HEAD acoustics

DATA SHEET

APPLICATION EXAMPLES INCLUDED

Code 7970

VOCAS

Voice Control Analysis System

OVERVIEW

VoCAS

Code 7970

Voice Control Analysis System

Automatic Speech Recognition (ASR) systems find widespread application in telephone networks, cars, mobile appliances, multimedia, and IoT devices such as smart speakers. Performance of the ASR mainly depends on the variety of talkers, languages, dialects, as well as environmental conditions. Thus, testing with a lot of variations is beneficial for developing and improving ASR systems. On the other hand, variation makes testing more time-consuming and exact reproduction more difficult.

The VoCAS voice control analysis system organizes and automates testing of ASR systems and devices. It offers easy setup and enables comprehensive, variation-rich testing under realistic and fully reproducible conditions.

KEY FEATURES

VoCAS Core

› Basic framework of VoCAS to built up the application with VoCOPTs according to user requirements
› Test sequence template for creating test cases to test and train ASR systems (application of customized audio sources requires VoCAS options)
› Remote control of appropriate background noise and reverberation simulation software
› Remote control of the HRT I turntable

VoCAS Options

› Recording and editing customized audio files
› Creating databases of customized audio sources for application in VoCAS
› Distracting and testing ASR systems with concurrent talkers
› Automated and accelerated test sequences with Python scripting
› Remote control of VoCAS via REST API
› Batch generation of audio files with different parameters to train or test ASR systems

APPLICATIONS

Reproducible testing of speech recognition systems in realistic simulated environments (complex scenarios with multiple talkers, background noise, acoustic environments)

Meaningful benchmarking of different speech recognition systems

Recording of audio files

DETAILS

VoCAS is a test software for speech recognition evaluation. It supports the optimization and benchmarking of any device incorporating speech signal processing and speech recognition.

Speech recognition systems are an essential component of voice-controlled IoT devices such as smart speakers, smartphones, tablets, and multimedia systems in vehicles. VoCAS helps evaluating the speech recognition quality realistically and objectively of appropriate devices in complex scenarios. The software considers all factors that significantly influence the quality of speech recognition systems: the acoustic environment, background noise, different talkers, and different languages or accents. Thus, with the help of reproducible test sequences adapted to the respective applications, manufacturers receive essential information for optimizing their speech recognition systems.

VoCAS provides multiple features for various applications. The availability of all features divides into the fundamental core version (VoCAS Core) and several optional features (VoCAS Options, VoCOPTs). The VoCAS Options extend and increase the scope of application.

PRODUCT STRUCTURE

VoCAS Core

VoCAS Core is the basic framework of VoCAS. It contains necessary functions to create and run testcases, specify playback conditions, and control background noise or reverberation simulation remotely. The presentation of results from the testcases can be customized according to the users preferences. VoCAS Core requires VoCAS options for recording, importing and applying customized audio sources (SQLite database), and other advanced applications.

VoCOPT – VoCAS Option

VoCAS Options – short VoCOPTs – extend VoCAS Core with additional features. These features fit to special user requirements. Using any VoCOPT requires VoCAS Core. VoCOPTs can be freely combined.

VoCOPT 1 – Signal Recorder

Record customized sentences in *.dat audio file format. Additionally, the recorder has editing functions for customization of recordings with individual tags (meta data) and individual file names. The recorder interface presents highlighted text consecutively to
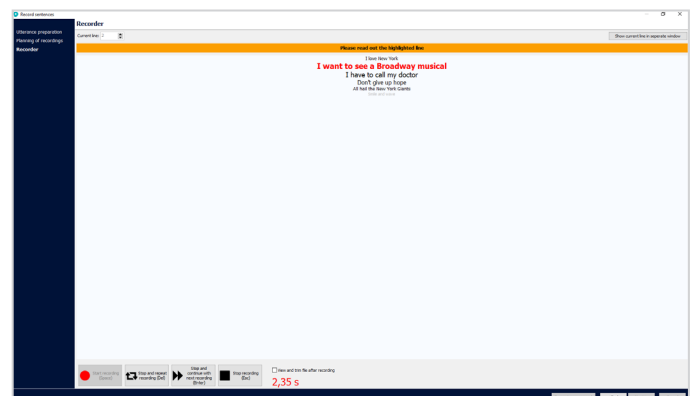

the speaker to read it out loud. There is an extra interface to present solely the text for reading out loud. Further, each recording may be reviewed and trimmed if necessary. The recorder wizard guides step by step from preparation and planning to the execution of the recordings.

