

COTOHA音声認識

なお、Developers環境アカウントでは、本APIをご利用いただけませんのでご了承ください。
音声認識APIは6つのAPIで構成されます。

1. [ファイル音声認識](#)
 - 短時間の音声ファイルの音声をテキストに変換するAPIです。
2. [ストリーミング音声認識](#)
 - 長時間の音声ファイルの音声やマイクからの入力などのストリーミング音声をテキストに変換するAPIです。
3. [音声認識ユーザ辞書更新](#)※
 - 音声認識APIの語彙に含まれていない単語を登録するAPIです。
 - 辞書は毎時0分時点のデータが反映されます。実際にデータが反映されるまでには、一定の時間を要します。
4. [音声認識ユーザ辞書クリア](#)※
 - 登録された単語のクリアを行うAPIです。
 - 辞書は毎時0分時点のデータが反映されます。実際にデータが反映されるまでには、一定の時間を要します。
5. [音声認識ユーザ辞書適用状態取得](#)※
 - 辞書のデータが反映されているかを確認するAPIです。
6. [音声認識ユーザ辞書ダウンロード](#)※
 - 最後にアップロードされた辞書をダウンロードするAPIです。

※これらのユーザ辞書を扱うAPIは、日本語モデルのみの提供となります。

GitHub

[GitHub](#)にて、サンプルコードや本リファレンスのPDFを公開しています。

- [Python](#)
- [Node.js](#)
- [JavaScript](#)
- [リファレンス](#)

こんな時はどうする？

- ファイルを音声認識させたい
 - 音声長が60秒を超える場合は、[ストリーミング音声認識](#)をご利用下さい。
 - それ以下の場合は、[ファイル音声認識](#)をご利用下さい。
- リアルタイムに音声認識させたい
 - [ストリーミング音声認識](#)をご利用下さい。
- wavファイルを音声認識させたい
 - wavファイルの音声認識を行いたい場合、ヘッダの除去などを適切に行いLinear PCMもしくはμ-Lawフォーマットに変換を行なって下さい。

- [GitHub](#)のサンプルコードでは、入力をwavファイルとしております。ご参照ください。
- 辞書登録を行いたい
 - 事前に単語を登録したい場合は、[音声認識ユーザ辞書更新](#)をご利用下さい。
 - 音声認識シーケンスごとに単語を変えたい場合は、[一時辞書](#)をご利用下さい。

ファイル音声認識

ファイル音声認識では、短時間の音声ファイルの音声を認識してテキストに変換します。
対象は60秒以内の音声ファイルとなります。60秒を超える場合は、[ストリーミング音声認識](#)を利用してください。

HTTPリクエスト

```
POST <API Base URL>/asr/v1/speech_recognition/<ASR Model ID>
```

ASR Model ID は音声認識に用いるモデルを識別するIDです。[モデル一覧](#)を参照して、適切なモデルを選択してください。

リクエストヘッダ

リクエストはマルチパートで記述します。マルチパートに用いる境界文字列を定めて、リクエストヘッダのboundaryで示してください。

キー名	説明
Content-Type	multipart/form-data; boundary=<境界文字列>
Authorization	Bearer [Access Token]

リクエストボディ

マルチパートの各パートに設定するデータの内容および構造は、「パート種別」毎に規定します。
「パート種別」は以下のものを使用します。なお、リクエスト内のパート種別は「Content-Disposition」のnameパラメータで設定します。
ボディのパートの順番は、必ず、パラメータパート、音声データパート、コマンドパートの順で設定してください。

パート種別	必須	name	Content-Type
パラメータパート	必須	parameter	application/json; charset=UTF-8
音声データパート	必須	audio	application/octet-stream
コマンドパート	必須	command	application/json; charset=UTF-8

パラメータパート

音声認識開始要求を行います。詳細は[ストリーミング音声認識](#) - [音声認識開始要求](#)の項を参照してください。

音声データパート

音声データは下記のいずれかの形式に変換をして、バイナリを音声データパートとしてください。

※wavファイルの音声認識を行いたい場合、ヘッダの除去などを適切に行いLinear PCMもしくはμ-Lawフォーマットに変換を行なって下さい。

フォーマット	サンプリングレート[Hz]	分解能 [bit]	チャネル	バイト順
Linear PCM	モデルの帯域(8000または16000)以上※	16	モノラル	リトルエンディアン
μ-Law	8000	8	モノラル	リトルエンディアン

※ ただし、8000, 11025, 16000, 22050, 32000, 44100, 48000, 88200, 96000[Hz]以外の音声に対しては、正常な動作を保証しません。

コマンドパート

音声認識停止要求を行います。詳細は[ストリーミング音声認識](#) - [音声認識停止要求](#)の項を参照してください。

リクエストサンプル

以下にリクエストのサンプルを示します。

HTTPヘッダ

```
Content-Type: multipart/form-data; boundary=<境界文字列>
Authorization: Bearer <Access Token>
```

HTTPボディ

```
--<境界文字列>
Content-Disposition: form-data; name="parameter"
Content-Type: application/json

{
  "msg":
  {
    "msgname": "start"
  },
  "param":
  {
    "baseParam.samplingRate": 16000,
    "recognizeParameter.domainId": "<ASR Domain ID>"
  }
}
--<境界文字列>
```

```
Content-Disposition: form-data; name="audio"
Content-Type: application/octet-stream
```

<バイナリ音声データ>

--<境界文字列>

```
Content-Disposition: form-data; name="command"
Content-Type: application/json
```

```
{
  "msg": {
    "msgname": "stop"
  }
}
--<境界文字列>--
```

