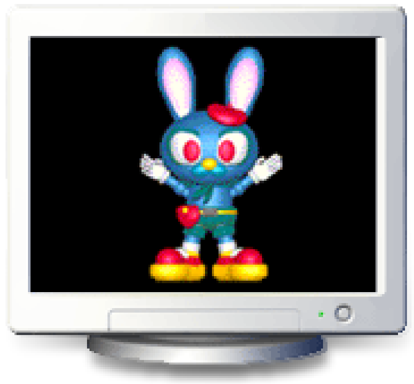
Speech recognition

For the human linguistic concept, see Speech perception

Speech recognition (also known as **automatic speech recognition** or **computer speech recognition**) converts spoken words to text. The term "voice recognition" is sometimes used to refer to recognition systems that must be trained to a particular speaker—as is the case for most desktop recognition software. Recognizing the speaker can simplify the task of translating speech.



The display of the Speech Recognition screensaver on a laptop, in which the character responds to questions, e.g. "Where are you?" or statements, e.g. "Hello."

Speech recognition is a broader solution which refers to technology that can recognize speech without being targeted at single speaker—such as a call center system that can recognize arbitrary voices.

Speech recognition applications include voice user interfaces such as voice dialing (e.g., "Call home"), call routing (e.g., "I would like to make a collect call"), domotic appliance control, search (e.g., find a podcast where particular words were spoken), simple data entry (e.g., entering a credit card number), preparation of structured documents (e.g., a radiology report), speech-to-text processing (e.g., word processors or emails), and aircraft (usually termed Direct Voice Input).

History

The first speech recognizer appeared in 1952 and consisted of a device for the recognition of single spoken digits [1] Another early device was the IBM Shoebox, exhibited at the 1964 New York World's Fair. Lately there have been numerous improvements like a high speed mass transcription capabilitiy on a single system like Sonic Extractor [2]

One of the most notable domains for the commercial application of speech recognition in the United States has been health care and in particular the work of the medical transcriptionist (MT). According to industry experts, at its inception, speech recognition (SR) was sold as a way to completely eliminate transcription rather than make the transcription process more efficient, hence it was not accepted. It was also the case that SR at that time was often technically deficient. Additionally, to be used effectively, it required changes to the ways physicians worked and documented clinical encounters, which many if not all were reluctant to do. The biggest limitation to speech recognition automating transcription, however, is seen as the software. The nature of narrative dictation is highly interpretive and often requires judgment that may be provided by a real human but not yet by an automated system. Another limitation has been the extensive amount of time required by the user and/or system provider to train the software.

A distinction in ASR is often made between "artificial syntax systems" which are usually domain-specific and "natural language processing" which is usually language-specific. Each of these types of application presents its own particular goals and challenges.

Applications

Health care

In the health care domain, even in the wake of improving speech recognition technologies, medical transcriptionists (MTs) have not yet become obsolete. The services provided may be redistributed rather than replaced.

Speech recognition can be implemented in front-end or back-end of the medical documentation process.

Front-End SR is where the provider dictates into a speech-recognition engine, the recognized words are displayed right after they are spoken, and the dictator is responsible for editing and signing off on the document. It never goes through an MT/editor.

Back-End SR or Deferred SR is where the provider dictates into a digital dictation system, and the voice is routed through a speech-recognition machine and the recognized draft document is routed along with the original voice file to the MT/editor, who edits the draft and finalizes the report. Deferred SR is being widely used in the industry currently.

Many Electronic Medical Records (EMR) applications can be more effective and may be performed more easily when deployed in conjunction with a speech-recognition engine. Searches, queries, and form filling may all be faster to perform by voice than by using a keyboard.

