

## データシート

含：  
アプリケーション例



Code 7970

# VOCAS

音声認識分析システム

# 概要

## VoCAS

### Code 7970

#### 音声認識分析システム

自動音声認識（ASR）システムは、電話ネットワーク、車、モバイル機器、マルチメディア、スマートスピーカー等のIoTデバイスに広く採用されています。ASRの性能は、様々な話者、言語、方言、環境条件に主に依存します。従って、豊富なバリエーションでのテストが、ASRシステムの開発と改善に有益です。しかしながら豊富なバリエーションでテストをするにはより時間がかかり、正確な再現もより困難です。

VoCAS音声認識分析システムは、ASRシステムとデバイスのテストをオーガナイズし、自動化します。セットアップは簡単で、リアルかつ完全に再現可能な条件下で、豊富なバリエーションによる包括的なテストを可能にします。

## 主なフィーチャー

### VoCAS Core

- ユーザーニーズに応じてオプションの VoCOPT を用いてアプリケーションを構築する為の VoCAS の基本フレームワーク
- ASRシステムをテスト及びトレーニングする為のテストケース作成用テストシーケンスプレート（カスタマイズ音声ソースの適用には 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ソフトウェアのリモートコントロール
- HRT I ターンテーブルのリモート操作

### VoCAS オプション

- カスタマイズ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音声ソースのデータベース作成
- 妨害話者によるASRシステムの妨害とテスト
- Python スクリプトを使用した自動、高速テストシーケンス
- REST API を介した VoCAS のリモート操作
- ASRシステムのトレーニングまたはテストの為の異なるパラメータを含む音声ファイルのバッチ生成

## アプリケーション

リアルなシミュレーション環境（複数話者、バックグラウンドノイズ、音響環境を含む）における音声認識システムの再現可能なテスト

異なる音声認識システムの有意義なベンチマーク

音声ファイルの録音

# 詳細

VoCASは音声認識評価の為にテストソフトウェアです。音声信号処理と音声認識機能を搭載したデバイスの最適化とベンチマーキングをサポートします。

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デバイスの重要なコンポーネントです。VoCASは、複雑なシナリオにおけるデバイスの音声認識品質を現実的かつ客観的に評価するのに役立ちます。このソフトウェアは、音声認識システムの品質に大きく影響を与えるすべての要因（音響環境、バックグラウンドノイズ、様々な話者、様々な言語やアクセント）を考慮します。従って、それぞれのアプリケーションに適応した再現可能なテストシーケンスにより、製造メーカーは音声認識システムを最適化する為の重要な情報を得ることができます。

VoCASは、様々なアプリケーション向けに複数の機能を搭載しています。基本的なコアバージョン（VoCAS Core）とアプリケーションにより機能拡張が可能な複数のオプション機能（VoCAS オプション、VoCOPTs）に分かれています。

## 製品構成

### VoCAS Core

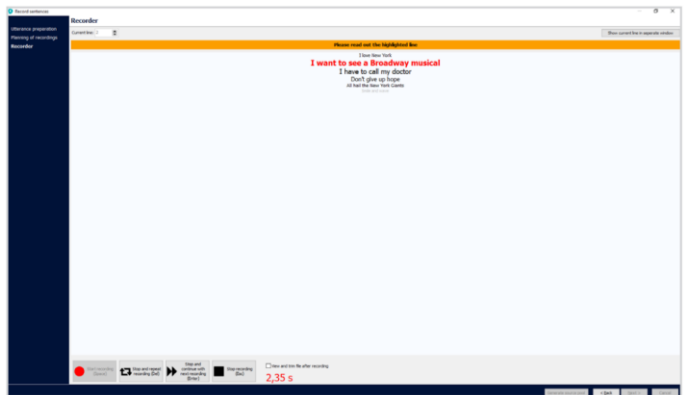
VoCAS CoreはVoCASの基本フレームワークです。テストケースを作成し実行する為に必要な機能、再生条件設定の機能、バックグラウンドノイズや残響シミュレーションをリモートで操作する機能を含んでいます。テスト結果の表示方法は、ユーザーの好みに応じてカスタマイズできます。VoCAS Coreは、録音、インポート、カスタマイズ音声ソース（SQLiteデータベース）を適用する為のVoCASオプションと、その他の高度なアプリケーションを必要とします。

