Let me Finish

Automatic Conflict Detection Using Speaker Overlap

Felix Grezes
Justin Richards
Prof. Andrew Rosenberg
Baseline Recap

- Kitchen-sink approach: ~3k low-level features
- Results fairly good — 79.1 % UAR*

How can we improve performance while clarifying the model?

* on train + dev data
Our Approach

• Motivated by prior work, linguistic knowledge

• Hypothesis:

  \textit{The proportion of overlapping speech in recorded discussions should be a strong predictor of conflict level.}

• Testable and interpretable
Gold Overlap

Added feature:

- Corpus contains hand-labeled speaker turns
- Gold (actual) speaker overlap ratios were computed from this meta-data
- Conflict classifier built upon gold overlap ratio

Results:

- 74.2 % UAR* using gold overlap + baseline features
- 79.4 % UAR* using ONLY gold overlap

* on train + dev data
Baseline feat. vectors Train

Gold Overlap Ratios Train

SMOreg alg.

Overlap Model

Overlap Predictions
Model: Predicting Conflict

SMO alg.

Conflict Model

Conflict Predictions

Baseline + PO

PO only

Gold Conflict Labels

Train

Train

Train
Experiment Results

Train + Dev Data:

- 80.5 % UAR using **only** predicted overlap
  - Baseline: 79.1% UAR using ~3k features
  - Gold overlap: 79.4% UAR using one feature
- Adding even one extra feature, performance **decreases.**
Pred Overlap > Gold Overlap?

Theories:

• Finds instances of overlap that were missed by hand-annotators.
  o Interjections, failed interruptions

• Identifies a pressured or competitive quality in speech that is independent of overlap
Performance

Test Data:

• 83.1 % UAR using only predicted overlap

• Baseline: 80.8 % UAR using ~3k features

We improved performance using a clearer, more intuitive model.
Segmentation & Degradation

- All segments
- 1st and 2nd segments
- 1st segment
Thank You!

Contact:

fgrezes@gc.cuny.edu
jrichards@gc.cuny.edu
andrew@cs.qc.cuny.edu

The Speech Lab at Queens College:
speech.cs.qc.cuny.edu