VOICE-ACTIVATED USER INTERFACES IN DRUG DELIVERY DEVICES CAN ENHANCE COMPLIANCE AND OUTCOMES

Here, Salvatore Forte, Innovation Manager at Flex, looks at the evolution of voice-activated interfaces in medical devices, particularly autoinjectors, and the advantages for patient compliance and quality of care.

One of the major challenges that medical device makers face is properly leveraging technological innovations to design medical devices that are more patient-centric and can ease the patient experience when self-administering at home. One of the main drivers for designing more intuitive medical devices has been the evolution of user interfaces, particularly for automated drug delivery devices. In the last decade, battery-operated autoinjectors have progressively transitioned from having bulky mechanical buttons to large touchscreen displays as the primary user interface in a conscious attempt to create a smartphone-like user experience. But graphical user interfaces do not necessarily lead to devices that are simpler to use and, if not carefully designed, can ultimately overwhelm the patient and limit device acceptance.

Many industries have been positively affected by voice technology, with voice assistants widely deployed in smart products across multiple markets. Everyone is accustomed to smart consumer products with voice recognition capability in their home. Moreover, the covid-19 pandemic has dramatically increased the demand for touchless interfaces to limit spreading infection, with voice clearly becoming the primary method by which to deploy hands-free interaction with any sort of device.

This is paving the way to implementation of voice-enabled drug delivery devices, with new autoinjectors that could be equipped with a voice user interface (VUI). Delivering a more user-friendly human-machine interaction would transform at-home care. The patient can talk to an autoinjector and expect it to execute specific tasks according to the spoken commands, such as starting

> "The covid-19 pandemic has dramatically increased the demand for touchless interfaces to limit spreading infection, with voice clearly becoming the primary method by which to deploy hands-free interaction with any sort of device."

or suspending the injection and adjusting speed settings. Additionally, the voice user-interface could be used to establish a dialogue with the autoinjector, which, via miniature speaker, could query the patient for health conditions either before or after the injection and collect the feedback by voice (Figure 1). This would set new autoinjectors aside from traditional drug delivery devices, supplementing them with patient monitoring capabilities. The autoinjector could record data on how the patient responds to therapy, and the physician could use those contextual insights to adjust the treatment, if necessary, in real-time.

Of course, there are challenges with voice-controlled autoinjectors, and this article addresses the technological solutions to achieve optimal audio performance by careful design. Machine learning, sensor fusion and audio edge-processing can, together, ensure accurate voice-detection as well as mitigate risks of errors in noisy environments.

CHALLENGES AND TECHNICAL SOLUTIONS

Embedded Voice-Recognition for Edge Processing

The are several challenges at the system-level to fulfil a successful voice-recognition implementation that not only works well

Salvatore Forte
Innovation Manager
T: +39 02 8732 9218
E: salvatore.forte@flex.com

Flex
176 Via Ernesto Breda
20126 Milan
Italy

www.flex.com

Figure 1: Autoinjector voice-interface. Assisted injection process and collection of patient feedback on health status.

"In recent years, advancements in artificial intelligence and machine-learning algorithms have been driving innovation in the voice processing space, becoming key to enabling embedded voice implementation."

from the user viewpoint but also meets the technical requirements for integration in battery-powered autoinjectors. These are normally built around small microcontrollers, with limited processing capability and memory resources. When analysing the spectrum of voice-enabled products currently available in the marketplace, one can easily confirm that the majority are based on complex voice solutions that can interpret full natural speech. They require high computational power and therefore can only operate remotely in the cloud.

However, in recent years, advancements in artificial intelligence and machine-learning algorithms have been driving innovation in the voice processing space, becoming key to enabling embedded voice implementation. Today, highly efficient keyword recognition models can fit in small microcontrollers (e.g. the ARM Cortex M) to assist domain-specific device tasks. They can recognise up to dozens of command words, giving them an adequate vocabulary with which to deploy meaningful use cases for injection devices, such as controlling the operation of the built-in device's motor.

"Voice-enabled autoinjectors must be able to recognise words spoken by the user with a high level of accuracy, even in the presence of background noise, and mitigate the risks of errors that might lead to hazardous consequences for the patient."

