

4.23 Speech and Software Technologies

Speech and Software Technologies (I) Pvt. Ltd., a TATA enterprise, started its research and development activities in the area of **speech technologies** in the year 1994. Today we are one of the leading product-engineering global IT service providers, providing value-added, speech-based IT solutions and services to large and medium-scale organizations.

All our products and services are built around our inherent capability to converge into all medium of communications i.e. telephone, Internet and wireless.

The testimony of our technological prowess lies in the state-of-art applications running for some of the world's leading organizations. Our clients have gained enormous mileage over their competitors, in customer satisfaction and in overall customer relation management.



Our presence in major countries gives us global reach and a vast rollout support capability. Together with our formidable team of high-calibre Doctorates, Post-Graduates, software and hardware professionals, we have successfully positioned ourselves at the vanguard of the global speech enabled - IT services revolution.

We continue to reiterate our key imperatives of focusing on our core technology competencies and keeping pace with continuous growth and learning.

Products:

Indian English Voice

Indian English Via Voice is a speech recognition software that replaces the need to type a document. The software developed was customized to the Indian accent to improve the recognition accuracy. After the adaptation process to the Indian accent the software shows accuracy levels of 98 to 99 %. The software can also be used for voice-based navigation and program management. that give our client a decisive competitive advantage.

There are two versions of Indian English Via Voice currently available:

(i) Indian English Voice -Executive Edition

Product features:

1. Powerfully speech recognition

- Dictate directly into most popular Microsoft window's applications.

- Gives the delivery status over the phone, thereby valuing the customers time.
- Create voice text dictation macros and voice form dictation templates to insert entire documents into your current work without having to stop the dictation.
- Control your PC by launching and executing programs with your voice.
- Create voice shortcuts for repetitive keyboard commands.

2. Natural language commands

- Use natural language to perform tasks in Microsoft word.
- Create documents; insert tables, change fonts, move paragraphs and print document.



3. Proof reading and editing made easy

- Have your text or e-mail read back to you with Via Voice text-to-speech system.
- Playback your dictation in your own voice to make corrections.
- Format and edit dictated text with your voice.
- Delay or delegate the document correction and modification; recorded voice remains with the file.

4. Industry's largest active vocabulary

- Expand the 64,000 word base vocabulary with up to 64,000 personal words.
- Use the included 2,60,000 word backup dictionary.

5. Simple and easy setup

- Move briskly through the setup and training with the Via Voice user wizard.
- Learn quickly by using multimedia quick tours, voice tips, extensive on-line help and comprehensive user guide.
- Find what commands you can say by just asking, "what can I say?"

(ii) Indian English via Voice-Home Edition

Product features:

1. Powerful speech recognition

- Dictate directly into most popular Microsoft window's applications.

- Create basic voice shortcuts for repetitive keyboard commands.

2. Natural language commands

- Use everyday language to perform tasks in Microsoft Word.

- Create documents, change fonts, move paragraphs and print the perfect document.

3. Proof reading and editing made easy

- Have your text or e-mail read back to you
- Format and edit dictated text with your voice
- Delay or delegate the document correction and modification; recorded voice remains with the file

4. Industry's largest active vocabulary

- 2,60,000 word dictionary.
- Expand the 64,000 base vocabulary with up to 64,000 personal words.

5. Simple and easy setup

- Move briskly through the setup and training with the Via Voice user wizard.
- Learn quickly by using multimedia quick tours, voice tips, extensive on-line help and comprehensive user guide.

Automatic Speech Recognition

Automatic Speech Recognition (ASR) is a system that users can converse with, as if they are conversing with another human being.

The ASR system is used to listen, recognize and understand what the user is trying to speak and then respond in a natural-sounding synthetic voice using an unrestricted vocabulary Text-to-Speech (TTS) synthesis system or a restricted vocabulary pre-recorded human voice prompts.

The ASR system understands the intention of the user based on the key word or the particular sequence of key words spotted in the fluent speech of the user, and initiates appropriate action.

Text-to-Speech Synthesizer

Text-to-Speech (TTS) synthesizer is a technology that can be used when an application needs to give



a natural sounding synthesized voice for dynamic information read-out like e-mails, and where full digital recordings is not appropriate.

Our Indian TTS has been customized to the Indian English accent for a natural sounding audible response with the proper nouns being articulated clearly.