VoCOPT 2 – Import/Source Pool Generator

Import and generate a customized audio source pool (SQLite database) from existing audio files. The provided wizard guides through preparation steps such as tagging and file conversion

parameters to the eventual import and compilation of the files into an audio source pool (database). The supported audio files range from exclusive HEAD acoustics formats (*.hdf/*.dat) to
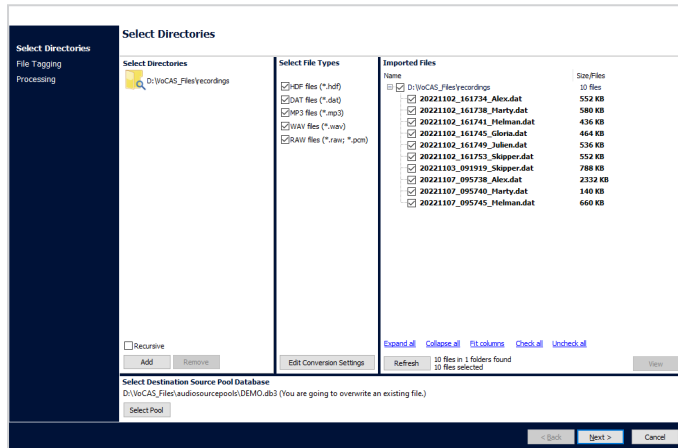

common audio file formates (*.mp3/*.wav/*.raw). A recursive search provides the quick presentation of all appropriate files in a specific directory. The tagging editor allows both manual and automatic tagging of the files. Before importing the files to the database, they may be altered by various available post-processing actions.

VoCOPT 3 – Python Scripting

VoCAS provides a script editor to write customized Python scripts. Apply and execute these scripts in the sequence template of a testcase. Apply Python scripts for automation of various processes in VoCAS.
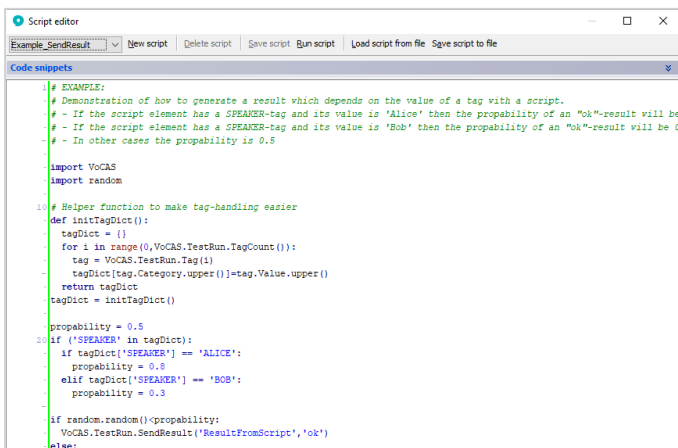

VoCOPT 4 – Multitalker
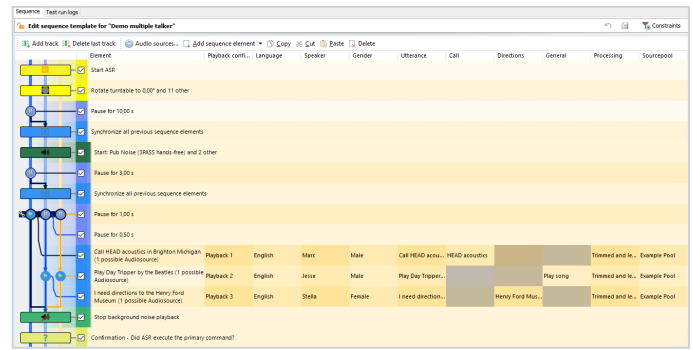
Apply up to four simultaneous talkers to the sequence template. The sequence template provides different tracks for each talker. Thus, the playback of different talkers can start and run shifted but concurrent. Specify different playback configurations, e.g., gain, Lombard effect for each talker.

