

# Story Segmentation Experiments at The University of Iowa

---

David Eichmann<sup>1,2</sup> & Dong-Jun Park<sup>2</sup>

<sup>1</sup>School of Library and Information Science

<sup>2</sup>Computer Science Department

# Focus of Work

- For video data, just use a shot boundary run
- For text data:
  - Speech pauses longer than a certain threshold
    - $t = 1.25$  sec &  $t = 1.50$  sec
  - Trigger phrases in transcript

# News Typing

- Very direct approach:
  - Declare everything news...
  - Unless we're using trigger phrases and someone says 'network', then declare it misc.

# Trigger Phrases

- Successful TDT segmentation systems not only tried to analyze ASR content, they looked for particular artifacts in the text stream
  - A story-terminating trigger phrase (story wrap):

```
<Word stime="348.75" dur="0.22" conf="0.981"> BROOKS </Word>  
<Word stime="348.97" dur="0.52" conf="0.981"> JACKSON </Word>  
<Word stime="349.52" dur="0.19" conf="0.981"> C. </Word>  
<Word stime="349.71" dur="0.19" conf="0.981"> N. </Word>  
<Word stime="349.91" dur="0.19" conf="0.981"> N. </Word>  
<Word stime="350.10" dur="0.35" conf="0.981"> WASHINGTON </Word>  
</SpeechSegment>
```

- The end time of the segment is used as the boundary

# Trigger Phrases

- A story-initiating trigger phrase (story lead):

```
<Word stime="246.53" dur="0.23" conf="0.983"> BROOKS </Word>  
<Word stime="246.76" dur="0.35" conf="0.989"> JACKSON </Word>  
<Word stime="247.23" dur="0.44" conf="0.989"> JACKSON </Word>  
<Word stime="247.67" dur="0.75" conf="0.989"> EXPLAINS </Word>  
</SpeechSegment>
```

- Here the start time of the segment is used as the boundary

# Trigger Phrases

- We also keyed on network IDs:

```
<Word stime="758.61" dur="0.37" conf="0.967"> THIS </Word>  
<Word stime="758.98" dur="0.16" conf="0.976"> IS </Word>  
<Word stime="759.14" dur="0.11" conf="0.975"> THE </Word>  
<Word stime="759.25" dur="0.16" conf="0.983"> C. </Word>  
<Word stime="759.41" dur="0.16" conf="0.983"> N. </Word>  
<Word stime="759.56" dur="0.16" conf="0.983"> N. </Word>  
<Word stime="759.72" dur="0.41" conf="0.985"> HEADLINE </Word>  
<Word stime="760.13" dur="0.26" conf="0.982"> NEWS </Word>  
<Word stime="760.39" dur="0.37" conf="0.983"> NETWORK </Word>  
</SpeechSegment>
```