### VoCOPT – VoCAS オプション

VoCAS オプション – VoCOPT – はVoCAS Core を追加機能により拡張します。これらの機能が異なるユーザーニーズに対応します。いずれのVoCOPTの利用にも前提条件としてVoCAS Coreが必要です。VoCOPTは自由に組み合わせることができます。

### VoCOPT 1 – 信号レコーダー

カスタマイズされた文章を\*.datオーディオファイル形式で録音します。さらに、レコーダーには個別のタグ（メタデータ）や個別のファイル名で録音データをカスタマイズする為の編集機能があります。レコーダーインターフェースは、スピーカーが読み上げるテキストを次々にハイライト表示します。読み上げるテキストのみを表示する為の

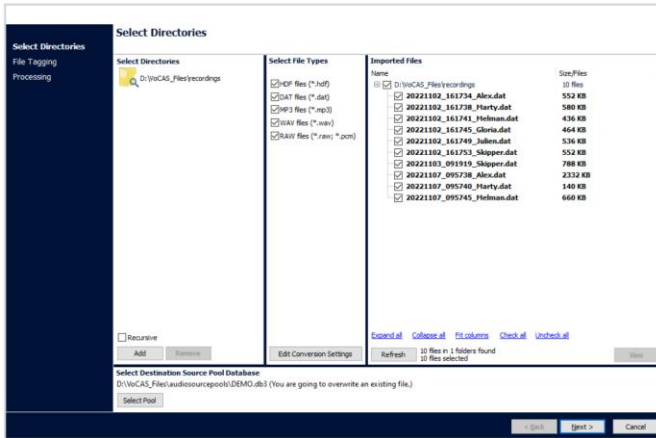


追加インターフェースもあります。さらに、各録音データは必要に応じてレビューしたりトリミングできます。レコーダーウィザードは、準備とプログラミングから録音の実行まで、ステップバイステップでガイドします。

### VoCOPT 2 – インポート / ソース Pool ジェネレーター

既存のオーディオファイルからカスタマイズ音声ソースPool（SQLiteデータベース）をインポートし生成します。提供されたウィザードは、タグ付けやファイル変換パラメータ等の準備ステップをガイドし、最終的にファイルをオーディオソースPool（データベース）にインポートおよびコンパイルします。サポートされている音声ファイルは、HEAD acoustics専用フォーマット（\*.hdf/\*.dat）から一般的な音声ファイルフォーマット（\*.mp3/\*.wav/\*.raw）まで多岐にわたります。リカーシブ検索により、特定のディレクトリ内のすべての適切なファイルが素早く表示されます。

タグ付けエディタは、ファイルのマニュアル及び自動によるタグ付けの両方が可能です。ファイルをデータベースにインポートする前に、利用可能な様々なポストプロセッシングアクションによって変更することができます。

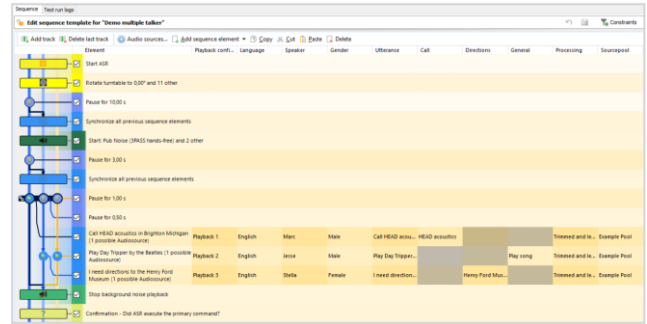


## VoCOPT 3 – Python スクリプト作成

VoCASはカスタマイズドPythonスクリプトを書くためのスクリプトエディタが搭載されています。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て実行します。VoCASの様々なプロセスの自動化にPythonスクリプトを適用します。



## VoCOPT 4 – 複数話者



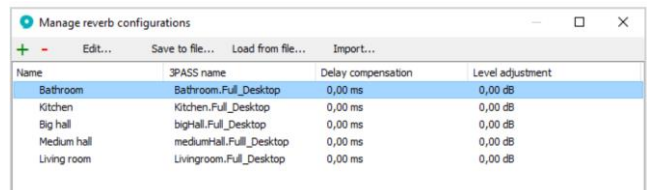
最大4人の同時話者をシーケンステンプレートに適用します。シーケンステンプレートには、話者ごとに異なるトラックがある為、異なる話者の再生を、互いに重ねたり、ずらしながら同時再生できます。ゲインやロンバード効果等の再生条件を話者ごとに設定できます。