レスポンス

詳細は「ストリーミング音声認識 - レスポンス」の項を参照してください。

レスポンスサンプル

以下にレスポンスの例を示します。

```
[ {
  "msg" : {
    "msgname" : "started",
    "uniqueId" : "4d97031a-cfa9-4d66-968a-be708644a893"
  }
}, {
  "msg" : {
    "msgname" : "speechStartDetected",
    "uniqueId" : "4d97031a-cfa9-4d66-968a-be708644a893"
  },
  "timeinfo" : {
    "startDetectTime" : 0
  }
}, {
  "msg" : {
    "msgname" : "speechEndDetected",
    "uniqueId" : "4d97031a-cfa9-4d66-968a-be708644a893"
  },
  "timeinfo" : {
    "endDetectTime" : 3460
  }
}, {
  "msg" : {
    "msgname" : "recognized",
    "uniqueId" : "4d97031a-cfa9-4d66-968a-be708644a893"
  },
  "result" : {
    "type" : 1,
    "sentence" : [ {
      "surface" : "これは テスト 用 の 音声ファイル です",
      "score" : 0.975848,
      "startTime" : 0.0,
      "endTime" : 3.46
    }
  ]
}
```

```
    } ]
  }
}, {
  "msg" : {
    "msgname" : "recognized",
    "uniqueId" : "4d97031a-cfa9-4d66-968a-be708644a893"
  },
  "result" : {
    "type" : 2,
    "sentence" : [ ]
  }
}, {
  "msg" : {
    "msgname" : "completed",
    "uniqueId" : "4d97031a-cfa9-4d66-968a-be708644a893",
    "cause" : "STOP"
  }
} ]
```

エラーコード

リクエストの内容を確認してください

以下のエラーが発生した場合は、記載された対処をした上でリクエストをやり直してください。エラーが解決しない場合、[ご契約者様お問い合わせ](#)よりお問い合わせください。

code	message	detail	説明および対処
410	Invalid Parameter	msg.msgname is invalid -> XXX.	パラメータ不正。 msgnameを確認してください。
410	Invalid Parameter	param.baseParam.samplingRate is invalid -> XXX and over.	パラメータ不正。サンプリングレートがモデルの帯域以上となっているか確認してください。
410	Invalid Parameter	recognizeParameter.enableProgress must be false for file speech recognition.	パラメータ不正。ファイル音声認識では逐次認識結果出力機能をご利用いただけません。
410	Invalid Parameter	param.baseParam.samplingRate is invalid -> The sampling rate must be 8000 Hz.	パラメータ不正。サンプリングレートが8000Hzとなっているか確認してください。
411	Invalid State	Invalid Unique ID	実行不能ステート。音声認識APIの呼び出し順番を確認してください。

code	message	detail	説明および対処
413	Invalid Data	Data is not available	最大音声長の超過エラー。音声データ長が最大音声長を超える場合、複数回に分けて認識を行ってください。また、サンプリングレート、音声コーデック、モデルの指定が適切かご確認ください。
550	No Resource	VRG_RESPONSE_SERVICE_UNAVAILABLE	空きリソースが存在しません。ASR Domain IDが正しいかご確認ください。ASR Domain IDが正しい場合、APIへのアクセスが混雑している可能性がございます。しばらく時間を置いて、リクエストをやり直してください。
651	Session Timeout		セッションタイムアウト。APIの呼び出し順番を確認してください。
652	Excess Of Max Voice Length	Excess Of Max Voice Length -> 60 [s]	最大音声長の超過エラー。音声データ長が最大音声長を超える場合、複数回に分けて認識を行ってください。また、サンプリングレート、音声コーデック、モデルの指定が適切かご確認ください。

一時辞書の内容を確認してください

以下のエラーが発生した場合は、記載された対処をした上でリクエストをやり直してください。エラーが解決しない場合、[ご契約者様お問い合わせ](#)よりお問い合わせください。

code	message	detail	対処
410	Invalid Parameter	Invalid Parameter -> List is null	表記・読みを記述してください。
410	Invalid Parameter	Cascade words exceeded 1000	一時辞書では、追加単語を1000行以下で指定してください。
410	Invalid Parameter	Surface empty	表記を記述してください。
410	Invalid Parameter	Reading empty	読みを記述してください。
410	Invalid Parameter	Invalid domainid	ASR Domain IDを確認してください。
410	Invalid Parameter	Invalid Model Name	ASR Model IDを確認してください。
410	Invalid Parameter	Prob illegal	重みは1~100の整数値で指定してください。
410	Invalid Parameter	Surface illegal	表記として設定可能な文字列か確認してください。
410	Invalid Parameter	Reading illegal	読みとして設定可能な文字列か確認してください。
410	Invalid Parameter	Surface too long	表記と読みの合計を251byte以下で指定してください。

お問い合わせください

以下のエラーが発生した場合は、[ご契約者様お問い合わせ](#)よりお問い合わせください。

エラーコード	メッセージ
500	Internal Error
510	Out Of Memory
551	Recognition Timeout
552	Network Error
553	Network Timeout
601	Recognition Converter Error
610	Out Of Memory
611	Invalid License
612	Invalid Config
650	No Resource
690	External Command Execute Failed
691	External Command Fatal
692	External Command Error
693	External Command Warn

音声認識開始要求時のパラメータをご確認ください

音声認識開始要求時にパラメータ不正がある場合、HTTPステータスコード500で下記のレスポンスが返却される場合があります。

```
{
  "fault": {
    "faultstring": "NullPointerException",
    "detail": {
      "errorcode": "Internal Server Error"
    }
  }
}
```

ストリーミング音声認識

ストリーミング音声認識では、マイクからの入力などのストリーミング音声や長時間の音声ファイルの音声を認識してテキストに変換します。音声の全体の長さは最長3,000秒です。3,000秒を超える音声は、3,000秒以下に分割して、それぞれストリーミング音声認識を実行してください。

要求の種類

ストリーミング音声認識では、以下の4種類の要求を利用します。

- 音声認識開始要求
- 音声認識データ送信

- 音声認識停止要求
- 音声認識キャンセル要求

標準的な要求の流れ（シーケンス）は以下の通りです。

1. 音声認識開始要求
2. 音声データ送信（複数回）
3. 音声認識停止要求

音声認識開始要求

音声認識クライアントから音声認識機能を利用する際に、音声認識クライアントが最初に音声認識APIサーバへ要求するリクエストです。本リクエストで音声認識のパラメータ設定が必要です。

HTTPリクエスト

```
POST <API Base URL>/asr/v1/speech_recognition/<ASR Model ID>
```

ASR Model ID は音声認識に用いるモデルを識別するIDです。[モデル一覧](#)を参照して、適切なモデルを選択してください。

リクエストヘッダ

キー名	説明
Connection	Keep-Alive
Content-Type	application/json; charset=UTF-8
Authorization	Bearer [Access Token]

