

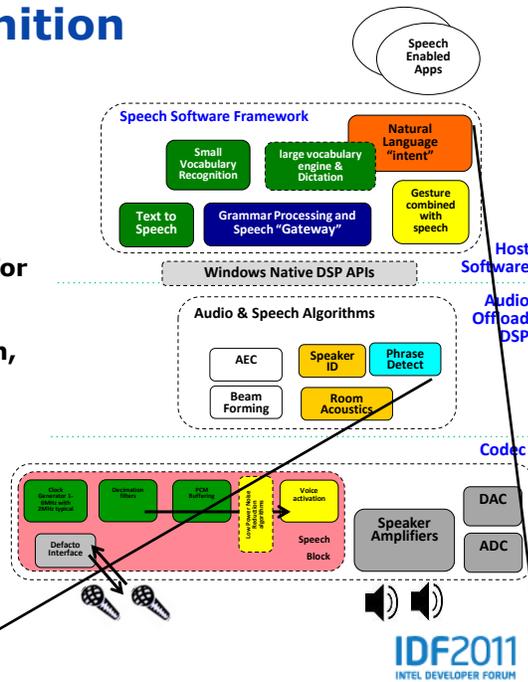
Summary:

SpeechFX embedded neural net (“SFX NN”) voice recognition software and voice recognition through cloud computing, such as the Apple® SIRI (“SIRI”), fundamentally approach voice recognition from opposite ends of the deliverable--it is an apples to oranges comparison. The basic distinguishing factor and inherent benefit is that the SFX NN may be embedded on an IC and SIRI is an in-the-cloud solution. The diagram¹ below illustrates this distinction (information contained in the two boxes connected by arrows is provided by SFX and is not included in the Intel® IDF slide prepared by Intel).

Speech Recognition

The “Full” Picture

- Digital MEMS array (2-4) attached to codec using 2-wire PDM interface
- Voice activation logic in Hardware Codec
- Ultra Low Power Clocking for DSP & Codec during S0ix
- Offloaded (firmware) Microphone noise reduction, Room acoustics, Phrase Detection and Speaker Identification
- Host based recognizers for both small and large grammars
- Speech Gateway as the master and owner



SFX NN voice recognition resides on the chip contained in the device and triggers the instantaneous VR process without a push-to-talk button.

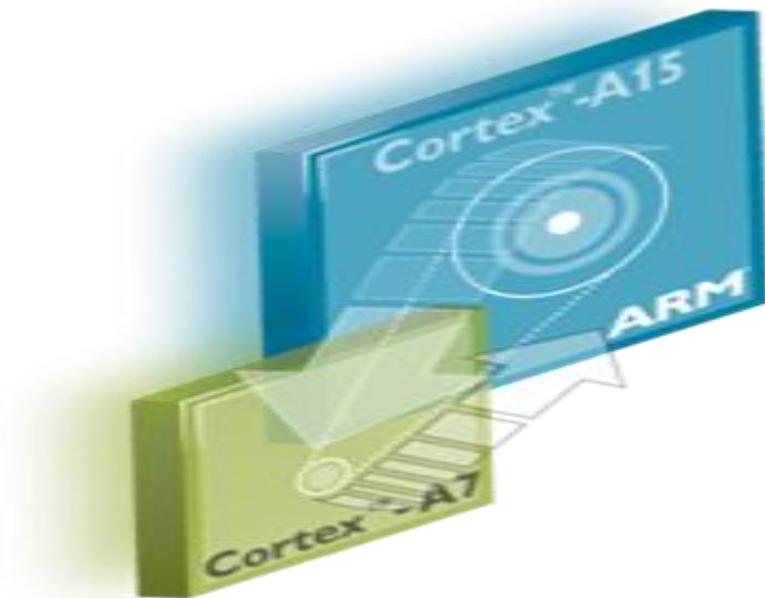
SIRI resides in the cloud, and is only accessible through a push-to-talk feature on a device that has network connection.

