

Voice Recognition Software

Conversational technologies are in demand. Speech interfaces are a natural way to control and navigate mobile devices for improved safety and ease of use. Using the Voiceln SDK, developers can quickly and easily integrate Voice Recognition (VR) into applications or onto devices. Soon, most applications will likely need a voice interface. SpeechFX (SFX) has it now!

SFX Voiceln

Voiceln is based on the SFX proprietary neural network technology. It provides highly accurate speaker-independent Automatic Speech Recognition (ASR). SFX Voiceln represents a significant leap forward in VR technology. The SFX neural network phoneme localization technology employs significantly more phonetic, linguistic and prosodic information about the speech signal. This improves recognition rates even in noisy environments and also enhances language acquisition.

Improved Features

Phonetic pronunciation analysis

Accurate phoneme estimates

Better speech detection and speech framing

Improved rejection of non-speech sounds

Benefits

Helps users learn correct language pronunciation

Increased recognition accuracy

Improved background noise rejection

An efficient, scalable, low-cost solution

With Voiceln, the software components and architecture operate directly on the main processor, eliminating the need for an additional CPU or DSP. The flexible, self-contained modules allow developers to design systems that operate within application memory and MIPS constraints, and are highly portable across many processor and operating system (OS) platforms.

With multiple OS and hardware platform support, plus over ten languages, Voiceln offers developers the best solution to meet their needs as well as the needs of end users. Voiceln is a state-of-the-art VR SDK that is flexible and simple to use, and provides the highest accuracy for applications to deploy on devices. This is truly a “write-once, deploy many times” solution, saving developers and manufacturers time and money.

Voiceln Features

- Phonetic Localization
- Improved phoneme estimates
- Improved rejection of non-speech sounds
- Improved speech detection and speech framing
- Improved demo applications
- Accurate with no training necessary
- Noise-tolerant
- Insensitive to environmental acoustics
- Performs with background speech and music
- Multi-pass recognition
- Optimized letter (spelling) and digit recognition (voice dialing)
- Phoneme Alignment

Flexible

- Speaker-independent, continuous speech
- Word spotting / Finite-state grammars
- Dynamic word spotter and grammar creation during run-time
- Create specific grammars using your own words and lists
- Unlimited vocabulary
- N-best support for word-spotting, grammars, and spelling
- Adjustable out-of-vocabulary rejection
- Normalized confidence scoring for phonemes, words or phrases
- Speaker dependent phonetic voice tags
- Far-field microphone capabilities
- Offline grammar and word-spotting node compilers
- User data association with recognition words
- Wide character support

Efficient

- Compact neural networks and small memory footprint
- Small computing requirements (starts at 20 MIPS)
- Dynamic optimization to minimize memory and MIPS
- Recognition time – 100 ms after end-of-speech detection

Vocabulary

- Active vocabulary size, only limited by memory and MIPS
- Customizable IPA user dictionary
- Supports spelling greater than 65 K words

Languages:

US English, UK English, Canadian French, European French, German, Italian, Japanese, Korean, Castilian Spanish, Latin American Spanish, Swedish, Mandarin Chinese

VoiceIn Embedded Programming Interface

	Win32	Qnx	Windows Mobile	Linux	OS X	iOS
C/C++	•	•	•	•	•	•
Java	•		•	•		
VB	•		•			
C#	•		•			

Future Ports - • Contact SFX sales

Memory Usage

- Base Engine: 352 KB ROM
- Dictionary: 0 – 2,323 KB ROM
- Neural Networks: 50 – 225 KB ROM
- 200 KB - 1 MB RAM based on number of words or grammar size
(approximately 1.2 KB per word ROM/RAM)

MIPS

- Word-spotting: 10 words – 25 MIPS (can vary based on the CPU architecture)
- Finite state grammar for continuous speech (not dictation): 20 words
- 70 MIPS (can vary based on the CPU architecture)

Audio Requirements

- 8 KHz – 8 bit μ law/alaw (mono channel)
- 8 KHz, 11 KHz, 16 KHz 10-16 bit PCM, mono-channel
- Bandwidth: 100 Hz to sample frequency/2 -- LPF 3dB skirt
- Spectral Flatness: +/- 6dB
- Total Harmonic Distortion (THD) < 2%
- Signal-to-noise ratio (SNR) > 0dB for best results

Platform Support Packages

Hardware Platform	Win32	Windows Mobile	Mac OSX	Apple iOS	QNX 6.3	Linux	No OS
Analog Device							
Blackfin 531/533/535/53							•
ARM							
Arm 7		•					•
ARM 9/V4/V4i		•		•	•	•	•
ARMv6/ARMv7				•			
Epson							
51C33 Family, GNU33							•
Freescale (Motorola)							
i.MXL		•		•	•	•	•
PowerPC 5100/5200				•	•		
Intel							
SA-1110		•		•	•	•	•
XScale		•		•	•		
X86	•			•	•	•	
X86_64	•	•	•	•		•	
Renesas (Hitachi)							
SH3							•
SH4		•		•	•		
Samsung							
S3C ARM Family		•		•	•	•	•
Texas Instruments							
OMAP 710 / 720		•					•
OMAP 5910		•		•	•	•	•

* Contact sales: sales@speechfxinc.com

Developer Support Program: SDK training, application development.

