Truly Natural™
Fluent Speech Engine

Embedded, Large Vocabulary Continuous Speech Recognition with NLU

Sensory's TrulyNatural features a state-of-the-art small-footprint, scalable in size and features, speaker-independent recognizer.

Many applications don't need or want to rely on cloud-based connection to do high-performance speech recognition. Others seek a client/cloud distributed system with optimal performance. The market concerns regarding privacy, performance and bandwidth are driving more processing to the edge. TrulyNatural (TNL) is the natural language speech recognition solution for these needs. Sometimes a cloud solution is needed that is smaller footprint so more channels can call in simultaneously, TrulyNatural can assist in that too.

Sensory is a pioneer in the use of embedded neural network-based speech recognition and has become the industry leader in optimizing and engineering speech recognition software with small footprints and minimal MIPS. This extensive experience and continuous innovation have led to the first **embedded** large vocabulary continuous-speech recognizer (LVCSR) with state-of-the-art cloud performance. Sensory's TrulyNatural engine is ideal for consumer electronics, home automation, mobile/telephony devices, internet-of-things, automotive, and PC-based applications. **Contact Sales at sensory.com to learn more about adding TrulyNatural™ to your product today!**

Fast

Embedded hence zero cloud communication delay. Fewer MIPS means lower latency and quicker response.

Accurate

Best-in-class for embedded applications. Outperforms cloud solutions in domain specific tasks.

Natural

Allows users to speak naturally to their devices. Avoids need to learn scripted preset command sets.

Flexible

Supports both simple grammar-based vocabularies and statistical language models to fit a variety of use cases and memory requirements.

Privacy/Secure

Embedded execution negates the transfer of private information to the cloud.

Small

Memory requirements are an order of magnitude smaller than comparable cloud-based solutions.

Customizable

Unlike many Cloud solutions, the recognition vocabulary can be tailored to a specific task or domain.

Adaptable

Able to run on platforms such as Raspberry Pi (Linux/Android) to Multicore ARM (iOS/Android) and compatible with most OS based systems.

Scalable

Ideal for smaller footprint device requiring vocabularies of hundreds of words in under 3 MB of memory, and more complex devices like smartphones, cars, and robots that require embedded natural language interfaces.

Optimized

Optimized engine is able to take advantage of available hardware acceleration and multiple cores (e.g., ARM Neon).

Noise Robust

A single solution works under a wide range of noise conditions, Acoustic models are trained under a wide range of noise and channel conditions.


Mobile

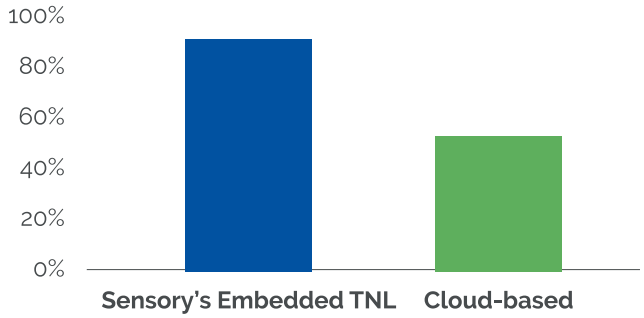

Automotive


IoT


Medical Equipment


Hearables/Bluetooth

Task Completion Rate: Voice Control Microwave



Domain Specific TNL Outperforms Cloud-based Equivalent

The advantage of utilizing TrulyNatural's **domain specific assistant** is seen when comparing task completion rates (TCR) between two microwaves that support a voice control feature. TCR is the rate at which a device is able to complete a requested task per presented task. Sensory's TrulyNatural task-tailored recognition vocabulary results in a greater TCR than the cloud-based voice control and therefore a more reliable experience for the user. A total of 40 utterances were evaluated for task completion.[1]

TrulyNatural SDK Technical Specifications

Technical Features

- TrulyNatural features a large-vocabulary continuous-speech recognition (LVCSR) engine, suitable for LVSCR tasks, including both grammar-based and language model vocabularies.
- TrulyNatural SDK features the following TrulyHandsFree phrase spotter technology:
  Fixed wake word / Enrolled fixed wake word / User-defined wake word
  Speaker verification / Speaker identification
- Speaker-independent recognition (No training or enrollment required).
- Includes a deep learned Voice Activity Detector, ideal for noisy environments.
- Supports out-of-vocabulary rejection.
- Prebuilt and custom vocabularies supported.
- Configurable data flow architecture supports concurrent and sequential recognition

Platforms

- C-based SDK allows for quick hardware and operating system portability.
- C API with support for live audio capture on iOS, macOS, Linux, and Windows.
- Java API with support for live audio capture on Android.
- Model output suitable for use with Sensory's DSP phrase spotter products
- TrulyNatural libraries for Sensory's standard ports are fully optimized. Vector instructions are used where available on x86 and ARM CPUs.
- **Webtools** available for fast and easy language vocabulary and grammar development.

Developer Notes

- SDK Samples provide source code coverage of common tasks.
- Includes command line tools to evaluate recognizer models, including support for both live audio and batch recognition.
- Suite of command-line tools and utilities (most with source code) provide a quick and convenient way to create and evaluate models.
- Natural Language Understanding: used to extract meaningful information from the recognition output. We utilize several NLU approaches depending on data availability and size requirements. **Contact Sensory to determine the best fit for your product.**

Foreign language support:

US & UK English, Mandarin, Korean, Japanese, French, Italian, German, Portuguese, Spanish, and Russian with many others in development. Contact us at Sensory.com to find out more about your language requirements.

Sensory, Inc., 4701 Patrick Henry Drive, BDLG 7, Santa Clara, CA 95054, United States, 408.625.3300, **sensory.com**

[1] Request the full study from sales@sensory.com