VoCOPT 5 – REST Remote Control and Result Access

Control fundamental VoCAS functions remotely via REST API.
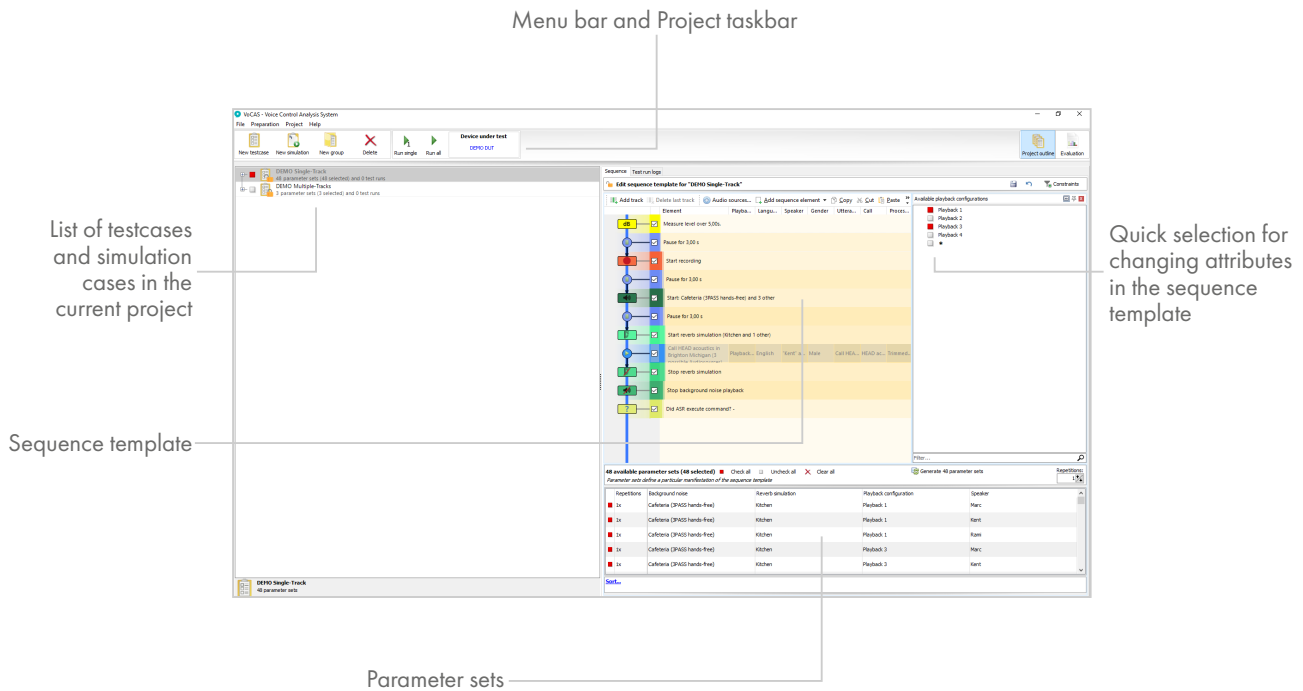
VoCOPT 6 – Simulation Mode

Create testcases for simulation. Simulated testcases run without actual acoustic audio output. Its use case is batching existing audio sources (speaker and background noise, reverberation simulation) into one file.
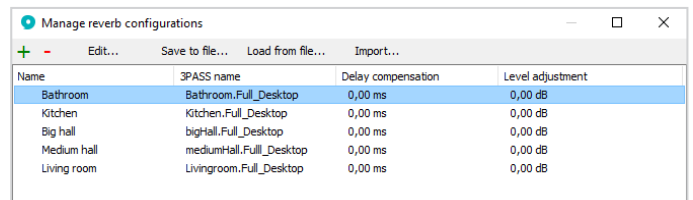
VoCAS simulates test environments. Based on an impulse response of a reverberant room, the software calculates the acoustic sum at the DUT position (talker, background noise, reverberation) and records the simulation in *.dat and/or *.wav file format. With the calculation of reverberant room's acoustics, these files are lifelike representations of actual acoustic testing in this room. Hence, VoCAS conveniently generates a large amount of clearly sorted audio test files in arbitrary variants.

The sequence template of the simulation mode is limited compared to normal testcases.

