

## Introduction

The Speech Synthesis Markup Language Service enables developers to customise the pronunciation of the speech output for their various use cases.

This document explains the general information regarding the components of the SSML and a sample code on the implementation.

## Speech Synthesis Markup Language (SSML)

Speech Synthesis Mark-Up Language (SSML) aids speech synthesizers by adding mark-up on input text. This enables the speech synthesizers to construct speech waveforms that sounds more natural.

It consists of two parts:

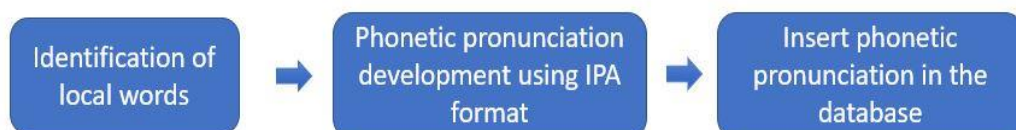
### - 1) Phonetic Pronunciation development

- o International Phonetic Alphabet (IPA)

The IPA is designed to represent only those qualities of speech that are part of oral language: phone, phonemes, intonation and the separation of words and syllables where each symbol is associated with a particular English sound.

Using the National Speech Corpus (NSC) Lexicon provided, developers can create custom pronunciation terms. The data provided were accustomed to Singapore English, any custom pronunciation outside of the current scope would require a linguist.

word	IPA
gongfu	gɔŋfu
Gu	gu
Gu Zao ...	gu tsa...
hakka	haka
ice	ais
istana	istana
jiang	dʒiang
kacang	katʃaŋ
kakak	kakakʔ
kaki	kaki
kambing	kamberŋ
kampong	kamponŋ
kang	kaŋ
kopi	kowp̩i
kueh	kue
kut	kut
laksa	lɔk'sa
lapis	lapis
Lemak	lemakʔ
long	loŋ
lontong	lontɔŋ
lor	lɔ:



- o Creation of a database with custom IPA to aid the SSML in a more accurate pronunciation of the local words.

## - 2) SSML Micro-service



- The SSML Micro-service replaces local words in the database with its corresponding IPA to produce a SSML text which would be passed into a supported TTS Service.
- With the aid of the NSC Lexicons, pronunciations of custom Singapore accented words such as “*Siew Dai*”, “*Kopi O*”, etc are achievable and are closer to the actual pronunciation of Singaporeans.

Below is a sample implementation of the SSML based on the following sentence:

- *I've travelled from Tanjong Pagar to Bedok just to have my favourite Ayam Rendang*

```
- <spea version='1.0' xmlns="https://www.w3.org/2001/10/synthesis" xml:lang='en-US'> I've travelled from
- <phoneme alphabet="ipa" ph="tandʒoŋ"> Tanjong</phoneme>
- <phoneme alphabet="ipa" ph="paɡa"> Pagar</phoneme> to
- <phoneme alphabet="ipa" ph="bədɔʔ"> Bedok</phoneme>
- just to have my favourite Ayam
- <phoneme alphabet="ipa" ph="rəndaŋ"> Rendang</phoneme>
- </spea>
```

Fig 1. Sample of a SSML code