

# 多語言發音特徵

## MULTILINGUAL ARTICULATORY FEATURES

*Authors: Sebastian Stüker, Tanja Schultz,  
Florian Metze, and Alex Waibel*

*Professor: 陳嘉平*

*Reporter: 吳柏鋒*

# 摘要

- 簡介
- 多語言發音特徵
- 實驗

# 簡介

- 我們可以透過發音特徵(AF)來作跨語言的語音辨識
- 找出各個語言中最佳的特徵偵測器來改善分類正確性

# 簡介

- 在最後實驗部分，分別針對**monolingual**、**crosslingual**、**multilingual**三者，在發音特徵來做比較

# 多語言發音特徵

- 我們找出多語言中共同的發音**phonetic**，並將這些共同**phonetic**，集合成為一個共同的單元集合
- 並以國際語音學學會的國際音標(**IPA**)為基礎，則不同語言在**IPA**中有共享的符號

# 多語言發音特徵

- 在**IPA**中，我們將聲音分成子音與母音兩種：
  - (1) 子音:主要依發音位置和發音方式來分
  - (2) 母音:主要依舌頭的垂直與水平位置來分

# 多語言發音特徵

Feature		Languages
CONSONANT	VOICED	CH GE EN JA SP
	UNVOICED	CH GE EN JA SP
<i>Manner</i>	ASPIRATED	CH EN
	PLOSIVE	CH GE EN JA SP
	NASAL	CH GE EN JA SP
	TRILL	GE SP
	FLAP	EN SP
	FRICATIVE	CH GE EN JA SP
	AFFRICATE	CH GE EN JA SP
	APPROXIMANT	CH GE EN JA SP
	LATERAL-APPROXIMANT	CH GE EN JA SP
<i>Place</i>	BILABIAL	CH GE EN JA SP
	LABIODENTAL	CH GE EN JA SP
	DENTAL	EN SP
	ALVEOLAR	CH GE EN JA SP
	POSTALVEOLAR	GE EN JA SP
	RETROFLEX	CH EN
	PALATAL	CH GE EN JA SP
	VELAR	CH GE EN JA SP
	UVULAR	JA
	GLOTTAL	GE EN JA
VOWEL		CH GE EN JA SP
<i>Vertical</i>	ROUND	CH GE EN JA SP
	UNROUND	CH GE EN JA SP
	TONAL1-5	CH
	CLOSE	CH GE EN JA SP
	CLOSE-MID	GE EN JA SP
<i>Horizontal</i>	OPEN	CH GE EN JA SP
	OPEN-MID	CH GE EN
	FRONT	CH GE EN JA SP
	CENTRAL	GE EN
	BACK	CH GE EN JA SP

# 多語言發音特徵

- 我們可以依據特徵對應的關係，計算出特徵共享因子

$$sf_{\Lambda} = \frac{\sum_{i \in \Lambda} |\Phi_{L_i}|}{|\Phi|}, |\Phi| = |\Phi_{LI}| + \sum_{i \in \Lambda} |\Phi_{LD_{L_i}}|$$