## VoCOPT 5 – REST リモートコントロールと結果アクセス

REST API を介してリモートで基本的なVoCAS機能进行操作

## 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、実際の音声出力なしで実行されます。この機能を利用して、既存の複数の音声ソース（話者音声とバックグラウンドノイズ、残響シミュレーション）を一連の処理用に1つのファイルにまとめることができます。



VoCASはテスト環境をシミュレートします。残響室内のインパルス応答に基づいて、ソフトウェアはDUT位置（話者、バックグラウンドノイズ、残響）での音響を合算し、シミュレーションデータを\*.dat及び/又は\*.wavファイル形式で収録します。残響室内の音響計算により、これらのファイルを用いることにより、この部屋で実際に音響テストを行うのと同じ音響条件でテストを再現することが可能となります。従って、VoCASは任意のバリエーションで明確に整理された大量の音声テストファイルを便利に生成することができます。

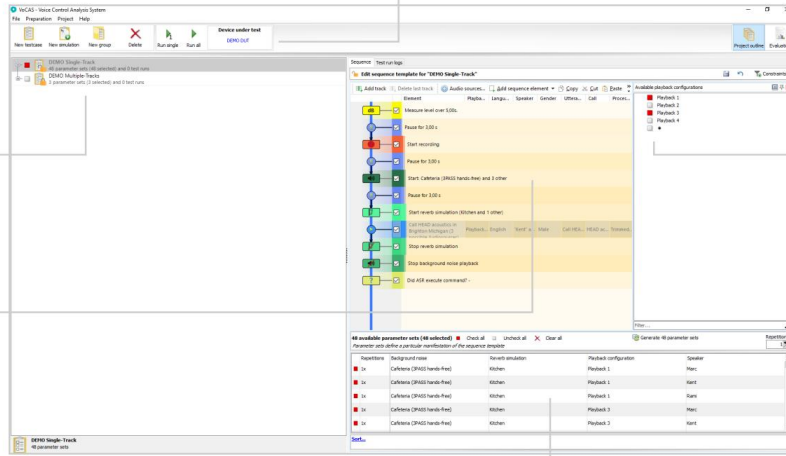
シミュレーションモードのシーケンステンプレートは、通常のテストケースと比較して制限されています。

# ユーザーインターフェイスと操作

## メニューバーとプロジェクトタスクバー

現在のプロジェクト内の  
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

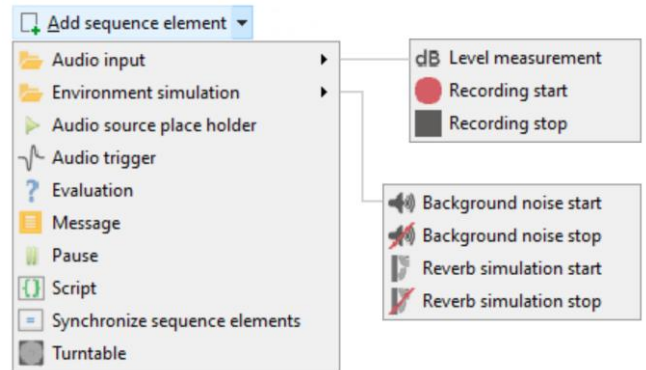


シーケンステンプレートで  
変更したい属性を  
素早く選択

パラメータセット

VoCASは、テストケースをフローダイアグラムとして縦に提示します。ユーザーは、シーケンス要素の追加や再配置ができ、それぞれの設定を変更して任意のテストシナリオを作成できます。音声ソースは中心的な要素であり、周辺データがテストと評価の条件を拡張します。音声ソースファイルには再生用のメタデータ（「タグ」）が含まれ、データベースに保存されます。音声ソースのシーケンス要素の周りには、テストケースを詳細条件で拡張する為の多くの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をリモート操作する為のインターフェース、HRT Iターンテーブルの制御インターフェース、重要な結果を生成するための評価エディタ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システムの性能評価の為の可能性が広がります。シーケンステンプレート内の要素の範囲は、利用可能なVoCASオプションに依存します。VoCASは、現在のシーケンステンプレートをベースに可能性のあるすべてのバリエーションをパラメータセットのリストとして自動的に生成します。パラメータセットには、スピーカー、再生設定、バックグラウンドノイズ、ターンテーブルの角度等、シーケンステンプレートをベースとしユニークに組み合わせたパラメータが含まれます。

