TimeBank, **762**
- TIMIT, 243
- Tishby, N., 678, 689, 690
- Titone, D., 377
- Titov, I., 491
- Tjong Kim Sang, E. F., 13
- ToBI, **267**
- boundary tones, 267
- tiers, 268
- Todaka, Y., 229
- Tokenization, **47**
- tokenization
- for N-grams, 86
- sentence, 70
- in TTS, 251
- word, **69**
- tokens
- word, **86**
- Tokuda, K., 285
- Tolle, R., 772
- Tomabechi, H., 524
- Tomalin, M., 364
- Tomita, M., 920
- Tomokiyo, L. M., 288, 357
- Tone unit, **285**
- Top-down, **433**
- top-down, 431, 433, 460
- top-down parser, **433**
- topic (information structure), 472
- Topic signature, **805**
- topic-based language models, **112**
- topicalization, **398**
- Topline, **271**
- Touretzky, D. S., 380
- Toutanova, K., 156, 160, 168, 170, 491, 532, 899
- Towell, G., 651
- trace, 398, **408**
- trachea, **218**
- trailing unit method of digit pronunciation, 256
- Training set, **91**
- training set, **91**, 153
- crossvalidation, 155
- how to choose, 92
- Tran, B.-H., 347
- transcription, 421
- time-aligned, **243**
- transducer
- composition, 58, **58**, 367
- inversion, **58**
- sequential, **59**
- subsequential, **59**
- weighted, **378**
- Transduction grammar, **912**
- transfer model, **884**
- Transformation Based Learning, *see* TBL
- transformation-based learning, *see* TBL

- transformation-based tagger, 135  
Transformations and Discourse Analysis Project (TDAP), 10, 169  
transition probability, 148, 182, 185, 321, 322  
transitions in finite-state automaton, 27  
transitive, 426  
transitive verbs, **405**  
translation  
    difficulty of literary, 874  
    impossibility of, 889  
translation divergences, **876**  
Translation model, **890**  
translation probability, 891  
Traum, D. R., 853–855, 871  
TREC, 790, 823  
Tree-Adjoining Grammar (TAG), 427  
    adjunction in, 427  
    probabilistic, 491  
    substitution in, 427  
Tree-Adjoining Grammar, 538  
tree-structured lexicon, **347**  
Treebank, **408**  
treebank, **408**, 471  
Treiman, R., 377  
trellis  
    Viterbi, 148  
trie, **382**  
trigger language model, **113**  
Trigram, **91**  
trigram, **91**  
triphone  
    HMM model for speech recognition, 348  
Trost, H., 387  
Trubetskoi, N. S., 531  
Truth-conditional semantics, **561**  
Tsakalidis, S., 343, 351  
Tsang, V., 691  
Tseng, H., 81, 160, 682  
Tsochantaridis, I., 733  
Tsuji, J., 532, 920  
Tsukahara, W., 832  
Tsuruoka, Y., 532  
TTS, **249**  
    diphone, 273  
    diphone vs unit selection, 278  
    HMM, 285  
    hourglass metaphor for, 250  
    unit selection, 278  
Tuantranont, J., 9  
Tufis, D., 162  
Tukey, J. W., 333, 687  
Tune, **267**  
tune  
    final fall, 267  
    question rise, 267  
tune:continuation rise, 267  
Tür, G., 161, 162, 360, 364  