USER INTERFACE AND OPERATION

Menu bar and Project taskbar

List of testcases and simulation cases in the current project

Sequence template

Quick selection for changing attributes in the sequence template

Parameter sets

VoCAS presents testcases as a top-down flow diagram. The user can add and rearrange sequence elements as well as alter their respective settings to create the desired test scenario. Audio sources are the central element while the surroundings extend the test and evaluation conditions. The audio source files include meta data ("tags") for playback and are stored in databases. Around the sequence element of the audio source, there are plenty of options to extend the complexity of the testcase. VoCAS provides remote control interfaces for the background noise simulation software, a control interface for the HRT I turntable, and an evaluation editor for generating significant results. Combining the available sequence elements provides various possibilities to assess and evaluate the performance of ASR systems. The range of elements within the sequence template depends on the available VoCAS Options. VoCAS automatically generates all possible variations resulting from the current sequence template in a list of parameter sets. A parameter set includes the unique combi-

nation parameters from the sequence template such as speaker, playback configuration, background noise, or turntable angle.
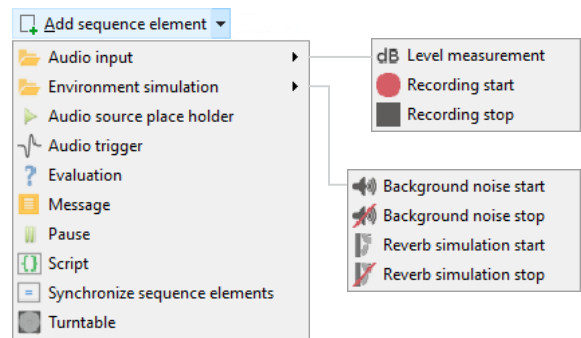
Thus, the list of parameter sets contains all respective variations of parameters from the sequence template. Once the sequence template is saved and locked, the testcase is ready-to-run. Each active parameter set runs after starting the testcase.

RESULTS AND EVALUATION

VoCAS provides a separate interface to sort and visualize the results from measurements in tabular form.

› Adjust the content of rows and columns.
› Distinguish results by allocating different colors from a broad range.
› Set constraints to view specific results isolated from others.
› Export and save the customized result table as Excel file.

OPTIONS

VoCAS Options

VoCOPT 1 (Code 7971)
› VoCAS Option Signal Recorder

VoCOPT 2 (Code 7972)
› VoCAS Option Import/Source Pool Generator

VoCOPT 3 (Code 7973)
› VoCAS Option Python Scripting

VoCOPT 4 (Code 7974)
› VoCAS Option Multitalker

VoCOPT 5 (Code 7975)
› VoCAS Option REST Remote Control and Result Access

VoCOPT 6 (Code 7976)
› VoCAS Option Simulation Mode

Hardware

HRT I (Code 6498)
› HEAD acoustics Remote-Operated Turntable

DSB IV.I (Code 2408.1)
› Digital Sound Board (substitutes *lab*CORE in the measurement configuration)

HMA V (Code 1420)
› Head Mouth Amplifier for HMS, 2 × 30 W, Class-D

Software

3PASS *lab* (Code 6990)
› Background noise simulation with fixed microphone position and loudspeaker arrangement

3PASS *flex* (Code 6995)
› Background noise simulation with flexible microphone position and loudspeaker arrangement

3PASS *reverb* (6996)
› Background noise simulation with reverberation simulation

HAE-BGN (Code 6970)
› Automated equalization for background noise simulation in labs

HAE-car (Code 6971)
› Automated equalization for background noise simulation in car cabins

GENERAL REQUIREMENTS

Hardware

*lab*CORE (Code 7700)
› Modular multi-channel hardware platform
*core*BUS (Code 7710)
› *lab*CORE I/O bus mainboard
*core*OUT-Amp2 (Code 7720)
› *lab*CORE power amplifier board
*core*IN-Mic4 (Code 7730)
› *lab*CORE microphone input board

One of the following Head Measurement Systems:

› HMS II.3 (Code 1703)
  » Head Measurement System, basic version with right ear simulator, 3.3 pinna, and artificial mouth
› HMS II.3 LN (Code 1703.1)
  » Head Measurement System, low-noise version with right ear simulator, 3.3 pinna, and artificial mouth
› HMS II.3 LN HEC (Code 1703.2)
  » Head Measurement System, low-noise version with human-like ear canal simulator right and artificial mouth
› HMS II.5 (Code 1705)
  » Head Measurement System, with 3.3 pinna and artificial mouth (w/o ear simulators)
› HMS II.6 (Code 1706)
  » Head Measurement System, with artificial mouth and free-field microphones (left and right)
› HMS II.7 (Code 1707)
  » Head Measurement System, with artificial mouth and free-field ICP® microphones (left and right)