¹ IDF 2011; CHWS004, page 12

SFX NN embedded voice recognition value vs. Cloud accessed voice recognition:

1. SFX NN eliminates any latency or delay in recognition.
2. SFX NN engages in any “hands-free” environment, without hand-held or touch requirement of push-to-talk. Push-to-talk is neither convenient nor user friendly.
3. SFX NN may always be in a “listening mode” for key word detection as the trigger. The constant “always-on mode” will reside on a low powered DSP chip, such as an ARM7.
4. SFX NN may be deployed with the “big.LITTLE” architecture² for key word detection on a low powered ARM7 chip with direct interaction to an ARM15 chip in the silicon.

big.LITTLE Processing



The big.LITTLE allows for several speech embedded applications to initiate from the device—on the local / resident chip, such as:

- Voice dialing
- Calendar and contact manager
- Dictate text message or email: SFX NN can develop and implement a language model to improve dictation. Language models for particular domains such as email and text messages can also be trained.
- Text-to-speech (TTS) output

² ARM The Architecture for the Digital World

- Hands free application control
- Hands-free email listening and response (audio attachment to email, similar to what Microsoft is doing in their latest Sync release).
- Hands-free game play

The SFX NN team has experience deploying speech software in the big.LITTLE environment, beginning with the TriCore processor developed by the Siemens® semiconductor group, later known as Infineon Technologies® which was acquired by Intel in August 2010.

5. SFX NN operates in noisy environments, with robust adaptation to multiple noisy conditions. The NNs are pre-trained (and can be re-trained) for a variety of noisy environments. We suspect SIRI will be susceptible to noise more than SFX NN because they are trying to do so much, between the device and the cloud. SFX NN processes noise on the device to allow for near instantaneous recognition.
6. Currently SFX NN offers more language support:
 - a. SIRI has only three: English, French, and German.
 - b. SFX has eight: English, French, German, Italian, Spanish, Japanese, Korean, and Chinese. For small key-word detection, SFX has tools to add support for an additional language in as quickly as several weeks (depending on language).
7. Additional languages for SFX NN key words can be developed very quickly, with a low cost to train NN keywords. SIRI new languages are costly, and must be developed and built out with language specific AI and Natural language processing. SFX NN could be tied to AI and Natural language capabilities through Windows native DSP APIs (for example) for more functionality. Approaching approximately 2,000 phoneme categories requires large language models; SIRI's lengthy language development process requires significant data for each language and natural language processing.
8. SIRI servers may overload with too many users.
9. SFX NN accesses standards based APIs, Natural language and AI through the cloud as may be provided by a Google® search engine or any other proven search engines which are available through other providers/vendors/carriers.
10. SFX NN as a standard feature on Intel's Atom SoC will be distributed through multiple channels vs. the limited and closed Apple channel. Google has a server/cloud solution supporting Android™. This is basically a server/cloud solution, not embedded on the device with the inherent features of the low-powered, always in listening mode that the SFX NN delivers. Board distribution of an embedded solution:
 - a. Multiple mobile carriers
 - b. Multiple mobile, handheld, tablet to PC device OEMs

11. SFX NN low powered device with an ARM7 component does not require the antenna to be on.
12. SFX NN, an embedded solution, does not require a service provider to process recognition, and does not need Wi-Fi or contract plans.
13. SIRI cost to end user may increase due to data transfer costs.
14. SFX NN has a very small footprint and low CPU usage. Furthermore, the SFX NN can be implemented in hardware to further reduce power consumption and space requirements.
15. SFX NN requires low power and thus can always be “on” listening for the key word and rejecting noise.
16. SFX NN flexibility on the chip can deploy multiple key word spotters to achieve the “open” command style of SIRI.
17. In addition to the enumerated embedded features, SFX NN may be deployed on servers for cloud applications. SFX NN can increase the efficiency and CPU utilization on servers, which is a direct cost reduction for users and providers.

Conclusion:

SFX NN currently provides the value of embedded speech detection of key-word and small to large vocabulary recognition. SFX NNs are market proven, providing highly accurate voice recognition in video games and mobile devices (see www.speechfxinc.com). SFX NN expects greater and more robust recognition utilizing the development of and implementation of future neural networks and voice recognition engines. Recognition is triggered on the device with low-power requirements in noisy environments. The dynamic flexibility of parallel processing with access to the cloud for more advanced AI and Natural language functionality begins at the speech gateway embedded on one or multiple processors in silicon providing latent free and hands free access vs. push-to-talk, delayed response through the cloud. SFX NN is a smart and compelling solution for not only embedded functionality, but will greatly enhance the use of a SIRI type cloud implementation.