Turian, J. P., 921  
Turing equivalent, 537  
Turing machine, 42, 536  
    as origin of finite automaton, 42  
Turing test, **6**, 7  
Turing, A. M., 6, 9, 80  
Turk  
    Mechanical, 249  
Turk, A., 119  
Turkish, 877  
    morphemes per word, 46  
    number of words, 46  
    part-of-speech tagging, 160  
    really long words in, 46  
turn, **827**  
turn correction ratio, 851  
Turn-taking, **827**  
turn-taking, **827**  
Turner, J., 823, 824  
Turney, P., 679  
Tutiya, S., 854  
two-level  
    morphology, 79, 365, 367, 386  
    feasible pair, 61  
    lexical tape, 60  
    surface tape, 60  
    rule, 367  
    and Optimality Theory, 374  
    compiling into FST, 368  
two-level morphology  
    feasible pairs, **61**  
Tyler, L. K., 78, 362  
type 0 grammar, 536  
Type hierarchy, **527**  
type hierarchy  
    example of for **agr**, 528  
    example of use for subcategorization frames, 529  
typed feature structures, 527  
    appropriateness conditions for, 527  
    atomic types, **527**  
    complex types, 527  
    fail type, **527**  
    simple types, 527  
    subtype, **527**  
    what good are they anyhow?, 527  
Types, **527**  
types  
    word, **86**  
Typology, **877**  
typology  
    linguistic, **877**  
Tyson, M., 766, 767, 774  
UCREL, 130  
Ueffing, N., 920  
uh  
    filled pause, 361, 422, **422**  
uh as filled pause, 423  
Ullman, J. D., 38, 39, 42, 81, 395, 434, 461, 462, 538, 539, 541, 549  
um  
    filled pause, **422**  
um as filled pause, 423  
Umeda, N., 284  
UMLS, 792  
unaspirated sound, **226**  
ungrammatical sentences, **395**  
unification, 497  
    [], 498  
    algorithm, 513  
    grammar, 501  
    negation in, **530**  
    path inequalities in, **530**  
    set-valued features in, **530**  
    subsumption in, **500**  
unification grammars, 404  
union, 40  
Unique beginner, **631**  
unit production, **441**  
Unit selection synthesis, **278**  
Universal, **842**  
universal of language, **877**  
UNIX, 17  
 $\text{UNK}_i$ , **95**  
unknown word  
    penalty in MT, **919**  
unknown words  
    in N-grams, **95**  
    in POS tagging, 158  
    open and closed-class words, **159**  
    use of capitalization, **159**  
    use of morphology, **158**  
    use of MaxEnt, **159**  
unrestricted, 536  
unvoiced sound, **219**  
upper tape, 60  
User generated content, **775**  
user initiative, **842**  
user-centered design, 850  
Usher, A. P., iv  
Uszkoreit, H., 532  
Utsuro, T., 775  
Utterance, **85**, **421**, **829**  
utterance, **85**, **421**, 426, **829**  
V (vocabulary), **164**  
Vagueness, **555**  
vagueness  
    tests to distinguish from ambiguity, 555  
Valence, **510**  
Valencia, A., 772

