

VoCon Hybrid Speech Recognition.

VoCon Hybrid v4.10 from Nuance is a unique SDK for embedded and connected speech recognition applications. This feature-rich development suite supports creating speech applications for cars, phones, watches, home appliances, and many more. Natural language understanding and speech signal enhancement are available to create robust and intuitive voice control solutions.

Feature	Benefit
Fully hybrid solution	Complete range of embedded and connected speech recognition services from embedded digit recognition to connected dictation and complex search functionality.
Scalable modular architecture	Supports a range of operating points so that smaller applications are not affected by higher CPU & RAM required for larger applications
Software Development Kit (SDK)	Includes sample applications and all tools required to build speech recognition applications
Large vocabulary support	Enables embedded recognition for large lists up to millions of items. Cloud recognition offers even larger and constantly updated recognition vocabularies
High accuracy across a wide range of noise environments	VoCon Hybrid is developed for mobile applications with high accuracy operation to SNR levels as low as 5dB
Part of the Nuance speech & language technology family	VoCon is part of an unrivaled portfolio of speech and language solutions, working on its own, or as fine-tuned chain of technologies including noise reduction, speech recognition, language understanding, dialogue, text-to-speech, voice biometrics, and more.
All Inclusive Main Menu (a1M)	Enables all commands from all domains to be spoken in a single utterance on the main menu
Wake-up word	Always listening mode with key-word activation removes the need for a "press to talk" button
Barge-in	Allows user to speak over spoken dialog prompts and be recognized
Natural language understanding	Recognizes natural speech, eliminating restriction to predefined commands for all VoCon Hybrid languages
Global language support	Global support for over 25 languages provides universal functionality
Full Flexible POI Recognition	Recognizes every possible word in a database in every possible order, especially suited for Points of Interests (POI)
Domain-specific solutions	Dedicated solutions for key domains such as music, destination entry, telephony, and more. Multilingual recognition e.g. for domains such as music and destination entry
Spelling module	Spelling module available as backup to whole word recognition.
Hardware Acceleration	Support Spansion and Neon Expansions to speed up the speech recognition tasks

This advanced speech recognition engine delivers a new level of speaker-independent and continuous speech recognition capabilities with unique features. It provides superior functionality, unmatched accuracy, and high performance for a variety of applications that benefit from speech control. Designed as a modular and scalable engine, VoCon Hybrid can accommodate a large range of embedded platforms. Deployments can be customized with optimized footprints as dictated by the required functionality for each application.

Applications

VoCon Hybrid is optimized for a wide variety of applications and the VoCon product family has been deployed in automotive, accessibility, industrial and consumer electronics solutions for over 10 years.

Automotive

Consumer electronic devices

- PNDs
- cell phones
- gaming
- smart watches

Industrial applications

- warehousing/distribution
- medical

Home appliances, utilities and entertainment

Accessibility products for the blind and disabled

VoCon Hybrid Speech Recognition.

The total footprint* of VoCon Hybrid can vary from 3.2MB upwards depending on task complexity and platform restrictions. Deployments can be customized with lowest possible footprint to support the desired functions. See the table below for examples of data size per language and RAM usage for specific components.

Code & Data	
Functionality	Code size (MB)
Basic command & control (C&C) application	3.2MB
Full functionality, largest acoustic models	7.5MB

Data – typical model sizes	
Component	Data size per language
Acoustic model – Gen 5 – per language	~4MB
Acoustic model – Gen 4, compact – per language	~900kB
CLC – mono-lingual – general purpose transcriptions	300-7300kB
CLC – multi-lingual - music collection compilation	700-3000kB

Use cases: Data size and RAM usage		
Component	Data size per language	Total RAM usage
Digit recognition	4kB	1.25MB
Basic C&C application 100/10,000 commands	10 / 500kB	1.3 / 1.8MB
Telephony (voice-activated dialling) with grammars + SLMs, including NLU. 1350 contacts.	0.52MB	12.6MB
1-shot voice destination entry POI & addresses (UDE) all USA, FST based, including NLU	300MB	56MB

* Footprint: Reference VoCon Hybrid v4.10 SDK Code Size Table, Dragon Drive Reference Design ASR/NLU v.1.0

Languages

Featuring advanced language capability, VoCon Hybrid supports complex dynamic content in over 30 languages throughout North America, Europe, the Middle East and Asia Pacific. This universal functionality facilitates the creation of global solutions using a single engine.

- Arabic (Gulf Region)
- Australian English
- Brazilian Portuguese
- Bulgarian
- Canadian French
- Cantonese
- Castilian Spanish
- Czech
- Danish
- Dutch
- EU Portuguese
- Finnish
- Flemish
- French
- German
- Greek
- Italian
- Indian English
- Japanese
- Korean
- Mandarin Chinese
- Mandarin Taiwan
- NA Spanish
- Norwegian
- Polish
- Russian
- Slovak
- Swedish
- Thai
- Turkish
- UK English
- US English