

Julius で遊ぶ

大語彙連続音声認識エンジン Julius で遊んでみた。
環境は、特に音声環境の設定とかしていない FreeBSD。

インストール

<http://julius.sourceforge.jp/index.php?q=newjulius.html>

から最新版をダウンロード。

そもそもマイクから入力を得るつもりがないので、

```
$ ./configure --with-mictype=oss
$ make
$ make install
```

と適当にコンパイル

で、辞書等々のファイルは、ディクテーションキットよりダウンロードして展開。

<http://julius.sourceforge.jp/index.php?q=juliuskit.html>

遊んでみる

Microsoft PCM, 16 bit, mono 16000 Hz な音声ファイル test.wav を作成して

```
$ julius -C dictation-kit-v3.0/light.jconf -input rawfile
```

とかして起動。

```
enter filename->
```

とプロンプトがでるので、用意した test.wav と入力。

で、めでたく解析結果が見られる。

MP3 から WAV 作成

```
mpg123 --mono --rate 16000 --wav hoge.wav hoge.mp3
```

とか

Microsoft Speech SDK 5.1

で遊んでみようとして、サンプルをコンパイルしようとしたら、

sphelper.h 内でのコンパイルエラー。

無理矢理キャストしてみたりして、とりあえずコンパイルできたけど ...

よくわからないけど、Microsoft .NET Speech SDK を使えってこと？

```

769c769
<     const uLenVendorPreferred = wcslen(pszVendorPreferred);
---
>     const size_t uLenVendorPreferred = wcslen(pszVendorPreferred);
1418c1418
<     static CoMemCopyWFEX(const WAVEFORMATEX * pSrc, WAVEFORMATEX ** ppCoMemWFEX)
---
>     static HRESULT CoMemCopyWFEX(const WAVEFORMATEX * pSrc, WAVEFORMATEX ** ppCoMemWFEX)
2372c2372,2373
<         for (const WCHAR * psz = (const WCHAR *)IPParam; *psz; psz++) {}
---
>         const WCHAR * psz = (const WCHAR *)IPParam;
>         for (; *psz; psz++) {}
2559c2560
<     SPPHONEID* pphoneId = dsPhoneId;
---
>     SPPHONEID* pphoneId = (SPPHONEID*)((void*)dsPhoneId);
2633c2634
<         pphoneId += wcslen(ophoneId) + 1;
---
>         pphoneId += wcslen((const wchar_t*)((void*)ophoneId)) + 1;

```