Military

High-performance fighter aircraft

Substantial efforts have been devoted in the last decade to the test and evaluation of speech recognition in fighter aircraft. Of particular note are the U.S. program in speech recognition for the Advanced Fighter Technology Integration (AFTI)/F-16 aircraft (F-16 VISTA), the program in France on installing speech recognition systems on Mirage aircraft, and programs in the UK dealing with a variety of aircraft platforms. In these programs, speech recognizers have been operated successfully in fighter aircraft with applications including: setting radio frequencies, commanding an autopilot system, setting steer-point coordinates and weapons release parameters, and controlling flight displays.

Working with Swedish pilots flying in the JAS-39 Gripen cockpit, Englund (2004) found recognition deteriorated with increasing G-loads. It was also concluded that adaptation greatly improved the results in all cases and introducing models for breathing was shown to improve recognition scores significantly. Contrary to what might be expected, no effects of the broken English of the speakers were found. It was evident that spontaneous speech caused problems for the recognizer, as could be expected. A restricted vocabulary, and above all, a proper syntax, could thus be expected to improve recognition accuracy substantially.[3]

The Eurofighter Typhoon currently in service with the UK RAF employs a speaker-dependent system, i.e. it requires each pilot to create a template. The system is not used for any safety critical or weapon critical tasks, such as weapon release or lowering of the undercarriage, but is used for a wide range of other cockpit functions. Voice commands are confirmed by visual and/or aural feedback. The system is seen as a major design feature in the reduction of pilot workload, and even allows the pilot to assign targets to himself with two simple voice commands or to any of his wingmen with only five commands.[4]

Speaker independent systems are also being developed and are in testing for The F35 Lightning II (JSF) and the Aermacchi M346 lead in fighter trainer. These systems have produced word accuracies in excess of 98%.

Helicopters

The problems of achieving high recognition accuracy under stress and noise pertain strongly to the helicopter environment as well as to the fighter environment. The acoustic noise problem is actually more severe in the helicopter environment, not only because of the high noise levels but also because the helicopter pilot generally does not wear a facemask, which would reduce acoustic noise in the microphone. Substantial test and evaluation programs have been carried out in the past decade in speech recognition systems applications in helicopters, notably by the U.S. Army Avionics Research and Development Activity (AVRADA) and by the Royal Aerospace Establishment (RAE) in the UK. Work in France has included speech recognition in the Puma helicopter. There has also been much useful work in Canada. Results have been encouraging, and voice applications have included: control of communication radios; setting of navigation systems; and control of an automated target handover system.

As in fighter applications, the overriding issue for voice in helicopters is the impact on pilot effectiveness. Encouraging results are reported for the AVRADA tests, although these represent only a feasibility demonstration in a test environment. Much remains to be done both in speech recognition and in overall speech recognition technology, in order to consistently achieve performance improvements in operational settings.

Battle management

Battle Management command centres generally require rapid access to and control of large, rapidly changing information databases. Commanders and system operators need to query these databases as conveniently as possible, in an eyes-busy environment where much of the information is presented in a display format. Human-machine interaction by voice has the potential to be very useful in these environments. A number of efforts have been undertaken to interface commercially available isolated-word recognizers into battle management environments. In one feasibility study speech recognition equipment was tested in conjunction with an integrated information display for naval battle management applications. Users were very optimistic about the potential of the system, although capabilities were limited.

Speech understanding programs sponsored by the Defense Advanced Research Projects Agency (DARPA) in the U.S. has focused on this problem of natural speech interface. Speech recognition efforts have focused on a database of continuous speech recognition (CSR), large-vocabulary speech which is designed to be representative of the naval resource management task. Significant advances in the state-of-the-art in CSR have been achieved, and current efforts are focused on integrating speech recognition and natural language processing to allow spoken language interaction with a naval resource management system.