# Trigger Phrase Profile

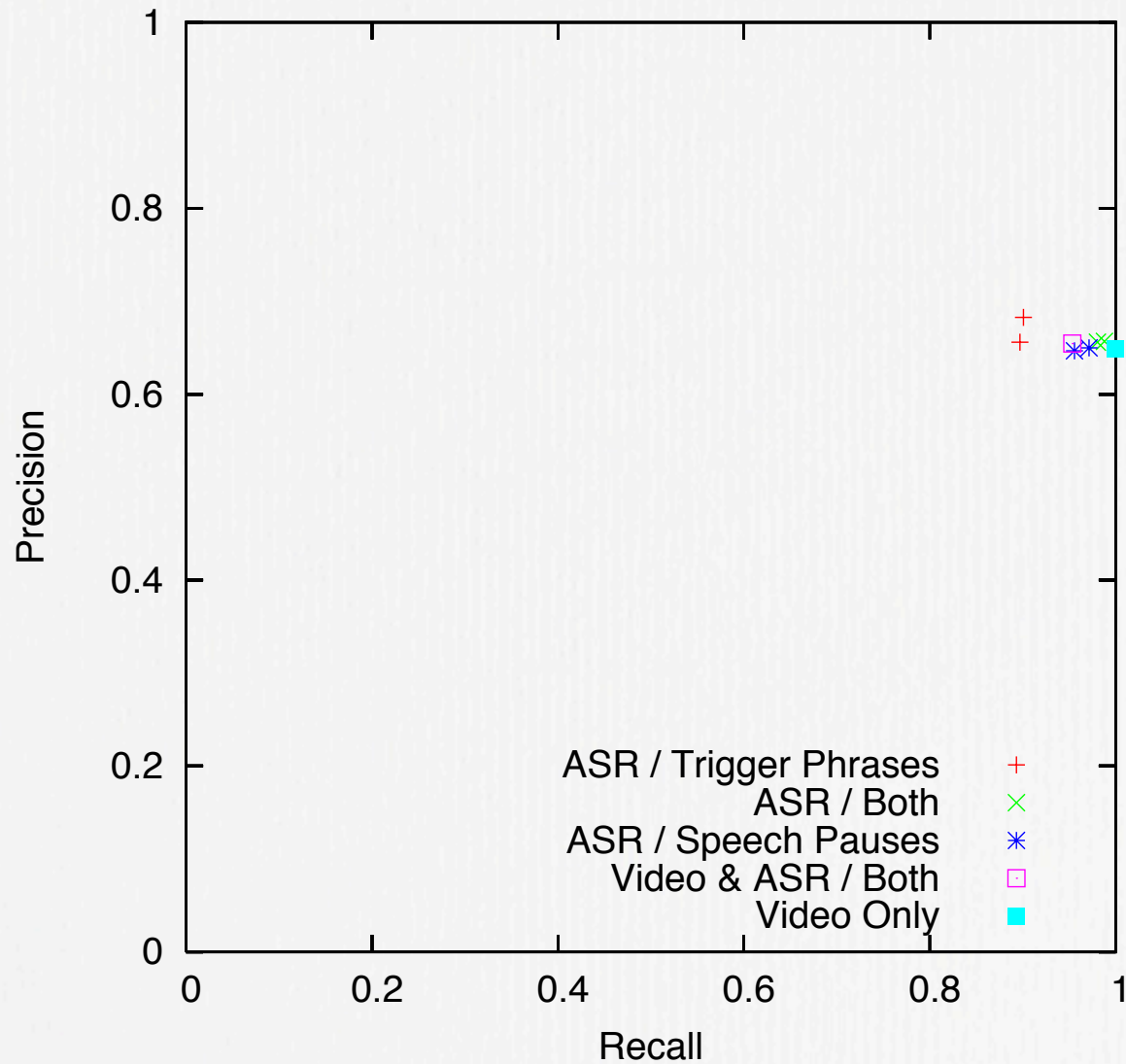
<b>Trigger Type</b>	<b>ABC</b>	<b>CNN</b>
Story Lead	4	4
Story Wrap	6	3
Network ID	1	3

