



Department of  
Computer Science  
Roinn na Ríomhoilíochta

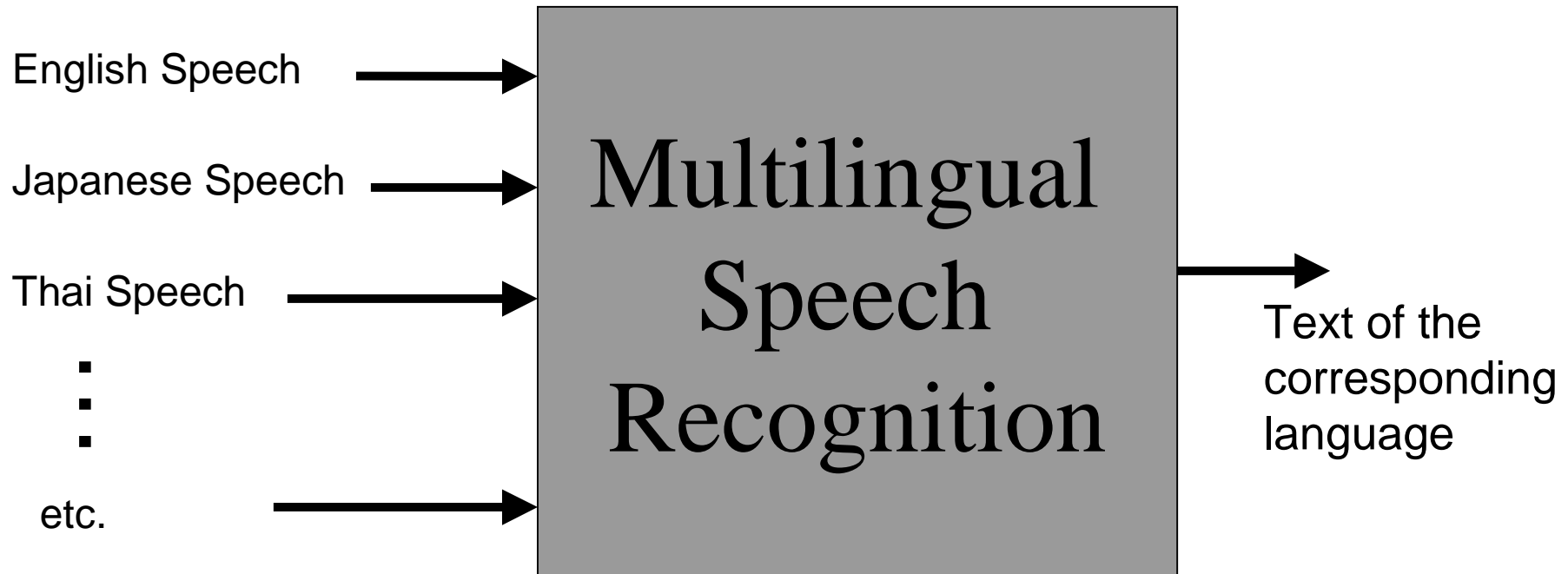


# Articulatory-Acoustic-Feature-based