リクエストボディ

キー名		必須	説明	設定可能範囲
msg				
	msgname	必須	メッセージ種別	start
param				
	baseParam.samplingRate	必須	音声データのサンプリングレート	pcmの場合はモデルの帯域(8000, 16000)以上の値※1、mulawの場合は8000
	recognizeParameter.domainId	必須	[ASR Domain ID]	8桁の半角英数字
	recognizeParameter.enableContinuous	任意 ※2	連続認識機能有効設定	true (default: true)
	recognizeParameter.enableProgress	任意	逐次認識結果出力有効設定 ※3	true, false (default: false)
	recognizeParameter.maxResults	任意	認識結果最大数	1~30 (default: 1)
	baseParam.audioCodec	任意	音声コーデック	pcm, mulaw (default: pcm)
	baseParam.filler	任意	言い淀みの出力有無	true, false (default: false)
	baseParam.punctuation	任意	句読点の出力有無	true, false (default: false)
	baseParam.reading	任意	読み仮名の出力有無	true, false (default: false)
	baseParam.delimiter	任意	区切り文字（半角スペース）の出力有無	true, false (default: true)
	words	任意	一時辞書	1000語以下

※1 ただし、8000, 11025, 16000, 22050, 32000, 44100, 48000, 88200, 96000[Hz]以外の音声に対しては、正常な動作を保証しません。

※2 2020/3/23 までは連続認識機能有効設定は必須となっていました、任意のパラメータとなりました。

※3 ただし、本機能はファイル音声認識APIではご利用いただけません。

音声コーデック

Linear PCM(pcm)およびμ-Law(mulaw)の指定が可能です。ただし、μ-Law形式の音声認識は、帯域が8kHzのモデル（ja-gen_tf-08, ja-md11_nrw-08, ja-md12_nrw-08）のみでご利用いただけます。wavファイルの音声認識を行いたい場合、ヘッダの除去などを適切に行いLinear PCMもしくはμ-Lawフォーマットに変換を行なって下さい。

言い淀みの出力有無

本パラメータをtrueに設定することで、言い淀み（「あー」「えー」といった発話）を[]付きの形で出力することが出来ます。

※日本語モデル（ja-gen_tf-16, ja-gen_sf-16, ja-gen_tf-08, ja-md11_nrw-08, ja-md12_nrw-08）でのみ有効です。

```
baseParam.filler=false (default)
```

これは テスト 用 の 音声ファイル です

```
baseParam.filler=true
```

[えーと] これは テスト 用 の 音声ファイル です

句読点の出力有無

本パラメータをtrueに設定することで、句読点を含む形で音声認識結果が出力されます。

※ ja-gen_tf-16, ja-gen_tf-08, ja-md11_nrw-08, ja-md12_nrw-08 モデルでのみ有効です。

※文末に句点がつく保証はございません。

```
baseParam.punctuation=false (default)
```

こんにちは よろしくお願ひします

```
baseParam.punctuation=true
```

こんにちは 。 よろしくお願ひします

読み仮名の出力有無

本パラメータをtrueに設定することで、認識結果を漢字仮名交じり文だけではなく、カタカナでも取得可能です。

※日本語モデル（ja-gen_tf-16, ja-gen_sf-16, ja-gen_tf-08, ja-md11_nrw-08, ja-md12_nrw-08）でのみ有効です。

```
'reading': 'コレ ワ テスト ヨウ ノ オンセイファイル デス'
```

readingを含む出力フォーマットについては[レスポンスサンプル](#)も参照してください。

区切り文字（半角スペース）の出力有無

本パラメータを `false` に設定することで、通常の認識結果で単語間に挿入される半角スペースを除去することが可能です。

※英語モデル（`en_en-gen_sf-16`）ご利用時は、`true` を推奨します。

```
baseParam.delimiter=true (default)
```

これはテスト用の音声ファイルです

```
baseParam.delimiter=false
```

これはテスト用の音声ファイルです

逐次認識結果出力有効設定

本パラメータを `true` に設定することで、終話に至っていない発話の認識途中結果を逐次的に取得することが可能です。

認識途中結果は認識結果タイプ `type=0` として、返却されます。

なお、認識途中結果に対しては、各出力有無パラメータの指定に関わらず、一律で読み仮名の出力はされず、言い淀み・句読点・区切り文字は出力されます。

認識結果最大数

`recognizeParameter.maxResults` として指定された数を上限として、音声認識結果を複数返却します。ただし、指定した数は上限値であり、必ずその数だけ返却されることは保証されていません。

`recognizeParameter.maxResults` を2以上とした場合の出力フォーマットについては[レスポンスサンプル](#)を参照してください。

一時辞書

音声認識シーケンスごとに設定可能な一時的な辞書です。

[音声認識ユーザ辞書更新](#)と併用してご利用いただけます。

一時辞書はシーケンスが終了すると破棄され、ユーザ辞書の更新は行われません。

キー名	説明	値の範囲	必須
surface	表記※1	<code> </code> , <code>/</code> , <code>:</code> , <code>;</code> , <code>[</code> , <code>]</code> , <code>\t</code> , <code>\n</code> を含まない、かつ空文字ではない文字列	必須
reading	読み※1	全角カタカナ（ 別表 に記載）	必須
prob	単語の重み※2	1~100の整数値（default: 30）	任意

※1 表記と読みの合計を251byte以下で指定してください。

※2 登録単語の出現確率を調整したい場合にご利用ください。

30より大きい値を設定することで、通常の辞書登録よりも高めの効果で辞書登録が可能です。

また、30より小さい値を設定することで、通常の辞書登録よりも低めの効果で辞書登録が可能です。

リクエストサンプル

以下にリクエストのサンプルを示します。

HTTPヘッダ

```
Connection: Keep-Alive
Content-Type: application/json; charset=UTF-8
Authorization: Bearer <Access Token>
```

