

## TravellingWave launches speech-augmented text entry on mobile devices

### *New company uses internally developed speech technology*

**TravellingWave** is an early-stage company that develops software using speech recognition technology for entering text into mobile devices. In August, the company announced its launch and investment from individual Seattle investors. The amount of the investment was undisclosed.

The company has developed technology it calls "predictive speech-to-text." Its initial product is VoicePredict, which allows users to input text using a combination of text and voice. The software, including speaker-independent speech recognition processing, is all on the phone. The user types the first letter or two of the word and presses a soft key and says it (or vice versa), and the system completes the word or provides a list of candidates. The system uses both predictive text algorithms and speech recognition (the latter developed internally). Since users will most likely be looking at the screen as they type, the speech recognition has been designed to allow speech at a distance from the microphone, so that the user need move the phone to their mouth for each word entered. In situations where speaking is undesirable, the approach turns into standard text prediction.

TravellingWave has recently concluded early trials of its system, and plans to deploy with mobile carriers and handset manufacturers in the coming year. The beta system is currently available for Windows Mobile 2005 Smartphones. Geoff Entress, an investor and a principal with venture capital firm Madrona Venture Group, said, "The TravellingWave proposition is fundamental to the future of the wireless industry, namely: how do you input text into the ever-shrinking mobile device?"

TravellingWave was founded by Dr. Ashwin Rao, who has held research positions at AT&T Bell-Labs and Dragon Systems. Other executives include Gregory Aronov and Marat Garafutdinov (formerly with Conversay and Microsoft) and Joseph O'Neill (formerly with Motorola).

The company got its name from the front-end processing used in the speech recognition, techniques based on published research on compact features that model the travelling wave phenomena in the human cochlea.\* The company said that these modulation features yield higher raw recognition accuracy, superior noise robustness, and reduced footprint for the speech recognition engine.

Rao said that the vocabulary of the speech engine is tens of thousands of words. The application uses what Rao considers proprietary acoustic and language modeling. He notes that, unlike dictation of standard documents (like letters, memos, and medical/legal reports), text messages (SMS, email, IM) tend to be very short; e.g., SMS text is restricted to 140-160 characters per message. He said that the company's techniques may be viewed as an optimization of standard modeling techniques for the messaging application. Rao said that the technology can be deployed in the network, in which case the vocabulary can be increased to hundreds of thousands of words.

---

\* References: (a) "On Decomposing Speech into Modulated Components," A. Rao, R. Kumaresan, *Journal of the IEEE Trans. On Speech and Audio Processing*, May 2000; (b) "Model Based Approach to Envelope and Positive Instantaneous Frequency Estimation of Signals," R. Kumaresan and A. Rao, *Journal of the Acoustical Society of America*, March 1999.