# Official Runs

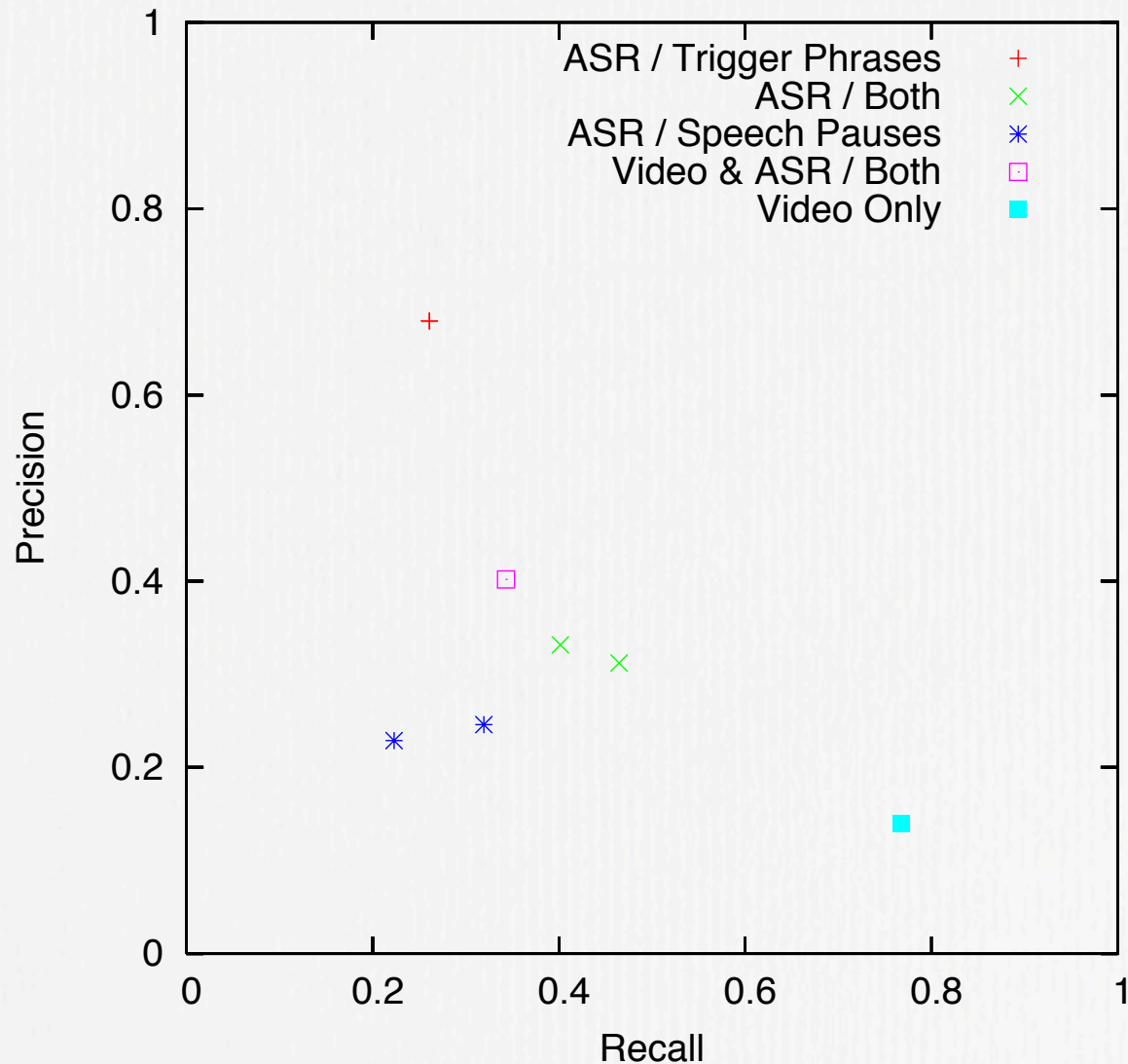
Run	Text Method	Thresh. (sec.)	Video Method	Cond.	Story Boundary		News Class.	
					Rec	Prec	Rec	Prec
UIowaSSo301	trigger	–	–	3	0.261	0.679	0.901	0.683
UIowaSSo302	both	1.50	–	3	0.402	0.332	0.980	0.656
UIowaSSo303	pause	1.50	–	3	0.223	0.229	0.956	0.647
UIowaSSo304	trigger	–	–	3	0.261	0.679	0.897	0.656
UIowaSSo305	both	1.25	–	3	0.465	0.312	0.988	0.657
UIowaSSo306	pause	1.25	–	3	0.319	0.246	0.971	0.650
UIowaSSo307	both	1.50	product	2	0.343	0.402	0.953	0.654
UIowaSSo308	–	–	product	1	0.767	0.140	1.000	0.648