HTTPボディ

```
{
  "msg": {
    "msgname": "start"
  },
  "param": {
    "baseParam.samplingRate": 16000,
    "recognizeParameter.domainId": "<ASR Domain ID>",
    "baseParam.audioCodec": "pcm",
    "baseParam.filler": true,
    "baseParam.reading": true,
    "baseParam.delimiter": false,
    "baseParam.punctuation": true,
    "recognizeParameter.enableProgress": true,
    "recognizeParameter.maxResults": 2
  },
  "words": [
    {
      "surface": "エヌ・ティ・ティ・コミュニケーションズ株式会社",
      "reading": "エヌティティコミュニケーションズカブシキガイシャ",
      "prob": "20"
    },
    {
      "surface": "COTOHA",
      "reading": "コトハ",
      "prob": "40"
    }
  ]
}
```

音声認識データ送信

音声認識を行うデータを音声認識APIサーバへ送信するリクエストです。リクエストを送ってから次のリクエストを送るまでの間隔を240ミリ秒[ms]※としてください。音声認識データ送信のレスポンスを待ち、エラーがないことを確認して次のデータを送信してください。なお、最初の音声認識データ送信は音声認識開始要求のレスポンスを受けてから1秒以内※に行なってください。

※通信環境等の問題で、所定の時間内に次のリクエストを送信できない場合は、レスポンスを受信次第、次のリクエストを送信して下さい。

HTTPリクエスト

```
POST <API Base URL>/asr/v1/speech_recognition/<ASR Model ID>
```

リクエストヘッダ

キー名	説明
Connection	Keep-Alive
Content-Type	application/octet-stream
Unique-Id	音声認識開始要求のレスポンスヘッダで返されたUniqueID
Authorization	Bearer [Access Token]
Cookie	- token（リクエストの順番を識別するためのユニークな文字列。直近のレスポンスで受け取ったtokenを設定します。） - GCLB（音声認識のシーケンスごとに設定されるユニークな文字列。音声認識開始要求のレスポンスで受け取ったGCLBを設定します。）

リクエストボディ

音声データは下記のいずれかの形式に変換をして、バイナリをリクエストボディとしてください。

※wavファイルの音声認識を行いたい場合、ヘッダの除去などを適切に行いLinear PCMもしくはμ-Lawフォーマットに変換を行なって下さい。

フォーマット	サンプリングレート[Hz]	分解能[bit]	チャネル	バイト順	音声データ長[ms]
Linear PCM	モデルの帯域(8000または16000)以上※1	16	モノラル	リトルエンディアン	240※2
μ-Law	8000	8	モノラル	リトルエンディアン	240※2

すなわち、1回のリクエストボディのサイズは、Linear PCMフォーマットの場合は8kHzのモデルの場合は3840バイト、16kHzのモデルの場合は7680バイト、μ-Lawフォーマットの場合は1720バイトとなります。

※1 ただし、8000, 11025, 16000, 22050, 32000, 44100, 48000, 88200, 96000[Hz]以外の音声に対しては、正常な動作を保証しません。

※2 ただし、音声認識データ送信の最後のリクエストに限り、240ミリ秒[ms]未満のデータも可

リクエストサンプル

以下にリクエストのサンプルを示します。

HTTPヘッダ

```
Connection: Keep-Alive
Content-Type: application/octet-stream
Authorization: Bearer <Access Token>
Cookie: GCLB=<音声認識開始要求のレスポンスで受け取ったGCLBを設定>; token=<直近のレスポンスで受け取ったtoken>
```

HTTPボディ

<バイナリ音声データ>

音声認識停止要求

音声認識クライアントから音声認識処理の停止要求するリクエストです。音声認識APIサーバは全ての音声認識が完了したのち、200(OK)を返します。なお、音声認識停止要求は音声認識データ送信の最後のレスポンスを受けてから、1秒以内に行なってください。

HTTPリクエスト

POST <API Base URL>/asr/v1/speech_recognition/<ASR Model ID>

リクエストヘッダ

キー名	説明
Connection	Keep-Alive
Content-Type	application/json; charset=UTF-8
Unique-Id	音声認識開始要求のレスポンスヘッダで返されたUniqueID
Authorization	Bearer [Access Token]
Cookie	- token（リクエストの順番を識別するためのユニークな文字列。直近のレスポンスで受け取ったtokenを設定します。） - GCLB（音声認識のシーケンスごとに設定されるユニークな文字列。音声認識開始要求のレスポンスで受け取ったGCLBを設定します。）

リクエストボディ

キー名		説明	設定可能範囲
msg			
	msgname	メッセージ種別	stop

リクエストサンプル

以下にリクエストのサンプルを示します。

HTTPヘッダ

Connection: Keep-Alive
Content-Type: application/json; charset=UTF-8
Authorization: Bearer <Access Token>
Cookie: GCLB=<音声認識開始要求のレスポンスで受け取ったGCLBを設定>; token=<直近のレスポンスで受け取ったtoken>

HTTPボディ

```
{
  "msg": {
    "msgname": "stop"
  }
}
```

音声認識キャンセル要求

音声認識クライアントから音声認識処理のキャンセルを要求するリクエストです。キャンセルでは送信済み音声データの認識処理は行わず、200(OK)を返します。なお、音声認識キャンセル要求は音声認識データ送信の最後のレスポンスを受けてから、1秒以内に行なってください。

HTTPリクエスト

```
POST <API Base URL>/asr/v1/speech_recognition/<ASR Model ID>
```

リクエストヘッダ

キー名	説明
Connection	Keep-Alive
Content-Type	application/json; charset=UTF-8
Unique-Id	音声認識開始要求のレスポンスヘッダで返されたUniqueID
Authorization	Bearer [Access Token]
Cookie	- token（リクエストの順番を識別するためのユニークな文字列。直近のレスポンスで受け取ったtokenを設定します。） - GCLB（音声認識のシーケンスごとに設定されるユニークな文字列。音声認識開始要求のレスポンスで受け取ったGCLBを設定します。）

リクエストボディ

キー名		説明	設定可能範囲
msg			
	msgname	メッセージ種別	cancel

リクエストサンプル

以下にリクエストのサンプルを示します。

HTTPヘッダ

```
Connection: Keep-Alive
Content-Type: application/json; charset=UTF-8
Authorization: Bearer <Access Token>
Cookie: GCLB=<音声認識開始要求のレスポンスで受け取ったGCLBを設定>; token=<直近のレスポンスで受け取ったtoken>
```

