

# データシート

含：  
アプリケーション例



Code 7970

# VoCAS

**Voice Control Analysis System**

# 概要

## VoCAS

### Code 7970

#### Voice Control Analysis System

自動音声認識 (ASR) システムは電話網、自動車、モバイルデバイス、マルチメディア、スマートスピーカーのような IoT デバイス等様々なアプリケーションで利用されています。ASR の性能は主に様々な話者、言語、方言、環境条件に依存します。その為、ASR システムの開発や改善の為に多岐に亘るテストが必要となります。しかしながらそうしたテストには時間がかかり、正確な再現も難しい課題となっています。

ボイズコントロール分析システム VoCAS は HEAD acoustics が ASR システムやデバイスのテストの設計・自動化の為に開発しました。簡単にテストのセットアップができ、リアルかつ完全に再現性のある方法で包括的なバリエーションに富んだテストを行うことが可能です。

## 主なフィーチャー

### VoCAS Core

- VoCAS Core は VoCAS の基本フレームワークです。ニーズに応じて VoCOPT を組み合わせアプリケーションを構築可能です。
- ASR システムのテストとトレーニングの為にテストケースを作成する為のテストシーケンスプレート (カスタマイズされた音源を適用するには VoCAS オプションが必要です。)
- バックグラウンドノイズと残響シミュレーション用ソフトウェアのリモートコントロール
- ターンテーブル HRT I のリモートコントロール

### アーティフィシャルマウス

- カスタマイズド音声ファイルの録音と編集
- VoCAS に適用する為のカスタマイズド音源のデータベース作成
- 妨害話者を伴った ASR システムのテスト
- Python スクリプトによるテストシーケンスの自動化
- ASR システムのトレーニングとテストの為に複数パラメーターによる音声ファイルのバッチ作成

## 主なフィーチャー

リアルにシミュレーションされた環境 (複数話者、バックグラウンドノイズ、音響環境の複合的なシナリオ) における音声認識システムの再現性のあるテスト

複数の音声認識システムのベンチマーク

音声ファイルの録音



# 詳細

VoCAS は音声認識テストの為にソフトウェアです。音声信号処理と音声認識機能を備えたデバイスの最適化とベンチマークをサポートします。

音声認識システムはスマートスピーカー、スマートフォン、タブレット、車載マルチメディアシステム等の音声操作による IoTデバイスの基本機能となっています。

VoCAS は複合的なシナリオにおいてデバイスの音声認識品質のリアルかつ客観的な評価をサポートします。音声認識システムの品質に重大な影響を及ぼすすべての要素を考慮します。即ち、音響環境、バックグラウンドノイズ、話者の種類、言語やアクセントの違い等です。個々のアプリケーションに適用された再現性のあるテストシーケンスにより、音声認識システムの最適化に必要なテスト結果を獲得することができます。

VoCAS は様々なアプリケーションに対応する様々な機能を搭載しています。すべての機能は基本コアの部分（VoCAS Core）と複数のオプション機能（VoCAS オプション、VoCOPT）に含まれています。VoCOPT はアプリケーションを拡張します。

## 製品の構造

### VoCAS Core

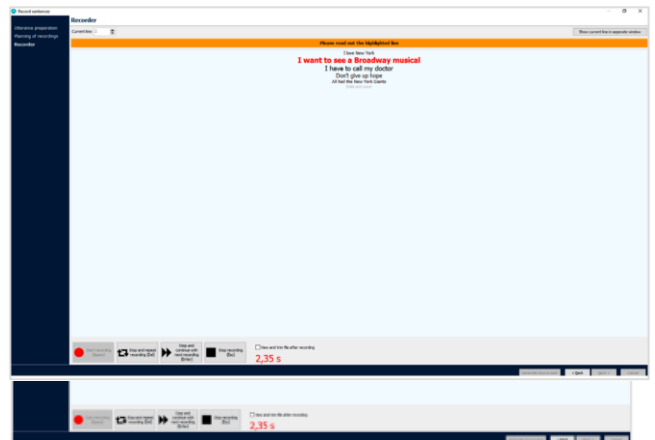
VoCAS Core は VoCAS の基本フレームワークです。テストケースの作成と実行、再生の条件設定、バックグラウンドノイズや残響シミュレーションの遠隔制御のために必要な機能を含みます。テストケースの結果表示はニーズに応じてカスタマイズ可能です。録音、カスタマイズ音源（SQLite データベース）や他の高度なアプリケーションのインポートと適用の為に VoCAS オプションが必要です。

