Glossary

ACCOR: a language corpus
ACH: Association For Computing In The Humanities
ACL: Association For Computational Linguistics
ACL/DCI: Association For Computational Linguistics' Data Collection Initiative
ADPCM: adaptive differential pulse-code modulation (PCM)
AI: Artificial Intelligence; also Articulation Index
ALBAYZIN: a language corpus
ALLC: Association for Literary and Linguistic Computing
AM: acoustic model
ANDOSL: Australian National Database of Spoken Language
ANN: artificial neural network, also NN
AR: autoregressive (model)
ARPA: Advanced Research Projects Agency (U.S.), predecessor of DARPA
ARS: a language corpus
ASCII: American Standard Code for Information Interchange
ASR: automatic speech recognition
ASR: automatic speech recognition
ATIS: Air Travel Information Service (task)
ATLAS II: Japanese-English translation system by Fujitsu
ATM: asynchronous transfer mode
ATR: a language corpus

BMFT: Bundesministerium für Forschung und Technik, German Federal Ministry for Research and Technology
BNC: British National Corpus
bps: bits per second
BRA: Basic Research Actions
BRAMSHILL: a language corpus
BREF: a language corpus

C_0: first cepstral coefficient
CAD: Computer-Assisted (or -Aided) Design
CADCAM: Computer-Assisted Design / Computer-Assisted Manufacture
CAR: a language corpus
CART: Classification and Regression Tree
CCG: combinatory categorial grammar
CD-ROM: compact disk—read-only memory
CELP: code-excited linear prediction
CFE: Caterpillar Fundamental English
CFG: context-free grammar
CKY: Cocke, Kasami, and Younger (algorithm)
CL: controlled language
CLAWS: Constituent-Likelihood Automatic Word-Tagging System
CLE: core language engine
CLR: Consortium for Lexical Research
CMN: cepstral mean normalization
CMU: Carnegie Mellon University
CPU: central processing unit (of a computer)
CSC: character shape code
CSLU: Center for Spoken Language Understanding
CSR: continuous speech recognition
C-STAR: Consortium for Speech TrAnslation Research
CUG: categorial unification grammar
CV: consonant-vowel phoneme sequence
CVC: consonant-vowel-consonant phoneme sequence

DARPA: Defense Advanced Research Projects Agency (U.S.), successor to ARPA
dB: deciBel
DBMS: data base management system
DBMT: dialogue-based machine translation (MT)
DK: Dansk Korpus
DP: dynamic programming
DPCM: differential pulse-code modulation (PCM)
dpi: dots per inch (resolution quality)
DRT: discourse representation theory
DSP: digital signal processing
DTW: dynamic time warping
DoD: Department of Defense (U.S.)

EAGLES: Expert Advisory Group on Linguistic Engineering Standards

EBMT: example-based machine translation (MT)
ECI: European Corpus Initiative
EDI: electronic data interchange
EDR: electronic dictionary research
ELRA: European Language Resources Association
ELSNET: European Network In Language And Speech
EM: expectation maximization algorithm, also estimate-maximize
EMIR: European Multilingual Information Retrieval
ESCA: European Speech Communication Association
ESPRIT: European Strategic Programme for Research and Development in Information Technology
ETL: Electrotechnical Laboratory (Japan)

F_0: Fundamental Frequency
FIR: finite impulse response filter
FOPC: first order predicate calculus
FSA: finite state automaton
FSM: finite state machine
FST: finite state transducer
FUF: functional unification grammar framework
FUG: functional unification grammar

GIST: Generating InStructional Text
GPD: generalized probabilistic descent
GPSG: generalized phrase structure grammar

HAMT: human-aided machine translation (MT)
HCI: human-computer interaction
HG: head grammar
HLT: human language technology
HMM: hidden Markov model
HPSG: head-driven phrase-structure grammar
Hz: Hertz (cycles per second)

ICAME: International Computer Archive of Modern English
ICASSP: International Conference on Acoustics, Speech, and Signal Processing
ICR: intelligent character recognition
ICSLP: International Conference on Spoken Language Processing
IDS: Institut für deutsche Sprache
IIR: infinite impulse response filter