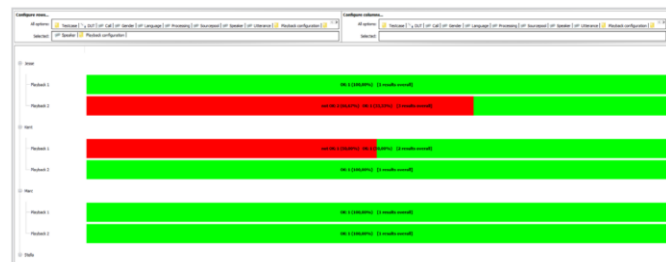
従って、パラメータセットのリストには、シーケンステンプレートをベースにしたすべてのバリエーションのパラメータが含まれています。シーケンステンプレートが保存されロックされると、テストケースは実行準備完了です。アクティブパラメータセットは、テストケースを起動後に実行されます。



## 結果と評価

測定結果を表形式でソート、視覚化する為のインターフェース

- 行と列の内容調整
- 豊富な選択肢から異なる色を割り当てて結果を区別
- 他の結果とは切り離して特定の結果を表示する為の条件設定
- カスタマイズド結果テーブルをExcelファイルでエクスポートし保存



## オプション

### VoCAS オプション

#### VoCOPT 1 (Code 7971)

- › VoCAS オプション信号レコーダー

#### VoCOPT 2 (Code 7972)

- › VoCAS オプション インポート/ソース Pool ジェネレーター

#### VoCOPT 3 (Code 7973)

- › VoCAS オプション Python スクリプティング

#### VoCOPT 4 (Code 7974)

- › VoCASオプション 複数話者

#### VoCOPT 5 (Code 7975)

- › VoCAS オプション REST リモートコントロールと結果アクセス

#### VoCOPT 6 (Code 7976)

- › VoCAS オプション シミュレーションモード

## ハードウェア

### HRT I (Code 6498)

- › HEAD acoustics リモート操作式ターンテーブル

### DSB IV.I (Code 2408.1)

- › デジタルサウンドボード (*labCORE* の代替)

### HMA V (Code 1420)

- › HMS用マウスアンプ、2 × 30 W、クラス-D

## ソフトウェア

### 3PASS *lab* (Code 6990)

- › バックグラウンドノイズシミュレーション (固定マイク、スピーカー)

### 3PASS *flex* (Code 6995)

- › バックグラウンドノイズシミュレーション (非固定マイク、スピーカー)

### 3PASS *reverb* (6996)

- › バックグラウンドノイズシミュレーション (残響シミュレーション)

### HAE-BGN (Code 6970)

- › 実験室バックグラウンドノイズシミュレーション自動イコライゼーション

### HAE-car (Code 6971)

- › 車室内バックグラウンドノイズシミュレーション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)

- › *labCORE* I/O BUSメインボード

#### *coreOUT-Amp2* (Code 7720)

- › *labCORE* パワーアンプボード

#### *coreIN-Mic4* (Code 7730)

- › *labCORE* マイクロホン入力ボード

以下のHEAD 測定システムのいずれか：

- › HMS II.3 (Code 1703)

- › HEAD 測定システム、基本バージョン、右耳シミュレーター、3.3 耳介、疑似マウス

- › HMS II.3 LN (Code 1703.1)

- › HEAD 測定システム、低ノイズバージョン、右耳シミュレーター、3.3 耳介、疑似マウス

- › HMS II.3 LN HEC (Code 1703.2)

- › HEAD 測定システム、低ノイズバージョン、ヒューマンライク右耳外耳道、疑似マウス

- › HMS II.5 (Code 1705)

- › HEAD 測定システム、3.3 耳介、疑似マウス (耳シミュレーター非搭載)

- › HMS II.6 (Code 1706)

