

## 音声認識プログラム SR0101 仕様書

### 概要

単語音声認識プログラムである。外部から音声認識の開始の指示により、マイクを ON にして音声認識を行い、認識結果をファイルに出力するものである。

### 仕様

|           |                         |
|-----------|-------------------------|
| 名称        | 単語音声認識プログラム SR0101      |
| OS        | Windows XP professional |
| SR エンジン   | Microsoft 音声認識エンジンを利用   |
| エンジン IF   | SAPI Ver.5.1            |
| 対応言語      | 日本語(英語その他の言語は別途対応)      |
| ON/OFF 制御 | ファイルの内容を定期的に監視          |
| 結果出力      | ファイルに書き出し               |

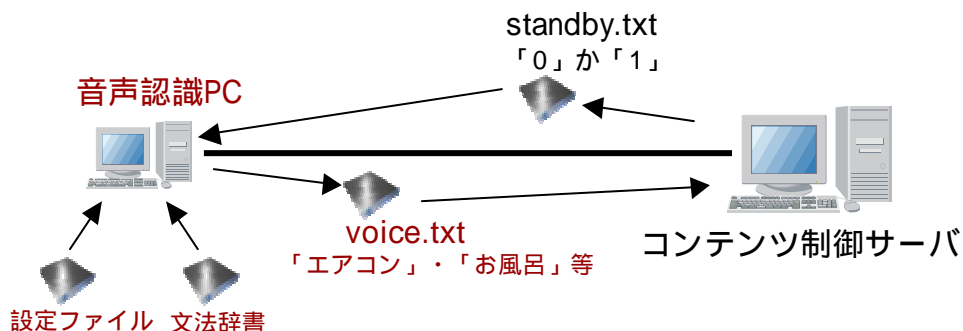
### 基本動作

- 1) 制御サーバの管理する ON/OFF 制御用ファイルを1秒以下毎に監視

「0」なら、マイクを OFF にして音声入力を受け付けない

「1」なら、マイクを ON にして音声入力を受け付ける

- 2) マイク ON 時に単語を認識したら、認識結果出力ファイルに結果を出力



## 機能

### 文法辞書ファイル

認識する単語を指定するファイル。認識単語の変更は、このファイルの内容を変更することで行う。

また、「認識結果出力ファイル」に出力する文字もここで指定する。

認識する言葉と、出力する文字は任意に対応させることが可能であり、

例えば、『スタート』という発声に対して、「101」などの数字を書き出せる。

なお、認識単語を追加・変更するためのユーザーインターフェースは無し。

文法辞書は UNICODE で記述し、記述方法に若干のクセがあるため、変更の際はお問い合わせください。

### 任意のファイルを設定可能

「ON/OFF 制御用ファイル」、「認識結果出力ファイル」は、任意のフォルダ、任意のファイル名に変更可能。

また、「文法辞書ファイル」も、任意のフォルダ、任意のファイル名に変更可能。

これらの指定は、SR0101.INI ファイルに指定する(設定方法は『INI ファイル』参照)。

初期設定では、それぞれ、カレントディレクトリの「stand-by.txt」、「speech.txt」、「gram.j.xml」となっている。

コンテンツ制御サーバ上のファイルを指定する場合は、あらかじめネットワークドライブの設定をしておき、それ上のファイルを指定すればよい。

### 「認識結果出力ファイル」のクリア

「ON/OFF 制御用ファイル」が「0」になったのを検出したときに、「認識結果出力ファイル」を内容の無いカラのファイルにすることが可能。ただし、「ON/OFF 制御用ファイル」を「0」にしてから、最大1秒かかることに注意する必要がある。SR0101.INI ファイルで「する/しない」を設定。初期設定では「OFF」になっている。

## 起動方法

SR0101.exe を実行すると、モニタ画面が表示され、音声認識処理が開始される。

Windows の立ち上げと共に起動する場合には、SR0101.exe のショートカットをスタートメニューに置く。

## 終了方法

モニタ画面の右上の終了ボタンをクリックすることで終了する。

本プログラムを終了せずに Windows を終了しても問題ない。

## 注意事項

コントロールパネルのサウンド、音声認識の各設定項目は本プログラムの動作に直接かわり、すでに調整済みであり、変更できない。

## モニタ画面 (SR0101.exe)

起動すると、次のようなモニタ画面が常に表示される。この画面は調整等における確認用のものである。

通常動作時に操作することはない。

The screenshot shows the 'Speech Recognition SR0101' application window. The window title is 'Speech Recognition SR0101'. The menu bar includes 'File', 'C&C', 'Dictation', and 'Help!'. The main area contains several sections:

- Options:** A list of checkboxes including 'Create Recognition Context', 'Retain Reco Audio', 'Activate Microphone', 'Load Dictation', 'Activate Dictation', 'Load C&C', 'Activate C&C', 'Load Spelling', and 'Activate Spelling'. The 'Activate Microphone' checkbox is circled in red.
- Grammars:** A dropdown menu showing 'Microsoft Japanese Recognizer v5.1'.
- Recognition Log:** A list of recognition events with timestamps and text. The entries are: [11:00:04.361] -START, [11:00:05.321] -マイコンテンツ [Hypothesis], [11:00:05.321] -マイコンテンツ, [11:00:08.671] -ナビゲーション [Hypothesis], and [11:00:08.659] -ナビゲーション.
- Confidence:** A section showing 'RULE="LIVING" (00000)', 'ナビゲーション', and 'Overall confidence = 1.'.
- Properties:** A section showing 'PROPERTIES: [ 0, 0 ] "VALUE" = "ナビゲーション" (1 = 3)'.
- Current C&C Grammar:** A text field showing 'grammar (Japanese)'.
- Buttons:** 'Alternates' and 'Submit' buttons.

Red annotations with arrows point to various elements:

- 'マイク OFF 時はグレイ化' points to the 'Activate Microphone' checkbox.
- '使用中の SR エンジン' points to the 'Microsoft Japanese Recognizer v5.1' dropdown.
- '認識した時刻、認識候補単語' points to the first two entries in the recognition log.
- '認識結果' points to the entire recognition log section.
- '確信度' points to the 'Overall confidence = 1.' text.
- '1: 正解' points to the 'Overall confidence = 1.' text.
- '0: 正解かも' points to the 'Overall confidence = 1.' text.
- '-1: おそらく異なる' points to the 'Overall confidence = 1.' text.
- '確信度が1のときだけ結果出力する' points to the 'Overall confidence = 1.' text.
- '認識言語' points to the 'grammar (Japanese)' text field.

## 文法ファイル (gram\_j.XML)

このファイルで認識する単語を指定する。

文法ファイルは、必ず **UNICODE** で編集・保存する必要があり、サフィックスは"XML"にする。

認識単語を変更する場合は、下図の四角で囲まれた部分だけを修正する。

各行は、認識する各単語であり、以下のような書式で記述する(順不同)。

```
<P VAL="nn">/ / ;</P>
```

- nn 任意の数値。ただし他の単語とは重複して指定できない  
(認識結果出力ファイルに)出力する文字列。任意の文字列が指定できる。  
認識する単語のよみをひらがなで指定する。

```
<GRAMMAR LANGID="411">
  <DEFINE>
    <ID NAME="VALUE" VAL="1" />
  </DEFINE>
  <RULE NAME="LIVING" TOPLEVEL="ACTIVE">
    <L PROGRAM="VALUE" PROPI="VALUE">
      <P VAL="1">/START/スタート;</P>
      <P VAL="2">/マイコンテック/まいこんてんく;</P>
      <P VAL="3">/ナビゲーション/なびげーしょん;</P>
      <P VAL="4">/バイタルセンサ/ばいたるせんさー;</P>
      <P VAL="5">/インフォメーション/いんふあめーしょん;</P>
      <P VAL="6">/メニュー/めいぬ;</P>
      <P VAL="38">/メニュー/めいぬ;</P>
      <P VAL="7">/戻る/もどる;</P>
      <P VAL="8">/ランチ/らんち;</P>
      <P VAL="9">/レストラン/れすとらん;</P>
      <P VAL="10">/イタリアン/いたりあん;</P>
      <P VAL="11">/オリーブの木/おりーぶのき;</P>
      <P VAL="12">/マクドナルド/まくどなると;</P>
      <P VAL="13">/らんぷ亭/らんぷてい;</P>
      <P VAL="14">/決定/けつてい;</P>
      <P VAL="24">/決定/けつていします;</P>
      <P VAL="34">/決定/けつていです;</P>
      <P VAL="44">/決定/けつて;</P>
      <P VAL="15">/帰ります/かえります;</P>
      <P VAL="16">/冷蔵庫/れいぞうこ;</P>
      <P VAL="17">/電子レンジ/でんしれんじ;</P>
      <P VAL="18">/ラジカセ/らじかせ;</P>
      <P VAL="19">/辞まります/おわかります;</P>
      <P VAL="20">/暑い/あつい;</P>
      <P VAL="21">/寒い/さむい;</P>
      <P VAL="22">/それ/それ;</P>
      <P VAL="23">/あれ/あれ;</P>
      <P VAL="24">/きみがよ/きみがよ;</P>
      <P VAL="25">/きよしこのよる/きよしこのよる;</P>
    </L>
  </RULE>
</GRAMMAR>
```

出力する文字列は、英数字でも構わない

同じ出力文字列に、複数のよみを与えることもできる

## INI ファイル (SR0101.INI)

使用するファイルのディレクトリ、ファイル名およびオプション設定を指定する。

[使用するファイル]

|           |               |
|-----------|---------------|
| VOICE     | 認識結果出力ファイル    |
| STANDBY   | ON/OFF 制御ファイル |
| GRAMMAR01 | 文法ファイル        |