Training air traffic controllers

Training for air traffic controllers (ATC) represents an excellent application for speech recognition systems. Many ATC training systems currently require a person to act as a "pseudo-pilot", engaging in a voice dialog with the trainee controller, which simulates the dialog which the controller would have to conduct with pilots in a real ATC situation. Speech recognition and synthesis techniques offer the potential to eliminate the need for a person to act as pseudo-pilot, thus reducing training and support personnel. In theory, Air controller tasks are also characterized by highly structured speech as the primary output of the controller, hence reducing the difficulty of the speech recognition task should be possible. In practice this is rarely the case. The FAA document 7110.65 details the phrases that should be used by air traffic controllers. While this document gives less than 150 examples of such phrases, the number of phrases supported by one of the simulation vendors speech recognition systems is in excess of 500,000.

The USAF, USMC, US Army,US Navy and FAA as well as a number of international ATC training organizations such as Air Services Australia, Royal Australian Air Force and Civil Aviation Authorities in Italy, Brazil, Canada are currently using ATC simulators with speech recognition from a number of different vendors.

Telephony and other domains

ASR in the field of telephony is now commonplace and in the field of computer gaming and simulation is becoming more widespread. Despite the high level of integration with word processing in general personal computing, however, ASR in the field of document production has not seen the expected increases in use.

The improvement of mobile processor speeds made feasible the speech-enabled Symbian and Windows Mobile Smartphones. Speech is used mostly as a part of User Interface, for creating pre-defined or custom speech commands. Leading software vendors in this field are: Microsoft Corporation (Microsoft Voice Command), Nuance Communications (Nuance Voice Control), Vito Technology (VITO Voice2Go), Speereo Software (Speereo Voice Translator), Digital Syphon [5](Sonic Messager appliance) and SVOX.

Further applications

- Automatic translation;
- Automotive speech recognition (e.g., Ford Sync);
- Telematics (e.g. vehicle Navigation Systems);
- Court reporting (Realtime Voice Writing);
- Hands-free computing: voice command recognition computer user interface;
- Home automation;
- Interactive voice response;
- Mobile telephony, including mobile email;
- Multimodal interaction;
- Pronunciation evaluation in computer-aided language learning applications;
- Robotics;
- Video games, with Tom Clancy's EndWar and Lifeline as working examples;
- Transcription (digital speech-to-text);
- Speech-to-text (transcription of speech into mobile text messages);
- Air Traffic Control Speech Recognition [6].

Performance

The performance of speech recognition systems is usually specified in terms of accuracy and speed. Accuracy is usually rated with word error rate (WER), whereas speed is measured with the real time factor. Other measures of accuracy include Single Word Error Rate (SWER) and Command Success Rate (CSR).

In 1982, Kurzweil Applied Intelligence and Dragon Systems released speech recognition products. By 1985, Kurzweil’s software had a vocabulary of 1,000 words—if uttered one word at a time. Two years later, in 1987, its lexicon reached 20,000 words, entering the realm of human vocabularies, which range from 10,000 to 150,000 words. But recognition accuracy was only 10% in 1993. Two years later, the error rate crossed below 50%. Dragon Systems released "Naturally Speaking" in 1997 which recognized normal human speech. Progress mainly came from improved computer performance and larger source text databases. The Brown Corpus was the first major database available, containing several million words. In 2001, recognition accuracy reached its current plateau of 80%, no longer growing with data or computing power. In 2006, Google published a trillion-word corpus, while Carnegie Mellon University researchers found no significant increase in recognition accuracy.[7]

"

Algorithms

Both acoustic modeling and language modeling are important parts of modern statistically-based speech recognition algorithms. Hidden Markov models (HMMs) are widely used in many systems. Language modeling has many other applications such as smart keyboard and document classification.

Hidden Markov models

Modern general-purpose speech recognition systems are based on Hidden Markov Models. These are statistical models which output a sequence of symbols or quantities. HMMs are used in speech recognition because a speech signal can be viewed as a piecewise stationary signal or a short-time stationary signal. In a short-time (e.g., 10 milliseconds)), speech can be approximated as a stationary process. Speech can be thought of as a Markov model for many stochastic purposes.