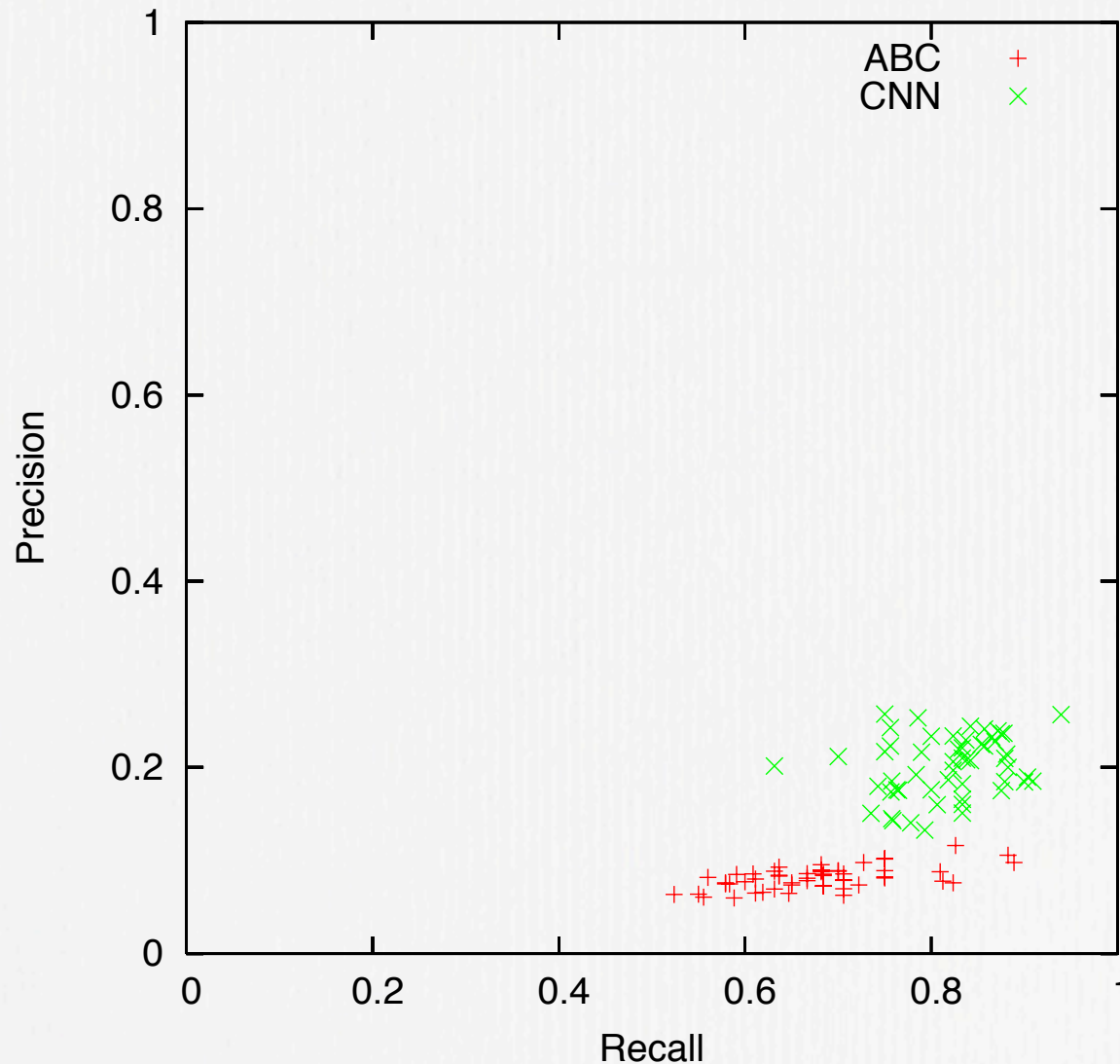
# News Typing



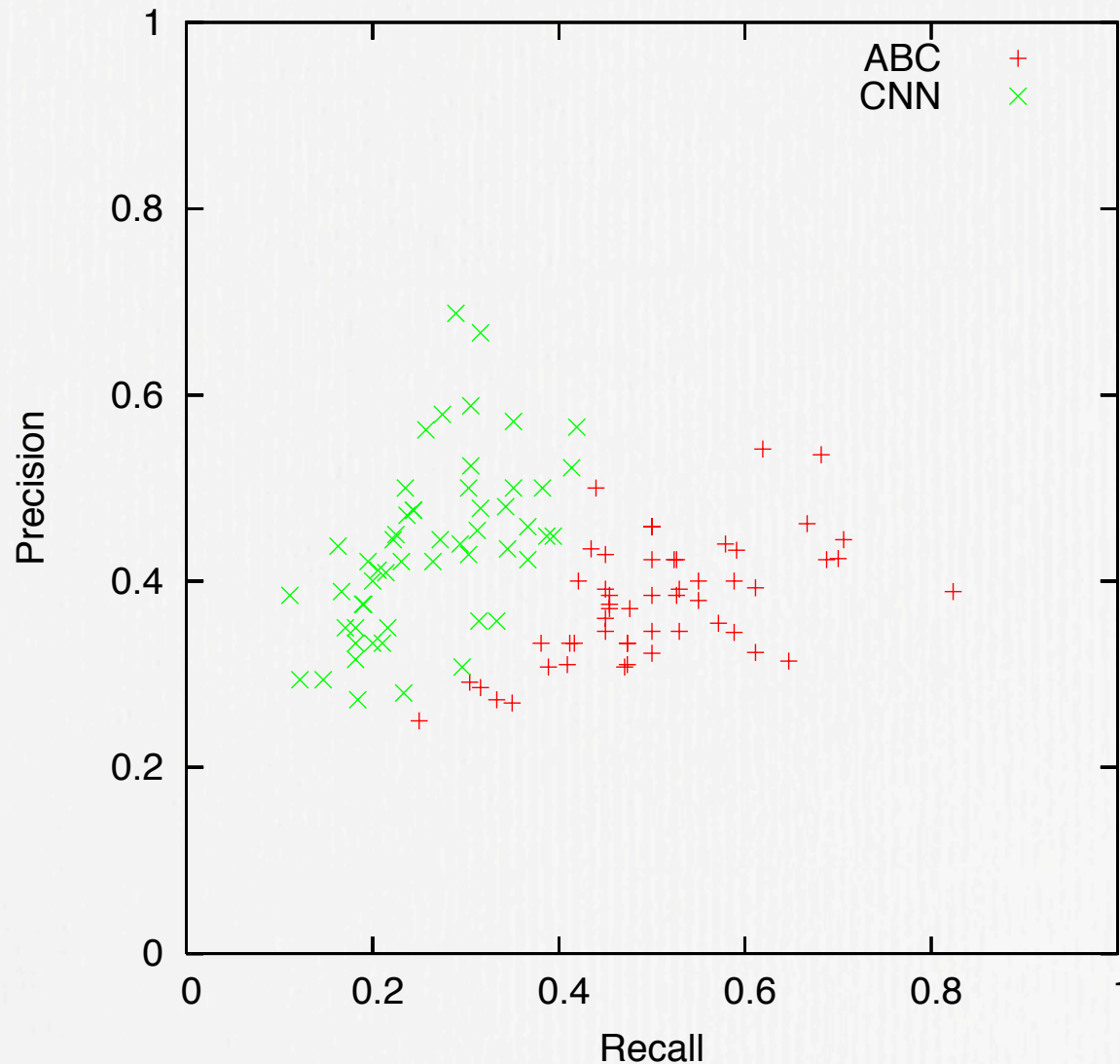