- Valtchev, V., 316, 341, 343  
 Value iteration, **863**  
 van Benthem, J., 587  
 van Deemter, K., 709  
 Van Ess-Dykema, C., 857, 858, 871  
 van Halteren, H., 170, 824  
 van Harmelen, F., 584  
 van Lehn, K., 619  
 van Noord, G., 387  
 van Rijsbergen, C. J., 459, 483, 744, 822  
 van Santen, J. P. H., 39, 270, 274, 285  
 van Vuuren, S., 9  
 van Wijnagaarden, A., 10, 427  
 van Zaanen, M., 492  
 Vanderwende, L., 620, 810, 823  
 VanLehn, K., 827  
 van den Berg, M., 734  
 van den Bosch, A., 284, 378, 379, 381  
 Van Son, R. J. J. H., 118  
 Van Valin, Jr., 428  
 Vapnik, V. N., 13, 213  
 Várad, T., 887  
 Varges, S., 871  
 Variable, **562**  
 variable  
   existentially quantified, 565  
   universally quantified, 565  
 variable-length N-gram, **113**  
 variables  
   in FOL, **562**  
 Variance, **308**  
 variation  
   sociolinguistic, **229**  
 Vasilescu, F., 659  
 Vasserman, A., 484, 491, 532  
 Vauquois triangle, **881**  
 Vauquois, B., 10, 427, 882  
 Veale, T., 646  
 Veblen, T., 72  
 Vector length, **784**  
 vector length, **677, 784**  
 vector model, 822  
 Vector quantization, **305**  
 Vector space model, **782**  
 vector space model, 779, **782**, 821  
 Veenstra, J., 458  
 Veilleux, N., 264  
 velar sound, **221**  
 Veldhuis, R., 276  
 Velichko, V. M., 333  
 velocity feature, **304**  
 velum, **221**  
 Venditti, J. J., 268  
 Vendl, Z., 575  
 Venkataraman, A., 335  
 Venkataramani, V., 243  
 Venugopal, A., 920  
 verb, **125**  
   modal, **128, 406**  
   phrasal, **126**  
 Verb alternation, **635**  
 Verb group, **407**  
 verb phrase, **393, 404**  
   semantic attachments, 611  
 verb-framed language, **878**  
 verbs  
   copula, **128**  
   irregular, 49  
   irregular, number in English, 49  
   main verb, 49  
   modal verb, 49  
   primary verbs, 49  
   subclasses, 125  
 Verhagen, M., 762, 775  
 Verifiability, **554**  
 Vermeulen, P. J. E., 850  
 Veronis, J., 689  
 Véronis, J., 652  
 Verspoor, C. M., 773  
 vertex, 27  
 vertices in directed graphs, 27  
 Vidal, E., 380  
 Viegas, E., 645  
 Vieira, R., 733  
 Vietnamese, 877  
 Vijay-Shanker, K., 151, 538  
 Vilain, M., 724  
 Vintsyuk, T. K., 146, 212, 213, 333  
 Vitale, A. J., 256  
 Viterbi  
   decoding in MEMM, 210  
   for semantics in dialog, 839  
   trellis, 148  
 Viterbi algorithm, 74, **184**  
   and stack decoder, 344  
   backtrace in, **186**  
   exercises modifying, 336  
   for letter-to-sound, 261  
   for POS tagging, **145**  
   history of, 213  
   in HMM alignment in MT, 899  
   in unit-selection synthesis, 280  
   limitations, 338  
 Viterbi approximation, **338**  
 Viterbi backtrace, **186**  
 Viterbi training, **329**  
 Viterbi, A. J., 213  
 VITERBI ALGORITHM, 147, 186, 323  
 vocabulary  
   growth as function of token count, 86  
   vocal  
     cords, *see* vocal folds  
     folds, **218**  
     tract, **219**  
   vocal cords, *see* vocal folds  
   vocal folds, **218**  
   vocal tract, **219**  
   Vogel, I., 285  
   Vogel, S., 920  
   Voice, **275**  
   Voice talent, **274**  
   voiced sound, **219**  
   voiceless sound, **219**  
   VoiceXML, **846**  
   Voiers, W., 282  
   Volkmann, J., 235, 301  
   von Kempelen, W., 249  
   von Neumann, J., 42  
   Voorhees, E. M., 13, 651, 790  
   Vossen, P., 688  
   Voutilainen, A., 135, 137–139, 155, 418, 419, 428  
   vowel, **219**  
     back, **222**  
     front, **222**  
     harmony, **369**  
     height, **222**  
     high, **222**  
     low, **222**  
     mid, **222**  
     reduced, **224**  
     rounded, **222**  
   vowel harmony, **369**  
   VP attachment, **473**  
   VQ (vector quantization), **305**  
   VSO language, **877**  
   VTLN, **357**  
   VUI, **850**  
   vxml, **846**  
  