Automatic Language Identification

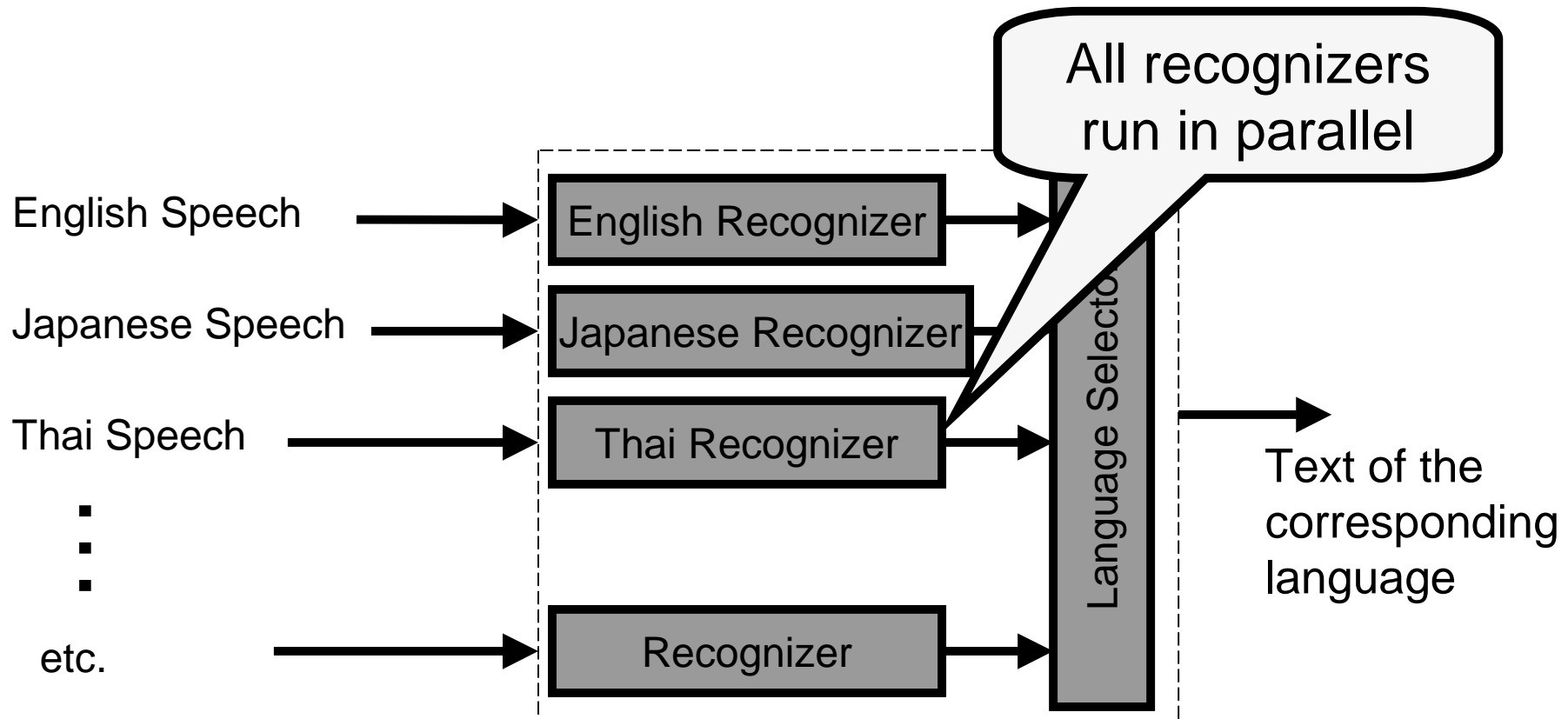
Supphanat Kanokphara  
and  
Julie Carson-Berndsen

