SPEECHFX™

FonixTalk SDK

The multi-language TTS SDK featuring the industry's smallest memory footprint and highest vocal quality.

FonixTalk

FonixTalk is multi-language TTS software featuring the industry's smallest memory footprint, making it ideal for embedded applications. FonixTalk builds upon the legacy DECtalk® product, while making significant improvements in the technology's ability to produce more natural, less robotic-sounding synthetic speech.

FonixTalk introduces a new articulatory vocal tract model that creates a more natural-sounding voice. Developers and clinicians can create modified voices in minutes. FonixTalk's small memory footprint and extensive operating system (OS) support allows developers to add speech output to many devices with small CPU capacity or minimal memory.

FonixTalk also includes a simplified API that enables developers to integrate speech into their products or applications more quickly.

FonixTalk's nine TTS voices feature improved naturalness, inflection, intonation, pauses and changes in pitch, speed and emphasis that create a more normal speech pattern, which is easier to listen to and understand.

Features and Benefits

- Highly intelligible, natural sounding synthesized voices
- Intelligible at fast speaking rates
- Improved male, female and child voices
- Improved sentence-level phrasing and intonation
- Sophisticated pre-processing rules for greater pronunciation and prosodic accuracy
- Multiple languages: US English, German, Italian, Castilian Spanish, Latin American Spanish, Korean, and Mandarin Chinese
- Developers and clinicians may create custom voices (less than 0.5 KB per voice)
- Expanded set of speech parameters with dynamic control over voice, pitch, speaking rate, stress and language
- Able to tune voices for specific acoustic environments
- Change voices and languages on demand
- Correctly pronounce non-native words using the same voice
- Full-featured TTS with a small memory footprint (less than 500 KB for a single language with up to nine voices)
- Less than 2.5 MB for all seven languages, with nine voices per language
- NAND Flash support
- Flexibility to operate on many hardware and OS platforms
- C/C++/C#/Java/Visual Basic (VB) Interfaces

An efficient, scalable, low-cost solution

FonixTalk software components and architecture are designed to operate directly on the main processor, eliminating the need for an additional CPU or DSP. The flexible, self-contained modules enable developers to design systems to operate within memory and MIPS constraints and to be highly portable across many processor and OS platforms. The TTS engine operates across a full array of hardware and software combinations and includes a new simpler API, making time-to-market faster and cheaper for developers.

New FonixTalk Features

- New vocal tract model for increased naturalness
- Greater control over speaker parameters, custom voices
- Improved Spanish, German and Italian
- Improved prosody
  - -More accurate pronunciation
  - -Sophisticated pre-processing includes new parser
- Better singing voice
- Additional, simplified API
- Asian language support
- Pronounce non-native words correctly
- Change language on word-by-word basis using the same voice

Classic FonixTalk TTS Product Features

- Highly intelligible TTS voices
- Natural-sounding voices
- Pauses, inflections and emphasis
- Male and female voice support
- User-programmable pronunciation dictionaries
- Flexible voice parameter controls (custom voice capability)
- Universal phoneme set allows multi-language pronunciation without changing language modes
- Email reading
- SMS dictionary available
- Efficient use of MIPS and memory changing language modes
- NAND Flash support

Available Languages

US English, German, Italian, Castilian Spanish, Latin American Spanish, Korean, and Mandarin Chinese

Voices (9)

- 4 male • 4 female • 1 child

Programming Language Interfaces

- C/C++
- C#
- VB
- Java

CPU Requirements

- 10 - 30 MIPS (MIPS dependent on CPU architecture. These numbers are general cases estimates.)

Audio Requirements

- 8 KHz (8 bit µlaw or 16 bit PCM) buffered or streamed
- 11 KHz (16 bit PCM) buffered or streamed

Deployment Environment & Memory Usage

- Base engine: 1243 KB
- 128 KB – 2 MB ROM (depending on dictionary size)
- 64 KB – 256 KB RAM

Additional ROM usage - approximate (KB)

Supported Languages	Linguistic Components	Dictionary			
		Large		Small	
		Size (KB)	Words	Size (KB)	Words
US English	170	1875	73699	363	8241
German	1138	-	2153	51	-
Italian	372	167	2179	-	-
Castilian Spanish	338	10	422	-	-
Latin American Spanish	335	6	268	-	-
Korean	322	2	50	-	-
Mandarin Chinese	615	-	20902	58	-
All languages total	3290 KB	2060 KB	99673	472 KB	8241

Dictionary sizes can be adjusted as needed.

ROM calculation:

ROM = base engine + linguistic component + dictionary for each language

US and UK English small dictionary example:

342 + (75 + 75) + (100 + 120) = 712 KB RO

Hardware supported

- Analog Devices - Blackfin 533/535/525/527
- ARM - Arm 7 / Arm 9
- Epson - S1C33 Family
- Freescale - PowerPC 5200
- Intel – XScale / x86
- MIPS - R4xxx
- Renasas (Hitachi) – SH3 / SH4
- Samsung – S3C ARM Family
- Texas Instruments - OMAP 710 / 720 / 5910

Operating systems supported

- Win32
- Windows Mobile
- Linux
- MAC OSX
- No OS
- Contact SpeechFX for other operating systems

SPEECH FX™
FREEDOM OF SPEECH®