Another reason why HMMs are popular is because they can be trained automatically and are simple and computationally feasible to use. In speech recognition, the hidden Markov model would output a sequence of n-dimensional real-valued vectors (with n being a small integer, such as 10), outputting one of these every 10 milliseconds. The vectors would consist of cepstral coefficients, which are obtained by taking a Fourier transform of a short time window of speech and decorrelating the spectrum using a cosine transform, then taking the first (most significant) coefficients. The hidden Markov model will tend to have in each state a statistical distribution that is a mixture of diagonal covariance Gaussians which will give a likelihood for each observed vector. Each word, or (for more general speech recognition systems), each phoneme, will have a different output distribution; a hidden Markov model for a sequence of words or phonemes is made by concatenating the individual trained hidden Markov models for the separate words and phonemes.

Described above are the core elements of the most common, HMM-based approach to speech recognition. Modern speech recognition systems use various combinations of a number of standard techniques in order to improve results over the basic approach described above. A typical large-vocabulary system would need context dependency for the phonemes (so phonemes with different left and right context have different realizations as HMM states); it would use cepstral normalization to normalize for different speaker and recording conditions; for further speaker normalization it might use vocal tract length normalization (VTLN) for male-female normalization and maximum likelihood linear regression (MLLR) for more general speaker adaptation. The features would have so-called delta and delta-delta coefficients to capture speech dynamics and in addition might use heteroscedastic linear discriminant analysis (HLDA); or might skip the delta and delta-delta coefficients and use splicing and an LDA-based projection followed perhaps by heteroscedastic linear discriminant analysis or a global semitied covariance transform (also known as maximum likelihood linear transform, or MLLT). Many systems use so-called discriminative training techniques which dispense with a purely statistical approach to HMM parameter estimation and instead optimize some classification-related measure of the training data. Examples are maximum mutual information (MMI), minimum classification error (MCE) and minimum phone error (MPE).

Decoding of the speech (the term for what happens when the system is presented with a new utterance and must compute the most likely source sentence) would probably use the Viterbi algorithm to find the best path, and here there is a choice between dynamically creating a combination hidden Markov model which includes both the acoustic and language model information, or combining it statically beforehand (the finite state transducer, or FST, approach).

Dynamic time warping (DTW)-based speech recognition

Dynamic time warping is an approach that was historically used for speech recognition but has now largely been displaced by the more successful HMM-based approach. Dynamic time warping is an algorithm for measuring similarity between two sequences which may vary in time or speed. For instance, similarities in walking patterns would be detected, even if in one video the person was walking slowly and if in another they were walking more quickly, or even if there were accelerations and decelerations during the course of one observation. DTW has been applied to video, audio, and graphics – indeed, any data which can be turned into a linear representation can be analyzed with DTW.

A well known application has been automatic speech recognition, to cope with different speaking speeds. In general, it is a method that allows a computer to find an optimal match between two given sequences (e.g. time series) with certain restrictions, i.e. the sequences are "warped" non-linearly to match each other. This sequence alignment method is often used in the context of hidden Markov models.

Further information

Popular speech recognition conferences held each year or two include SpeechTEK and SpeechTEK Europe, ICASSP, Eurospeech/ICSLP (now named Interspeech) and the IEEE ASRU. Conferences in the field of Natural language processing, such as ACL, NAACL, EMNLP, and HLT, are beginning to include papers on speech processing. Important journals include the IEEE Transactions on Speech and Audio Processing (now named IEEE Transactions on Audio, Speech and Language Processing), Computer Speech and Language, and Speech Communication. Books like "Fundamentals of Speech Recognition" by Lawrence Rabiner can be useful to acquire basic knowledge but may not be fully up to date (1993). Another good source can be "Statistical Methods for Speech Recognition" by Frederick Jelinek and "Spoken Language Processing (2001)" by Xuedong Huang etc. More up to date is "Computer Speech", by Manfred R. Schroeder, second edition published in 2004. The recently updated textbook of "Speech and Language Processing (2008)" by Jurafsky and Martin presents the basics and the state of the art for ASR. A good insight into the techniques used in the best modern systems can be gained by paying attention to government sponsored evaluations such as those organised by DARPA (the largest speech recognition-related project ongoing as of 2007 is the GALE project, which involves both speech recognition and translation components).