# Story Segmentation, Overall Results



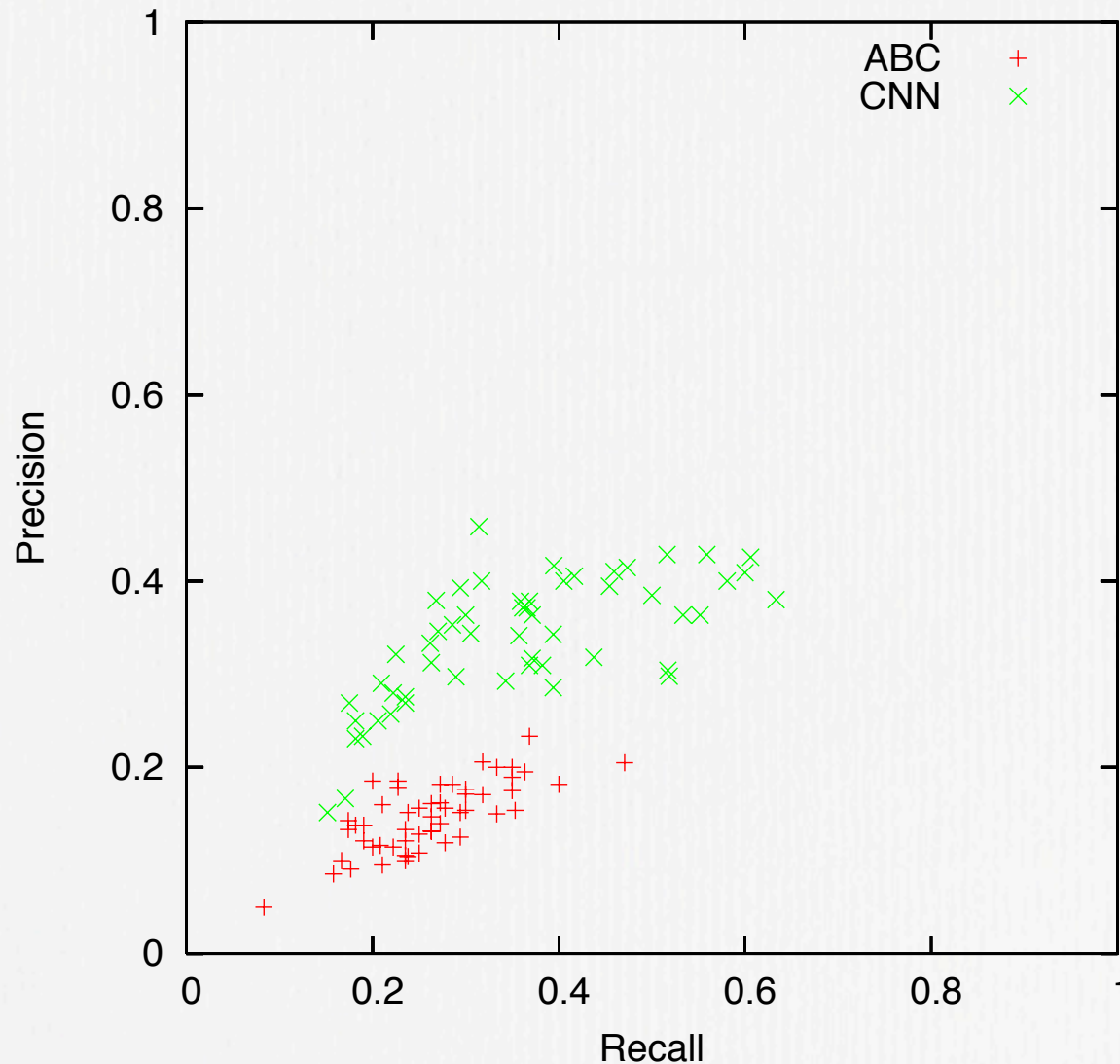
# Story Segmentation, Cond. 1, Video Only (Product)



