

THE MOBILE COGNITIVE ASSISTANT: BRIDGING THE GAP BETWEEN IN-CAR AND OUTSIDE-THE-CAR EXPERIENCES

Dr. Stefan Ortmanns, Dr. Udo Haiber

Nuance Communications, Inc.

By 2020, there will be 26 billion intelligent, capable, connected devices armed with conversational virtual assistants that manage nearly every possible consumer experience. But already today one of the biggest consumer challenges in our more and more connected world is the need to learn and remember the specific capabilities and vocabularies of multiple assistants spread across different services and devices.

The next generation of the mobile cognitive assistants for connected (and eventually autonomous, electric and shared) cars are determined by number of indispensable prerequisites. These so-called automotive assistants need to be always available and not limited by connectivity constraints, but based on a robust hybrid, embedded-cloud architecture, taking consideration for data privacy needs. At the same time, they need to be intelligent and knowledgeable, leveraging machine learning and contextual reasoning that enables more personalized and contextual results as well as collaborative interactions with assistants in order to complete more complex, realistic tasks with higher accuracy. They need to interact with drivers as well as passengers using latest advancements in speech signal enhancement technologies paired with reliable voice biometrics technology differentiating between multiple users. And last, but not least the automotive assistant needs to intelligently give drivers seamless access and control to third party assistants and chatbots inside and outside the car using latest cognitive arbitration capabilities. In the following speech we will elaborate the underlying technologies and capabilities to bring the mobile cognitive assistant to live.

1. Hybrid AI Stack

For the automotive environment, we believe in a holistic experience for the driver – in a world with multiple apps, assistants, services and content providers, consumers need the ability to have seamless access to the information and content that matters inside and outside the car. The user experience in the car relies on a tight integration with the car itself.

The basis for the mobile cognitive assistant is Dragon Drive, our customizable, conversational AI platform for the connected car. It's hybrid AI technology stack tightly integrates the required speech and interaction technologies, including speech signal enhancement (SSE), embedded and cloud speech recognition and natural language understanding (NLU) technologies, text-to-speech (TTS) synthesis, further multimodal interaction technologies as well as dialogue management.

Powered by artificial intelligence, Dragon Drive's conversational automotive assistant listens, understands and responds to drivers. It understands and learns the needs and preferences of drivers and passengers over time to provide a personalized experience, including access to entertainment, navigation, points of interest, news feeds and in-car features such as heating and air conditioning.

