

title: 音声サービスとは titleSuffix: Azure Cognitive Services description: Speech Service は、音声テキスト変換、テキスト読み上げ、音声翻訳が 1 つの Azure サブスクリプションにまとめられたものです。Speech SDK、Speech Devices SDK、または REST API を使用して、アプリケーション、ツール、デバイスに音声を追加します。services: cognitive-services author: trevorbye manager: nitinme ms.service: cognitive-services ms.subservice: speech-service ms.topic: overview ms.date: 11/23/2020 ms.author: trbye ms.openlocfilehash: ddb61a79e6af402b06fe8c18df8c62f9287437a3 ms.sourcegitcommit: c3739cb161a6f39a9c3d1666ba5ee946e62a7ac3 ms.translationtype: HT ms.contentlocale: ja-JP ms.lasthandoff: 04/08/2021 ms.locfileid: "107209716"

# 音声サービスとは

音声サービスは、音声テキスト変換、テキスト読み上げ、音声翻訳が 1 つの Azure サブスクリプションにまとめられたものです。[Speech CLI](#)、[Speech SDK](#)、[Speech Devices SDK](#)、[Speech Studio](#)、または [REST API](#) を使用して、アプリケーション、ツール、デバイスを簡単に音声対応にできます。

[!IMPORTANT] Speech Service は、Bing Speech API と Translator Speech に代わるものです。移行の手順については、"移行"に関するセクションを参照してください。

以下の機能は、Speech Service に属しています。各機能の一般的なユースケースの詳細を学習したり、API リファレンスを参照したりするには、次の表のリンクを使用してください。

サービス	機能	説明	SDK	REST
<a href="#">音声テキスト変換</a>	リアルタイムの音声テキスト変換	音声テキスト変換では、オーディオ ストリームまたはローカル ファイルからテキストへの文字起こしまたは翻訳がリアルタイムで行われ、結果のテキストをアプリケーション、ツール、またはデバイスで使用したり表示したりできます。音声テキスト変換を <a href="#">Language Understanding (LUIS)</a> で使用し、文字起こしされた音声からユーザーの意図を抽出し、音声コマンドで対応します。	<a href="#">はい</a>	<a href="#">はい</a>
	<a href="#">一括音声テキスト変換</a>	一括音声テキスト変換では、Azure Blob Storage に格納された大量の音声データを対象に、音声からテキストへの文字起こしを非同期で行うことができます。音声からテキストへの変換に加え、一括音声テキスト変換は、ダイアライゼーションや感情分析にも対応します。	<a href="#">いいえ</a>	<a href="#">はい</a>
	<a href="#">マルチデバイスの会話</a>	会話の中で複数のデバイスまたはクライアントを接続して、音声ベースまたはテキストベースのメッセージを送信し、文字起こしや翻訳を容易に実現します。	<a href="#">はい</a>	<a href="#">いいえ</a>
	<a href="#">会話の文字起こし</a>	リアルタイムの音声認識、話者識別、およびダイアライゼーションが有効になります。話者を区別することができるため、対面会議の文字起こしに最適です。	<a href="#">はい</a>	<a href="#">いいえ</a>
<a href="#">音声合成</a>	<a href="#">カスタム音声モデルの作成</a>	独自環境での認識と文字起こしに音声テキスト変換を使用している場合は、カスタムの音響、言語、発音モデルを作成してトレーニングし、周囲の雑音や業界固有の語彙に対応できます。	<a href="#">いいえ</a>	<a href="#">はい</a>
	テキスト読み上げ	テキスト読み上げでは、 <a href="#">音声合成マークアップ言語 (SSML)</a> が使用され、入力テキストが人間のような合成音声に変換されます。ディープ ニューラル ネットワークを利用した、人間に似た音声であるニューラル音声を使用します。 <a href="#">言語のサポート</a> に関するページを参照してください。	<a href="#">あり</a>	<a href="#">はい</a>
	<a href="#">Custom Voice の作成</a>	ブランドや製品に固有のカスタム音声フォントを作成します。	<a href="#">いいえ</a>	<a href="#">はい</a>
<a href="#">音声翻訳</a>	音声翻訳	音声翻訳を使用すると、音声のリアルタイムの多言語翻訳がアプリケーション、ツール、デバイスで可能になります。音声間翻訳や音声テキスト翻訳にはこのサービスを使用します。	<a href="#">はい</a>	<a href="#">いいえ</a>
<a href="#">音声アシスタント</a>	音声アシスタント	Speech Service を使用した音声アシスタントにより、開発者は、そのアプリケーションとエクスペリエンスのための自然で人間のような会話型インターフェイスを作成できます。音声アシスタント サービスは、Bot Framework の Direct Line Speech チャネルまたは統合カスタム コマンド サービスを使用してタスクを遂行するアシスタントの実装とデバイスとの間に、高速で信頼性の高い対話を実現します。	<a href="#">はい</a>	<a href="#">いいえ</a>
<a href="#">Speaker Recognition</a>	話者の検証と識別	Speaker Recognition サービスは、固有の音声特性によって話者を検証および識別するアルゴリズムを提供します。Speaker Recognition は、"だれが話しているのか"という質問に回答するために使用されます。	<a href="#">はい</a>	<a href="#">はい</a>