Measurement Microphone
› For mouth equalization, measurement recordings, level measurement, or audio trigger

Continued on next page

GENERAL REQUIREMENTS

System Requirements

Computer
›   Multi-core processor
›   8 GB RAM (recommended: 16 GB RAM)
›   40 GB free disk space

Operating system (one of the listed)

›   Windows 11 x64
    »   Pro, Enterprise, Education; version 21H2 or newer; languages: US, Western European
›   Windows 10 x64
   »   Pro, Enterprise, Education; version 1809 or newer; languages: US, Western European

SCOPE OF DELIVERY

VoCAS Core (Code 7970) includes the following components
›   Setup DVD, including demo project and demo audio database
›   USB dongle

IN PRACTICE

APPLICATION EXAMPLES

Single Talker in Anechoic Room with Background Noise Simulation

Basic measurement configuration in an anechoic room. The presented configuration has a HMS II.3 (HATS) for playback of audio sources. *lab*CORE transfers the audio data from VoCAS to the HATS mouth loudspeaker. The DUT (e.g., smart speaker) is in front of the HATS. It stands on the HRT I turntable. HRT I rotates the DUT into different angles respective to the HATS. Therefore, different positions of the DUT microphone can be tested. 3PASS *lab* executes playback of realistic background noise via *lab*BGN and the loudspeaker arrangement according to ETSI TS 103 224.