### VoCOPT – VoCAS オプション

VoCAS オプション – VoCOPT – は機能追加により VoCAS Core を拡張します。これらの機能は具体的なユーザーニーズに対応します。VoCOPT するには VoCAS Core が必要です。VoCOPT を任意に組み合わせることができます。

### VoCOPT 1 – Signal Recorder

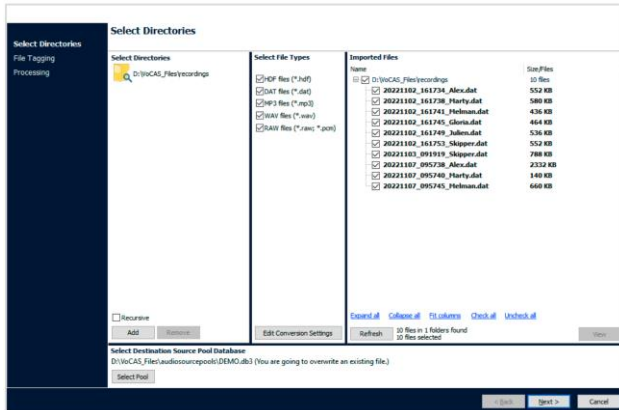
.dat オーディオファイル形式でカスタマイズド文章を録音します。Recorder は録音ファイルにタグ（メタデータ）やファイル名を加え編集可能です。Recorder インターフェイスは話者が読み上げるべきテキストを順次ハイライトしていきます。読み上げるテキストのみを表示することもできます。さらに個々の録音ファイルは必要に応じてレビューやトリミングが可能です。Recorder Wizard が録音の準備から実行までのステップをガイドしていきます。



### VoCAS 2 – インポート / Source Pool 生成

カスタマイズド音源の Source Pool（SQLite データベース）を既存のオーディオファイルからインポート、生成します。Wizard がタグ付けやファイル変換パラメータ等の準備からファイルのインポート、音声 Source Pool（データベース）への編集までの一連のステップをガイドします。

オーディオファイルは HEAD acoustics の専用フォーマット (.hdf / .dat) から一般的なオーディオファイルフォーマット (.mp3 / .wav / .raw) までサポートされています。



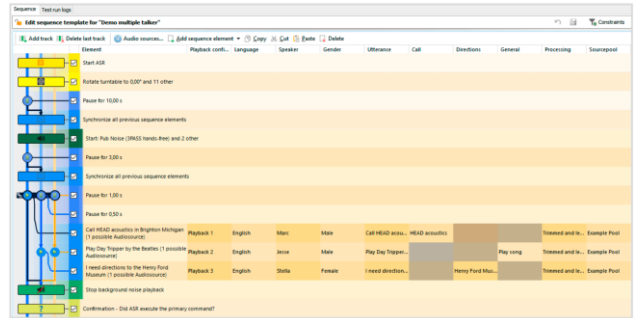
リカーシブ検索機能によりすべての関連ファイルを特定のディレクトリに適用可能です。Tagging Editor はファイルをマニュアルでも自動でもタグ付け可能です。ファイルをデータベースにインポートする前に利用可能な様々な後処理オプションにより編集が可能です。

## VoCOPT 3 – Python スクリプティング



VoCAS はカスタマイズ Python スクリプトを書く Script Editor を搭載しています。テストケースのシーケンステンプレートにこれらのスクリプトを適用し実行します。VoCAS の様々なプロセスの自動化の為に Python スクリプトを適用します。

## VoCOPT 4 - 複数話者



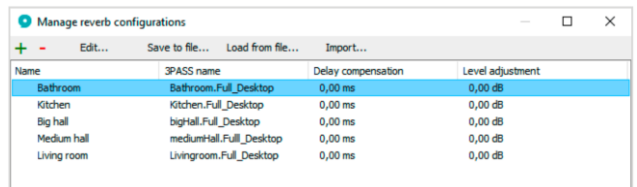
シーケンステンプレートに話者を最大4人まで適用可能です。シーケンステンプレートは話者 1 人に複数のトラックを割り当てます。こうして、複数話者の再生が開始となります。発話の切り替えや、同時発話ができます。話者ごとにゲインやロンバード効果の再生条件の設定が可能です。

## VoCOPT 5 – REST リモートコントロール & 結果へのアクセス

REST API を用いて VoCAS の基本機能をリモート操作

## VoCOPT 6 – シミュレーション・モード

シミュレーションの為にテストケースを作成します。シミュレーションされたテストケースは実際のオーディオ出力なしで実行されます。ユースケースは既存のオーディオソース（話者 & バックグラウンドノイズ、残響シミュレーション）を一つのファイルにまとめます。



