

Hybrid Spoken Term Discovery-DNN System

Finding words to learn segments

Roland Thiollière*, Ewan Dunbar*, Gabriel Synnaeve*[†]
Maarten Versteegh*, Emmanuel Dupoux*

*LSCP, École Normale Supérieure / EHESS / CNRS, Paris, France

[†]now at Facebook AI Research

rolthiolliere@gmail.com, emd@umd.edu,
gabrielsynnaeve@gmail.com, maartenversteegh@gmail.com,
emmanuel.dupoux@gmail.com

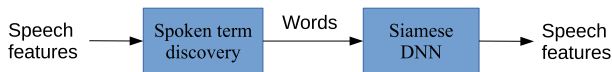
Interspeech, 2015



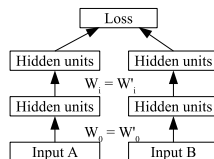
Finding words to learn phonemes

System overview and results

- Use Track 2 to help with Track 1.
- Two components: spoken term discovery and siamese deep neural network.



Similarity Matrix (source [1])



ABnet

[1] A. Jansen and B. van Durme, Efficient spoken term discovery using randomized algorithms, in Proceedings of ASRU, 2011

Track 1 evaluation: ABX

- Improved over baseline a lot.
- Close to topline in english!

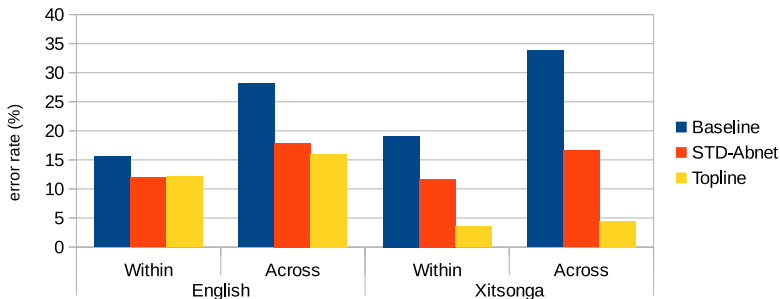


Figure: Within and across speaker Minimal Pair ABX error rates