Multiple Talkers in Anechoic Room with Background Noise Simulation

Advanced measurement configuration in an anechoic room. The presented configuration has three HMS II.3 for playback of audio sources. *lab*CORE transfers the audio data from VoCAS to the HATS mouth loudspeakers. The DUT (e.g., smart speaker) is in the center of the HATS arrangement. It stands on the HRT I turntable. HRT I rotates the DUT into different angles respective to the primary HATS. Therefore, different positions of the DUT microphone can be tested. The additional HATS simulate concurrent talkers to distract the ASR. 3PASS *lab* executes playback of realistic background noise via *lab*BGN and the loudspeaker arrangement according to ETSI TS 103 224.
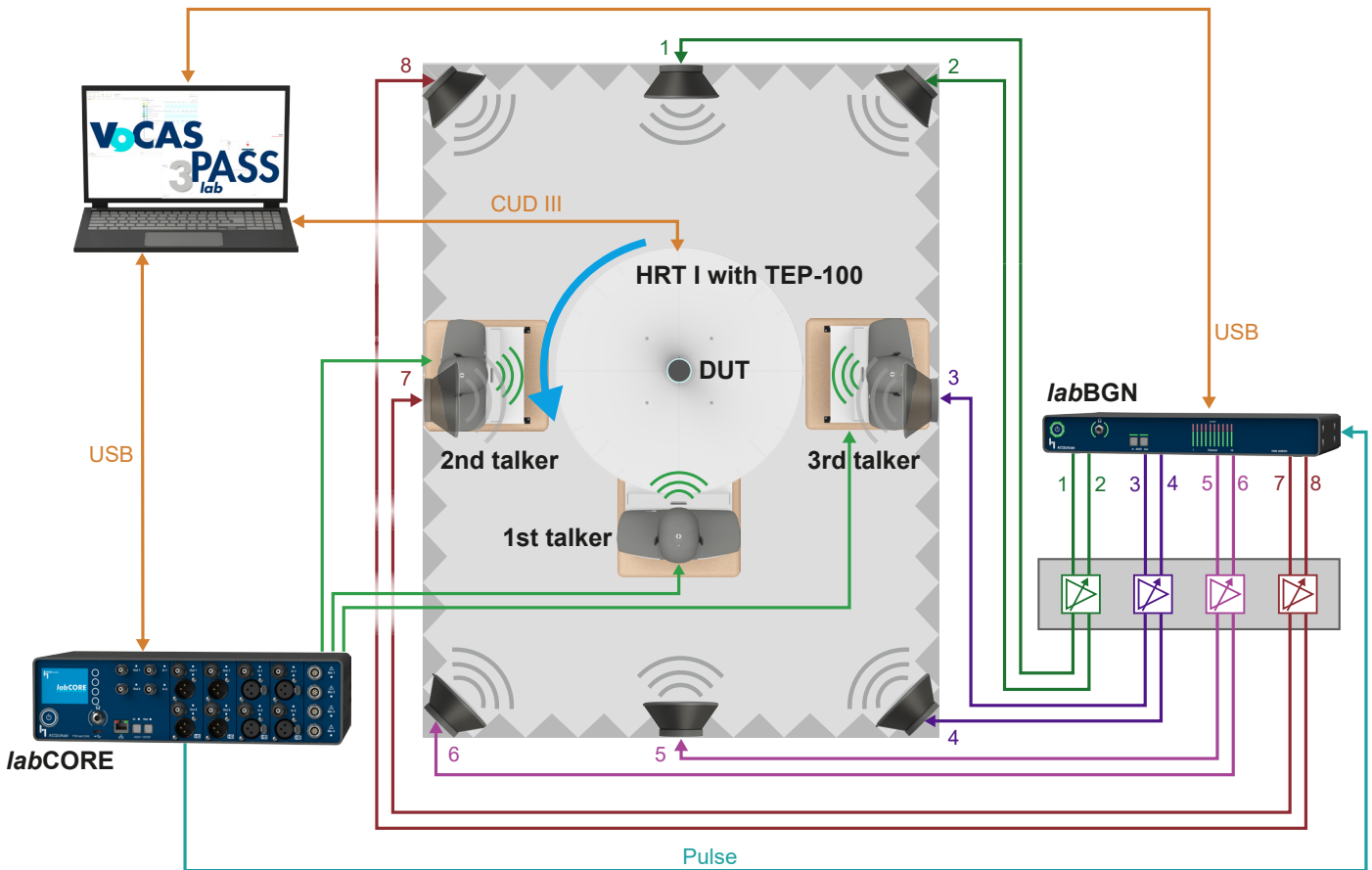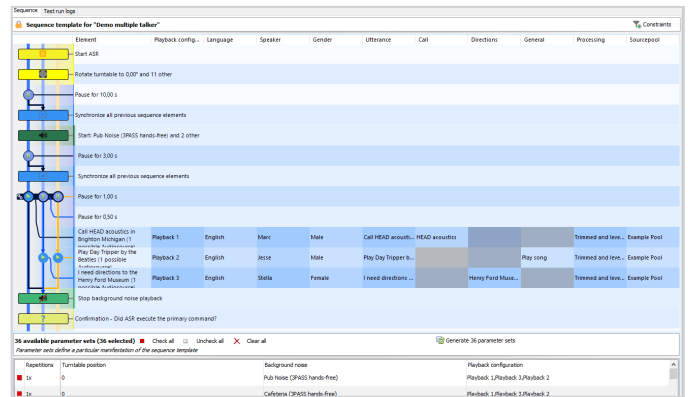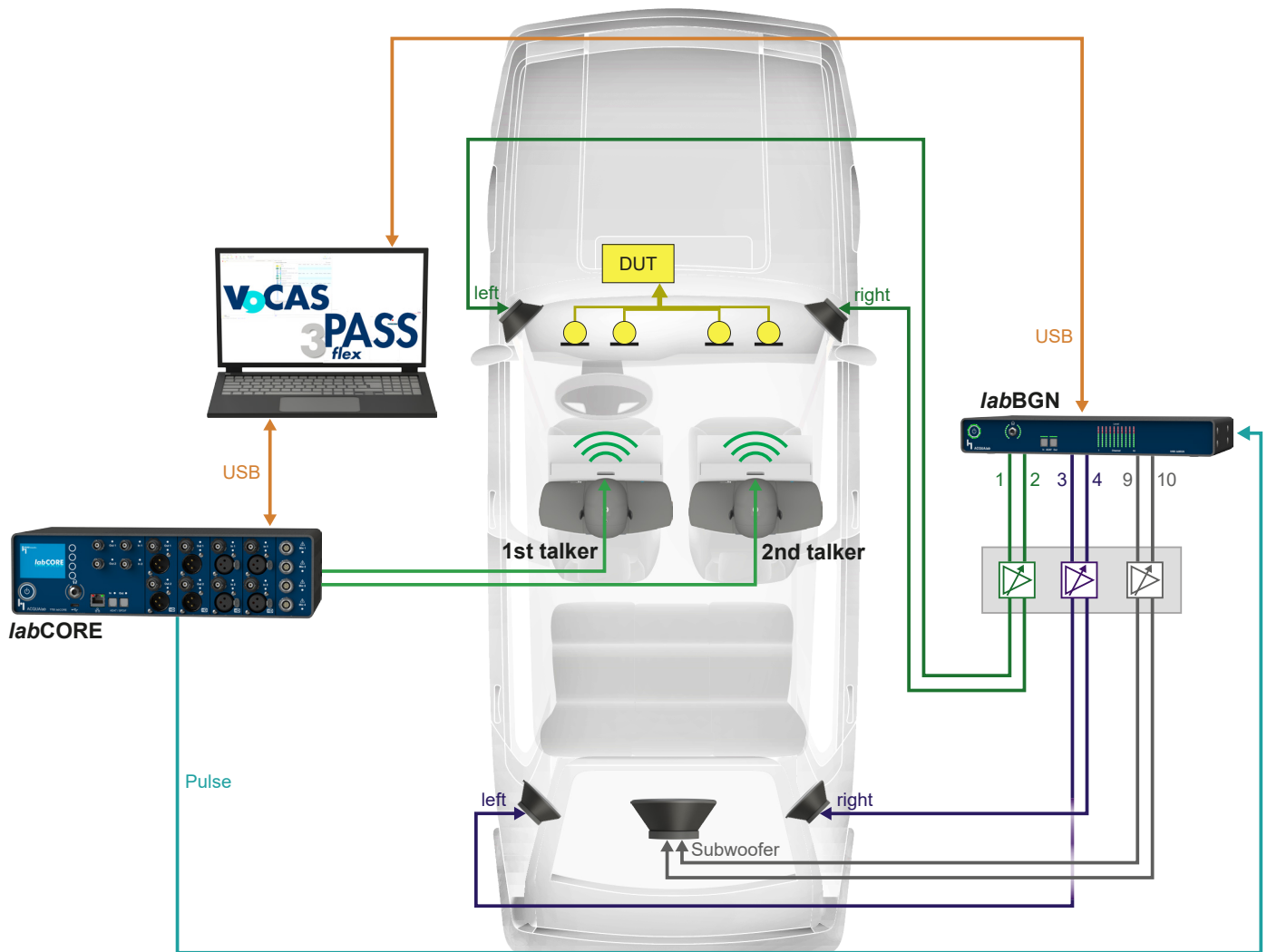
Single Talker In-Vehicle with Background Noise Simulation