VoCAS はテスト環境をシミュレーションします。残響室のインパルス応答に基づいて DUT 位置での全体の音響（話者、バックグラウンドノイズ、残響の総和）を計算し、シミュレーション内容を .dat 及び .wav ファイル形式で録音します。これらのファイルは残響室の音響の計算に基づき、この部屋に音響テストの為に実環境を再現します。このように、VoCAS を用いて、任意の変数を用いて、明確に定義された大量のオーディオテストファイルの生成を簡単に行うことが可能です。

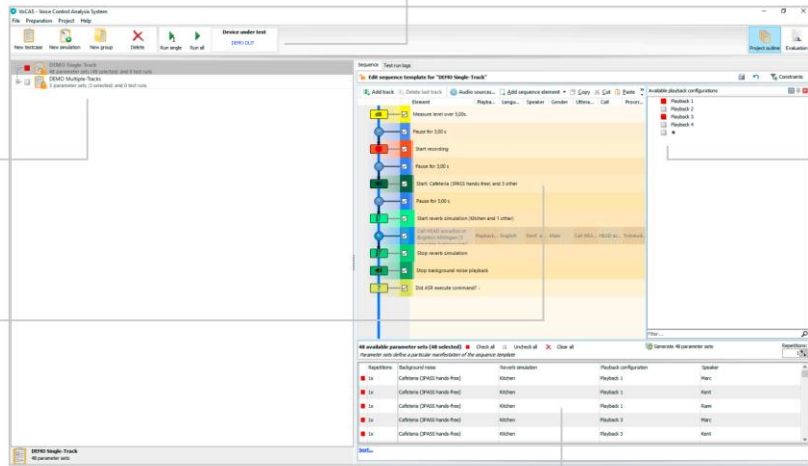
シミュレーションモードのシーケンステンプレートは通常のテストケースと比較すると限定的です。

# ユーザーインターフェイスと操作

メニューバー & プロジェクトタスクバー

現在のプロジェクトにおける  
テストケースと  
シミュレーションケース  
のリスト

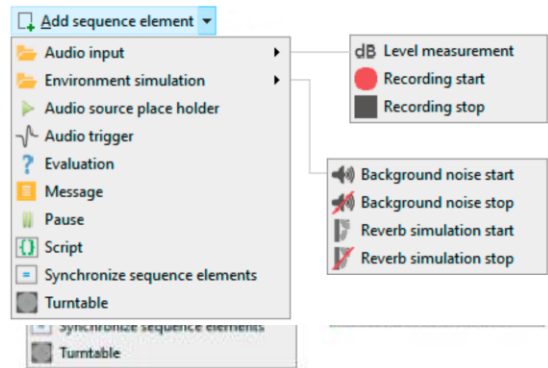
シーケンステンプレート



シーケンステンプレートから  
選択したパラメータの  
クイックチョイス

パラメータセット

VoCAS はテストケースを縦にフローチャートで表します。ユーザーは意図したテストシナリオになるように、シーケンスの要素を加えたり並べ替えたりすることができます。個々の設定内容の編集も可能です。音源は中心的な要素です。その周辺にテストや評価の条件を拡張する要素があります。音源ファイルは再生の為のメタデータ (“タグ”) を含みデータベースに保存されます。音源のシーケンス要素の周りにはテストケースを高度化する為の豊富なオプションがあります。VoCAS はバックグラウンドノイズシミュレーションソフトウェア 3PASS lab / flexの為のリモートコントロールインターフェイス、ターンテーブル HRT I の為の操作インターフェイス、重要な結果を生成する為の評価 Editor が搭載されています。利用可能なシーケンス要素を組み合わせることで、ASR システムの性能を様々な方たちで評価することができます。シーケンステンプレート内の要素は所有している VoCAS オプションにより利用可否が決まります。VoCAS はパラメータセットのリスト上の現在のシーケンステンプレートにより可能なすべてのバリエーションを自動生成します。一つのパラメータセットは話者、再生コンフィギュレーション、バックグラウンドノイズやターンテーブルの角度等シーケンステンプレートから任意に組み合わせられたパラメータが含まれます。



パラメータセットのリストはシーケンステンプレートのすべてのパラメータのバリエーションを含みます。シーケンステンプレートが保存されロックされたら、テストケースは実行準備完了です。テストケースを開始するとパラメータセットが一つずつアクティブとなり実行されます。

