

OOV OOV spotting OOV OOV OOV

[f r ah m] OOV spotting [t uw] OOV OOV

[f r ah m] [w er d] spotting [t uw] OOV [m aa d el ih ng]

From Word Spotting to OOV Modeling

Paul Fitzpatrick (6345g11)

Goal

To automatically extract filler vocabulary for word-spotting

Why?

So language model has something to work with

May improve recognition accuracy on keywords

Gives earlier payoff in domain-specific training

Scenario

Start with small lexicon (e.g. 5-50 words)

Start with weak language model

Bootstrap by clustering filler vocabulary from large collection of untranscribed data