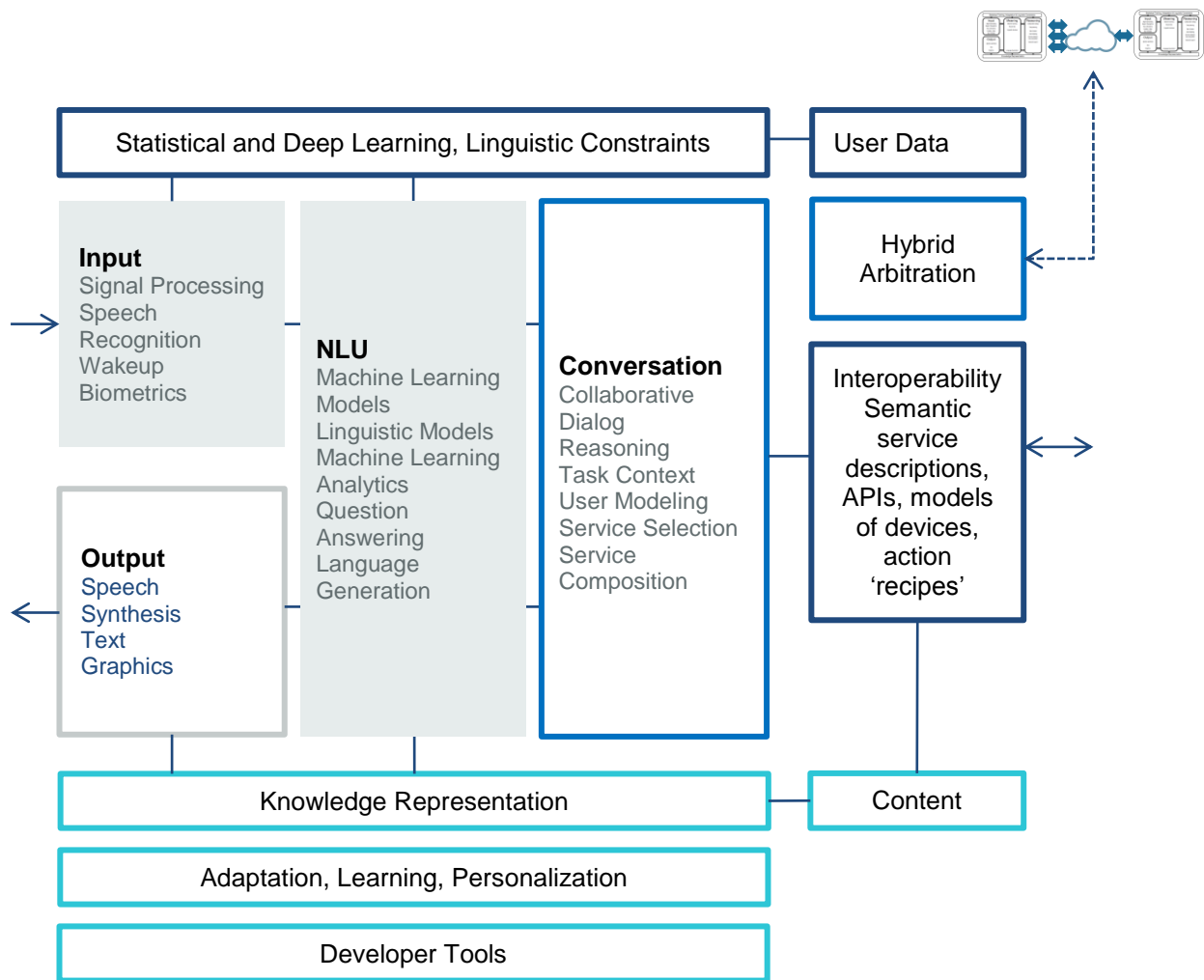


Image 1: The Foundation: Nuance Conversational AI Platform

Several Artificial Intelligence (AI) technologies come into play.

- Personalization means that the system learns certain likes, dislikes, and preferences of individual users, and uses this knowledge to make smart recommendations that match the driver’s needs.
- Contextualization means that the system considers the current situation and circumstances together with general world knowledge, which also often affects the decision for or against a certain option. To realize contextualization, Nuance uses a knowledge- and rule-based approach based on reasoning — an AI technology that is different from the machine learning approach deployed for personalization.

Together, both technologies significantly improve the user experience and the driving safety by delivering more relevant results and thus helping the driver to make better and faster decisions.

More and more, car sensors provide the most relevant context information for all kinds of driver interactions and driving related tasks and the integration of car sensor information into the interaction will significantly advance the user experience. For example, multimodal experiences combining gaze, gesture and voice control require tight integration with the vehicle interior.

2. Cognitive arbitration

Hybrid, cognitive arbitration describes a new artificial intelligence (AI)-powered capability that solves one of the biggest consumer challenges in today's connected world: the need to learn and remember the specific capabilities and vocabularies of multiple assistants spread across different services and devices. Nuance's cognitive arbitrator solves this challenge by seamlessly connecting and integrating disparate virtual assistants, third-party services, and content via a single interface that spans the automotive, smart home, and Internet of Things (IoT) ecosystem to complete complex tasks and enhance the user experience.

This principle of a master virtual assistant is central to Nuance's vision of a world in which multiple virtual assistants work together to benefit end users through intelligent and effective conversational AI as an integral part of the OEM brand and user experience. The cognitive arbitrator creates a singular, intelligent hub that listens, understands and routes a user's request to the specialized assistant or content service best suited to accomplish the task.

For example, a driver can talk to the in-car assistant to request driving directions and streaming music, and make requests that will be routed to other third-party assistants that handle tasks such as shopping, food ordering, personal banking, and more.

The cognitive arbitrator which can be operated embedded in the vehicle, in the cloud or as a hybrid solution, supports both explicit and implicit understanding. In the case of explicit invocation, users trigger a specific assistant by calling the name. In an environment of multiple assistants, however, it may be preferable for the user to not have to think about what assistant to call for a given request. The system can understand user preferences and the environment to route to the most appropriate virtual assistant or bot for a given command and context.

Regardless of the use of implicit or explicit command forms, Nuance's Wake-Up-Word or Just Talk technologies can be used for an improved user experience such that the user no longer has to press a push-to-talk button to initiate an interaction with the system.

The cognitive arbitrator learns user preferences to route the user to the assistant that best matches his personal needs (personalization) as well as contextual reasoning in order to route the user to the most appropriate assistant in the specific situational context.

We distinguish three levels of cognitive arbitration:

- Request driven: The user activates a specific assistant by explicitly invoking the assistant using a key word or an NLU intent. The routing is based on user preferences.
- Capability driven: The routing is based on a deeper understanding of the virtual assistant capabilities and strengths. The cognitive arbitrator fuses multiple applications together as needed.
- Consciousness: The cognitive assistant acts as a single repository of all task and dialog contexts spanning multiple virtual assistants and fuses them together

3. Multi-zone automotive assistants

Automotive assistants are expected to assist both drivers and passengers. Even more so, the assistant should be able to identify different users in the car, allow them to interact simultaneously with the automotive assistant and support them accomplishing different tasks independently from each other.

Dragon Drive identifies different users in the car and where they are seated and offers each user a personalized experience. For example, any passenger can say, “it’s a little cold in here.” Thanks to seating awareness and voice biometrics, Dragon Drive can respond by adjusting the climate only in the relevant zone, applying preferences of the identified user.