Such machine learning models can run as embedded software while requiring minimal computational power (as low as 100 MHz) and memory footprint (down to hundreds of kB) to function. Moreover, data can be processed entirely at the device-edge, requiring no connection to external voice services in the cloud. This solves data privacy and security concerns by keeping users' voice data locally on the device and never transferring it out. This eliminates the need to embed connectivity hardware and services in the device, which ultimately leads to lower system cost.

Command sets and languages can be personalised upon training the machine-learning model to fulfil the needs of different market regions. It is also possible to have a single VUI that natively implements multiple language models, rather than having single product variants that only support one language. Command models corresponding to each language can be deployed as binary files that are compiled together with the main application software and executed as single firmware on the device's microcontroller.

Audio Front-End Design for Ambient and Motor Noise Cancelling

Reliability is equally important, and commitment to safe products remains essential in the healthcare industry. Voice-enabled autoinjectors must be able to recognise words spoken by the user with a high level of accuracy, even in the presence of background noise, and mitigate the risks of errors that might lead to hazardous consequences for the patient. This is not trivial because electromechanical autoinjectors are normally exposed to a high level of interference noise, coming not only from the environment (home settings are often far from quiet) but also that generated by the device itself – mostly the motors and the movement of any associated mechanical gears. The latter adds to environmental noise interference and can drastically overwhelm the voice content. Obviously, VUIs need a clean speech signal to correctly recognise voice commands.

To address this, it is crucial to look at the solution holistically. Designers must figure out how to improve the acoustic performance of the audio-capturing system as a whole. This can be accomplished with proper arrangement of both hardware and signal processing software technologies, which must operate seamlessly on the device to ensure accurate voice recognition performance. Starting with the hardware (Figure 2), the autoinjector can accommodate multiple microphones to implement audio beam forming and enable
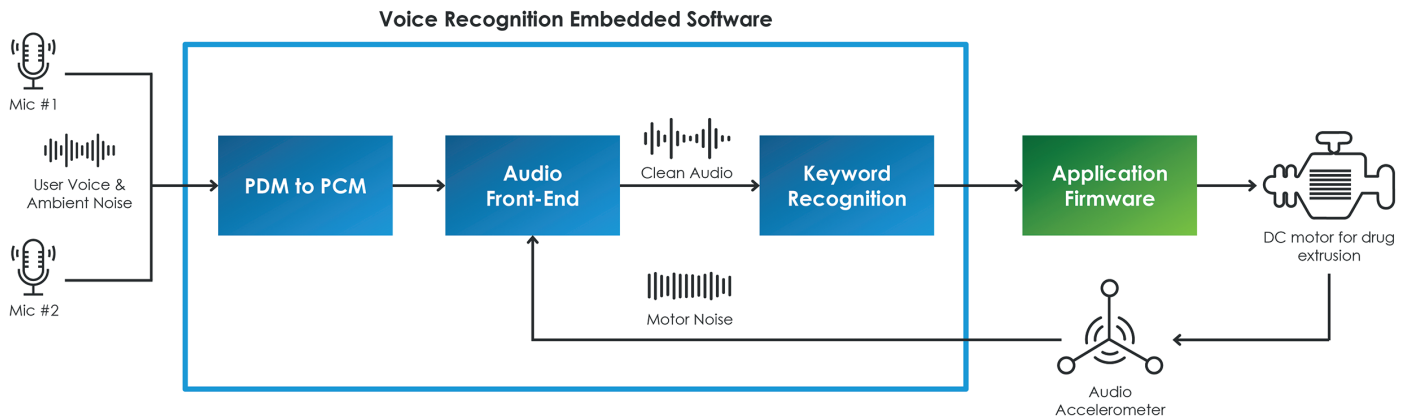
Figure 2: Block diagram of the voice-processing system.

directional voice capturing. There is no need to push for complicated hardware architecture, and two microphones are more than sufficient to achieve reliable source localisation to help to discriminate voice from sounds coming from other directions. In addition, a micro-electromechanical accelerometer can be used to perform some sensor fusion and implement a robust noise-cancelling strategy. An accelerometer can pick out the noise that propagates as mechanical vibrations through the autoinjector's plastic enclosure, and which is generated by the motorised needle and drug extrusion systems.