其中  $sf_{\Lambda}$  : 語言  $\Lambda$  的特徵共享因子

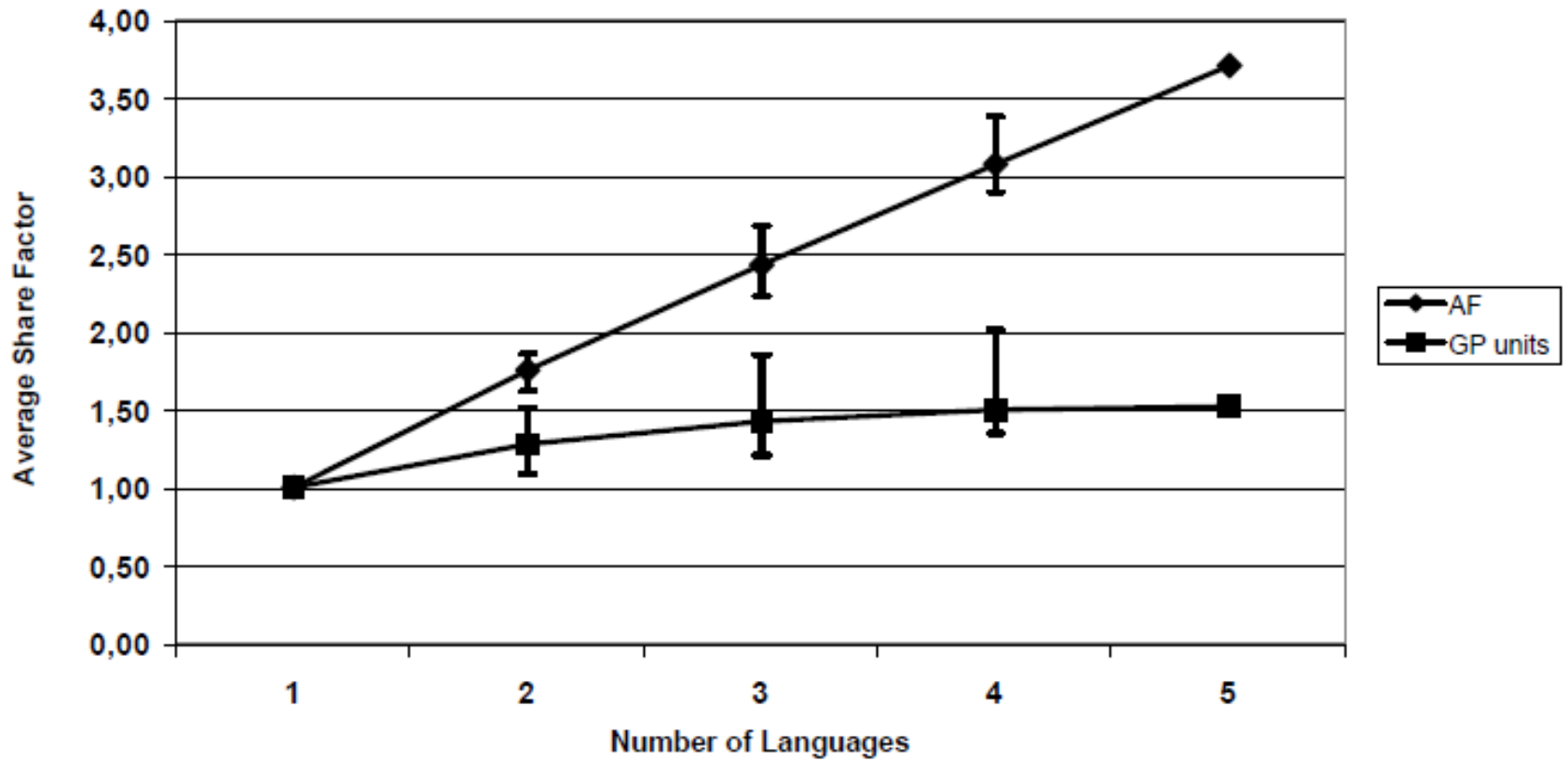
$\Phi_{L_i}$  : 語言相依特徵集合

$\Phi_{LI}$  : 超過一個語系的獨立發音特徵

$\Phi_{LD_{L_i}}$  : 語言  $L_i$  的特徵集合



# 多語言發音特徵



# 實驗

- 使用**GlobalPhone** 語料庫

錄音環境	在密閉空間使用麥克風錄音
取樣頻率	16kHz
取樣格式	16-bits PCM
音標	International Phonetic Alphabet (IPA)
語系	阿拉伯語、中文、上海話、克羅地亞、捷克語、法語、德語、 日語、 韓語、葡萄牙語、俄語、西班牙語、瑞典語、 泰米爾語、土耳其語， 共 15 種
語料量	每個語系由 100 個當地語者，讀 100 個報紙上政治與經濟相關句子 語音檔共 300 小時