HTTPボディ

```
{
  "msg": {
    "msgname": "cancel"
  }
}
```

レスポンス

HTTPリクエストに対して、以下のサーバ応答の組み合わせからなるHTTPレスポンスが返却されます。サーバ応答が1個以上含まれる場合はHTTPステータスコード200(OK)でHTTPボディ部にJSONが含まれるHTTPレスポンスが返却されます。サーバ応答が0個の場合、HTTPステータスコード204(No Content)が返却されます。

サーバ応答の種類	msgname	備考
音声認識準備完了応答	started	音声認識開始リクエストの応答
発話検出応答	speechStartDetected	音声データの発話開始を検出後のリクエストの応答
終話検出応答	speechEndDetected	音声データの発話終了を検出後のリクエストの応答
音声認識結果応答	recognized	音声認識結果
音声認識終了応答	completed	音声認識が終了、または音声認識処理中にエラーが発生

全てのサーバ応答に「msgname」と「uniqueId」が含まれます。「msgname」はサーバ応答の種類を表し、「uniqueId」は音声認識ごとにユニークな識別子です。

音声認識準備完了応答

音声認識開始要求に対して、音声認識APIサーバから音声認識クライアントに準備の完了を通知するメッセージです。

レスポンスヘッダ

キー名	説明
Content-Type	application/json; charset=UTF-8
Set-Cookie	- token（リクエストの順番を識別するためのユニークな文字列） - GCLB（音声認識のシーケンスごとに設定されるユニークな文字列）

レスポンスボディ

キー名	説明	備考
msg		
msgname	メッセージ種別	started
uniqueId	音声認識毎にユニークな識別子	音声認識データ送信や音声認識停止要求・音声認識キャンセル要求に必要なパラメータ

レスポンスサンプル

```
[
  {
    "msg": {
      "msgname": "started",
      "uniqueId": "3bfbe5de-eee7-4824-a661-3750d8cb9328"
    }
  }
]
```

発話検出応答

転送された音声データから発話を検出した際に返されるメッセージです。

レスポンスヘッダ

キー名	説明
Content-Type	application/json; charset=UTF-8
Set-Cookie	token（リクエストの順番を識別するためのユニークな文字列）

レスポンスボディ

キー名		説明	備考
msg			
	msgname	メッセージ種別	speechStartDetected
	uniqueId	音声認識毎にユニークな識別子	音声認識APIサーバが付与した識別子
timeinfo			
	startDetectTime	発話検出時間[ms]	音声データの先頭から発話検出までの時間

レスポンスサンプル

```
[ {
  "msg" : {
    "msgname" : "speechStartDetected",
    "uniqueId" : "3bfbe5de-eee7-4824-a661-3750d8cb9328"
  },
  "timeinfo" : {
    "startDetectTime" : 0
  }
} ]
```

終話検出応答

転送された音声データで終話（発話の終了）を検出した際に返されるメッセージです。

レスポンスヘッダ

キー名	説明
Content-Type	application/json; charset=UTF-8
Set-Cookie	token（リクエストの順番を識別するためのユニークな文字列）

レスポンスボディ

キー名		説明	備考
msg			
	msgname	メッセージ種別	speechEndDetected
	uniqueId	音声認識毎にユニークな識別子	音声認識APIサーバが付与した識別子
timeinfo			
	endDetectTime	終話検出時間[ms]	音声データの先頭から終話検出までの時間

レスポンスサンプル

```
[ {
  "msg" : {
    "msgname" : "speechEndDetected",
    "uniqueId" : "4b96875f-2137-48ed-8b49-1e20483a7c86"
  },
  "timeinfo" : {
    "endDetectTime" : 3200
  }
} ]
```

音声認識結果応答

音声認識の結果を音声認識APIサーバから音声認識クライアントに送信するためのメッセージです。

レスポンスヘッダ

キー名	説明
Content-Type	application/json; charset=UTF-8
Set-Cookie	token（リクエストの順番を識別するためのユニークな文字列）

レスポンスボディ

キー名		説明	備考
msg			
	msgname	メッセージ種別	recognized
	uniqueId	音声認識毎にユニークな識別子	音声認識APIサーバが付与した識別子
result			
	type	認識結果タイプ	<ul style="list-style-type: none">0：認識途中結果1：認識結果（サーバ終話検知）2：認識結果（最終）（音声認識停止要求）
sentence			
	surface	音声認識結果テキスト	
	score	スコア	0~1。数字が大きいほど確からしさが高い
	startTime	文の開始時刻[秒]	音声データの先頭からの時間
	endTime	文の終了時刻[秒]	音声データの先頭からの時間
	reading	認識結果の読み	`baseParam.reading=true`の時のみ

レスポンスサンプル

音声認識開始要球時のパラメータ

```
baseParam.filler: false(デフォルト値)
baseParam.reading: false(デフォルト値)
baseParam.delimiter: true(デフォルト値)
baseParam.punctuation: false(デフォルト値)
recognizeParameter.maxResults: 1(デフォルト値)
```

レスポンス

```
{
  "msg": {
    "msgname": "recognized",
    "uniqueId": "49bbc297-6a8a-4354-9679-2866cc27385a"
  },
  "result": {
    "type": 1,
    "sentence": [
      {
        "surface": "これ は テスト 用 の 音声ファイル です",
        "score": 0.851436,
        "startTime": 0,
        "endTime": 2.99
      }
    ]
  }
}
```

レスポンスサンプル

言い淀みの出力あり、読み仮名の出力あり、区切り文字の出力なし、句読点の出力あり、認識結果最大数2

音声認識開始要球時のパラメータ

```
baseParam.filler: true
baseParam.reading: true
baseParam.delimiter: false
baseParam.punctuation: true
recognizeParameter.maxResults: 2
```

レスポンス

```
{
  "msg": {
    "msgname": "recognized",
    "uniqueId": "a6342b20-6b0d-41b4-9c71-86af4f123369"
  },
  "result": {
    "type": 1,
    "sentence": [
      {
        "surface": "[えーと]、これはテスト用の音声ファイルです",
        "score": 0.851436,
        "startTime": 0,
        "endTime": 2.99,
        "reading": "エット、コレワテストヨウノオンセイファイルデス"
      },

