

## **Voice powered Telematics: a human factors view**

Since the invention of the telephone, technology has transmitted the human voice to more corners of the planet than any other form of communication. Now with the convergence of data-enabled cars, voice seems to be leading the way toward improved co-ordination of voice and data information on four wheels. But the difference between a good voice interface and a bad one can mean the difference between delighting or insulting your customers.

While human speech is a natural component of our everyday lives, voice interfaces are only just catching up. The learning curve for speech recognition and voice interfaces is steep. The list of challenges includes dealing with a) background noise, b) multiple sound sources, c) variations in accent, d) natural conversation, e) manners

and etiquette, f) error handling, g) listening and learning to name a few. Sound, voice and language are emerging as a tricky problem due to the large number of variables involved.

Voice activated information on demand is set to transform the way customers drive their data needs in the automotive environment. New breakthroughs in speech recognition and voice interfaces from companies such as One Voice Technologies, IBM, Lernaut and Hauspie Speech Products, Conversay, i-Voice, MobileAria, Speech Works, Nuance, Sensory Inc., Advanced Recognition Technologies, Visteon and Delphi Automotive Systems among others are on the horizon. Already companies such as Fonix have cemented partnerships with Motorola and telematics provider Sensoria Corporation that are expected to saturate the telematics market with their voice interface products in 2002. Currently voice products by Nuance and General Magic power GM OnStar's popular

Virtual Advisor, while Visteon outfitted Jaguar S-Type vehicles as early as 1999, with plans for more voice control in the 2002 Infiniti Q45 vehicle.

While all of this looks most promising, voice interfaces of the future will have to undergo user acceptance. And the track record to date has not been without criticism as many can attest to the familiar "I'm sorry, I didn't understand you, please try again" response. A report last year by Forrester Research echoed the concern. "Speech recognition averages one error per sentence," according to analyst John Dalton in his study called "The Web's Speech Impediments." Moreover, voice telematic providers view the presence of background noise from traffic, wind, engine, radio, rain and passengers as the biggest practical issue facing the industry today.

The background noise issue is being tackled in different ways. Wavemakers of Vancouver, BC approaches the problem by

canceling out background noise and reducing the amount of interference from the voice signal. The Wavemakers voice optimization software is called Clearstreams and powers Intel's StrongARM processing platform. According to Wavemakers' CEO Peter van der Gracht, the next challenge is to target specific speech while canceling out unwanted background voices.

Another company, Clarity of Troy, Michigan takes a completely different approach to the speech interface problem. Clarity focuses on extracting the voice signal from the host of sound sources drawing out speech as opposed to blocking out unwanted noise. Clarity's proprietary CVC technology extracts the users voice from a mixture of sounds, including background noise, music and multiple speakers. Clarity's president and CEO Raymond Gunn believes that microphones passively hear sounds as opposed to actively listening for selective sounds that provide interpretive clues.

Voice interfaces that actively listen to human conversation turn out to be extremely valuable in the telematic quest to navigate by voice. In a format more suited to human speech, virtual personal assistants listen and "adjust" to the needs of the user. Companies such as Webley and Wildfire offer a "hands free, eyes free" way to navigate voicemail, read and send email and retrieve faxes. Wildfire (a sister company of Orange) even adapts to your needs gradually training "herself" to better assist. "She learns with you first by prompting you, then gradually by shortening options and offering hints and new features related to your needs along the way," says Steve Rothschild, Senior Director of Marketing for Wildfire. The service is also sensitive to local cultural aspects such as speech mannerism, type of voice, local grammar and politeness.

Telematics providers are now looking to their partners to offer them the most user-friendly and

pleasurable voice interfaces. There is no substitute for bad navigation with speech recognition. Frustration and confusion are likely to be intensified more so than say with bad website navigation. As Forrester put it: "Voice technologies, whether automated or live, face a host of technical and human factors barriers that limit the range of viable applications". Fortunately the complicated issues involved in human speech are being addressed by innovative approaches that will keep the information conversation flowing.



Frank Spillers, MS is principal and Co-CEO of Experience Dynamics, a leading human factors and design



Dec-Jan 2001

consultancy that specializes in usability and ease of use for web, wireless and telematics design. Experience Dynamics is currently conducting exclusive telematics user needs field studies. The company is based in Portland, Oregon. Contact Frank at (503) 978-9183. [frank@experiencedynamics.com](mailto:frank@experiencedynamics.com)



Jan 2001

Dec-