## 結果と評価

VoCAS は測定結果をソートし、テーブル表示する別のインターフェイスを搭載しています。

- 行と欄の内容調整
- 結果の色分け表示
- 特定の結果のみに限定した表示設定
- カスタマイズした結果テーブルを Excel 形式でエクスポート、保存



# オプション

## VoCAS オプション

VoCOPT 1 (Code 7971)

- VoCAS オプション Signal Recorder

VoCOPT 2 (Code 7972)

- VoCAS オプション Import / Source Pool Generator

VoCOPT 3 (Code 7973)

- VoCAS オプション Python Scripting

VoCOPT 4 (Code 7974)

- VoCAS オプション Multitalker

VoCOPT 5 (Code 7975)

- VoCAS オプション REST Remote Control & Result Access

VoCOPT 6 (Code 7976)

- VoCAS オプション Simulation Mode

## ハードウェア

HRT I (Code 6498)

- HEAD acoustics 遠隔操作ターンテーブル

DSB IV.1 (Code 2408.1)

- デジタルサウンドボード (測定コンフィグレーションにおける *labCORE* の代替)

## ソフトウェア

3PASS *lab* (Code 6990)

- マイク位置とスピーカー配置が固定のバックグラウンドノイズシミュレーション

3PASS *flex* (Code 6995)

- マイク位置とスピーカー配置が任意のバックグラウンドノイズシミュレーション

3PASS *reverb* (6996)

- 残響シミュレーションを伴ったバックグラウンドノイズシミュレーション

HAE-BGN (Code 6970)

- 実験室におけるバックグラウンドノイズシミュレーションのための自動イコライゼーション

HAE-*car* (Code 6971)

- 車室内におけるバックグラウンドノイズシミュレーションのための自動イコライゼーション

# 一般要件

## ハードウェア

*labCORE* (Code 7700)

- モジュール式多チャンネルハードウェアプラットフォーム

*coreBUS* (Code 7710)

- I/O BUSメインボード

*coreOUT-Amp2* (Code 7720)

- パワーアンプボード

*coreIN-Mic4* (Code 7730)

- マイクホン入力ボード

HEAD 測定システム、以下うちいずれか：

- HMS II.3 (Code 1703)
- HMS II.5 (Code 1705)
- HMS II.6 (Code 1706)

測定マイクロホン

- マウスイコライゼーション、測定録音、音圧測定、音声トリガー用

マルチコアプロセッサ搭載PC (1.6 GHz以上)、4GB RAM、40 GB ディスク空きスペース、4 x USB ポート

## ソフトウェア

VoCAS (Code 7970)

- Voice Control Analysis System、Core バージョン  
マイクロソフト Windows 10 Pro  
または Windows 8.1 Pro