Product Features:

- Multiple channel synthesis capabilities. Supports from one to hundreds of simultaneous usable TTS channels.
- Gives the option to set different parameters of the TTS Engine like pitch, speaking rate, volume etc. All the parameters could be dynamically configurable during run time to give the best quality output.
- Vocabulary size of more than 100,000 words.
- Voice Quality is 1.8000 Hz/16bit PCM output speech for Telephony Applications. 2.11025 Hz/16bit PCM High Quality output speech for Desktop Applications.
- Accepts inputs in the form of raw text data, HTML text, URLs, Acronyms such as ASAP, :-), etc.
- Output could be obtained as a buffer or written to a file or redirected to an audio device
- Automatic Speech Recognition

Automatic Speech Recognition (ASR) is a system that users can converse with, as if they are conversing with another human being.

The ASR system is used to listen, recognize and understand what the user is trying to speak and then respond in a natural-sounding synthetic voice using an unrestricted vocabulary Text-to-Speech (TTS) synthesis system or a restricted vocabulary pre-recorded human voice prompts.

The ASR system understands the intention of the user based on the key word or the particular sequence of key words spotted in the fluent speech of the user, and initiates appropriate action.

Better Caller Experience: Provides hands-free and hassle free access over telephone

Recognition Accuracy: Under real conditions we can expect 97% +/- 2% accuracy for both cellular and landline telephone calls.

Response Time: Typical recognition time is less than 1 second.

Multiple Languages Interface: Our ASR algorithms can be trained in-house for any language like English in India, Tamil, Hindi, Telugu, Malayalam, US English, Australian English, UK English, Spanish etc.

Echo cancellation: In a voice barge-in situation, when a user interrupts the system prompt, the SLI system detects the user's speech and stops the play back of its own prompt immediately. But at the same time, our SLI system is insensitive to background (ambient) noise and the speech of people talking in the background.

Automatic ASR improvement: A Call Completion Rate (CCR) parameter is used in all of our Spoken Language Interface (SLI) systems to measure the percentage of calls successfully completed by the system. If the CCR is not satisfactory, the user interface is continuously modified through self-training of ASR algorithms to improve the CCR.

ASR Applications

ASR variants are:

- Speaker independent, automatic speech recognition with Keyword spotting capabilities, for continuous speech inputs, enabling friendly interface between user & application.
- Speaker independent speech recognition with limited vocabulary for telephony applications. This can be adapted for voice based selection & navigation commands.

Functional areas:

- Tele-banking
- Call centers or a 24x7 Customer Support systems
- Unified Messaging Solutions (UMS)
- Value added services over phone
- General enquiry systems
- CRM/SCM/ERP applications

Embedded Speech Recognition

It is no longer a fantasy of science fiction movies that one can send an email through the mobile while navigate a car with voice commands. In fact, the world of embedded software and systems has made way not only in consumer electronics but also in critical applications of medical equipments. Features like remote management, web enabling, etc., are possible, thanks to embedded software. In-fact embedded software will be needed in future to drive most applications. For instance, ovens and fuel

injections in cars are controlled by embedded software driving the micro controllers. This trend of automation to achieve better efficiency and sophistication will pave a way for better scope/opportunities in this area.

We have developed embedded systems competencies in hardware initialization; porting, device drivers; diagnostics; board design and optimized codes to meet resource constraints through the tested methodologies of integrated engineering. For specific requirements, we provide development and usage of customized tools to ensure full compliance to specific coding guidelines. We address the complete lifecycle of embedded software by developing software and hardware simulators for product testing.

Service offerings

The embedded division has expertise in all aspects of products design (specification, design, development, validation, documentation, packaging, technology, transfer, etc.), in the area of telecommunication. The embedded division also has expertise in Computer Telephony Integration (CTI) and embedded system design. The skills set available with us are:

- Fixed/floating point DSP design
- 8/16 bit micro controller design
- FPGA
- CPLD
- PAL based design
- Device Drivers for Windows NT

This technological expertise can be used to develop software for voice authentication and voice access/security of premises. We have the capability to speech enable automotive applications and consumer appliances. If you are imagining about a car that has its door locks on when you say "Locks on", then you are thinking about our line of work. In the future all applications will be speech enabled and you just have to talk to access them.

*(Source Website/Courtesy :
Speech and Software Technologies
3rd Floor, Ceebross Arcade, #1, 3rd Cross Road,
Kasturibai Nagar, Adyar, Chennai - 600 020
INDIA Tel : +91-44-24425370/71
Fax : +91-44-24425369 E-mail:sales@sstil.com
Website : http://www.sstil.com)*