[!INCLUDE [TLS 1.2 enforcement](#)]

## Speech Service を無料で試す

次の手順では、Microsoft アカウントと Azure アカウントの両方が必要です。Microsoft アカウントがない場合は、[Microsoft アカウント ポータル](#)で無料のものにサインアップすることができます。[Microsoft アカウントでサインイン] を選択し、サインインを求められたら [Microsoft アカウントの作成] を選択します。新しい Microsoft アカウントを作成して確認する手順に従います。

Azure アカウントを作成したら、[Azure サインアップ ページ](#)に移動し、[無料で始める] を選択し、Microsoft アカウントを使用して新しい Azure アカウントを作成します。[Azure 無料アカウントにサインアップする方法](#)については、こちらの動画をご覧ください。

[!NOTE] 無料の Azure アカウントにサインアップすると、最大 30 日間有効な、有料の音声サービス サブスクリプションに充当できる 200 ドルのサービス クレジットが付属しています。クレジットの残高がなくなるか、30 日が経過して有効期限が切れると、ご利用の Azure サービスは無効になります。Azure サービスを引き続き使用するには、アカウントをアップグレードする必要があります。詳細については、[Azure 無料アカウントのアップグレード方法](#)に関するページを参照してください。

音声サービスには、無料 (f0) とサブスクリプション (s0) の 2 つのサービス レベルがあります。これには、異なる制限事項と特典があります。無料の小規模の音声サービス レベルを使用する場合は、無料試用版またはサービス クレジットの有効期限が切れた後でも、この無料のサブスクリプションを保持できます。詳細については、「[Cognitive Services の価格 - Speech Services](#)」を参照してください。

## Azure リソースを作成する

Speech Service リソース (無料または有料レベル) を Azure アカウントに追加するには:

1. お使いの Microsoft アカウントを使用して [Azure portal](#) にサインインします。
  2. ポータル左上にある **[Create a resource](リソースの作成)** を選択します。[リソースの作成] が表示されない場合は、画面左上の折りたたまれたメニューを選択することで、いつでも見つけることができます。
  3. **新規** ウィンドウで、検索ボックスに「speech」と入力し、Enter キーを押します。
  4. 検索結果で、**[Speech]** を選択します。
- 
1. **[作成]** を選択して、次のことを行います。
  2. 新しいリソースに一意の名前を指定します。この名前は、同じサービスに関連付けられた複数のサブスクリプションを区別するのに役立ちます。
  3. 新しいリソースが関連付けられている Azure サブスクリプションを選択して、料金の課金方法を決定します。Azure portal で [Azure サブスクリプションを作成する方法](#)の概要はこちらにあります。
  4. リソースが使用される[リージョン](#)を選択します。Azure は、世界中のさまざまな地域で一般的に利用できるグローバル クラウド プラットフォームです。パフォーマンスを最適にするには、ユーザーまたはアプリケーションが実行されている場所に最も近いリージョンを選択します。音声サービスの可用性は、リージョンによって異なります。サポートされているリージョンにリソースが作成されていることを確認してください。[音声サービスがサポートされているリージョン](#)に関するページを参照してください。
  5. 無料 (F0) または有料 (S0) の価格レベルのどちらかを選択します。各レベルの価格と使用量クォータの完全な情報については、**[価格の詳細を表示]** を選択するか、[音声サービスの価格](#)に関するページを参照してください。リソースの制限については、「[Azure Cognitive Services の制限](#)」を参照してください。
  6. この Speech サブスクリプションの新しいリソース グループを作成するか、既存のリソース グループにサブスクリプションを割り当てます。リソース グループは、さまざまな Azure サブスクリプションを整理しておくのに役立ちます。
  7. **[作成]** を選択します これでデプロイの概要に移動し、デプロイの進行状況を示すメッセージが表示されます。
- <!--

[!NOTE] You can create an unlimited number of standard-tier subscriptions in one or multiple regions. However, you can create only one free-tier subscription. Model deployments on the free tier that remain unused for 7 days will be decommissioned automatically. -->  
> 新しい音声リソースを展開するまでに少し時間がかかります。

## キーとリージョンを見つける

完成したデプロイのキーとリージョンを見つけるには、次の手順に従います。

1. お使いの Microsoft アカウントを使用して [Azure portal](#) にサインインします。
2. **[すべてのリソース]** を選択し、Cognitive Services リソースの名前を選択します。
3. 左ペインの **[リソース管理]** から **[Keys and Endpoint](キーとエンドポイント)** を選択します。