Basic measurement configuration in a vehicle. The presented configuration has a HMS II.3 for playback of audio sources. *lab*CORE transfers the audio data from VoCAS to the HATS mouth loudspeaker. The DUT microphones (e.g., automatic speech recognition in-vehicle) are positioned at the interior mirror and the sun visors. 3PASS *flex* executes playback of realistic background noise via *lab*BGN and the loudspeaker arrangement in the car cabin.
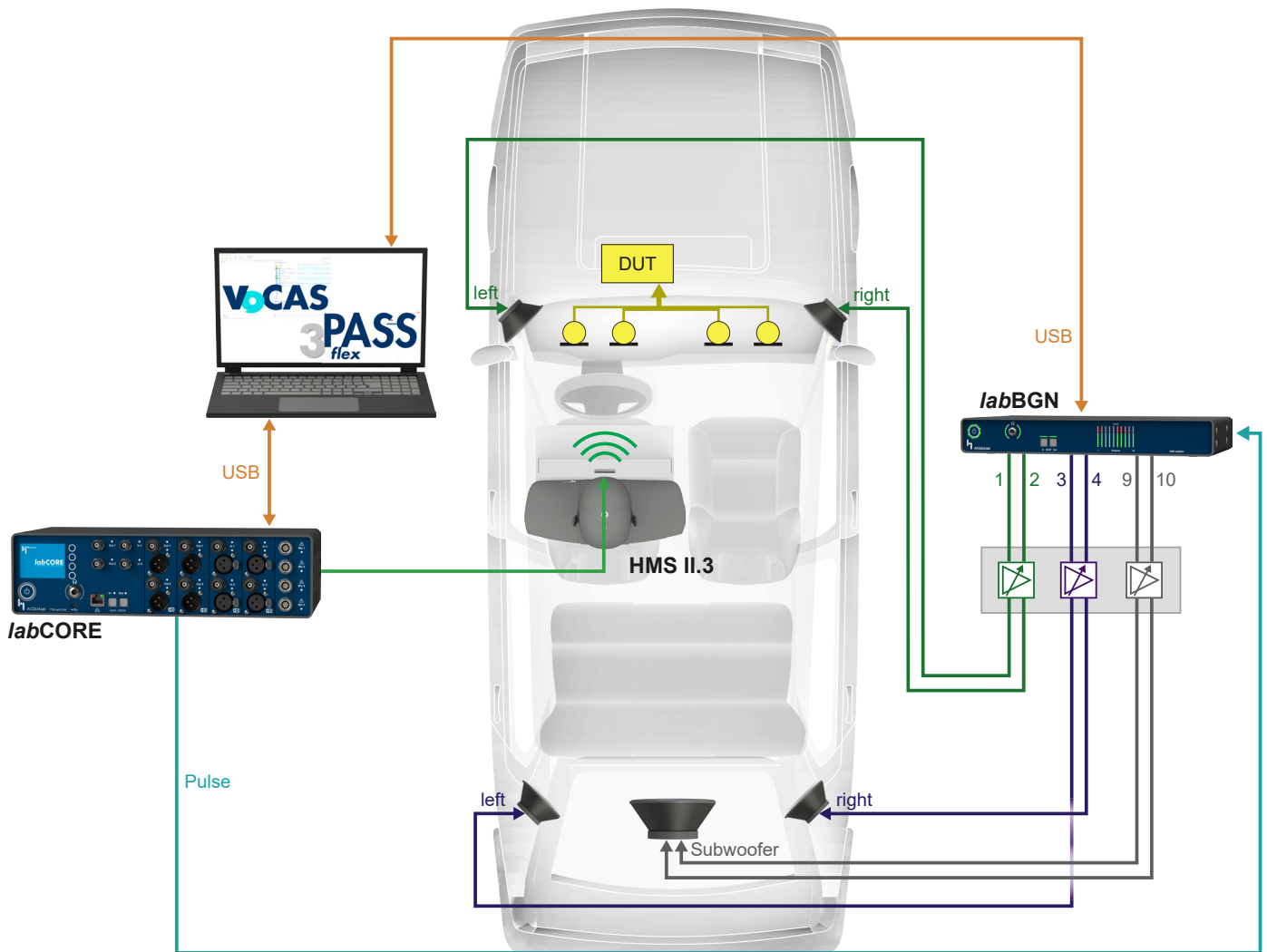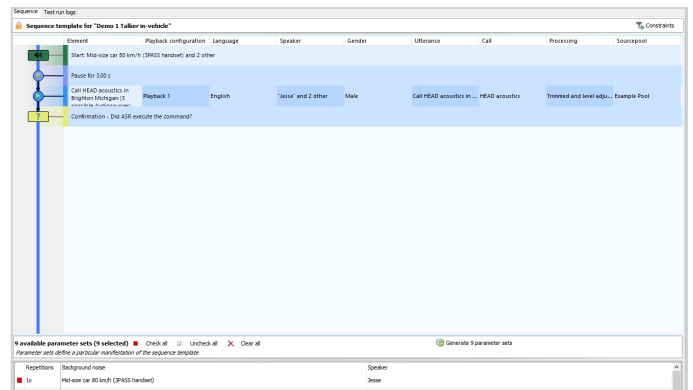
Multiple Talkers In-Vehicle with Background Noise Simulation

Advanced measurement configuration in a vehicle. The presented configuration has two HMS II.3 for playback of audio sources. *lab*CORE transfers the audio data from VoCAS to the HATS mouth loudspeaker. The DUT microphones (e.g., automatic speech recognition in-vehicle) are positioned at the interior mirror and the sun visors. The additional HATS simulates a concurrent talker to distract the ASR. 3PASS *flex* executes playback of realistic background noise via *lab*BGN and the loudspeaker arrangement in the car cabin.

HEAD acoustics

Contact Information