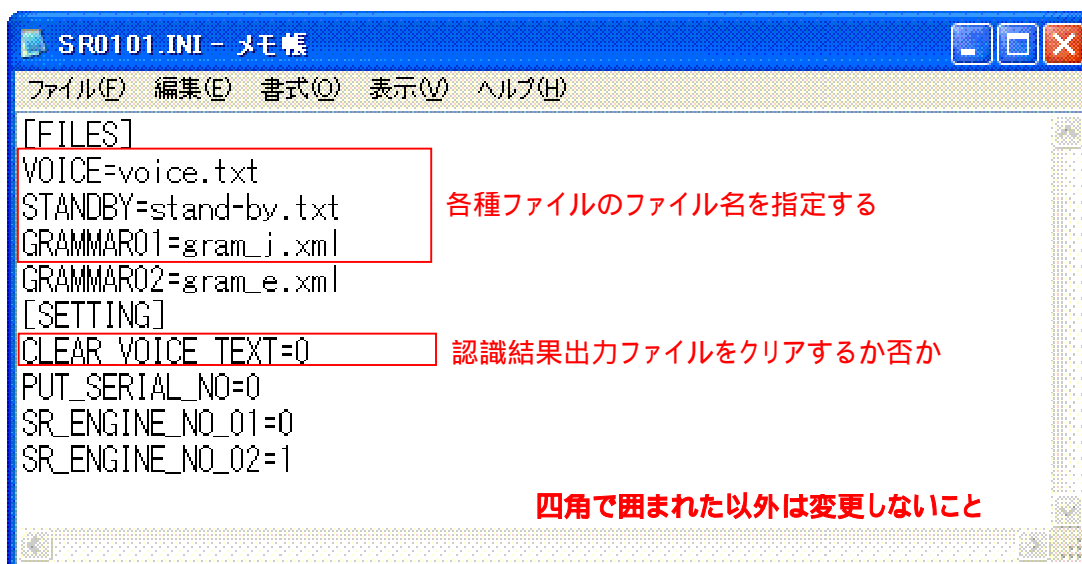
次のように、絶対パスで指定することもできる。なお、ディレクトリを指定しないときはカレントディレクトリとなる。

絶対パス指定の例: STANDBY=Z:\speech\standby.txt

[オプション設定]

CLEAR\_VOICE\_TXT

「ON/OFF 制御用ファイル」が「0」になったのを検出したときに、「認識結果出力ファイル」の内容をクリアする機能のフラグ。「1」を指定すると、クリアするようになる。

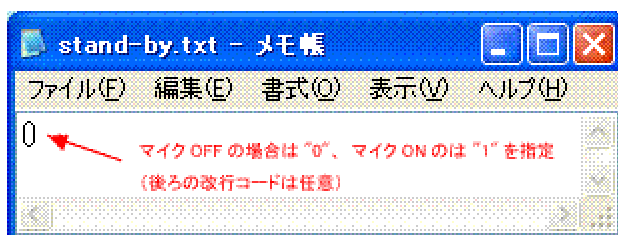


## ON/OFF 制御ファイル (stand-by.txt)

音声認識の ON/OFF を制御するためのファイル。

コンテンツサーバ側で書き出し、本プログラムでは定期的に読み込みを行う。

“0”(半角)の指定で OFF(認識しない)、“1”(半角)の指定で ON(認識する)



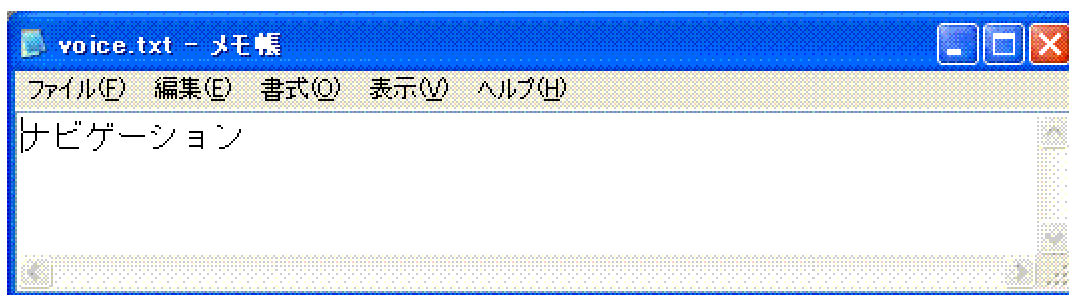
#### 認識結果出力ファイル (voice.txt)

本プログラムが音声を認識すると、このファイルに認識結果を1行出力する。

出力文字列は、文法ファイルの"出力する文字列"で指定したもの。

文字コードは SJIS である。なお、文字列の後に改行コード(0x0d,0x0a)が含まれる。

本プログラムが不定のタイミングで書き出すため、コンテンツサーバ側でこのファイルを排他的にオープンしてはならない。



以上