# [ Introduction ]

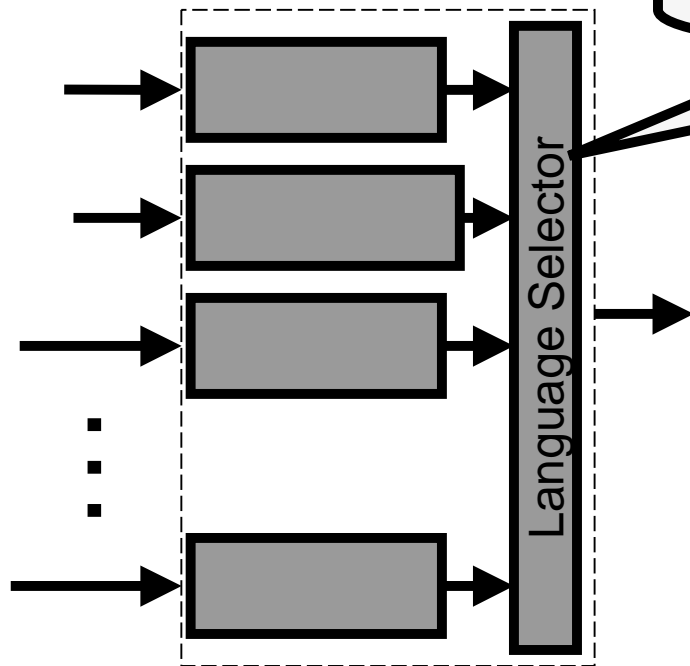
---



# [ Full Recognition ]



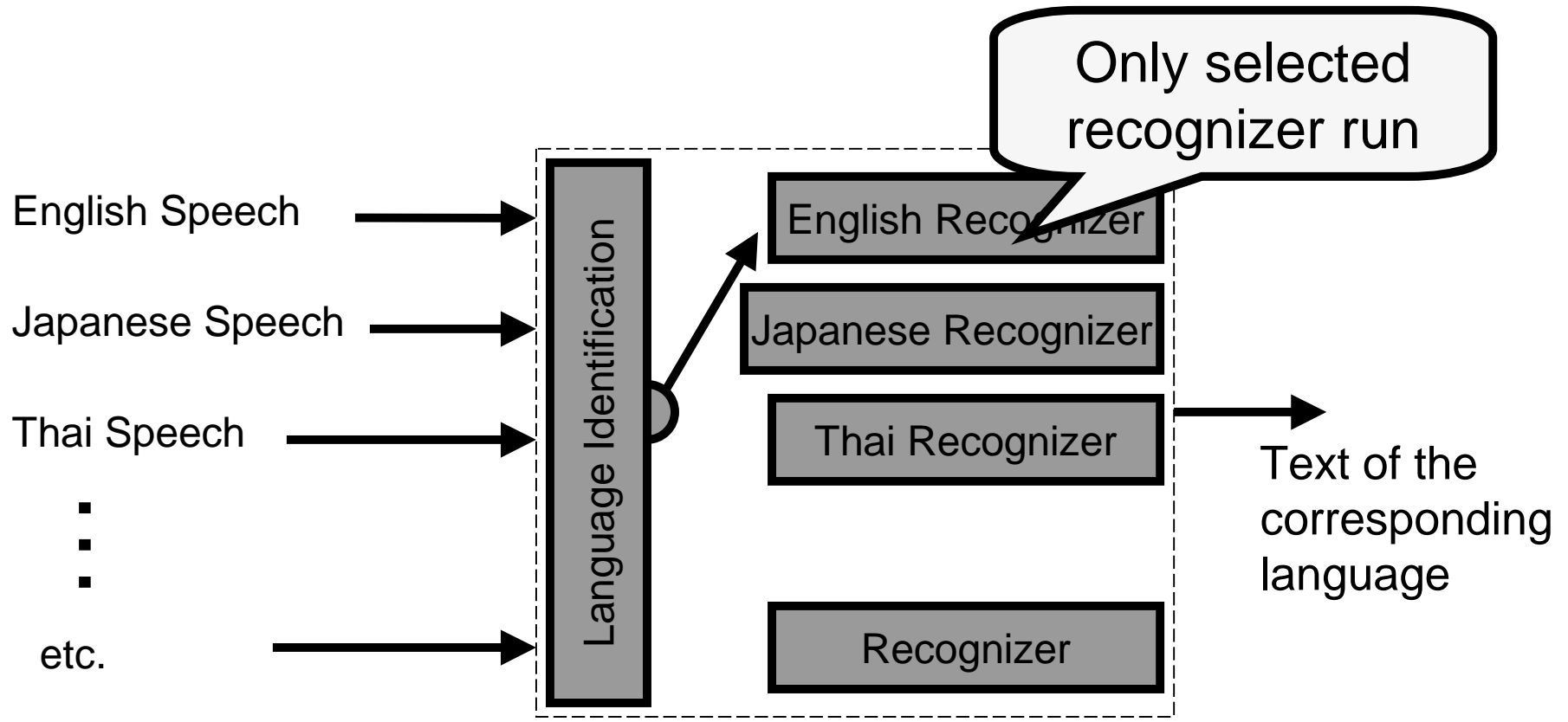
# [ Full Recognition ]