In terms of freely available resources, Carnegie Mellon University's SPHINX toolkit is one place to start to both learn about speech recognition and to start experimenting. Another resource (free as in free beer, not free software) is the HTK book (and the accompanying HTK toolkit). The AT&T libraries GRM library [8], and DCD library [9] are also general software libraries for large-vocabulary speech recognition.

For more software resources, see List of speech recognition software.

A useful review of the area of robustness in ASR is provided by Junqua and Haton (1995).

People with disabilities

People with disabilities can benefit from speech recognition programs. Speech recognition is especially useful for people who have difficulty using their hands, ranging from mild repetitive stress injuries to involved disabilities that preclude using conventional computer input devices. In fact, people who used the keyboard a lot and developed RSI became an urgent early market for speech recognition.[10] [11] Speech recognition is used in deaf telephony, such as voicemail to text, relay services, and captioned telephone. Individuals with learning disabilities who have problems with thought-to-paper communication (essentially they think of an idea but it is processed incorrectly causing it to end up differently on paper) can benefit from the software.

Current research and funding

Measuring progress in speech recognition performance is difficult and controversial. Some speech recognition tasks are much more difficult than others. Word error rates on some tasks are less that one percent. On others they can be as high as 50%. Sometimes it even appears that performance is going backwards as researchers undertake harder tasks that have higher error rates.

Because progress is slow and is difficult to measure, there is some perception that performance has plateaued and that funding has dried up or shifted priorities. Such perceptions are not new. In 1969, John Pierce wrote an open letter that did cause much funding to dry up for several years.[12] In 1993 there was a strong feeling that performance had plateaued and there were workshops dedicated to the issue. However, in the 1990s funding continued more or less uninterrupted and performance continued to slowly but steadily improve.

For the past thirty years, speech recognition research has been characterized by the steady accumulation of small incremental improvements. There has also been a trend to continually change focus to more difficult tasks due both to progress in speech recognition performance and to the availability of faster computers. In particular, this shifting to more difficult tasks has characterized DARPA funding of speech recognition since the 1980s. In the last decade it has continued with the EARS project, which undertook recognition of Mandarin and Arabic in addition to English, and the GALE project, which focused solely on Mandarin and Arabic and required translation simultaneously with speech recognition.

Commercial research and other academic research also continue to focus on increasingly difficult problems. One key area is to improve robustness of speech recognition performance, not just robustness against noise but robustness against any condition that causes a major degradation in performance. Another key area of research is focused on an opportunity rather than a problem. This research attempts to take advantage of the fact that in many applications there is a large quantity of speech data available, up to millions of hours. It is too expensive to have humans transcribe such large quantities of speech, so the research focus is on developing new methods of machine learning that can effectively utilize large quantities of unlabeled data. Another area of research is better understanding of human capabilities and to use this understanding to improve machine recognition performance.[13]

See also

- Audio mining
- Audio visual speech recognition
- Acoustic Model
- Digital dictation
- Direct Voice Input
- Keyword spotting
- List of speech recognition software
- Microphone
- Mondegreen
- Multimodal interaction
- OpenDocument
- Phonetic search technology
- Speech Analytics
- Speaker identification
- Speaker diarisation
- Speech corpus
- Speech processing
- Speech recognition in Linux
- Speech synthesis

- Speech verification
- Text-to-speech (TTS)
- VoiceXML
- Voxforge
- Windows Speech Recognition
- Speech technology

References

- Karat, Clare-Marie; Vergo, John; Nahamoo, David (2007). "Conversational Interface Technologies". In Sears, Andrew; Jacko, Julie A.. *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications (Human Factors and Ergonomics)*. Lawrence Erlbaum Associates Inc. ISBN 978-0805858709.
- Managing editors: Giovanni Battista Varile, Antonio Zampolli. (1997). Cole, Annie; Mariani, Joseph; Uszkoreit, Hans et al.. eds. *Survey of the state of the art in human language technology*. Cambridge Studies In Natural Language Processing. **XII–XIII**. Cambridge University Press. ISBN 0-521-59277-1.
- Junqua, J.-C.; Haton, J.-P. (1995). *Robustness in Automatic Speech Recognition: Fundamentals and Applications*. Kluwer Academic Publishers. ISBN 978-0792396468.