ILC-CNR: Istituto di Linguistica Computazionale
ILSAM: International Language for Serving and Maintenance
INaLF-CNRS: Frantext of Institut National de la Langue Francaise
INL: Instituut voor Nederlandse Lexicologie
Interlingua: language independent semantic representation
IR: information retrieval
ISDN: integrated services digital network
ISO: International Standards Organization

JEIDA: Japan Electronic Industry Development Association
J-RASTA: Relative Spectra PLP with Jah smoothing

LC: Lambek Calculus
LCD: liquid crystal display
LDA: linear discriminant analysis
LDC: Linguistic Data Consortium
LD-CELP: low-delay code-excited linear prediction (CELP)
LE: Language Engineering
LE-MLAP: Language Engineering Multilingual Action Plan
LFG: lexical-Functional grammar
LIG: linear indexed grammar
LKB: lexical knowledge base
LM: language model
LPC: linear predictive coding
LR: Left-to-right scanning, rightmost derivation parsing
LRE: Linguistic Research and Engineering
LV-CSR: large vocabulary continuous speech recognition

MAHT: Machine-Aided Human Translation, see also HAMT
MAP: maximum a-posteriori probability
MAR: multivariate auto-regression (AR)
MAT: machine-assisted translation
METEO: Canadian meteorological bulletin translator
MFCC: mel-frequency cepstral coefficient
MIT: Massachusetts Institute of Technology
MITalk: Dennis Klatt's MIT speech synthesizer
MLDB: multilingual lexical data base
MLP: multi-layer perceptron (ANN)
MMSE: minimum mean square error
MMSEE: minimum mean square error estimator

MOS: mean opinion score/scoring
MRD: machine-readable dictionary
MSR: multiple split regression
MT: machine translation
MTDB: multilingual terminological data base
MUC: Message Understanding Conference

NIST: National Institute of Standards and Technology (U.S.)

OGI: Oregon Graduate Institute of Science & Technology
OGI_TS: OGI Multilanguage Telephone Speech Corpus

PACE: Perkins Approved Clear English
PAHO: Pan American Health Organization
PCA: principal components analysis
PCFG: probabilistic context-free grammar
PCM: pulse-code modulation
PDA: personal digital assistant
PDL: page description language
PECO: Pays d'Europe Centrale et Orientale
PEP: Plain English Program
PLP: perceptual linear prediction

RASTA: relative spectra perceptual linear prediction (PLP)
RELATOR: a language corpus
RM: Resource Management task
RPS: root power sum
RST: rhetorical structure theory

SAM: Speech Technology Assessment Methodology
SAM-A: Speech Technology Assessment Methodology in Multilingual Applications
SGML: Standard Generalized Markup Language
SLS: spoken language system
SLU: spoken language understanding
SNR: signal-to-noise ratio
SR: speech recognition, also ASR
SRI: Stanford Research Institute
SSC: speech synthesis from concept
STI: speech transmission index
SUC: Stockholm-Umea Corpus

SUNDIAL: a language corpus
SUNSTAR: a language corpus
SYSTRAN: SYSTRAN language translation system

TAG: tree adjunction grammar
TED: Translanguage English Database
TEI: Text Encoding Initiative
TES: Telephone Enquiry System
TFS: typed feature structure
TI: Texas Instruments
TIDE: Telematics Initiative for Disabled and Elderly people, sponsored by the European Commission
TIF: terminology interchange format
TIMIT: Texas Instruments—MIT speech corpus
TKE: terminological knowledge engineering
TREC: Text Retrieval Evaluation Conference
TtS: Text-to-Speech

UKA: University of Karslruhe
USPS: U.S. Postal Service
UW: University of Washington (Seattle)

VC: vowel-consonant phoneme sequence
VERBMOBIL: a european language translation project
VLSI: very large scale integration (microchip technology)
VQ: vector quantization

WOZ: Wizard of Oz (data collection method)
WSJ: Wall Street Journal corpus
WSJCAM0: Wall Street Journal, Cambridge, language corpus zero
WST: word shape token
WWB: Writer's Workbench
WWW: World Wide Web

XUNET: Xperimental University NETwork

ZIP: Zone Improvement Code (U.S. Postal Service)

Citation Index

Index