The software framework is just as important to get best performance from the voice-enabled system. The acoustic front-end (AFE) must reject all ambient and motor noise to deliver a clean audio signal with enhanced voice clarity to the keyword recognition engine. The AFE is made up of a suite of algorithms that are used together to pre-process the audio input stream (both the microphone and accelerometer data sets), including data-format conversion, digital filtering, automatic gain control and adaptive interference cancelling, among others. Response time is crucial for accurate performance, and the AFE can be

parametrised and hand-tuned to adjust the response according to the specific interference sounds that the device would normally experience throughout its operation. These algorithms must quickly adapt to changing noise conditions, and recover quickly from instantaneous transition that happens, for instance, when the device is changing injection speed, which obviously would result in a different signal's frequency spectrum.

ENGINEERING CHARACTERISATION: KEYWORDS-RECOGNITION ACCURACY PERFORMANCE

Figure 3 shows the spectrograms of the audio signals recorded by a reference platform of an electromechanical autoinjector that embeds a complete VUI. The device can respond to a predetermined set of keywords. In this example, the autoinjector was being interrogated with a set of six different vocal commands (*Start Injection, Suspend, Continue, Slow Down, Speed Up, Abort*) that were spoken by the user while the device's drug-extrusion motor was powered to run at full speed. This verified the accuracy of the system in recognising the keywords while the device was exposed to the worst-case noise scenario. The measured sound pressure level of such background noise was 68 db(A). Moreover, the tests were performed with the AFE both disabled and enabled to evaluate how it affects the statistics while it is executed to clean up the audio input stream and reject background noise. Figure 3A shows what the microphones sense when the AFE software is disabled, with noise spread across the full spectrum and non-distinguishable voice content. The keyword recognition performance is negatively affected by the presence of such high-level noise, with a success rate of 52%
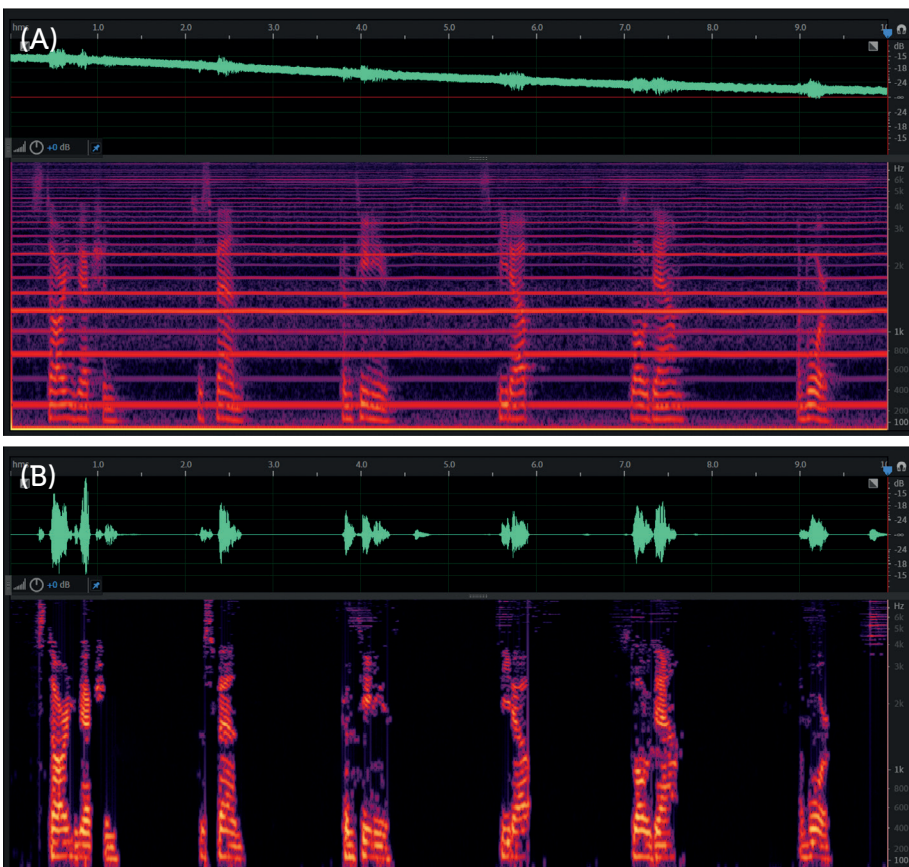