- 英語またはドイツ語バージョン、含：すべての最新のサービスパック

# 納品アイテム

VoCAS (Code 7970) は以下のコンポーネンツが同梱されています。

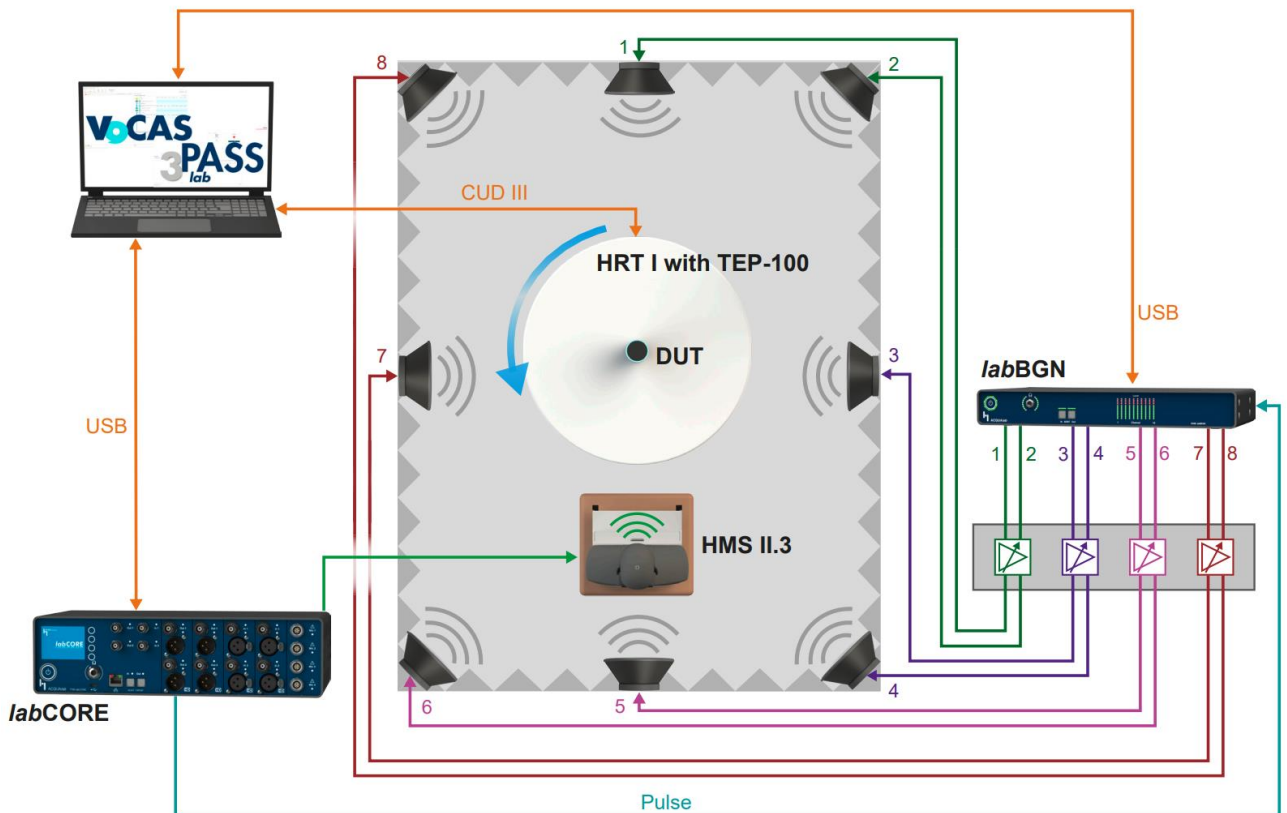
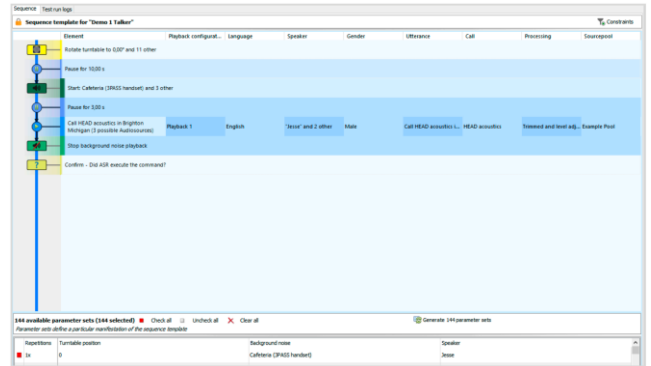
- セットアップDVD、  
含：デモプロジェクト、デモオーディオデータベース
- USB ドングル

