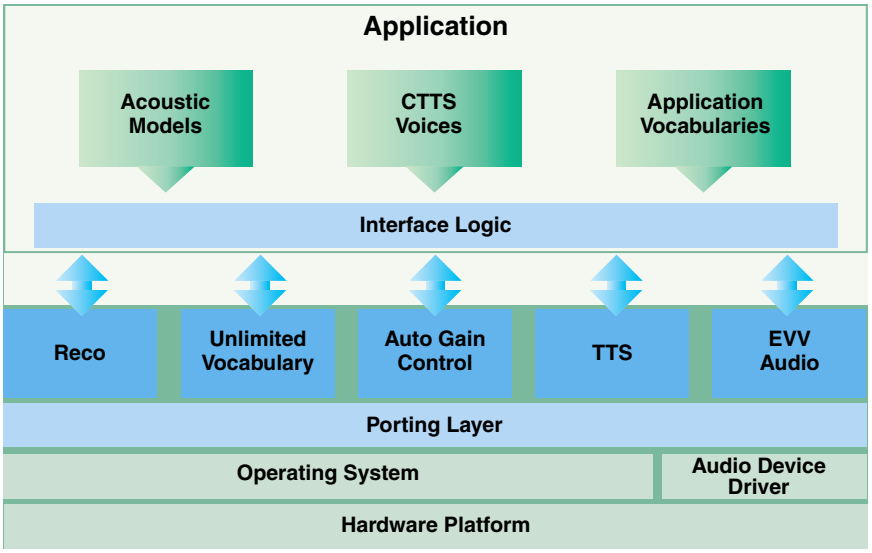
IBM Embedded ViaVoice, Version 4.2.

Highlights

- ***Provides fully integrated automatic speech recognition and text-to-speech in small mobile devices including automotive telematics systems and hands-free phones***
- ***Helps to minimize the skills and time needed to develop high-tech applications for devices and remote systems***
- ***Supports speech recognition vocabulary lists of up to 100,000 words in real time and a broad range of languages***
- ***IBM provides porting, integration, testing and consulting services, along with customized development workshops***



Embedded ViaVoice Architecure

Speech technology for mobile devices

IBM Embedded ViaVoice V4.2 delivers a higher level of speech technology to mobile devices such as automobile navigation systems, hands-free phones, personal digital assistants (PDAs) and other smart devices. Embedded device applications can use IBM speech technology in two forms:

- *Automatic speech recognition (ASR) – uses human speech to input commands into a mobile device*
- *Text-to-speech (TTS) – uses synthesized human speech to output text and other information from a mobile device.*

With Embedded ViaVoice and the e-business technology behind today's small mobile devices and automotive telematics systems, such as hands-free phones and navigation systems, developers can easily provide users with voice access to information from work, home, school or on the road.

Single, fully integrated architecture

The modular Embedded ViaVoice architecture provides fully-integrated speech recognition, speech synthesis, and other technology engines supporting the full feature requirements of an application with minimal central processing unit (CPU) utilization and memory requirements.

A single architecture with consistent application interfaces allows Embedded ViaVoice to support solutions from low resource PDAs through high performance in-car solutions to Java-based or markup language-based connected solutions. This single architecture implementation is a particular advantage to applications that need to span a broad range of platform capacities as well as solutions where significant growth in capacity is a requirement.

Broad language base

Embedded ViaVoice is available in a broad set of languages both for speech recognition and speech synthesis through the support of a worldwide network of IBM speech research and development laboratories.

With the latest release of Embedded ViaVoice, high quality embedded concatenative text-to-speech (eCTTS) was introduced to provide more human sounding speech synthesis for more advanced applications. Development of additional language models for ASR and voices for TTS as well as the continuous improvement of existing languages continues. For the most current language plans, please contact your local sales representative.

High recognition accuracy

The Embedded ViaVoice recognition engine is phoneme-based using finite state grammars to support highly accurate and noise robust continuous speech recognition. Through a comprehensive and vigorous research and development effort, IBM has reduced the word error rate of Embedded ViaVoice by an average of 35 percent annually over the last three years.

Large vocabulary recognition

The maximum supportable vocabulary has grown by a factor of 25 over the last four years. The most recent version of Embedded ViaVoice supports recognition of lists of up to 100,000 words in real time—allowing, for example, unconstrained recognition of any street in California or any place name within Germany.

Services and workshops

Porting and Integration Services includes porting to a new operating system, recompilation for a different processor architecture, or modification of the embedded audio layer to use new driver or codec. Alternatively with the Device Adaptation Kit, we supply the tools for the customer to perform and test the audio adaptation themselves.