- › HEAD 測定システム、疑似マウス、自由音場マイク (左右)

- › HMS II.7 (Code 1707)

- › HEAD 測定システム、疑似マウス、自由音場 ICP® マイクロホン (左右)

### 測定マイク

- › マウスイコライゼーション、音データ測定、レベル測定、または音声トリガー用

次ページへ続く

## 一般要件

---

### システム要件

PC

- ＞ マルチコアプロセサー
- ＞ 8 GB RAM (推奨 : 16 GB RAM)
- ＞ 40 GB 空ディスク容量

オペレーティングシステム (以下のいずれか)

- ＞ Windows 11 x64
  - ≫ Pro, Enterprise, Education; バージョン 21H2  
以降 ; 言語 : 米語、西洋言語
- ＞ Windows 10 x64
  - ≫ Pro, Enterprise, Education; バージョン 1809  
以降 ; 言語 : 米語、西洋言語

## 納品アイテム

---

VoCAS Core (Code 7970) 以下のコンポーネントを含みます。

- ＞ セットアップ DVD、含 : デモプロジェクト、デモ音声データベース
- ＞ USB ドングル

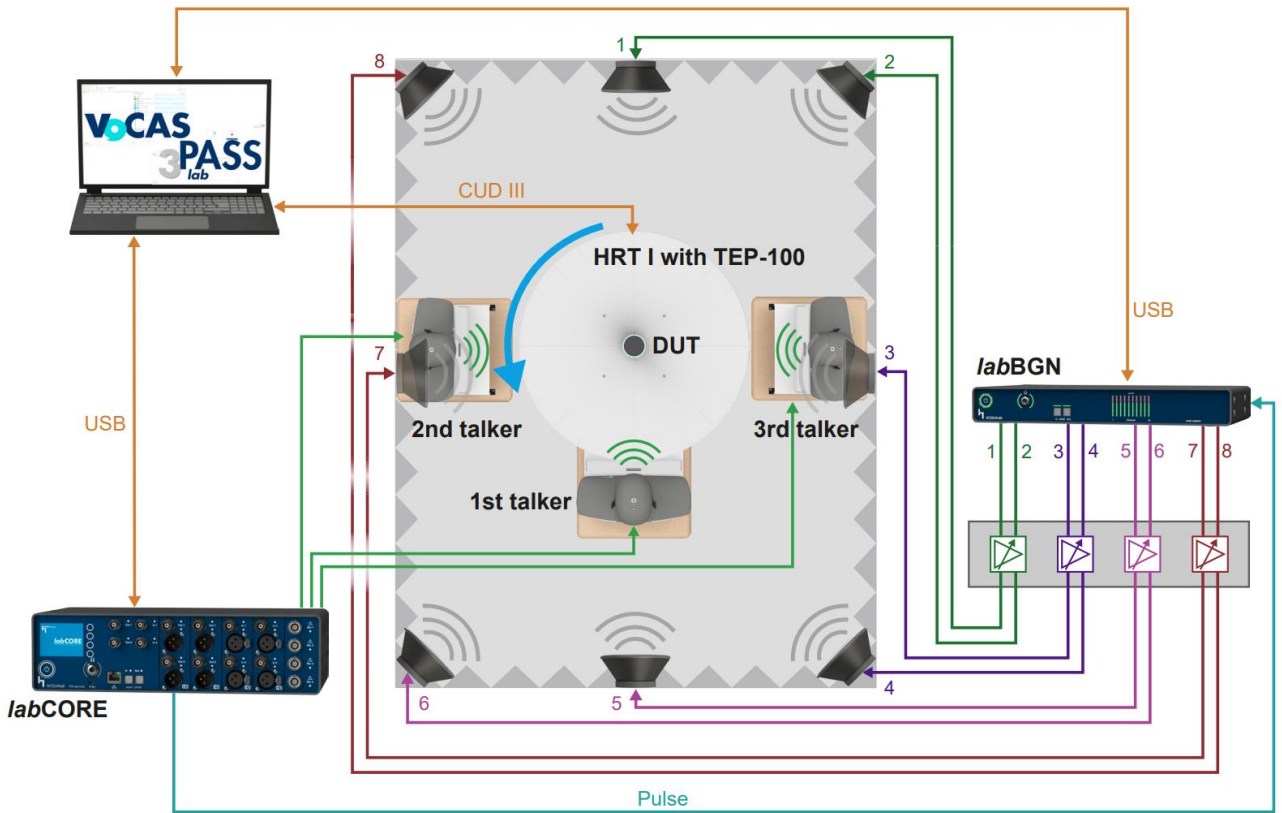
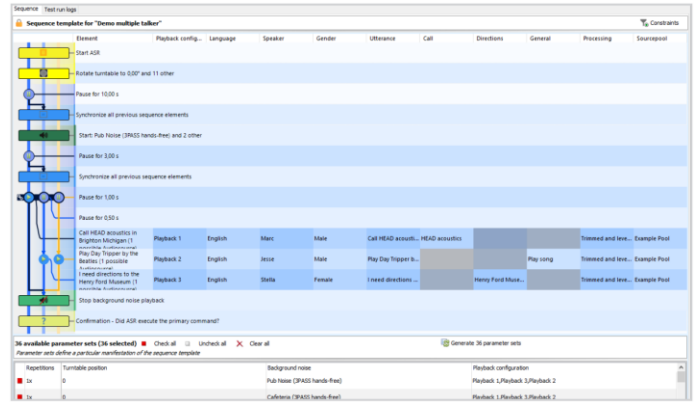