# 実際のテスト

## アプリケーション例

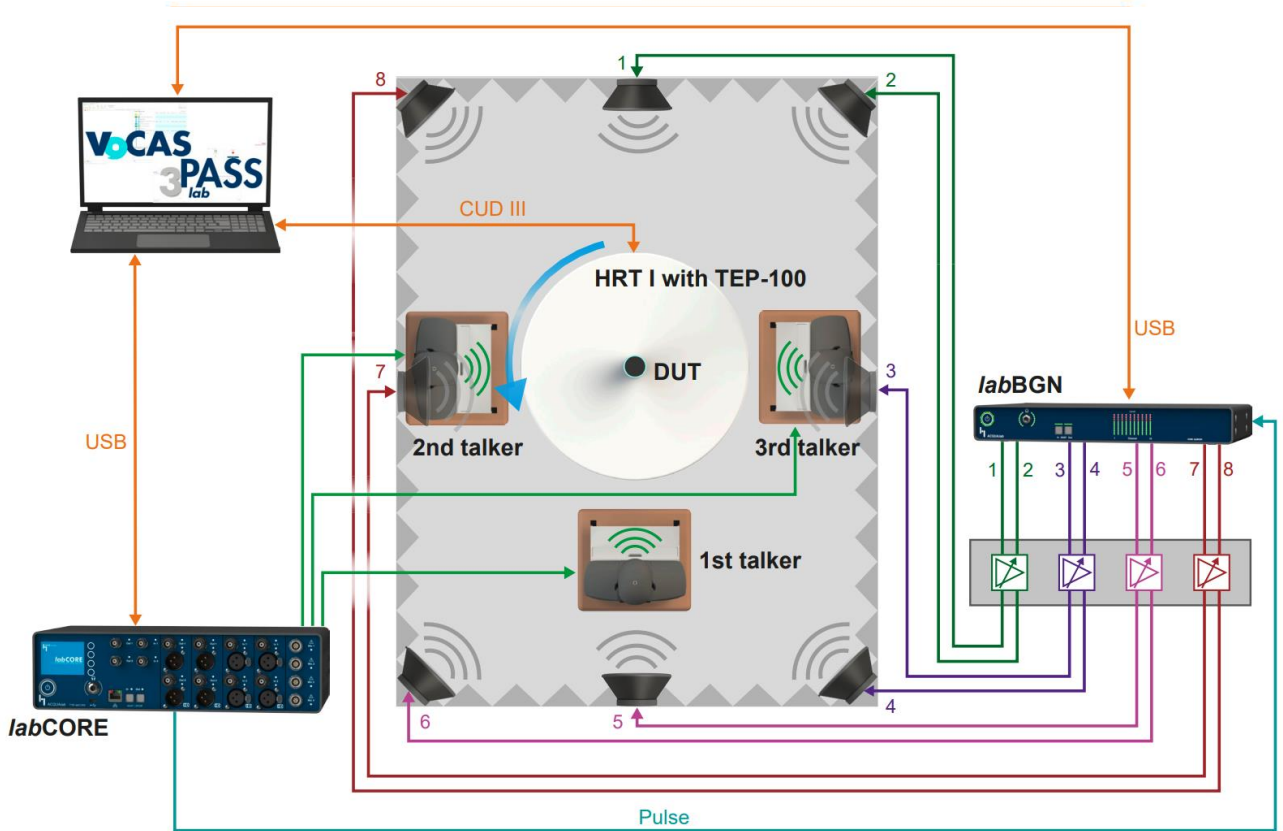
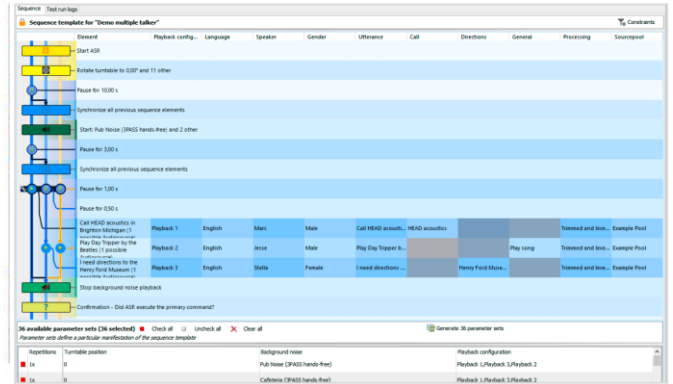
### 単独話者、無響室、バックグラウンドノイズシミュレーション

無響室における基本的な測定コンフィグレーションです。このコンフィグレーションは音源再生に HMS II.3 を用いています。labCORE が VoCAS から HATS マウスピーカーに音声データを転送します。測定対象のスマートスピーカーはターンテーブル HRT I に置かれ HATS の前に配置されます。HRT I は HATS に対して様々な角度になるように測定対象物を回転させる為、測定対象物のマイクロホンを異なる位置にしてテストすることができます。3PASS lab が ETSI TS 103 224 に基づいて labBGN とスピーカースystemを通じリアルなバックグラウンドノイズを再生します。