```

```
{
  "surface": "[えーと]。これはテスト用の音声ファイルです",
  "score": 0.532213,
  "startTime": 0,
  "endTime": 2.99,
  "reading": "エット。コレワテストヨウノオンセイファイルデス"
}
```

音声認識終了応答

音声認識処理の終了を音声認識APIサーバから音声認識クライアント宛に送信するメッセージです。音声認識処理中にエラーが発生した場合も本メッセージが返されます。

レスポンスヘッダ

キー名	説明
Content-Type	application/json; charset=UTF-8
Set-Cookie	token（リクエストの順番を識別するためのユニークな文字列）

レスポンスボディ

キー名	説明	備考
msg		
msgname	メッセージ種別	completed
uniqueId	音声認識毎にユニークな識別子	音声認識APIサーバが付与した識別子
cause	停止要因	以下のいずれか <ul style="list-style-type: none">STOP(音声認識停止要求)CANCEL(音声認識キャンセル要求)ERROR(エラー)END DETECTION※
errorinfo		
code	エラーコード	
message	エラーメッセージ	
level	エラーレベル	以下のいずれか <ul style="list-style-type: none">WARNERRORFATAL
detail	エラー詳細情報	

※通常は発生しない停止要因です。発生時は、音声認識開始要求時に`recognizeParameter.enableContinuous`が適切に設定されているかご確認ください。

レスポンスサンプル

```
[ {
  "msg" : {
    "msgname" : "completed",
    "uniqueId" : "4b96875f-2137-48ed-8b49-1e20483a7c86",
    "cause" : "STOP"
  }
} ]
```

エラーコード

リクエストの内容を確認してください

以下のエラーが発生した場合は、記載された対処をした上でリクエストをやり直してください。なお、ストリーム音声認識APIにおけるエラーの場合は、音声認識のシーケンスの最初、音声認識開始要求からやり直していただきます。その際、音声認識停止要求や音声認識キャンセル要求のリクエストをする必要はありません。エラーが解決しない場合、[ご契約者様お問い合わせ](#)よりお問い合わせください。

code	message	detail	説明および対処
410	Invalid Parameter	msg.msgname is invalid -> XXX.	パラメータ不正。 msgnameを確認してください。
410	Invalid Parameter	param.baseParam.samplingRate is invalid -> XXX and over.	パラメータ不正。サンプリングレートがモデルの帯域以上となっているか確認してください。
410	Invalid Parameter	param.baseParam.samplingRate is invalid -> The sampling rate must be 8000 Hz.	パラメータ不正。サンプリングレートが8000Hzとなっているか確認してください。
411	Invalid State	Invalid Unique ID	実行不能ステート。音声認識APIの呼び出し順番を確認してください。
412	Interval Too Brief		音声データの送信間隔が短すぎます。音声データの送信間隔を正しくしてください。

code	message	detail	説明および対処
450	Invalid Token	token Error	(Access Tokenではなく) Cookieのtokenが不正です。リクエストはレスポンス受信を待ってから行ってください。
550	No Resource	VRG_RESPONSE_SERVICE_UNAVAILABLE	空きリソースが存在しません。ASR Domain IDが正しいかご確認ください。ASR Domain IDが正しい場合、APIへのアクセスが混雑している可能性がございます。しばらく時間を置いて、リクエストをやり直してください。
552	Network Error	Network	レスポンスを受けてから所定の時間内にリクエストを行なっているかを確認してください。
600	Internal Error		レスポンスを受けてから所定の時間内にリクエストを行なっているかを確認してください。
651	Session Timeout		セッションタイムアウト。APIの呼び出し順番、呼び出し間隔を確認してください。

一時辞書の内容を確認してください

以下のエラーが発生した場合は、記載された対処をした上でリクエストをやり直してください。なお、ストリーム音声認識APIにおけるエラーの場合は、音声認識のシーケンスの最初、音声認識開始要求からやり直していただきます。その際、音声認識停止要求や音声認識キャンセル要求のリクエストをする必要はありません。エラーが解決しない場合、[ご契約者様お問い合わせ](#)よりお問い合わせください。

code	message	detail	対処
410	Invalid Parameter	Invalid Parameter -> List is null	表記・読みを記述してください。
410	Invalid Parameter	Cascade words exceeded 1000	一時辞書では、追加単語を1000行以下で指定してください。
410	Invalid Parameter	Surface empty	表記を記述してください。
410	Invalid Parameter	Reading empty	読みを記述してください。
410	Invalid Parameter	Invalid domainid	ASR Domain IDを確認してください。
410	Invalid Parameter	Invalid Model Name	ASR Model IDを確認してください。
410	Invalid Parameter	Prob illegal	重みは1~100の整数値で指定してください。
410	Invalid Parameter	Surface illegal	表記として設定可能な文字列か確認して下さい
410	Invalid Parameter	Reading illegal	読みとして設定可能な文字列か確認して下さい
410	Invalid Parameter	Surface too long	表記と読みの合計を251byte以下で指定してください。

お問い合わせください

以下のエラーが発生した場合は、[ご契約者様お問い合わせ](#)よりお問い合わせください。

エラーコード	メッセージ
500	Internal Error
510	Out Of Memory
551	Recognition Timeout
553	Network Timeout
601	Recognition Converter Error
610	Out Of Memory
611	Invalid License
612	Invalid Config
650	No Resource
690	External Command Execute Failed
691	External Command Fatal
692	External Command Error
693	External Command Warn

音声認識ユーザ辞書更新

※日本語モデルのみ

音声認識クライアントからユーザ辞書更新のリクエストを要求します。ユーザ辞書は音声認識に利用するモデル毎に登録が必要です。日本語の音声認識モデルに対してのみユーザ辞書を登録することが可能です。ユーザ辞書更新により、すでに登録済みのユーザ辞書の内容は上書きされますのでご注意ください。

HTTPリクエスト

```
POST <API Base URL>/asr/v1/speech_words/<ASR Model ID>/upload?domainid=<ASR Domain ID>
```

リクエストヘッダ

リクエストはマルチパートで記述します。マルチパートに用いる境界文字列を定めて、リクエストヘッダのboundaryで示してください。

キー名	説明
Content-Type	multipart/form-data; boundary=<境界文字列>
Authorization	Bearer [Access Token]