Figure 3: Spectrogram of user's voice without (A) and with (B) AFE for noise cancelling. Autoinjector motor operating at maximum speed (68 db(A) noise level).
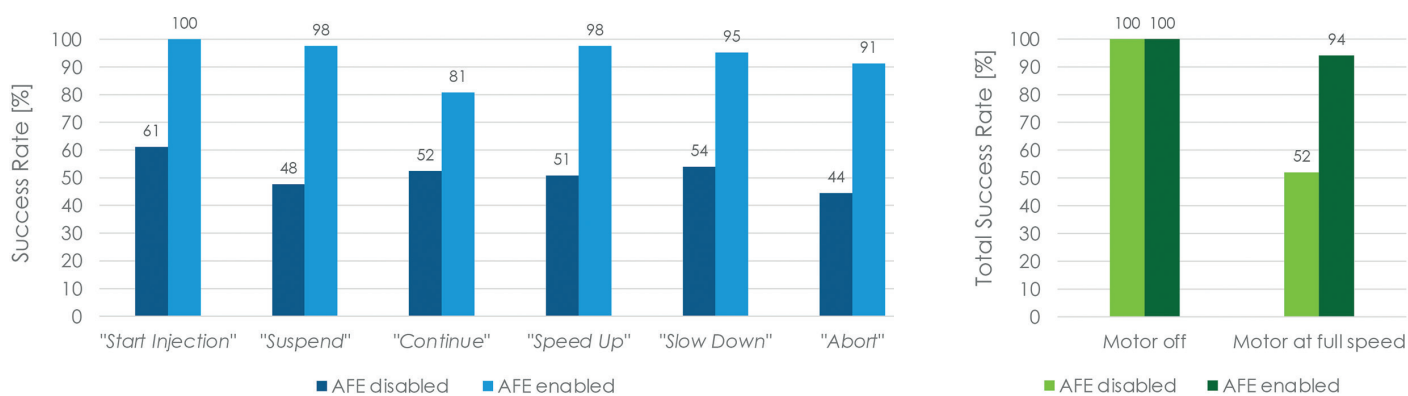
Figure 4: Keywords recognition accuracy. Test performed on 20 US-native subjects (ten male, ten female) with autoinjector's motor operated at full speed (noise level 68 dB(A)).

if no signal processing is performed. Conversely, Figure 3B shows what happens when the AFE is enabled to clean up the audio input stream, with the noise components that were visible before now completely eliminated, thereby providing the keyword recognition model with intelligible user voice. The keyword recognition success rate of the voice engine greatly improves when the AFE is enabled, resulting in a score of 94% (Figure 4).

CONCLUSIONS

Autoinjectors are often perceived as cumbersome devices and difficult to deal with for self-administration therapy at home. This tends to limit compliance and adherence to the treatment, so there is a clear demand for simpler-to-use devices with more user-friendly interfaces. Voice-recognition technology can address this demand. Using a VUI enhances device usability and

delivers a more engaging patient experience, which, in turn, promises to improve patient compliance and the overall quality of care. Technology advancements in artificial intelligence and low-power processing have enabled lightweight implementations that fit minimal hardware platforms with small microcontrollers. This sets the stage for successful implementation of new voice-enabled autoinjectors.

ABOUT THE COMPANY

Flex provides sketch-to-scale solutions, delivering innovative design, engineering, manufacturing, real-time supply chain insight and logistics services to a wide range of industries. Flex Health Solutions is a global leader in the design and manufacture of medical products for pharmaceutical and medtech companies. This includes the design and commercialisation of more than 75 regulated medical devices, from pens

and autoinjectors to pumps and inhalers. The company's approach is supported by FDA-registered and ISO 13485-compliant facilities and a world-class quality system.

ABOUT THE AUTHOR

Salvatore Forte is Innovation Manager at Flex, leading an R&D engineering team in the company's Design Center in Milan. Mr Forte has a solid technical background in analogue and digital electronics, specialising in embedded system design for power-sensitive medical devices, such as wearable health monitors, disposable point-of-care and automated drug delivery. He holds an MS degree in Electrical Engineering from University of Naples Federico II (Italy).