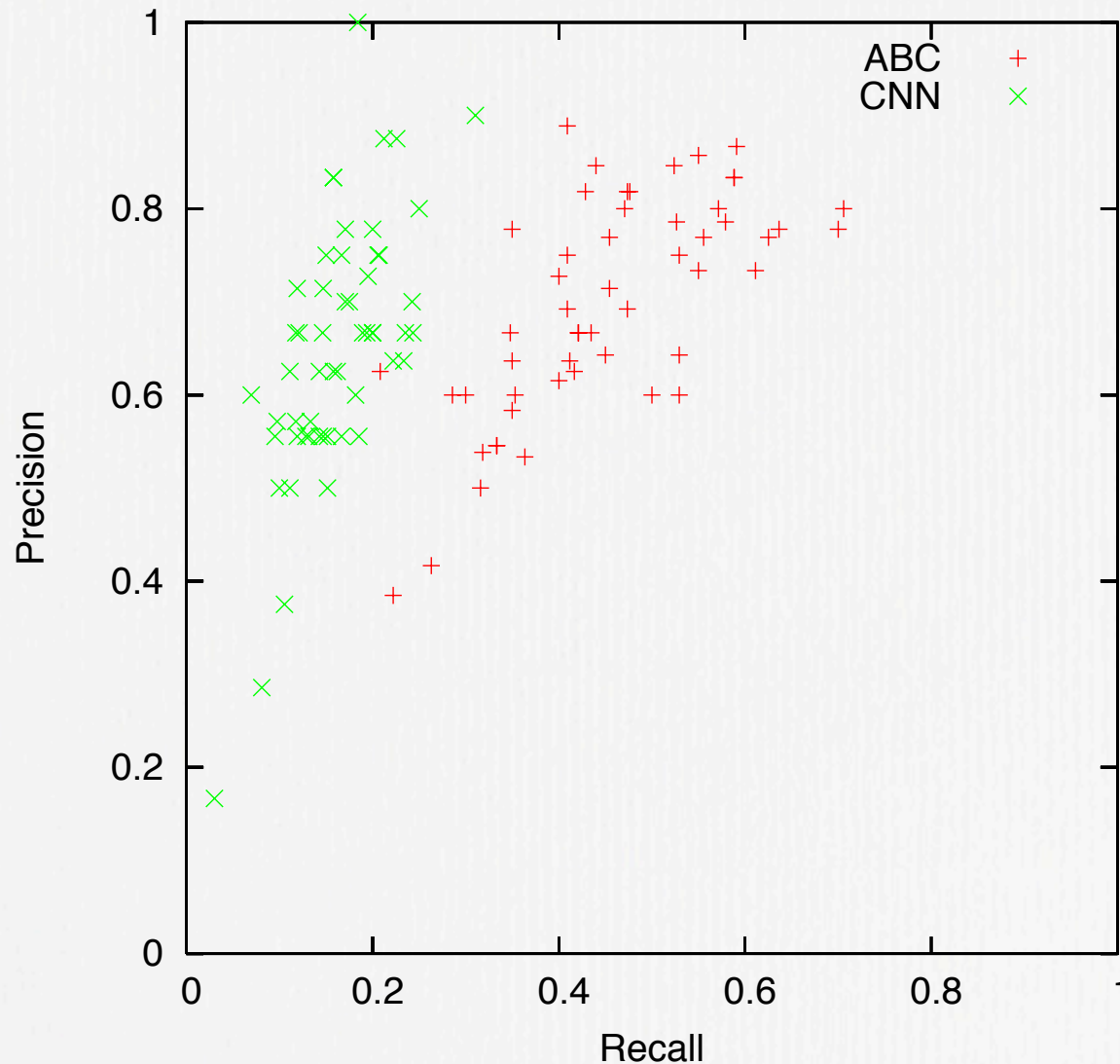
# Story Segmentation, Cond. 2, Video & Comb. Text