# 無響室でのバックグラウンドノイズシミュレーション環境下の複数話者テスト

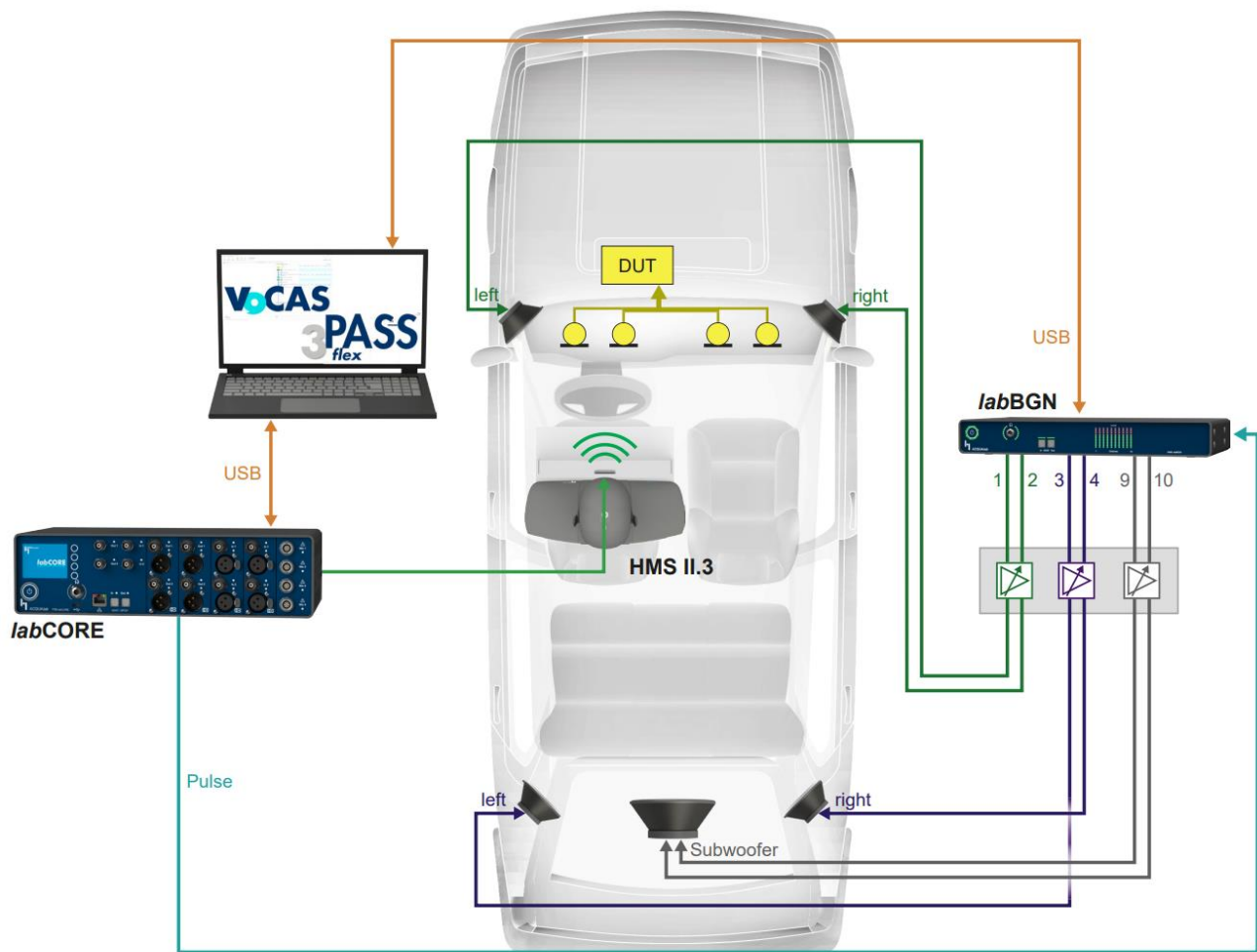
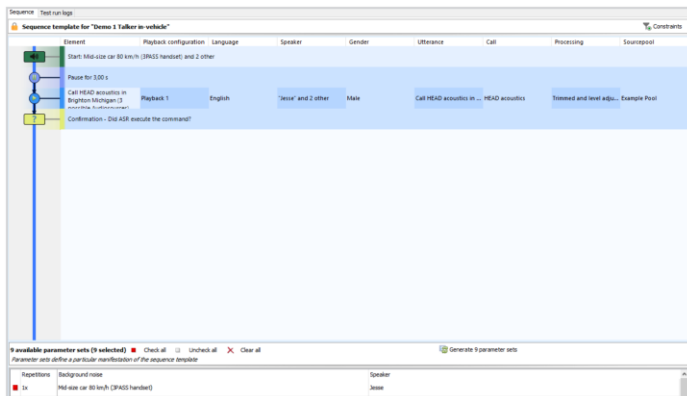
無響室における高度な測定システム構成。

