

Vocalizer Text-To-Speech.

Nuance's Vocalizer transforms the speech dialog experience by offering highly natural text-to-speech for every cloud and embedded use case. Vocalizer Cloud is offered through Nuance Cloud Services (NCS). Vocalizer Embedded is available as SDK for Windows, Linux, Android, and iOS-OSX platforms. Additional platforms are supported through professional services.

Feature	Benefit
Expressivity and naturalness	Improved expressive speech gives voices personality for the most natural and engaging experience possible
Multi-lingual support	Automatic language identification, foreign language dictionaries, and high-quality acoustic extensions provide unparalleled multi-lingual readout
Built-in domain intelligence	Optimization settings provide extra control options for special use cases such as SMS reading
Flexible speech generation	Volume, pitch, and speaking rate can be changed at run time for more dynamic and lively effects
Direct phonetic input	Optimal read-out of off-line phonetic databases such as navigation map data or music titles
User text rules	Customized read-out of application-specific abbreviations and text pattern using a user text processing rule set
User dictionaries	Application specific lexica can be phonetically optimized for accurate readout of exceptional pronunciations
Improved prompt tuning	With off-line tuning options any prompt set can be further optimized and customized for maximum quality
Seamless prompt insertion	Recorded audio prompts or tuned prompts are seamlessly blended with dynamic text-to-speech using automatic text matching (active prompt mechanism)
Vocalizer Studio	A comprehensive user-friendly suite of tools to prototype and optimize speech output applications. Easily create optimization data such as user text rules, user dictionaries and prompt databases
Languages and Voices	A truly universal voice portfolio offers 58 languages and 116 voices to facilitate the creation of global solutions using a single engine. The portfolio is continually expanding
Accuracy	High linguistic accuracy offers correct readout for all types of text including a large set of personal names
Scalability	A wide range of footprints scaling from 8 to over 900 MB ensures optimal performance from small mobile devices to powerful multi-media systems
Speech Synthesis Markup Language (SSML)	Support of SSML allows for TTS vendor-independent markup which is correctly interpreted by Vocalizer
Daily dictionary updates	NEW Daily dictionary updates are offered through Vocalizer Cloud, guaranteeing accurate pronunciation of trending words and names in the news domain. Currently 6 languages (EEFIGS) are supported, with more to be added in 2017
Morphed and multi-style voices	NEW Offline voice modifications are offered through professional services to change the perceived age and speaking style of a portfolio voice

Nuance's Vocalizer complete speech output solution generates high quality speech through a seamless blending of dynamic text-to-speech, pre-recorded audio and optimized tuned text-to-speech. The new Vocalizer engine is optimized for reading long text in a much more natural sounding way. New signal processing algorithms improve overall smoothness of the voice output and advanced syntactical analysis gives the spoken text a natural prosody, resulting in a unique experience.

Applications

Nuance TTS technology has been deployed successfully in numerous demanding applications ranging from navigation and automotive UI systems and consumer electronics to assistive technologies and industrial applications.

Automotive

- route guidance
- turn-by-turn directions
- infotainment

Consumer electronics

- smartphones
- e-book readers
- toys and game characters
- electronic dictionaries

Accessibility products for blind and disabled

- screen readers for PCs and mobile phones
- daisy book readers
- talking kiosks and ATMs

Industrial

- warehousing stock picking
- transportation

Vocalizer Embedded 3.0

With a broad range of options, Vocalizer Embedded offers an excellent quality trade-off for a variety of platforms and applications.

Code

The code size for a fully featured Vocalizer Embedded engine is 3 to 5 MB depending on the platform. This can be optimized based on the required language set, features and compiler choices.

Voice and Language Data

Voice Operating Point (VOP)	Flash size (excl. Code)	RAM usage
Embedded Compact – small versatile TTS suited for constrained platforms	average: 8 MB max: 22 MB	average: 5 MB max: 18 MB
Embedded Pro – high quality TTS optimized for navigation, in-car infotainment readout; basic SMS reading capabilities	average: 45 MB max: 71 MB	average: 11 MB max: 23 MB
Embedded High – high quality TTS read-out for SMS, news, and e-mail reading on embedded targets, suitable for all types of applications and use cases	average: 90 MB max: 283 MB	average: 17 MB max: 45 MB
Embedded Premium – highest quality concatenative synthesis, only available for selected voices	average: 300 MB max: 550 MB	average: 120 MB* max: 250 MB

(*) RAM usage for Embedded Premium assumes platform supports memory mapping

Multi-lingual voices include recorded material for one or more foreign languages. They are released for Embedded Pro and Embedded High and require 50% more memory (flash and RAM) for Embedded Pro and 100% more memory for Embedded High.

RAM usage includes code, language data, selected voice data and dynamic RAM. RAM requirement for Embedded Premium is 120 MB if platform supports memory mapping, otherwise this increases to 180 MB.

SDK Platforms

PC	Windows: 32 and 64 bit Linux x86: 32 and 64 bit OS X: 32 and 64 bit, Intel only, OSX 10.9+, Xcode 6.0+
Devices	Linux ARM: ARM32 Hardfp, ARM32 Softfp, ARM64 Android v4.0 (Ice Cream Sandwich), API level 14+, ARM32-v7a Android v7.0 (Nougat), API level 24+, ARM64-v8a iOS: arm64, armv6, armv7, armv7s, i386 and x86_64 simulator

Languages

Vocalizer offers the world's largest language and voice portfolio with 58 languages and 116 voices. This universal coverage facilitates the creation of global solutions using a single engine.

- Arabic
- Argentinean Spanish
- Australian English
- Basque
- Belgian Dutch
- Bengali
- Bhojpuri
- Brazilian Portuguese
- British English
- Bulgarian
- Canadian French
- Cantonese
- Catalan
- Chilean Spanish
- Colombian Spanish
- Croatian
- Czech
- Danish
- Dutch
- Farsi
- Finnish
- French
- Galician
- German
- Greek
- Hebrew
- Hindi
- Hungarian
- Indian English
- Indonesian
- Irish English
- Italian
- Japanese
- Kannada
- Korean
- Malay
- Mandarin
- Marathi
- Mexican Spanish
- Norwegian
- Polish
- Portuguese
- Romanian
- Russian
- Scottish English
- Shanghainese
- Sichuanese
- Slovak
- South African English
- Spanish
- Swedish
- Taiwanese Mandarin
- Tamil
- Telugu
- Thai
- Turkish
- US English
- Valencian

codefactory | We help human
technology interaction

info@codefactoryglobal.com

C/ de la Rutlla 56, local A
08221 Terrassa (Barcelona) Spain
Tel: +34 937 33 70 66
www.codefactoryglobal.com