   W. A., 769, 771  
   Wade, E., 850, 858  
   Wagner, R. A., 74, 170, 213  
   Wagstaff, K., 717, 733  
   Wahlster, W., 584  
   Waibel, A., 357, 871, 920  
   Wakao, T., 766  
   Waksler, R., 78  
   Wald, B., 230  
   Walker, K., 334  
   Walker, M. A. *et al.*, 825  
   Walker, M. A., 718, 733, 840, 842, 844, 845, 851, 852, 854, 860, 863, 871  
   Wall Street Journal  
     speech recognition of, 334  
   Wall, R. E., 16, 541, 542, 549, 550, 587

- Wallace, D. L., 11, 121, 139, 290  
Waltz, D. L., 646  
Wang, M. Q., 264, 364, 850  
Wang, W., 359, 364  
Wang, W. S.-Y., 283  
Wang, Y. Y., 920  
Wang, Z., 357  
Wanner, E., 547  
WANT() in BDI model, 865  
Ward, N., 832, 920  
Ward, T., 910, 921  
Ward, W., 356, 358, 690, 837  
Warnke, V., 871  
Warping, **333**  
warping, **333**  
Warren, D. H. D., 11, 531  
Warren, R. M., 362  
Warwick, S., 920  
Wasow, T., 428, 511, 530  
Wattarujeekrit, T., 773  
Waugh, L. R., 879  
wavefile format, 233  
waveform synthesis (stage of TTS), 250  
weak equivalence  
of grammars, 416  
Weaver, W., 652, 687, 919  
Web Ontology Language, **584**  
web site for book, xxv  
Webber, B. L., 13, 16, 532, 708, 709, 732, 734  
Weber, D. J., 80  
Weber, S., 870  
Weeds, J., 670, 689, 690  
Weenink, D., 236, 244, 247  
Wegmann, S., 364  
Wegstein, J. H., 10, 427  
weighted, **144**  
weighted finite-state automaton, 144, 174  
weighted FSA, **174**  
weighted FST, **378**  
weighted mutual information, **689**  
weighted transducer, **378**  
Weinschenk, S., 846  
Weinstein, C., 359  
Weinstein, J. N., 772  
Weinstein, S., 695, 718, 733  
Weintraub, M., 339, 340, 357  
Weir, D. J., 538, 689  
Weischedel, R., 150, 159, 170, 482, 688, 766, 775, 839  
Weiss, S. F., 382  
Weizenbaum, J., 7, 17, 25  
Weld, D. S., 690, 754, 755  
well-formed substring table, 461  
well-formedness constraint, 386  
Welling, L., 364  
Wellner, B., 771  
Wells, J. C., 224, 247  
Welsh, A., 362, 363  
Weng, F., 871  
Wessels, L. F. A., 850  
Wester, M., 364  
WFST, **461**  
Wh, **128**  
wh-non-subject question, **397**  
wh-non-subject-question, **511**  
semantic attachments, 607  
wh-phrase, 397, **397**  
wh-pronoun, **128**  
wh-questions, 396, 426  
wh-subject-questions, **397**  
semantic attachments, 607  
Wh-word, **397**  
wh-word, **397**  
Wheeler, D. W., 380, 428  
White, E. K., 771  
Whiteside, J. A., 850  
Whittaker, E. W. D., 111  
Whittaker, S., 842, 844, 852, 854, 871  
Whitton, L., 266  
Wicentowski, R., 384  
Wideband, **297**  
wideband speech, **232**  
Wiebe, J., 651, 775  
Wiemer-Hastings, P., 698, 733  
Wierzbicka, A., 641, 643, 646  
Wightman, C. W., 267, 284  
Wilcoxon signed-rank test, **156**  
wildcard  
regular expression, **20**  
Wilde, O., 72  
Wilensky, R., 11, 620, 864, 870  
Wilkes-Gibbs, D., 871  
Wilks, Y., 11, 587, 620, 646, 660, 688, 689, 766, 921  
Willett, P., 822  
Williams, C. E., 282  
Williams, E., 359  
Williams, J., 690  
Williams, J. D., 871  
Williams, R., 620, 774  
Wilson, C., 387  
Wilson, G., 756, 759  
Wilson, R., 827  
Winnnow algorithm, 170  
Winograd, T., 11, 585–587, 725, 732, 735, 825, 836  
Wirén, M., 837  
Wise, B., 9  
Witten, I. H., 120, 213, 482  
Wittgenstein, L., 829  
Wixon, D. R., 850  
Wizard-of-Oz, **850**  
Wolf, A. K., 586, 823  
Wolf, F., 359, 734  
Wolfram, W. A., 229  
Wong, A. K. C., 213  
Wong, W., 920  
Woodger, M., 10, 427  
Woodland, P. C., 111, 113, 281, 316, 341, 343, 348–352, 354, 356–358, 364  
Woods, W. A., 12, 586, 587, 608, 619, 732, 823  
Wooters, C., 90, 487  
word  
boundaries  
regular expression notation for, 21  
classes, *see* part-of-speech, **123**  
closed class, **124**  
count in Shakespeare, 86  
definition of, 84  
error  
evaluation for speech recognition, **330**  
fragment, **85**, 422  
function, **125**, 168  
how many in English, 86  
lattice, **340**  
open class, **124**  
prediction, **83**  
punctuation as, 85  
tokens, **86**  
types, **86**  
types in Hungarian, Turkish and English, 161  
Word alignment, **894**  
word alignment, **894**  
in MT, **893**  
Word association, **362**  
word boundary, **61**  
Word error, **330**  
word error rate, 288  
Word Grammar, 428  
word graph in ASR, **342**  
Word insertion penalty, **317**  
word insertion penalty, **317**  
word penalty  
in MT, **919**  
word relatedness, **665**  
word segmentation, **69**  
Word sense, **624**  
Word sense disambiguation, **556**  
word sense disambiguation, **650**, *see* WSD, *see* WSD  
Word senses, **556**  
word similarity, **664**  
word tokenization, **69**  
Wordform, **85**  
wordform  
and lemma, **623**  
versus lemma, 85  
WordNet, **629**, 630, 809  
Wordnet::Similarity, 669, 691

