

Nuance's Vocalizer Expressive total 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 Expressive engine has been 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, providing a unique user 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 systems

Consumer electronics

- cell phones
- e-book readers
- toys and game characters
- electronic dictionaries

Accessibility products for the blind and disabled

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

Industrial

- warehousing stock picking
- transportation



Vocalizer Expressive

Nuance's Vocalizer Expressive transforms the text-to-speech experience with enhanced expressivity. Advanced technology features improved multilingual support and optimized read-out of long text to enable the highest quality speech output for a wide range of applications and markets.

Feature	Benefit
Enhanced expressivity	Improved expressive speech gives the voice personality for the most natural and engaging user experience possible
Improved multi-lingual support	More accurate language identification and high-quality acoustic extensions provide unparalleled foreign language readout
Naturalness	Natural sounding human-like speech output guarantees an exceptional end user listening experience
Built-in domain intelligence	Optimization settings provide extra control options for special use cases such as SMS reading
Flexible speech generation	Volume and speaking rate can be changed at run time for more dynamic and lively effects
Direct phonetic input	Allows for optimal and seamless read out of off-line phonetic databases such as navigation map data
User text rules	Customized read out of application specific abbreviations and text pattern is possible 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 flexibility
Seamless prompt insertion	Recorded audio prompts or tuned prompts can be blended with dynamic text to speech seamlessly by active prompt matching
Vocalizer Expressive studio tool	A comprehensive user-friendly tools suite to prototype and optimize speech output applications by easily creating optimization data such as user text rules, user dictionaries and prompt databases
Universality	A truly universal voice portfolio offers 42 languages and 71 voices to facilitate the creation of global solutions using a single engine
Accuracy	High linguistic accuracy offers correct readout for all types of text input including a large dictionary of person names
Scalability	A wide range of footprints scaling from 2MB to 900 MB ensures optimal performance on embedded platforms from very small mobile devices to powerful multi-media systems
SSML (Speech Synthesis Markup Language)	Support of SSML allows for TTS vendor-independent markup which is correctly interpreted by Vocalizer Expressive

Languages

Vocalizer offers the world's largest language and voice portfolio with 42 languages and 71 voices. This universal coverage facilitates the creation of global solutions using a single engine.

- Arabic
- Argentinean Spanish
- Australian English
- Basque
- Belgian Dutch
- Brazilian Portuguese
- British English
- Canadian French
- Cantonese
- Catalan
- Colombian Spanish
- Czech
- Danish
- Dutch
- Finnish
- French
- Galician
- German
- Greek
- Hebrew
- Hungarian
- Indonesian
- Irish English
- Italian
- Japanese
- Korean
- Mandarin
- Mexican Spanish
- Norwegian
- Polish
- Portuguese
- Romanian
- Russian
- Scottish English
- Slovak
- South African English
- Spanish
- Swedish
- Taiwanese Mandarin
- Thai
- Turkish
- US English

Vocalizer Expressive

With approximately 2 MB engine size, the total footprint of a Vocalizer solution can vary from less than 4 MB to over 900 MB depending on voice, language and choice of voice model. With a broad range of options, Vocalizer offers an excellent quality/trade-off for a variety of platforms and applications.

Code

The code size for a fully featured Vocalizer engine is ~2 MB. This can be optimized depending on required language set, features and compiler choices.

Language Data

The code size for a fully featured Vocalizer engine is ~2 MB. This can be optimized depending on required language set, features and compiler choices.

Voice Data

Voice Model	Data size per voice	Total RAM usage
Compact – small versatile TTS suited for constrained platforms	average: 3 MB max: 8 MB	average: 6700 kB max: 9800 kB
Standard – natural sounding TTS with attractive footprint	average: 25 MB max: 32 MB	average: 7000 kB max: 8300 kB
Plus – high quality TTS optimized for navigation, in-car infotainment readout; has basic capabilities for SMS reading	average: 55 MB max: 97 MB	average: 9500 kB max: 13500 kB
Premium – high quality TTS read-out for SMS and e-mail reading on embedded targets, suitable for all types of applications and use cases	average: 95 MB max: 300 MB	average: 23000 kB max: 75800 kB
Premium High** - Server solution for highest quality readout	average: 306 MB max: 985 MB	average: 23000 kB max: 75800 kB

* total RAM usage includes code, language data, voice data and dynamic RAM
 ** Uses same voice repository as premium but different audio data compression