This sets considerable requirements to the speech signal enhancement (SSE) solution:

- Our comprehensive all-in-one solution improves hands-free speech communication and speech recognition results in noisy and adverse environments. The SSE eliminates off-axis interferences, supports passenger interference cancellation, barge-in, multi-zone processing, beamforming and 360° far-talk recognition.
- Beamforming and post-filtering separates unwanted signals such as interfering speech and background noises from the desired speech signal to achieve a high degree of directionality with a low number of microphones.
- Deep learning improves ‘always on’ and multi-zone processing by reducing background noise significantly w/o missing true speech

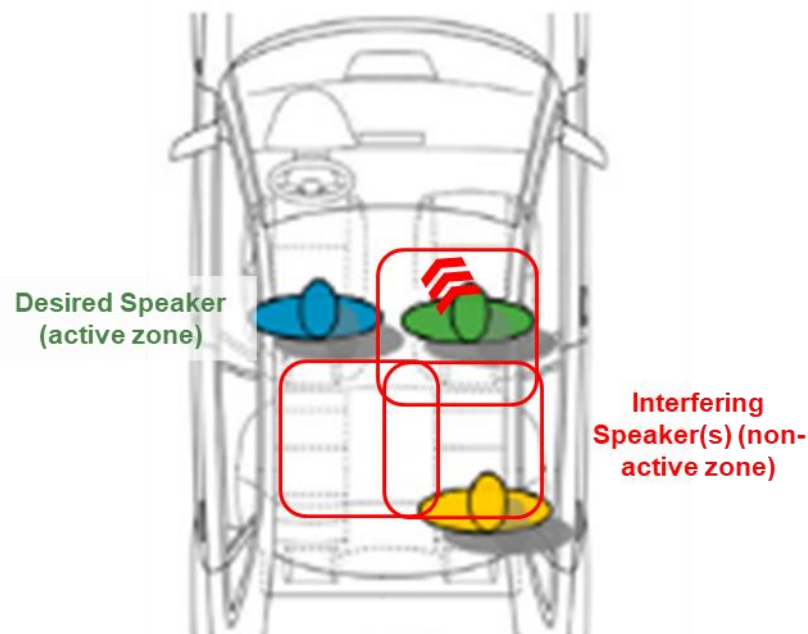


Image 2: Multi-zone processing

As multiple users interact with the cognitive assistant, it is also important to define smart techniques for the users to activate the assistant.

- **Wake-up Word:** Just like for a single user, a so-called wake-up word can be defined for the multi-user setup. When the automotive assistant ‘hears’ the wake-up word, it becomes active immediately and starts a speech session. For example: “Hello Dragon, what’s the weather in Frankfurt?”
- **Barge-in:** The barge-in feature, is closely connected to the wake-up word: The automotive assistant is alert even while speaking, so users can interrupt long prompts and thus shorten the interaction. As with the wake-up word, all passengers have access to this feature and can take over the lead in a voice interaction.
- **Just Talk:** With “Just Talk” mode activated, the cognitive assistant is standing-by for a command from the driver or passengers, without needing to be explicitly addressed by name or wake-up word or invoked by the push of a button.

4. Multimodality

A truly multimodal solution implies that the user can use multiple modalities that work together to accomplish one task. Dragon Drive now tightly integrates conversational artificial intelligence with non-verbal modalities enabling drivers to get information about and interact with places outside the car simply by looking at them. They can ask questions like, “What are the opening hours of that shop?” “Call that restaurant” or inquire about a parking garage asking, “How much does it cost to park there?” Dragon Drive can resolve the point of interest and provide a meaningful, human-like response.

The new technology, eye gaze detection based on a camera tracking of the eyes, enables users to simply look at the object in question and refer to it through conversational AI. The technology is imitating human behavior, as humans are very good at guessing where somebody is looking just by observing his or her eyes. Machines can also do this based on image recognition and Deep Learning, capabilities which will bring us into the age of truly multimodal assistants.

About Nuance Communications, Inc.

Nuance Communications, Inc. (NASDAQ: NUAN) is a leading provider of voice and language solutions for businesses and consumers around the world. Its technologies, applications and services make the user experience more compelling by transforming the way people interact with devices and systems. Every day, millions of users and thousands of businesses experience Nuance’s proven applications. For more information, please visit www.nuance.com.

About Nuance Automotive

Speech recognition, NLU, AI and predictive touch solutions from Nuance have pioneered many of the personal assistant technologies and intelligent systems in the devices we use every day from the world’s leading brands – including mobile devices, cars, televisions, wearable devices, and the emerging ecosystem of the Internet of Things. We deliver a more human experience with technology, keeping consumers better connected and informed – consistently adapting to and predicting their needs. The Nuance Automotive business delivers automotive-grade solutions

enabling drivers all over the world access to information and services and providing the safest, smartest and most natural user experience.

Dragon Drive, Nuances connected car platform, powers more than 200 million cars on the road today across more than 40 languages, creating conversational experiences for the world's leading car brands. To learn more about the ways in which Dragon Drive is amplifying the intelligence of the connected car, visit [Dragon Drive on nuance.com](https://www.nuance.com/connected-car).