リクエストボディ

リクエストボディに後述の「ユーザ辞書リスト」を記載します。

パート種別	必須	name	Content-Type
ユーザ辞書リスト	必須	cascadeword	text/plain; charset=UTF-8

音声認識ユーザ辞書リスト

リストの各行に「単語の表記」「水平タブ (HT:0x09)」「単語の読み」を記載してください。

```
<HYOKI><HT><YOMI>
<HYOKI><HT><YOMI>
...
<HYOKI><HT><YOMI>
```

単語の重み※1を指定する場合は下記のように記載してください。単語ごとに、重みの指定を行うことができます。

```
<HYOKI><HT><YOMI><HT><PROB>
<HYOKI><HT><YOMI>
...
<HYOKI><HT><YOMI><HT><PROB>
```

項目	説明	値の範囲	必須
HYOKI	表記※2	[, / , : , ; , [,] , \t , \" を含まない、かつ空文字ではない文字列	必須
YOMI	読み※2	全角カタカナ（別表に記載）	必須
PROB	単語の重み ※1	1~100の整数値（default: 30）	任意

※1 登録単語の出現確率を調整したい場合にご利用ください。

※2 表記と読みの合計を251byte以下で指定してください。

30より大きい値を設定することで、通常の辞書登録よりも高めの効果で辞書登録が可能です。
また、30より小さい値を設定することで、通常の辞書登録よりも低めの効果で辞書登録が可能です。

ユーザ辞書リストサンプル

```
エヌ・ティ・ティ・コミュニケーションズ株式会社 エヌティティコミュニケーションズカブシキガイシ  
ヤ  
COTOHA コトハ 20
```

リクエストサンプル

以下にリクエストのサンプルを示します。

HTTPヘッダ

```
Content-Type: multipart/form-data; boundary=<境界文字列>  
Authorization: Bearer <Access Token>
```

HTTPボディ

```
--<境界文字列>  
Content-Disposition: form-data; name="cascadeword"; filename="sample.tsv"  
Content-Type: text/plain  
  
エヌ・ティ・ティ・コミュニケーションズ株式会社 エヌティティコミュニケーションズカブシキガイシ  
ヤ  
COTOHA コトハ 20  
  
--<境界文字列>--
```

cURL

```
curl -H "Authorization:Bearer <Access Token>" -X POST -F cascadeword=@<音声認識  
ユーザ辞書リストを記載したtsvファイル> <API Base URL>/asr/v1/speech_words/<ASR Model  
ID>/upload?domainid=<ASR Domain ID>
```

レスポンス

指定した「表記」、指定した「読み」、「M」（固定）、「単語の重み」が返ります。

「M」はサーバが自動的に付与するものですので、通常は気にする必要はありません。また、「単語の重み」を指定しなかった場合は、デフォルト値である30が返却されます。

レスポンスヘッダ

キー名	説明
Content-Type	multipart/form-data; boundary=<境界文字列>

レスポンスサンプル

以下にレスポンスの例を示します。

```
--<境界文字列>  
Content-Type: text/plain
```

Content-Disposition: form-data; name="status"

code : 200

message : OK

detail : success

--<境界文字列>

Content-Type: text/plain

Content-Disposition: form-data; name="cascadeword"

エヌ・ティ・ティコミュニケーションズ株式会社 エヌティティコミュニケーションズカブシキガイシャ
M 30

COTOHA コトハ M 20

--<境界文字列>--

エラーコード

リクエストの内容を確認してください

以下のエラーが発生した場合は、記載された対処をした上でリクエストをやり直してください。エラーが解決しない場合、[ご契約者様お問い合わせ](#)よりお問い合わせください。

code	message	detail	対処
410	Invalid Parameter	Invalid Parameter -> List is null	表記・読みを記述してください。
410	Invalid Parameter	Invalid Parameter -> List Exceed 5000 lines	ユーザ辞書更新では、追加単語を5000行以下で指定してください。
410	Invalid Parameter	Surface empty	表記を記述してください。
410	Invalid Parameter	Reading empty	読みを記述してください。
410	Invalid Parameter	Invalid domainid	ASR Domain IDを確認してください。
410	Invalid Parameter	Invalid Model Name	ASR Model IDを確認してください。
410	Invalid Parameter	Prob illegal	重みは1~100の整数値で指定してください。

code	message	detail	対処
410	Invalid Parameter	Surface illegal	表記として設定可能な文字列か確認して下さい
410	Invalid Parameter	Reading illegal	読みとして設定可能な文字列か確認して下さい
410	Invalid Parameter	Surface too long	表記と読みの合計を251byte以下で指定してください。

お問い合わせください

以下のエラーが発生した場合は、[ご契約者様お問い合わせ](#)よりお問い合わせください。

code	message	detail
410	Invalid Parameter	Upload Error
410	Invalid Parameter	Download Error
600	Internal Error	-

音声認識ユーザ辞書クリア

※日本語モデルのみ

音声認識クライアントからユーザ辞書のクリアのリクエストを要求します。ユーザ辞書は音声認識に利用するモデル毎にクリアが必要です。日本語の音声認識モデルに対してのみユーザ辞書をクリアすることが可能です。

HTTPリクエスト

```
GET <API Base URL>/asr/v1/speech_words/<ASR Model ID>/clear?domainid=<ASR Domain ID>
```

リクエストヘッダ

キー名	説明
Authorization	Bearer [Access Token]

リクエストサンプル

以下にリクエストのサンプルを示します。

HTTPヘッダ

Authorization: Bearer <Access Token>

cURL

```
curl -H "Authorization:Bearer <Access Token>" <API Base  
URL>/asr/v1/speech_words/<ASR Model ID>/clear?domainid=<ASR Domain ID>
```

レスポンス

レスポンスヘッダ

キー名	説明
Content-Type	multipart/form-data; boundary=<境界文字列>

レスポンスサンプル

以下にレスポンスの例を示します。

```
--<境界文字列>  
Content-Type: text/plain  
Content-Disposition: form-data; name="status"  
  
code : 200  
message : OK  
detail : success  
  
--<境界文字列>--
```