- world knowledge, 551  
 Woszczyna, M., 871  
 Wouters, J., 282  
 WOZ, *see* Wizard-of-Oz system  
 Wright, H., 871  
 Wright, J. H., 826  
 Wright, R. N., 428  
 WSD, 649, **650**, 687  
   AI-oriented efforts, 688  
   all-words task, **651**  
   bootstrapping, **662**, 688  
   connectionist approaches, 688  
   decision list approach, 653, 655  
   decision tree approach, 688  
   evaluation of, 656  
   history of, 689  
   in IR, 790  
   in Machine Translation, 876  
   lexical sample task, **651**  
   limitations of selectional restrictions, 660  
   naive Bayes approach, 653  
   robust approach, 688  
   supervised machine learning, 688  
   unsupervised machine learning, 686  
   use of bag-of-words features, **652**  
   use of collocational features, **652**  
   use of machine readable dictionaries, 688, 689  
   use of selectional association, 661  
 Wu, D., 620, 689, 914, 920  
 Wu, J., 163  
 Wu, Z., 666  
 Wundt, W., 391, 426  
 Wunsch, C. D., 213  
  
 X-bar schemata, **427**  
  
 X-schema, **585**  
 Xia, F., 419  
 Xiang, B., 359  
 Xiuyang, Y., 358  
 Xu, K., 871  
 Xu, P., 487  
 Xu, W., 826  
 Xue, N., 81, 690  
  
 Y., 10, 247  
 Yale School of AI, 11  
 Yallop, C., 247  
 Yamada, H., 491  
 Yamada, K., 883, 886, 912, 914, 920  
 Yamada, T., 285  
 Yang, B., 792  
 Yang, Y., 250  
 Yankelovich, N., 829, 832, 846, 850  
 Yarowsky algorithm, **662**  
 Yarowsky, D., 155, 171, 257, 384, 655–657, 662, 663, 688  
 Yates, A., 690, 754, 755  
 Yates, J., 730  
 Yawelmani, 369, 371  
 Yeh, A., 771  
 yes-no question, **397**  
 yes-no-questions, 396, 403, 426, 855  
   semantic attachments, 607  
 Yi, L., 243  
 Yield, **467**  
 Yngve, V. H., 434, 461, 542, 546, 547, 832  
 Yokuts, 369  
 Yonkers Racetrack, 114  
 Yoon, S.-Y., 357  
 Yoshimura, T., 285  
 Young, M., 529  
 Young, S. J., 316, 335, 339, 343, 348–350, 352, 364, 471, 492, 871  
  
 Younger, D. H., 431, 461  
 Yuan, J., 266  
 Yupik, 877  
 Yuret, D., 492, 660  
 Yvon, F., 284  
  
 Z  
   normalization factor in  
     MaxEnt, 201  
 Zacharski, R., 712, 733  
 Zaenen, A., 532, 823  
 Zagoruyko, N. G., 333  
 Zajic, D., 810, 823  
 Zamora, A., 823  
 Zappa, F., 216  
 Zappa, M. U., 216  
 Zavrel, J., 162, 170, 491  
 Zelenko, D., 170  
 Zelle, J., 689  
 Zen, H., 285  
 Zens, R., 920  
 Zernik, U., 620, 689, 823  
 zeugma, **626**  
 Zhang, H., 914  
 Zhang, J., 748  
 Zhang, M., 748  
 Zhang, Q., 871  
 Zhang, Z., 823  
 Zhao, S., 690  
 Zheng, F., 243  
 Zheng, Y., 357  
 Zhou, G., 113, 748  
 Zhou, H., 357  
 Zhou, J., 357  
 Zhou, M., 690  
 Zhu, W.-J., 910, 921  
 Zhu, X., 112  
 Ziff-Davis corpus, 807, 809  
 Zimak, D., 462  
 Zissman, M., 359  
 Zue, V. W., 243, 850, 851, 871  
 Zweig, G., 364  
 Zwicky, A., 229, 555