Text is recognized together with the language identification

- More accurate
- but
- Slower

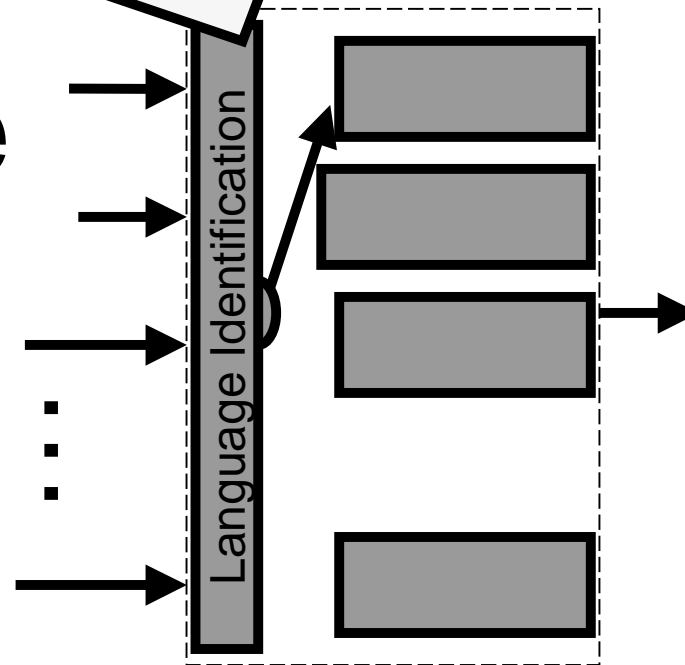
# Language Identification



# Language Identification

If LID selects the wrong recognizer,  
result will never be correct

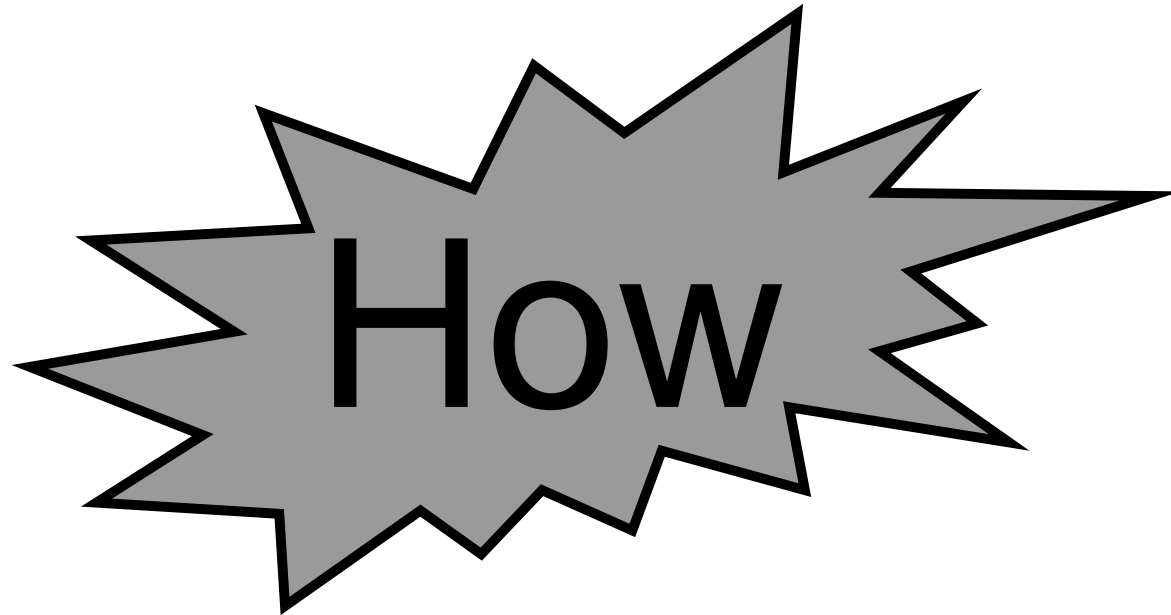
- Less accurate
- but