[1] Davies , K.H., Biddulph, R. and Balashek, S. (1952) *Automatic Speech Recognition of Spoken Digits*, J. Acoust. Soc. Am. **24**(6) pp.637 - 642
[2] http://www.digitalsyphon.com/services_sonicex.asp?contentpage=services_softsound&bodyid=services&services=services
[3] http://www.speech.kth.se/prod/publications/files/1664.pdf
[4] Eurofighter Direct Voice Input (http://www.eurofighter.com/et_as_vt_dv.asp)
[5] http://www.digitalsyphon.com
[6] http://supremis.co.uk
[7] "The History of Automatic Speech Recognition Evaluations at NIST" (http://www.itl.nist.gov/iad/mig/publications/ASRhistory/index.html). National Institute of Standards and Technology. May, 2009. . Retrieved May, 2010.
[8] http://www.research.att.com/projects/mohri/grm
[9] http://www.cs.nyu.edu/~mohri
[10] Speech recognition for disabled people (http://www.businessweek.com/1998/08/b3566022.htm)
[11] Friends international support group
[12] John Pierce (1969). "Whither Speech Recognition". *Journal of the Acoustical Society of America*.
[13] "Research Developments and Directions in Speech Recognition and Understanding, Part 1" (http://research.microsoft.com/pubs/80528/SPM-MINDS-I.pdf). MAY, 2009. . Retrieved May, 2010.

External links

- Speech Technology (http://www.dmoz.org/Computers/Speech_Technology/) at the Open Directory Project