## 複数話者、無響室、バックグラウンドノイズシミュレーション

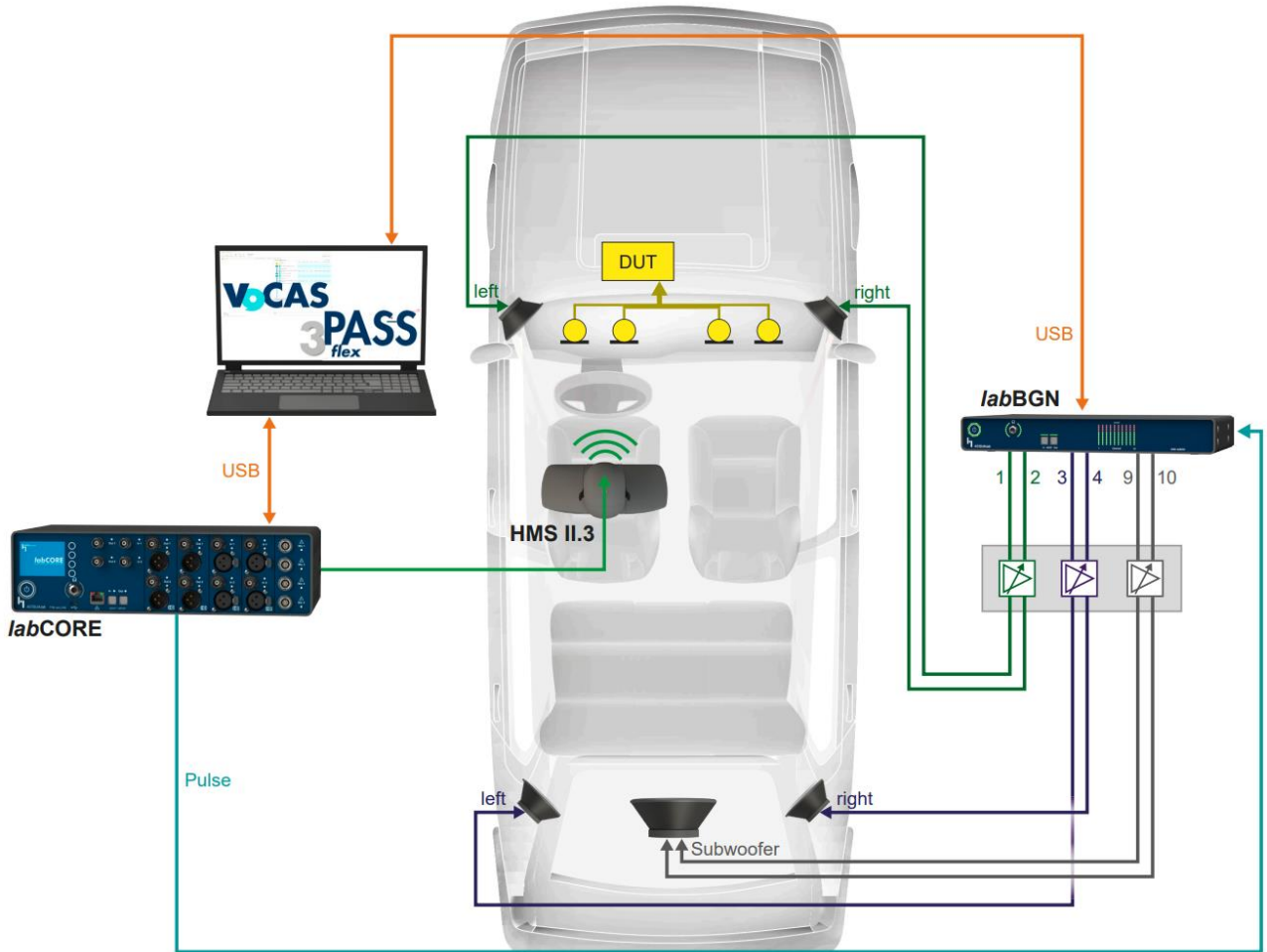
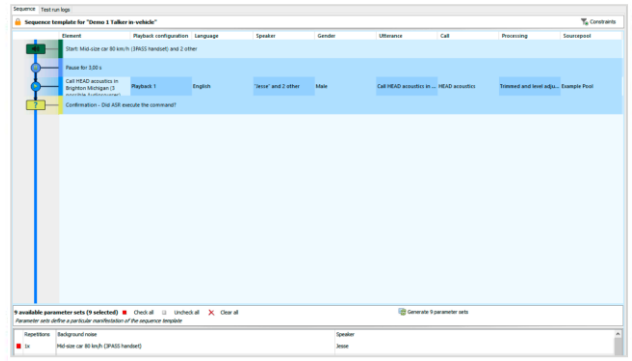
無響室における高度な測定コンフィグレーションです。このコンフィグレーションは音源再生に 3 台の HMS II.3 を用いています。*labCORE* が VoCAS から HATS マウスピーカーに音声データを転送します。測定対象のスマートスピーカーはターンテーブル HRT I に置かれ HATS に囲まれ中央に配置されます。HRT I はメインの HATS に対して様々な角度になるように測定対象物を回転させる為、測定対象物のマイクロホン異なる位置にしてテストすることができます。それ以外の HATS は ASR を妨害する話者をシミュレーションします。3PASS *lab* が ETSI TS 103 224 に基づいて *labBGN* とスピーカーシステムを通じリアルなバックグラウンドノイズを再生します。