- Faster



# [ Fast and Accurate System ]

---

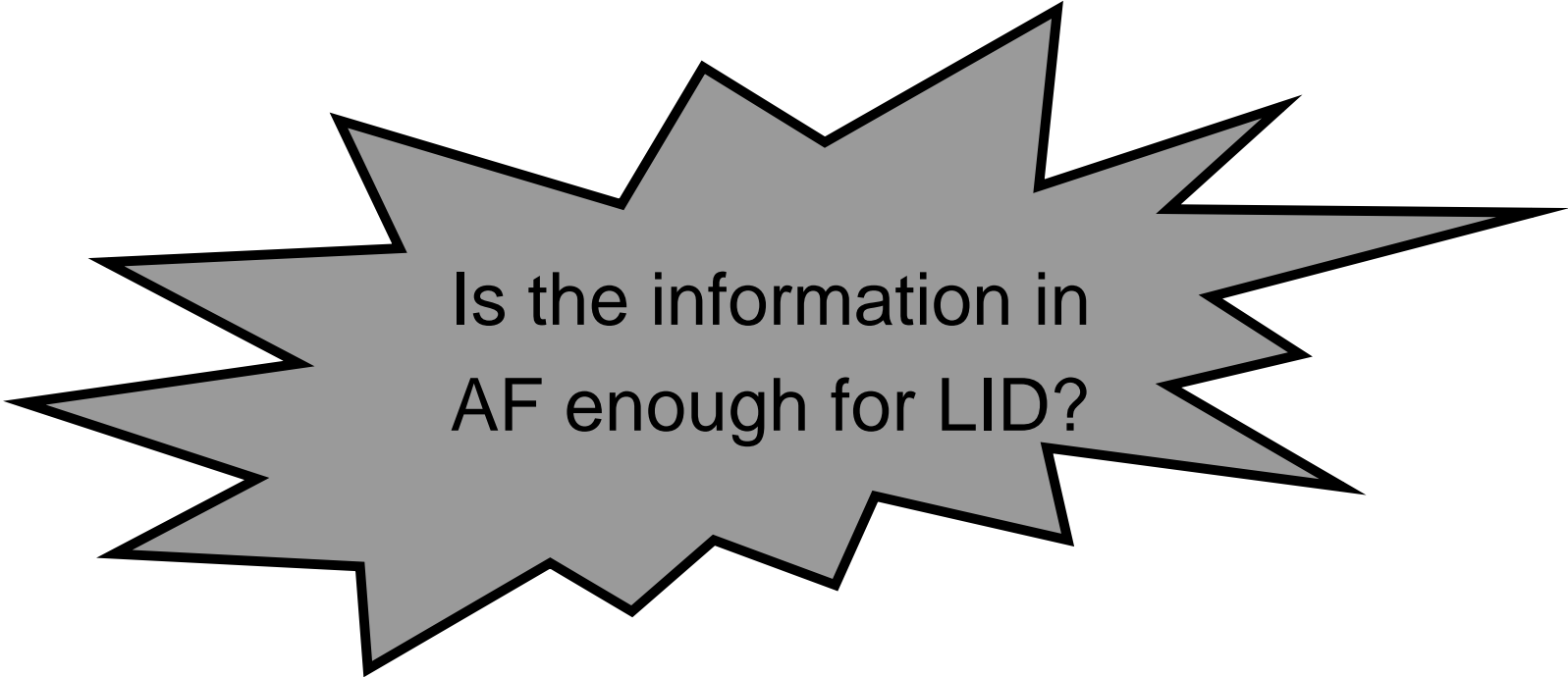
- Balance amount of information for LID



# [ Articulatory Feature Models ]

---

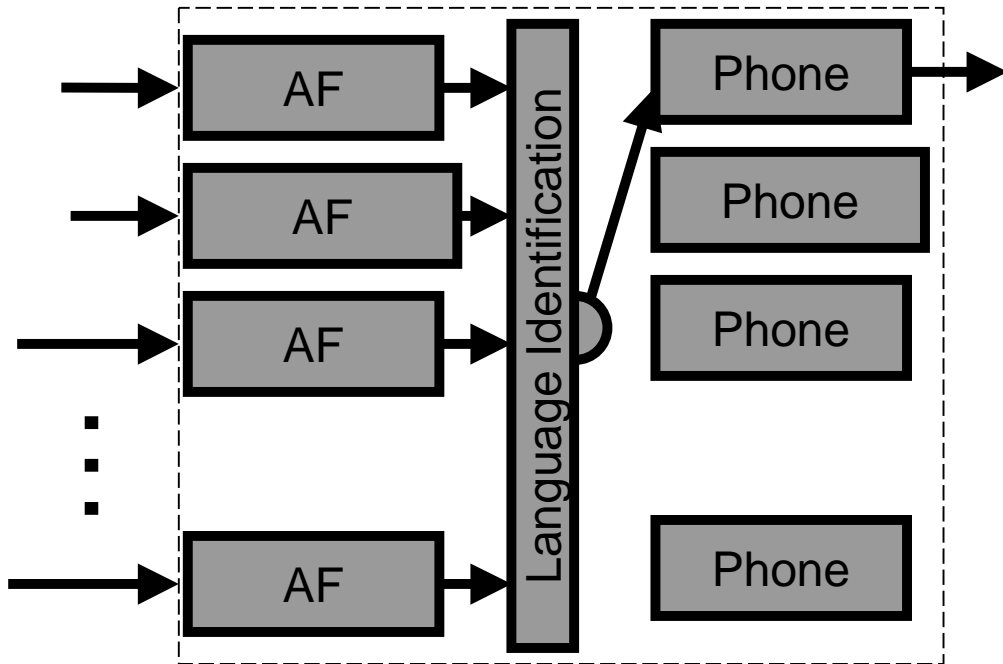
- Number of AF is less than phones'
- Faster than phone models



Is the information in  
AF enough for LID?