# 實驗

- **Monolingual AF**

- 針對個別每個語言中，每個特徵都分別訓練兩個模型，一個代表特徵存在，另一個代表特徵不存在

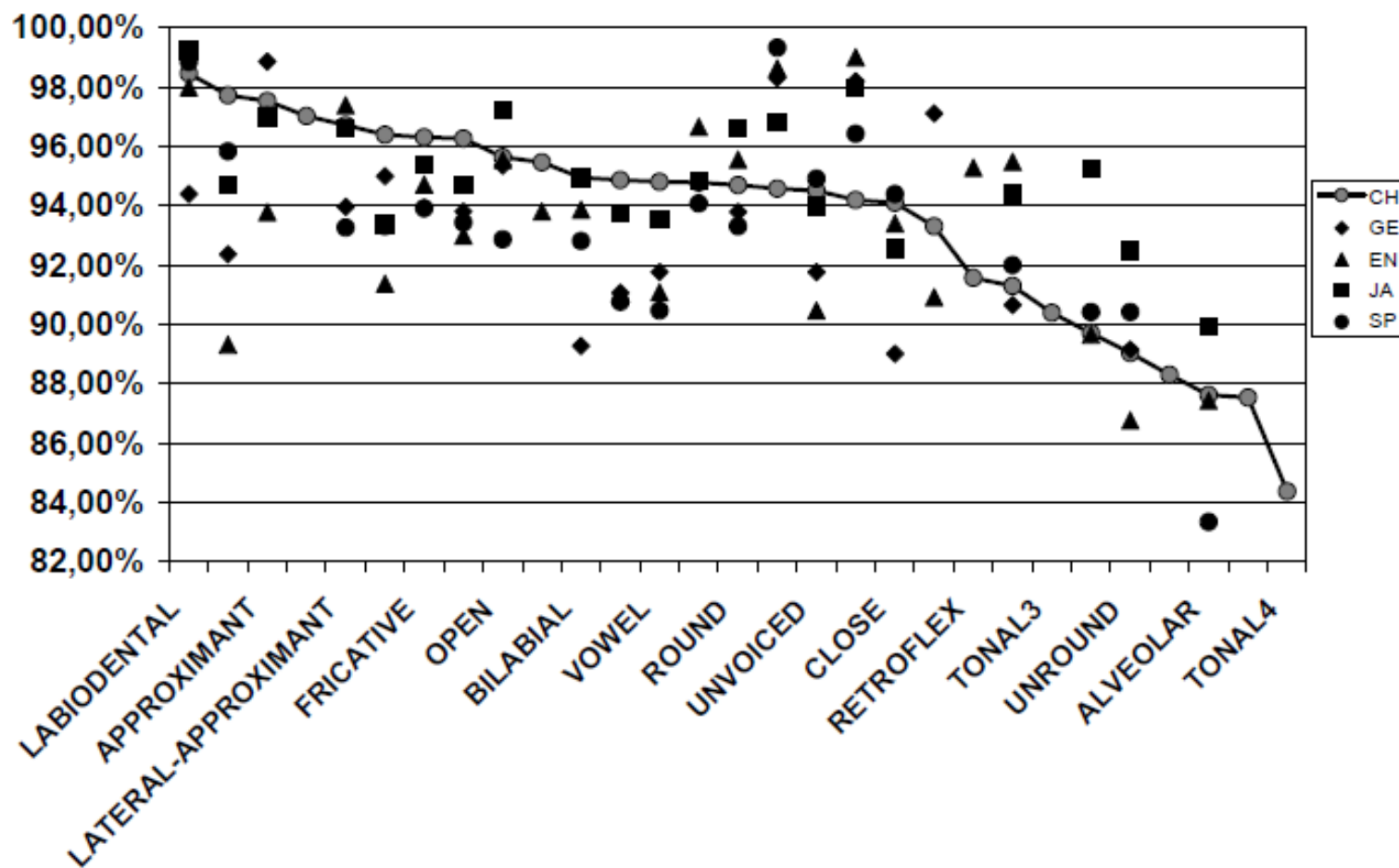
# 實驗

- **Crosslingual AF**

-實驗說明跨語言發音特徵偵測效果較單語言發音特徵偵測效果差

AF LID	Test Set				
	CH	EN	GE	JA	SP
CH	<del>93.52%</del>	87.42%	88.23%	86.45%	83.22%
EN	87.74%	<del>93.83%</del>	89.17%	88.41%	87.90%
GE	88.57%	87.90%	<del>92.94%</del>	86.46%	82.68%
JA	87.11%	87.65%	86.77%	<del>95.22%</del>	87.39%
SP	84.76%	86.36%	83.31%	87.76%	<del>93.46%</del>

# 實驗



# 實驗

- 實驗說明未必中文訓練出來的發音特徵偵測器未必在中文辨識中就是最好的
- 若是能夠挑選出每個語言最佳的中文特徵偵測器，則必能改善分類正確性

	Test Set				
AF	CH	EN	GE	JA	SP
native	93.52%	93.83%	92.94%	95.22%	93.46%
selected	95.04%	96.13%	96.12%	96.26%	96.36%

# 實驗

- Multilingual AF
  - 我們使用multilingual mixed(MM)技術，  
此技術是以所有語言具有的共同單元來訓練聲學模型且每個語言都提供相同數量的句子來作訓練

# 實驗

