

Native-to-Foreign (N2F) Pronunciation Converter (2nd prototype) User Manual

Overview.....	1
System requirements.....	2
Convert speech pronunciation by the software	3
Contact.....	5

Overview

The software converts native speech to foreign accented speech. We have confirmed that it nicely converts native English speech to Japanese-accented English speech, and expect it to work for other language pairs.

It supports two types of input (native pronunciation speech):

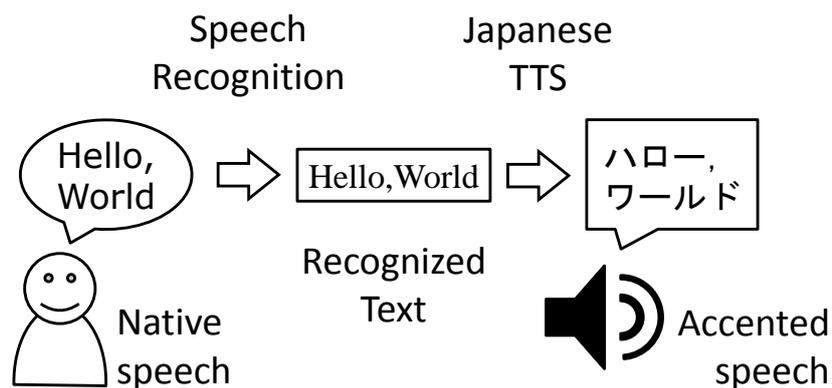
- Real-time speech (requires a microphone)
- Recorded speech (wav file)

And it supports two types of output (foreign accented speech):

- Real-time speech (requires a speaker)
- Save as wav file

How it works

As shown in the figure below, the software converts speech accent by reading aloud the speech recognized text by using foreign text-to-speech software (TTS). Conversion in the opposite way (removing foreign accent) doesn't work well because speech recognition technology does not accurately recognize accented speech.



System requirements

The software runs on a Windows PC with .NET Framework (ver. > 3?) installed. Language packs of the source and target language (pronunciation) are also required. Follow the instructions below if you don't have them yet. You also need a microphone/speaker for input/output. It is recommended that you use the best microphone you have and speak as clearly as you can.

Install or update the .NET framework runtime

The software runs in .NET framework 4.5.1 which is pre-installed in Windows 8.1. If you still use Windows 8 it is recommended to update to 8.1 from the Windows Store. If you use Windows XP/Vista/7 you can install .NET framework 4.5.1 from the following link, or upgrade to Windows 8.1.

Microsoft .NET Framework 4.5.1

<http://www.microsoft.com/ja-jp/download/details.aspx?id=40773>

Install the language packs for the source and target language

Installing the language pack to the PC will enable both speech recognition and text-to-speech for the language which are required for the N2F conversion software. Language packs can be installed from the control panel.

Navigate to “Add language” as follows, and add the language pack you need.

(In category view) Control Panel > Add language

(In classical view) Control Panel > Language

It is fun to test the N2F conversion software with different languages, but please be aware that the file size of the language packs can be quite large.

Convert speech pronunciation by the software

With .NET framework installed, you can execute PronunciationConverter2.exe.



Convert real-time speech from the microphone

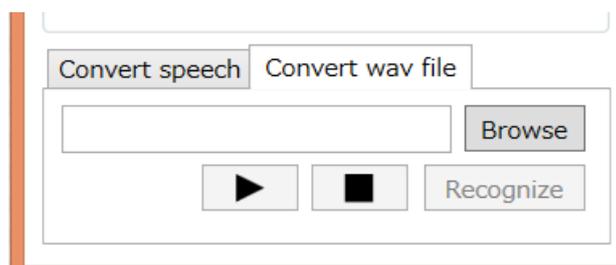
Press the microphone button (on the lower left) and speak to the PC microphone after the prompting sound. Converted speech will be played after you stop speaking. Press the button again to convert the next phrase. Recognized phrases will be displayed on the right part of the window. Try again if the system does not correctly recognize what you said.

Replay original / converted speech

Clicking the recognized phrase in the recognized phrases list will replay the converted speech. The two buttons on the lower right will replay all the phrases in the list.

Convert recorded speech

1. Switch to "Convert wav file" tab (located above the microphone button).
2. Press "Browse" button and select the .wav file you want to convert.
3. Play, stop or recognize as shown in the figure below.



Customize conversion parameters

You can customize the conversion parameters to explore variations of output speech, at the upper left part of the window. The parameters are briefly explained in the below table, but it is easier to just play with the app to find out how they work.

Table 1. Conversion Parameters

Input voice	Should match the user's original speech.
Output voice	Choose the foreign accent you want.
Speak speed	Speed of the output speech (right: fast, left slow)
Japanizer	(experimental feature; see below)
Word by word	Inserts noticeable pause between words.
Scenario	(explained in a later section)

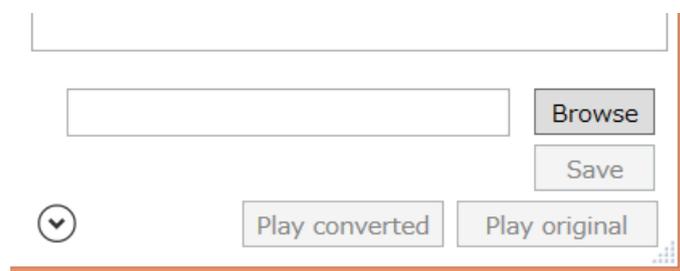
Japanizer is an experimental feature, which tries to generate Japanese-accented speech in a different way. For example, if you choose English(US) output voice with Japanizer checked, the output is kind of like “An US people trying to imitate Japanese-accented speech”.

Customized parameters can be saved to or loaded from file.

- To save settings, press “Save settings” button on the top-left part of the window. Settings will be saved to an xml file in the “settings” folder.
- To load settings, select a setting from the combo box on the left of the “Save settings” button. Saved settings will automatically appear in this list.

Save original / converted speech

Press  to access the Save / Browse button as shown in the figure below. Press “Browse” button to specify the folder where the files will be saved, and press “Save” button.



Provide a scenario (optional)

Speech recognition technology today is far from perfect and the N2F software based on it is no exception. Recognition accuracy can be significantly improved if you know what the speaker will say and provide the knowledge as a scenario to the software.

Scenarios should be written in Speech Recognition Grammar Specification (SRGS) Version 1.0, and placed in the “scenarios” folder. Please see the example scenarios in the “scenarios” folder and refer to the following documents for more details.

<http://www.w3.org/TR/speech-grammar/>

<https://msdn.microsoft.com/library/windows/apps/xaml/dn630428.aspx>

Contact

Takeshi Nishida

Associate professor at Graduate School of Intercultural Studies, Kobe University

<http://www2.kobe-u.ac.jp/~tnishida/index.html>

tnishida [at] people.kobe-u.ac.jp