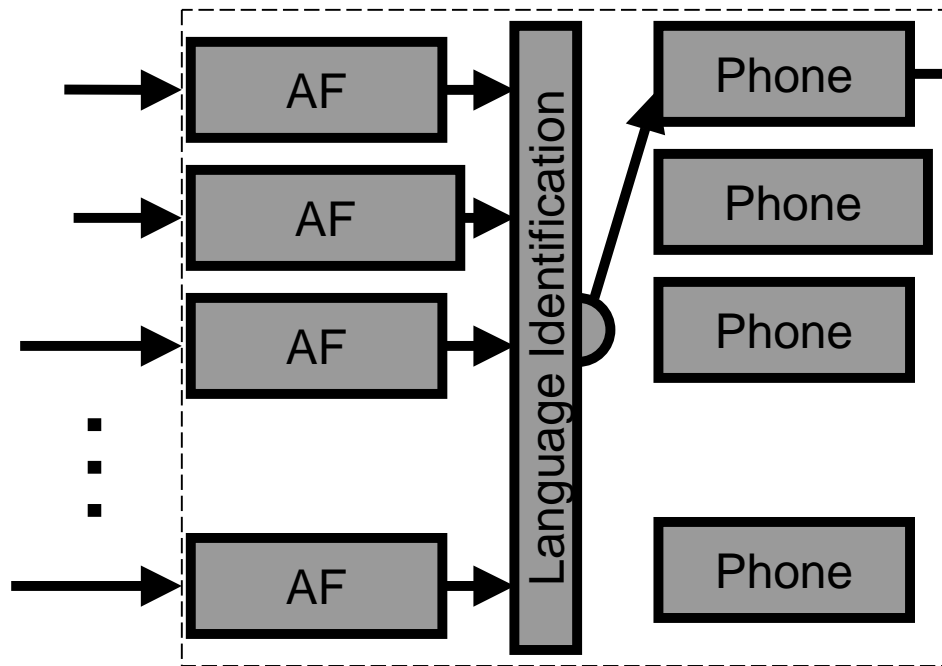


# System Overview



- $AF = AF_{AM} + AF_{LM}$
- $Phone = phone_{AM} + AF_{LM}$

# System Overview (2)



- MFCC, E,  $\Delta$ ,  $\Delta\Delta$
- CDHMM
- Bi-gram LM

# [ Experiment ]

tier	English (%)	Thai (%)	# states
phone	100	100	1,849+5,718
voice	100	97.3	84+108
manner	100	98.2	855+768
place	99.5	92.9	1,959+489
height	100	99.1	333+111
type	100	99.1	258+132

# [ Experiment ]

tier	English (%)	Thai (%)	# states
phone	100	100	1,849+5,718
voice	100	97.3	84+108
margin	100	98.2	855+768
place	99.5	92.9	1,959+489
height	100	99.1	333+111
type	100	99.1	258+132

Phone models

Best accuracy

High # states

Slow

# [ Experiment ]

tier	English (%)	Thai (%)	# states
phone	100	100	1,849+5,718
voice	100	97.3	84+108
manner	100	98.2	855+768
place	99.5	92.9	1,959+489
height	100	99.1	333+111
type	100	99.1	258+132

- **AF models**

- **Accuracy reduced**
- **Lower # states & Faster**

# [ Experiment ]

tier	English (%)	Thai (%)	# states
phone	100	100	1,849+5,718
voice	100	97.3	84+108 855+768
manner	100	98.2	1,959+489
place	99.9	92.9	
height	100	99.1	333+111
type	100	99.1	258+132

■ **AF models (vowel related tiers)**

**Good Accuracy**

**Lower # states & Faster**

# [ Conclusion ]

---

- AF models can be used as LID
- Vowel type tier contains more language information than other tiers.