エラーコード

リクエストの内容を確認してください

以下のエラーが発生した場合は、記載された対処をした上でリクエストをやり直してください。エラーが解決しない場合、[ご契約者様お問い合わせ](#)よりお問い合わせください。

code	message	detail	対処
410	Invalid Parameter	Invalid domainid	ASR Domain IDを確認してください。
410	Invalid Parameter	Invalid Model Name	ASR Model IDを確認してください。

お問い合わせください

以下のエラーが発生した場合は、[ご契約者様お問い合わせ](#)よりお問い合わせください。

code	message	detail
410	Invalid Parameter	Upload Error

code	Invalid Parameter message	Download Error detail
600	Internal Error	-

音声認識ユーザ辞書適用状態取得

※日本語モデルのみ

音声認識クライアントからユーザ辞書適用状態取得のリクエストを要求します。各モデルごとに、ユーザ辞書更新が完了したかどうかを取得することが可能です。

HTTPリクエスト

```
GET <API Base URL>/asr/v1/speech_words/<ASR Model ID>/isset?domainid=<ASR Domain ID>
```

リクエストサンプル

以下にリクエストのサンプルを示します。

HTTPヘッダ

```
Authorization: Bearer <Access Token>
```

cURL

```
curl -H "Authorization:Bearer <Access Token>" <API Base URL>/asr/v1/speech_words/<ASR Model ID>/isset?domainid=<ASR Domain ID>
```

レスポンス

ユーザ辞書更新が完了している場合はtrue、未完了の場合はfalseが返却されます。
エラー発生の有無によって、レスポンスの Content-Type が異なりますので、ご注意ください。

レスポンスヘッダ

キー名	説明
Content-Type	application/json

レスポンスボディ

キー名	範囲
isSet	true, false

レスポンスサンプル

以下にレスポンスの例を示します。

```
{
  "isSet" : true
}
```

エラー時のレスポンス

レスポンスヘッダ

キー名	説明
Content-Type	multipart/form-data; boundary=<境界文字列>

レスポンスサンプル

```
--<境界文字列>
Content-Type: text/plain
Content-Disposition: form-data; name="status"

code : 410
message : Invalid Parameter
detail : Invalid domainid

--<境界文字列>--
```

エラーコード

リクエストの内容を確認してください

以下のエラーが発生した場合は、記載された対処をした上でリクエストをやり直してください。エラーが解決しない場合、[ご契約者様お問い合わせ](#)よりお問い合わせください。

code	message	detail	対処
410	Invalid Parameter	Invalid domainid	ASR Domain IDを確認してください。
410	Invalid Parameter	Invalid Model Name	ASR Model IDを確認してください。

お問い合わせください

以下のエラーが発生した場合は、[ご契約者様お問い合わせ](#)よりお問い合わせください。

code	message	detail
410	Invalid Parameter	Upload Error
410	Invalid Parameter	Download Error

code	message	detail
------	---------	--------

音声認識ユーザ辞書ダウンロード

※日本語モデルのみ

音声認識クライアントからユーザ辞書のダウンロードのリクエストを要求します。各モデルごとに、最後にアップロードされた辞書をダウンロードすることが可能です。

HTTPリクエスト

```
GET <API Base URL>/asr/v1/speech_words/<ASR Model ID>/download?domainid=<ASR Domain ID>
```

リクエストヘッダ

キー名	説明
Authorization	Bearer [Access Token]

リクエストサンプル

以下にリクエストのサンプルを示します。

HTTPヘッダ

```
Authorization: Bearer <Access Token>
```

cURL

```
curl -H "Authorization:Bearer <Access Token>" <API Base URL>/asr/v1/speech_words/<ASR Model ID>/download?domainid=<ASR Domain ID>
```

レスポンス

最後にアップロードされた辞書内の単語が返却されます。各行には表記、読み、単語の重みが記載されています。

レスポンスヘッダ

キー名	説明
Content-Type	multipart/form-data; boundary=<境界文字列>

レスポンスサンプル

以下にレスポンスの例を示します。

```
--<境界文字列>  
Content-Type: text/plain  
Content-Disposition: form-data; name="status"
```

```
code : 200  
message : OK  
detail : success
```

```
--<境界文字列>  
Content-Type: text/plain  
Content-Disposition: form-data; name="cascadeword"
```

```
エヌ・ティ・ティコミュニケーションズ株式会社 エヌティティコミュニケーションズカブシキガイシャ  
30  
COTOHA コトハ 30
```

```
--<境界文字列>---
```

エラーコード

リクエストの内容を確認してください

以下のエラーが発生した場合は、記載された対処をした上でリクエストをやり直してください。エラーが解決しない場合、[ご契約者様お問い合わせ](#)よりお問い合わせください。

code	message	detail	対処
410	Invalid Parameter	Invalid domainid	ASR Domain IDを確認してください。
410	Invalid Parameter	Invalid Model Name	ASR Model IDを確認してください。

お問い合わせください

以下のエラーが発生した場合は、[ご契約者様お問い合わせ](#)よりお問い合わせください。

code	message	detail
410	Invalid Parameter	Upload Error
410	Invalid Parameter	Download Error
600	Internal Error	-

モデル一覧

モデル名	ASR Model ID
日本語汎用Talk&Free(16kHz)	ja-gen_tf-16
日本語汎用Short&Formal(16kHz)	ja-gen_sf-16
日本語汎用Talk&Free(8kHz)	ja-gen_tf-08
日本語通信業界向け(8kHz)	ja-mdl1_nrw-08
日本語保険業界向け(8kHz)	ja-mdl2_nrw-08
英語汎用_ネイティブShort&Formal(16kHz)	en_en-gen_sf-16

- 「Talk&Free」と「Short&Formal」に関して
 - Talk&Free：人間同士の自然な発話、言いよどみや言い間違いが頻出する音声認識に適したモデルです。（会議、雑談、コールセンタ、顧客対応など）
 - Short&Formal：事前にある程度話す内容を考えておくことができ、比較的明瞭に発話される単発の発話クエリ音声認識に適したモデルです。（検索クエリ、一問一答型システムとの対話など）
- 「8kHz」と「16kHz」に関して
 - 8kHz：電話回線を通しての音声に対してのご利用を推奨しております。
 - 16kHz：それ以外の音声に対してのご利用を推奨しております。