Classes for application developers are available for the Embedded ViaVoice Software Development Kit (SDK). In addition, customized development workshops provide skills transfer and instruction on application development, evaluation methodology and tools so customers can design and tune their system.

IBM usability experts can provide review of user interface design, vocabulary and grammar optimization (including creation and testing of alternative grammars) and execution of voice recognition accuracy tests. Using a state-of-the-art audio studio, consultation is also available on selection and placement of microphone and other components for optimizing the audio signal. Assistance is available in application design, implementation and testing, specialized tool creation, and creation of validation plans and programs. Support is available for your technical interactions with other solution partners and with your customers.

IBM expertise in voice

IBM's sustained research and development investment in speech recognition and synthesis for more than 30 years has resulted in multiple advances, including Embedded ViaVoice. For today's embedded applications, highly functional speech-based systems can be developed using this product.

Embedded ViaVoice supports multiple programming models. Many small footprint embedded applications use Embedded ViaVoice through its C/C++ language application interface. Alternatively, Java speech applications can be written to the standard Java Speech Application Programming Interface (JSAPI) as a component of a standards-based connected system. Finally, multimodal browser applications can be developed using the standards based Multimodal Toolkit and Browsers supported by Embedded ViaVoice for voice recognition and voice synthesis. IBM Embedded ViaVoice products enable customers to have a true competitive advantage in today's fast moving marketplace as well as a clear path for future growth through a single, fully integrated architecture.

For more information

Please contact your local IBM sales representative for more information or a demonstration of our embedded products.

Embedded ViaVice Features	
Functionality	• Portable event-driven architecture • Fully integrated ASR & TTS • Low CPU utilization • Small static & dynamic footprint • Scalable modular architecture • Single & Multi-threading support • Runtime event notification • Optional speaker enrollment/adaptation • Phoneme based • Speaker independent
Accuracy and Robustness	• Very large vocabulary recognition • Tunable rejection to address non-speech sounds and out of vocabulary words • Advanced front end noise suppression • Supports third party noise suppression • Enhanced speech/silence detection • Continuous and discrete digit recognition • Spell mode capable • Confidence scoring • Highly accurate for both male and female speakers • Pronunciation confusability reporting • N-Best and homonym support
Solution Development Tools	• Eclipse-based IDE • Application creation wizards • Grammar editor and templates • Vocabulary testing and analysis • Pronunciation compiler and variant generator • Evaluation toolkit • Gain control tuning tool • Tracing/debug interface • Device Adaptation Kit
Flexibility	• Broad language coverage • Additional languages in development • C API • JSAPI & extensions • X+V markup language Web browser • Automatic gain adjustment • Multiple listening modes Push to Talk; Push to Activate; Always listening • Run-time language switching • Run-time pronunciation manipulation • Scalable Acoustic Models • 11/16/22kHz sampling rates • SNR feedback
Grammar/Compiler Support	• Scalable vocabulary support • Built in grammar compiler • Finite state grammars • Multiple grammar formats: SRGS/BNF/JSGF • Annotations • Voice tags from text or acoustic input • Dynamic & unlimited vocabularies • Pre-compiled and runtime grammars
Speech Synthesis (TTS)	• Unlimited pronunciation domain • Multiple voices • Customizable voices • Dictionary support • Indexing support/pause and resume • Performance tuning parameters • API for phoneme generation • Manual override of automatic synthesis • SSML support

Functional Overview	
Processor MIPs required	• 25 MIPs minimum • 175 MIPs for 100k words
Random Access Memory	• 490k bytes
Flash/Read Only Memory	• 780k bytes
Audio	• 16-bit sample, mono • 11/16/22kHz sampling
Languages supported	• US English • UK English • European French • Canadian French • German • Italian • European Spanish • North American Spanish • Brazilian Portuguese • Japanese • Simplified Chinese (Mandarin) • Traditional Chinese (Mandarin) • Korean
Operating systems supported	• QNX • RTXC • Windows® CE for automotive • Pocket PC • Smart phone • Windows Mobile • µiTRON • Montevista Linux • Embedix Linux
Processor supported	• Hitachi SH4 • Motorola PowerPC • IBM PowerPC® • Intel® x86 • Intel StrongARM • Intel XScale™

Processor MIPs and memory kb example is for US English recognition. Support for additional languages, operating systems and processors are in development. Please contact your local IBM sales representative for more information.

IBM Corporation
8051 Congress Avenue
Boca Raton, Florida 33487
U.S.A.

Printed in the United States of America
05-04

e business on demand.

G507-1660-00