こちらの構成では、音源再生用の3台のHMS II.3があります。  
*labCORE*が、VoCASからHATSマウススピーカーに音声データを転送します。DUT（例：スマートスピーカー）はHATSの正面のHRT Iターンテーブル上に配置されます。HRT Iは、DUTをメイン話者のHATSに対して角度を変えながら回転させます。こうして、DUTマイクロフォンを異なる角度でテストできます。追加のHATSが、ASRを妨害する為の妨害話者をシミュレートします。  
 3PASS*lab*は、ETSI TS 103 224に準拠し、*labBGN*とスピーカーシステムを介してリアルなバックグラウンドノイズの再生を実行します。



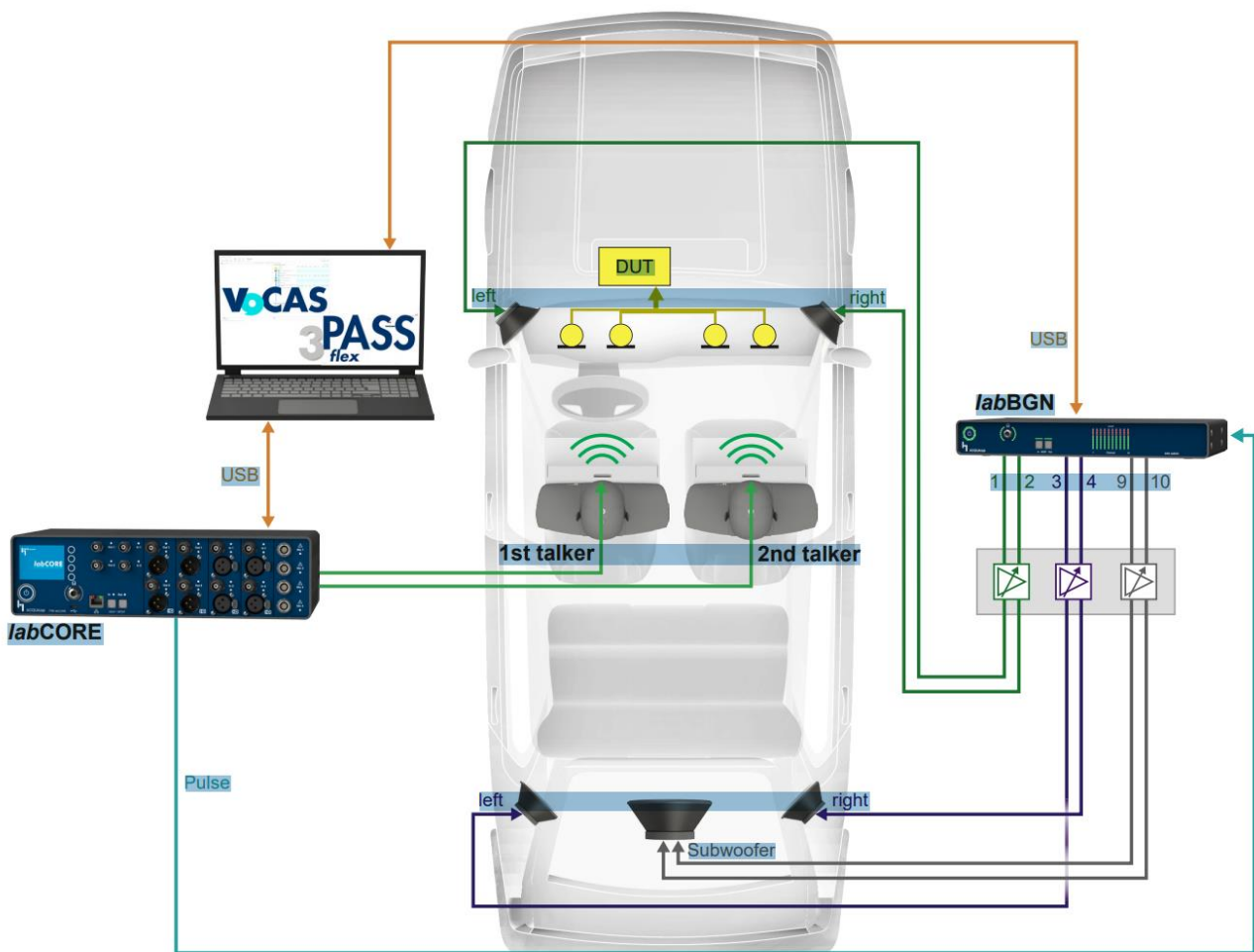
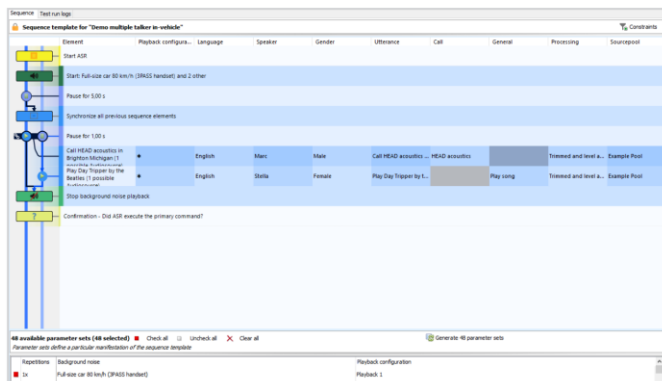
## 車室内でのバックグラウンドノイズシミュレーション環境下の単一話者テスト

車室内における基本的な測定システム構成。こちらの構成では、音源再生用にHMS II.3を利用します。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。DUTマイクروفオン（例：車載自動音声認識）が、車室内ミラーとサンバイザーに搭載されています。3PASSflexは、labBGNと車室内のスピーカーシステムを介してリアルなバックグラウンドノイズ再生を行います。



## 車室内でのバックグラウンドノイズシミュレーション環境下の複数話者テスト

車室内における高度な測定システム構成。こちらの構成では、音源再生用に2台のHMS II.3を用います。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。DUTマイクروفオン（例：車載自動音声認識）は、車室内ミラーとサンバイザーに搭載されています。2台目のHATSが、ASRを妨害する同時話者をシミュレートします。3PASS flexは、labBGNと車室内に配置したスピーカーシステムを介してリアルなバックグラウンドノイズ再生を行います。





「Python」とPythonのロゴは、Python Software Foundationの商標または登録商標であり、HEAD acoustics GmbH は Foundationの許可を得て使用しています。



#### 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)