



Vocalizer Speech Output Solutions

HIGH ACCURACY. FLEXIBILITY. SCALABILITY.

Nuance's Vocalizer 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 carefully selected top quality voice characters bring excellent speech output to all your applications through a powerful combination of state-of-the-art technology, user-friendly optimization and customization tools and an experienced service team. Designed as a highly flexible and scalable solution, Vocalizer can fit a wide range of embedded platforms and devices.

Applications

Vocalizer 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
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 order pick
transportation

Nuance Vocalizer is the leading speech output solution converting any written text into spoken output. Vocalizer combines state-of-the-art text-to-speech (TTS) technology with tools and services for enabling high quality speech output in a broad range of applications and markets.

Feature	Benefit
Naturalness	Natural sounding human speech output guarantees an exceptional end user listening experience
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 400 MB ensures optimal performance on embedded platforms from very small mobile devices to powerful multi-media systems
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 ruleset
User dictionaries	Application specific lexica can be phonetically optimized for accurate readout of exceptional pronunciations
Tuned prompts	With off line tuning options any prompt set can be further optimized and customized for readout
Seamless prompt insertion	Recorded audio prompts or tuned prompts can be blended with dynamic text to speech seamlessly by active prompt matching
Vocalizer 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
Universal portfolio	Offering 40 languages and more than 55 voices this truly universal voice portfolio facilitates the creation of global solutions using a single engine
Unrivaled multi-lingual readout	Using intelligent language identification and phonetic conversion capabilities, Vocalizer can interpret and correctly read out text and phonetic fragments from foreign languages
Data file only configuration	Configuration is possible through the simple addition of data files



Vocalizer Speech Output Solutions

HIGH ACCURACY. FLEXIBILITY. SCALABILITY.

Languages

Vocalizer offers the world's largest language and voice portfolio with 40 languages and more than 55 voices. This universal coverage facilitates the creation of global solutions using a single engine.

- Arabic
- Australian English
- Bahasa Indonesian
- Basque
- Brazilian Portuguese
- Canadian French
- Cantonese
- Castilian Spanish
- Catalan
- Czech
- Danish
- Dutch Belgium
- Dutch Netherlands
- EU Portuguese
- European French
- Finnish
- German
- Greek
- Hindi
- Hungarian
- Indian English
- Irish (English)
- Italian
- Japanese
- Korean
- Mandarin Chinese
- Mandarin Taiwanese
- Norwegian
- Polish
- Romanian
- Russian
- Scottish English
- Slovakian
- South African English
- Swedish
- Thai
- Turkish
- UK English
- US English
- US Spanish

With approximately 2 MB engine size, the total footprint of a Vocalizer solution can vary from less than 4 MB to over 400 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

Language data sizes range from 500 kB to 5 MB. Average size of language data across all languages is ~1.5 MB. When loaded at run time, language data requires ~300 kB of RAM.

Voice Data

Voice Model	Data size per voice	Total RAM usage*
Compact – small versatile TTS suited for constrained platforms	average: 650 kB max: 1100 kB	average: 3500 kB max: 4000 kB
Standard – natural sounding TTS with attractive footprint	average: 20 MB max: 30 MB	average: 4500 kB max: 6500 kB
Premium light – high quality TTS optimized for navigation and in-car infotainment readout	average: 48 MB max: 65 MB	average: 7900 kB max: 10100 kB
Premium – largest number of voices with highest quality readout for all types of applications and use cases	average: 110 MB max: 270 MB	average: 30000 kB max: 55000 kB
Standard high**	average: 55 MB max: 90 MB	average: 4500 kB max: 6500 kB
Premium high**	average: 360 MB max: 850 MB	average: 30000 kB max: 55000 kB

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

About Nuance Communications

Nuance is the leading provider of speech and imaging solutions for businesses and consumers around the world. Its technologies, applications and services make the user experience more compelling by transforming the way people interact with information and how they create, share and use documents. Every day, millions of users and thousands of businesses experience Nuance's proven applications and professional services. For more information, please visit www.nuance.com.

Please note that wherever future product information is mentioned it is intended to outline our general product direction and it should not be relied upon in making a purchasing decision. The development, release, and timing of any features or functionality described remains at our sole discretion, and the terms of any purchase will be limited to those documented in the contract. Nuance Communications Confidential.

Copyright © 2012, Nuance Communications, Inc. Nuance Proprietary Information. All rights reserved. Nuance, the Nuance logo and The experience speaks for itself are trademarks and/or registered trademarks of Nuance Communications or its affiliates in the United States and/or other countries. All other trademarks are the property of their respective owners.