各サブスクリプションには 2 つのキーがあります。アプリケーションでどちらのキーを使用しても構いません。キーをコード エディターやその他の場所にコピーして貼り付けるには、各キーの横にあるコピー ボタンを選択し、ウィンドウを切り替えてクリップボードの内容を目的の場所に貼り付けます。

さらに、SDK 呼び出しのリージョン ID (、など) である 値をコピーします。SDK 呼び出しのリージョン ID (westus、westeurope など) です。

[!IMPORTANT] これらのサブスクリプション キーは、Cognitive Service API にアクセスするために使用されます。キーを共有しないでください。Azure Key Vault を使用するなどして、安全に保管してください。これらのキーを定期的に再生成することもお勧めします。API 呼び出しを行うために必要なキーは 1 つだけです。最初のキーを再生成するときに、2 番目のキーを使用してサービスに継続的にアクセスすることができます。

# クイックスタートに取り組む

ほとんどの一般的なプログラミング言語向けのクイックスタートが提供されています。それぞれ基本的な設計パターンを学び、いずれも 10 分もかからずにコードを実行できるように作られています。それぞれの機能のクイックスタートについては、次の各記事を参照してください。

- [音声テキスト変換のクイックスタート](#)
- [テキスト読み上げのクイックスタート](#)
- [音声翻訳のクイックスタート](#)
- [意図認識のクイックスタート](#)
- [話者認識のクイックスタート](#)

Speech Service を使用する機会が得られたら、さまざまなシナリオを解決する方法を説明するチュートリアルを試してみてください。

- [チュートリアル:Speech SDK、LUIS、C# を使用して音声の意図を認識する](#)
- [チュートリアル:Speech SDK を使用してボットを音声対応にする \(C#\)](#)
- [チュートリアル:テキストを翻訳してセンチメントを分析し、翻訳済みテキストから音声を合成する Flask アプリを作成する \(REST\)](#)

## サンプルコードを取得する

Speech Service のサンプルコードは、GitHub 上で入手できます。これらのサンプルでは、ファイルやストリームからの音声の読み取り、連続的な認識と単発の認識、カスタムモデルの使用など、一般的なシナリオについて説明されています。SDK と REST のサンプルを見るには、次のリンクを使用してください。

- [音声変換、テキスト読み上げ、および音声翻訳のサンプル \(SDK\)](#)
- [バッチ文字起こしのサンプル \(REST\)](#)
- [テキスト読み上げのサンプル \(REST\)](#)
- [音声アシスタントのサンプル \(SDK\)](#)

## 音声エクスペリエンスをカスタマイズする

Speech Service は組み込みのモデルで問題なく動作しますが、製品や環境に合わせてエクスペリエンスをさらにカスタマイズおよび調整したいことがあります。カスタマイズ オプションは、音響モデルのチューニングから、独自ブランドに固有の音声フォントにまで及びます。

その他の製品では、医療や保険など特定の目的に合わせてチューニングされた音声モデルが用意されていますが、それらは、すべてのユーザーに対して一律に提供されるものです。Azure Speech のカスタマイズは、他のユーザーやカスタマーが手にすることのできない "独自" の競争優位をもたらします。つまり、カスタマイズしたモデルは自分だけのものであり、ユースケースに合わせて独自にチューニングされています。

Speech Service	プラットフォーム	説明
音声テキスト変換	<a href="#">カスタム音声</a>	音声認識モデルをニーズと使用可能なデータに合わせてカスタマイズします。話し方、ボキャブラリ、背景ノイズといった音声認識の障壁を克服可能です。
テキスト読み上げ	<a href="#">Custom Voice</a>	お客様のスピーチ データを使用して、テキスト読み上げアプリ用の認識可能な独自の音声を作成します。音声パラメーターのセットを調整することで、音声出力をさらに細かく調整できます。

## Docker コンテナを使用してオンプレミスにデプロイする

[Speech サービス コンテナ](#)を使用して、API 機能をオンプレミスにデプロイします。これらの Docker コンテナを使用すると、コンプライアンス、セキュリティ、またはその他の運用上の理由により、データにより近いところでサービスを使用できます。Speech サービスには、次のコンテナが用意されています。

- 標準音声変換
- カスタム音声変換
- 標準テキスト読み上げ
- ニューラル テキスト読み上げ
- カスタム テキスト読み上げ (プレビュー)
- 音声言語検出 (プレビュー)

## リファレンス ドキュメント

- [Speech SDK](#)
- [Speech Devices SDK](#)
- [REST API: 音声テキスト変換](#)

- [REST API: テキスト読み上げ](#)
- [REST API: 一括文字起こしとカスタマイズ](#)

## 次のステップ

[!div class="nextstepaction"] [音声テキスト変換の概要](#) [テキスト読み上げの概要](#)