

PRODUCT REVIEW



Separating voice from noise with Wavemakers

Not too long ago talking to yourself was considered taboo. Now therapists tell us it's a good thing, and we appear to be doing it all the time with our cell phones - Frank Spillers.

In the near to distant future voice commanded automotive interiors will be considered normal too. Voice activated interfaces will open up new frontiers in human flexibility, speed, responsiveness, and possibly overwhelm and frustration. At least that's the promise.

"The interface is always the defining factor in telematics, it has to be enjoyable and efficient or the user won't adopt or adapt well" says Peter Van der Gracht, CEO of Wavemakers (Vancouver B.C., Canada). But the

How does it work?

Wavemakers Clearstream™ technology clears background noises by detecting human speech and improving signal to noise ratio. What appears to be unique about Wavemakers is that the software imitates the human brain, extracting speech sources from an array of human and non-human speech much like how you focus on one speaker at a noisy party.

Wavemakers low cost software

New research funded by the Canadian government will allow Wavemakers to advance the state of the art of the technical side of the speech problem by "targeting" speech. This new technology will extend the intelligence of the microphone and SRE by not only localizing the direction of the speaker but by separating voice types and allowing a single speaker, amongst other speakers to control the speech interface.

Wavemakers takes the focus

end user interface is only half of the story. “Quality speech processing by the microphone and the speech recognition engine (SRE) can vastly impact the quality of experience for the end-user”.

Making the speech machine smarter?

A good interface in a poor acoustic environment with today’s microphone and SRE’s can be disheartening. The issue is not noise per say, but the strength of the algorithms and software technologies SRE’s and microphones employ to tackle the overall coherency and intelligibility problem. Wavemakers takes things a step further by boosting SRE intelligence with human voice detection including elevating microphone accuracy by forcing the microphone to “point” at the speech source of the speaker.

The resulting speech enhancements provide a low cost “mini-brain” for the speech engine. The practical

models the structure of conversation rather than the actual contents of the dialogue. The process involves what Co-Founder and CTO Pierre Zakarauskas calls “speech reconstruction”. Reconstruction refers to the way algorithms recreate the signals that are lost due to noise, gaps, pauses, echoes, and dead spaces in speech with “information” that results in giving the driver rapid SRE response, hence maintaining perception of coherency between driver and interface.

For improved fit, the software is optimized for the unique acoustic environment that characterizes different vehicle designs and the SREs – such as IBM, Philips, ART, Fonix, Microsoft and SpeechWorks engines.

Will Clearstream™ make waves in human to computer communication?

Technology such as Wavemakers solve one major piece of the speech interface puzzle. For instance, a

away from the technical problem of coherent speech input, noise and SRE processing. In essence, this frees the Tier One player to focus on the other side of the speech problem: branding an elegant and coordinated automotive user experience that blends ease of use and pleasure across speech, tactile and visual displays. Currently more attention is paid to how things look and feel in the cockpit and not how they sound. Lack of attention to the usability of the speech user interface can jeopardize the speech interface investment. Lessons from the last 15 years in speech interfaces accentuate the need to win consumer trust more than ever. “Wavemakers’ products have been shown to reduce the error rate of the speech recognition engines by 85%, says Richard Sones, Director of Sales and Marketing, but one should not lose sight of the human factors of the speech interface. In the end, it is the overall end-user experience that is the key to success”.

benefits for the driver or passengers include: stabilized volume anywhere in the car regardless of loudness or softness; filtered human speech ignoring all other non-human signals and decreased user wait time for SRE processing.

simple task such as hands-free phone number dialing in a noisy cab captures a 95% accuracy rate with Wavemakers enhancements. More complex tasks will require better interface design that account for human error.

*Frank Spillers is a Senior Human Factors Engineer and Principal of Experience Dynamics, a Portland, Oregon based usability and user interface design consultancy.
frank@experiencedynamics.com*