Article Sources and Contributors

Speech recognition *Source*: http://en.wikipedia.org/w/index.php?oldid=397236669 *Contributors*: 144.132.75.xxx, 1of3, A chameleon, Acerperi, Alan Liefting, Alexgru, Allen3, Altermike, Amatulic, Andrewrost3241981, Andycjp, Angelofwisdom, Ap, Apersontelisma, Apoc2400, Argon233, Ark25, Arthchan2003, Askeladden2006, Axistive, Az1568, Bb blogger, Beetstra, Beland, BenFrantzDale, Bernd in Japan, Betterusername, BettyeKeyes, Bird2000, Bisharbn, Bisqwit, Blackm0re, Blarrrgy, Blomskoe, Bluemoose, BokicaK, Bollinger, Borgx, Brandon, Brion VIBBER, Burschik, CMBJ, Calltech, Cameron Scott, Captain-n00dle, CaseKid, Ceefour, Centrx, Chairboy, Charles Matthews, Chexum, Chocolateboy, Christopher Parham, Closedmouth, Cnilep, Colanderman, Conversion script, Crazycomputers, Damian Yerrick, Dancter, Danpovey, DarthBatrous, DavesPlanet, David Eppstein, David Levy, David556wiki, DeadEyeArrow, Dennishc, Denoir, DerHexer, DeweyQ, Diberri, Dmaccormick, Dogposter, Double.reed, Dougher, Dpm64, Dragice, Dratman, Drewnoakes, Dyson1, Dysprosia, ESkog, EagleOne, Ed Poor, ElectronicHealthRecords, Elementaltoo, Elf, EndyFr, Enriquefoster, Equendil, Erencexor, Erxnmedia, Everyking, Excirial, FarmBoi, Fca neo, Forcery, Fritzpoll, Frymaster, Fumitol, Furrykef, FutureNJGov, Gaius Cornelius, Garyreece, Gdeyoe, Gemuetlich, Gene s, Giftlite, Ginsengbomb, Glenn, Goblin, Goransobacic, Gp2it, GraemeL, Graham87, Grayshi, Gregory Benoit, Guaka, Gudeldar, Guidogallopyn, Guy M, Gökhan, Hakeem.gadi, HealthIT.EHR, HelenaRiutortEN, HendrixEesti, Hephaestos, Hike395, Hlaw, Howard Wright, II MusLiM HyBRiD II, Ian Burnet, Ianbolland, Ijablokov, Into The Fray, Ipatrol, Irishguy, Itamar.evenzohar, Izrunas, JaGa, Jarkko Piiroinen, Javandyck, Jay Page, Jenliv2ryd, Jh51681, Jhumritaliya, Jimbo4mo, Jlcaune, Jorfer, Jorge Stolfi, Joshsteiner, Jtneill, Jtnelson, Juli@n, Kcbnac, Kesaloma, Ketil3, Khatru2, Kiore, Kku, Kmaclean, Krellmachine, Ksevio, Ksyrie, KyraVixen, Labnics, Latka, Laudaka, Lawrennd, LeCire, Lerdsuwa, Lfstevens, Lilo197, Lordvolton, Lorenbaker, M.m.77, Mac, Madmedea, Magiciandude, Magnolia Grandiflora, Maranlar, Martarius, Martinevans123, Masdmh, Mcsee, Mdrejhon, Michael Hardy, Michaelmdx, Michal Jurosz, Midgley, Minghong, Mischling, Modster, Mr Adequate, Musically ut, NJR ZA, Nad, Nandesuka, Nanshu, Nastajus, Neelix, Neumannkun, Ngmarle, Ninjakannon, Nirelan, No1lakersfan, Nobuyuki, Nohat, Nolset10, Notinasnaid, Notopia, Nrcprm2026, OccamzRazor, Oda Mari, Offtherails, Ofus, Oldgrowyoung, Oleg Alexandrov, One-dimensional Tangent, Onorem, Orinoco-w, Orpheus, OwenX, Oxymoron83, Pakada, Pakcw, Paranoid, Pearle, Pgillman, Pgr94, Phdceng, Phi beta, Pilotguy, Pinethicket, Pmlineditor, Pol098, Praveen pillay, Puchiko, Purplebooooks, R'n'B, RJASE1, Rajah, Randomchance, Reconsider the static, Rjwilmsi, Rmallins, Roadmr, Robert K S, Robert Merkel, Robsavoie, Ronz, Ruslan225, Rygelski, S.K., Sakurambo, Sander123, Sbowers3, Schmidtja, Schneelocke, Science4sail, Scjessey, Secretlondon, Setokaiba, Sharonxiv, Sidna, SimonP, Sinc, Singeman, Skittleys, Slicky, Snowolf, Splintercellguy, Stephen Keller, StephenReed, Stephenb, Stephenchou0722, Steven Zhang, Stevertigo, Swilliso, Switchercat, Swpb, Tarun1972, Tbackstr, Tbolotnikov, Tbutzon, Teapotgeorge, The Anome, Three-quarter-ten, Tinctorius, Tjbeach, Tjmage1, Tobias Bergemann, TonyRobinson, Trolleytimes, Unixer, Vangardvoice, Veinor, Vitaly, VoiceInteraction, W9000, Waggers, Wavelength, Wdustbuster, Wickethewok, Wikiuser3786, Wimt, Wittlessgenstein, Wumpus3000, Xpclient, Ylai, Zhenghua.tan, Zommaso, Zvika, محبوب عالم, 630 anonymous edits

Image Sources, Licenses and Contributors

Image:Toshiba-Speech-Systems-Rabbit-Mascot.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Toshiba-Speech-Systems-Rabbit-Mascot.png *License*: unknown *Contributors*: Double.reed, 1 anonymous edits

License