

Interview with Tom Freeman, VoiceBox Technologies
“Conversational Voice Platform” for mobile applications

*Tom Freeman, co-founder and senior vice president of marketing for **VoiceBox Technologies**, was interviewed by Bill Meisel in late July. Tom has over 20 years of software business management experience. He was a sales director at Interleaf, a document publishing company, and a Vice President at PCDOCS, a document management company. Tom was also the Vice President of Sales and Marketing for InfoAccess, where he was instrumental in selling the company to Stellant Corporation. After the sale, Tom became Stellant's Vice President of International Sales and Operations. Tom has a BA from the University of Washington and a Masters from the Executive MBA program at the University of Washington.*

Meisel: Please describe VoiceBox's corporate objectives and product line.

Freeman: The VoiceBox mission is to develop a clever and cooperative, next-generation Voice User

Experience that will catalyze the mass-market adoption of speech. VoiceBox Technologies delivers its Conversational Voice Search Platform, built on standard speech recognition and text-to-speech software, and enhanced with popular modular applications for embedded, server, and PC environments. The VoiceBox Conversational Voice Search Platform offers a best-in-class user experience for quickly searching and accessing digital content, taking full advantage of broadband, wired, and wireless networks. Modular applications may be mixed, matched, standalone, or integrated with other custom applications. This suite of applications includes hands-free dialing, iPod music search/access, off-board search, local search, voice destination entry, and control of XM satellite radio.

What distinguishes your technology from others in the marketplace?

From a consumer's perspective, the bottom line is that VoiceBox provides unparalleled usability:

- No need to learn a constrained grammar and command set;
- Works better in noisy environments and is tolerant of normal speaker imperfections – ums and ahs
- Minimizes the use of drill-down 'menus' or voice-command trees, which unquestionably causes user frustration; and
- Minimizes the need to constantly declare your query-topic or jump in and out of separate applications.

Technologically, VoiceBox enhances the efficiency and recognition rates of standard speech recognition engines through proprietary VoiceBox grammars and “Advanced Virtual Dictation.” VoiceBox applies proprietary semantic understanding technology that dynamically determines context and intent from freeform speech. Once VoiceBox determines the user intent, a context is established.

Speech Strategy News August 2007 (Continued)

The context is used to improve the accuracy of speech recognition by eliminating noise words, user mistakes, and verbal pauses—and thus sharpens output from the speech engine. Users can now interact with the systems using in-context dialogs called cooperative conversations. Smart search agents contain VoiceBox knowledge of a given content domain including knowledge about concepts, semantic patterns, conversational flow, environment, and previous conversations. The agent also can handle data retrieval and controls TTS and multi-modal interactions with the user.

Key intellectual property setting VoiceBox apart includes patent-pending technologies:

- Phonetic and spelling “Fuzzy Matcher” algorithms – determine the intent of utterances using a combination of grapheme and phonetic-based techniques, and corrects for speech recognition errors based on context and history.
- Context-based recognition – Corrects for human errors based on context and history.
- Context-based intelligent reference matching – Allows users to switch topics while remembering the context of previous topics. Also allows partially specified task queries based on the current conversational context.
- Adaptive correction – Automatically uses voice and psychological cues to learn how to correct misrecognitions.
- Application-specific sentence sharpening – Compensates for a number of human errors in speech, such as pauses to think, interjection of *uhhh*'s and other meaningless speech, repeated words, and stuttering.
- Smart Agents – Retrieve, check, personalize, filter, and integrate data from various data repositories to deliver the best possible response queries and allow humans to perform tasks with no knowledge of the data sources, to receive personalized responses, and allows cross-domain queries.

These technologies add substantial value to the underlying standard speech recognition engine, and work together to deliver a superior voice user experience and high task-completion rates in noisy, real-world environments.

VoiceBox has announced a number of development partnerships targeted at mobile search applications. Please outline the most important of these.

We believe all of our partnerships to be important, however the ones specifically focused on mobile search would include **Nuance** and **InfoSpace**.

VoiceBox and Nuance have a joint initiative to develop advanced voice navigation applications for personal navigation devices, automotive, and mobile markets. Our collective goal is to develop and deliver a voice user-experience that is so accurate and intuitive that it becomes the de-facto standard and preferred way to use navigation devices, no matter where consumers choose to use them.

VoiceBox successfully implemented eyes-free navigation in the newly upgraded version of InfoSpace/Sprint FindIt!, the first comprehensive location-based application that helps consumers easily and quickly find everything from nearby restaurants and movie times to maps and driving directions. VoiceBox added text-to-speech spoken turn-by-turn directions to the Maps & Directions category of InfoSpace Find It!, delivering on the first phase of a joint multi-year licensing agreement.

Is VoiceBox technology currently available to consumers in a deployment?

Yes. VoiceBox has implemented eyes-free navigation in the newly upgraded version of Sprint/InfoSpace Find It! as noted.

How fast do you see your market growing?

We define our market overall as queries away from the PC setting. As such, it is made up of mobile and portable devices, along with telematic systems. Research suggests the growth will be 700 million for non-cellphone devices in North America in 2010. Most analysts point to Portable Navigation Devices as the fastest growth segment among the mobile and portable markets.

Does your technology require cooperation with a speech technology vendor?

VoiceBox has strong partnerships with leading speech technology vendors, and works with all standard speech recognition and TTS technology, across most environments. Our announced relationships with Nuance and IBM indicate strong partnerships that are technologically and business oriented. Partnership goals and technological endeavors include joint development effort and cooperation on all sides.