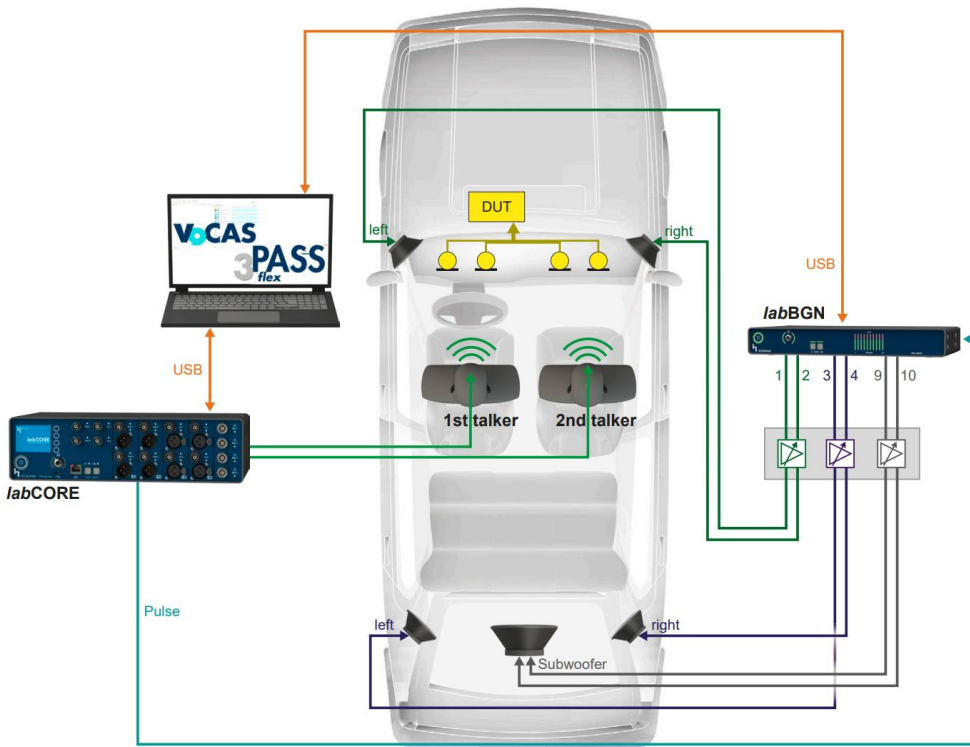
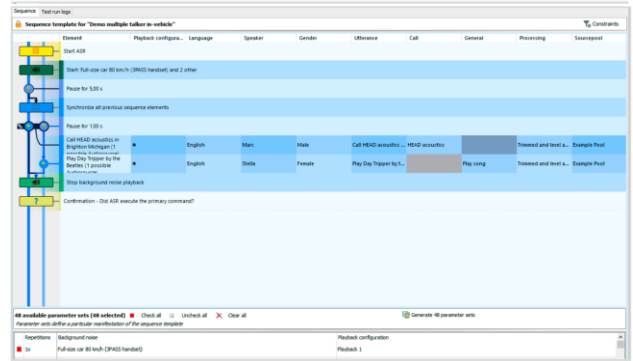
## 単独話者、車室内、バックグラウンドノイズシミュレーション

車室内における基本的な測定コンフィグレーションです。このコンフィグレーションは音源再生に HMS II.3 を用いています。labCORE が VoCAS から HATS マウスピーカーに音声データを転送します。測定対象の自動音声認識システムのマイクロホンはルームミラーとサンバイザーに搭載されます。3PASS flex が labBGN とスピーカーシステムを通じリアルなバックグラウンドノイズを再生します。



## 複数話者、車室内、バックグラウンドノイズシミュレーション

車室内における高度な測定コンフィグレーションです。このコンフィグレーションは音源再生に 2 台の HMS II.3 を用いています。*labCORE* が VoCAS から HATS マウスピーカーに音声データを転送します。測定対象の自動音声認識システムのマイクロホンはルームミラーとサンバイザーに搭載されます。2 台目の HATS は ASR を妨害する話者をシミュレーションします。3PASS *flex* が *labBGN* とスピーカーシステムを通じリアルなバックグラウンドノイズを再生します。



Microsoft、Excel、Windows はMicrosoft グループ各社の登録商標です。

Python と Python ロゴは Python Software Foundation の商標または登録商標です。HEAD acoustics GmbH は同 Foundation の許可により使用しています。



お問い合わせ：

〒240-0005 神奈川県横浜市保土ヶ谷区神戸町 134  
横浜ビジネスパークウエストタワー 8F

電話： 045-340-2236  
Eメール： [headjapan@head-acoustics.com](mailto:headjapan@head-acoustics.com)  
ウェブサイト： [www.head-acoustics.com](http://www.head-acoustics.com)