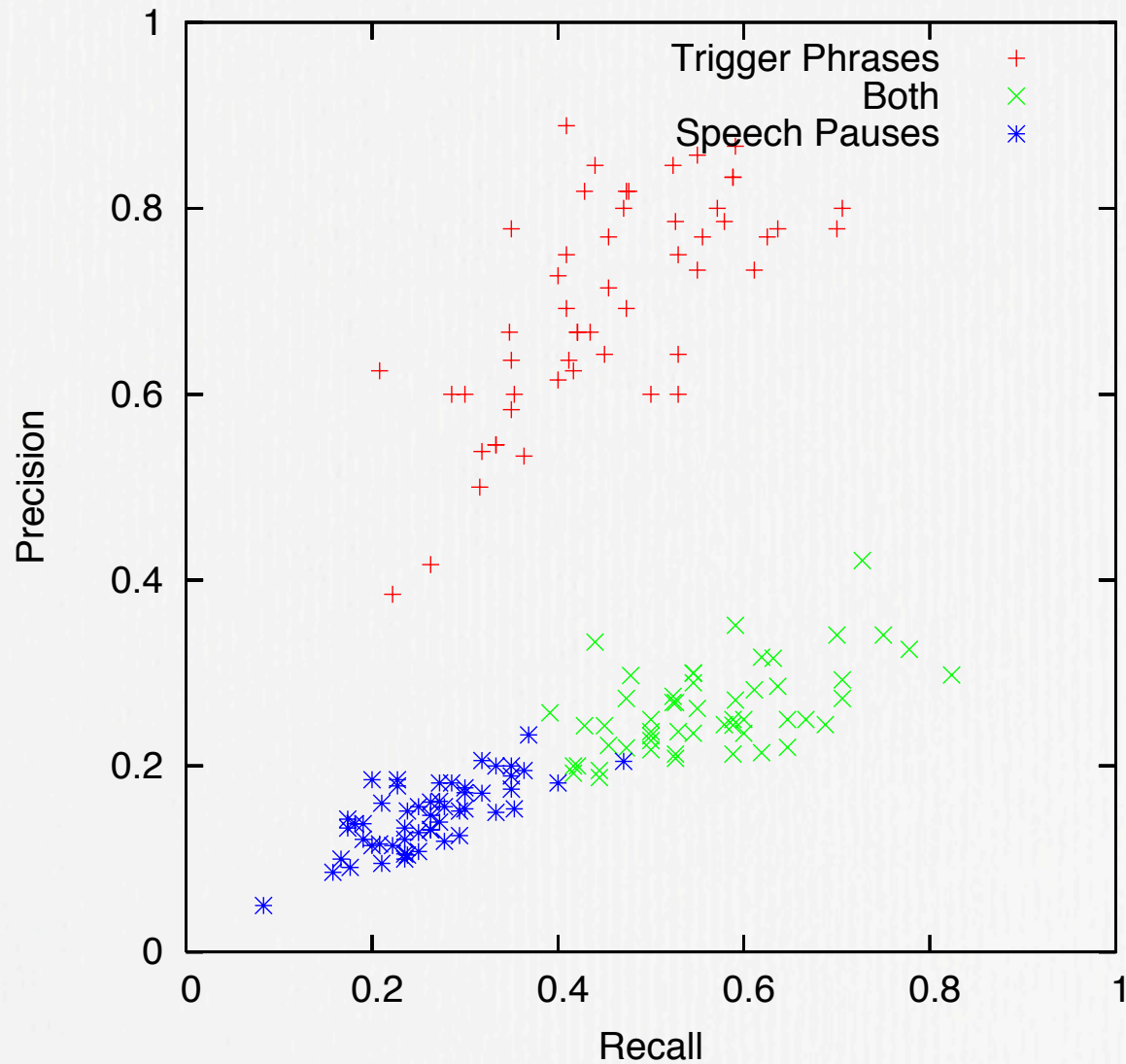
# Story Segmentation, Cond. 3, Speech Pauses



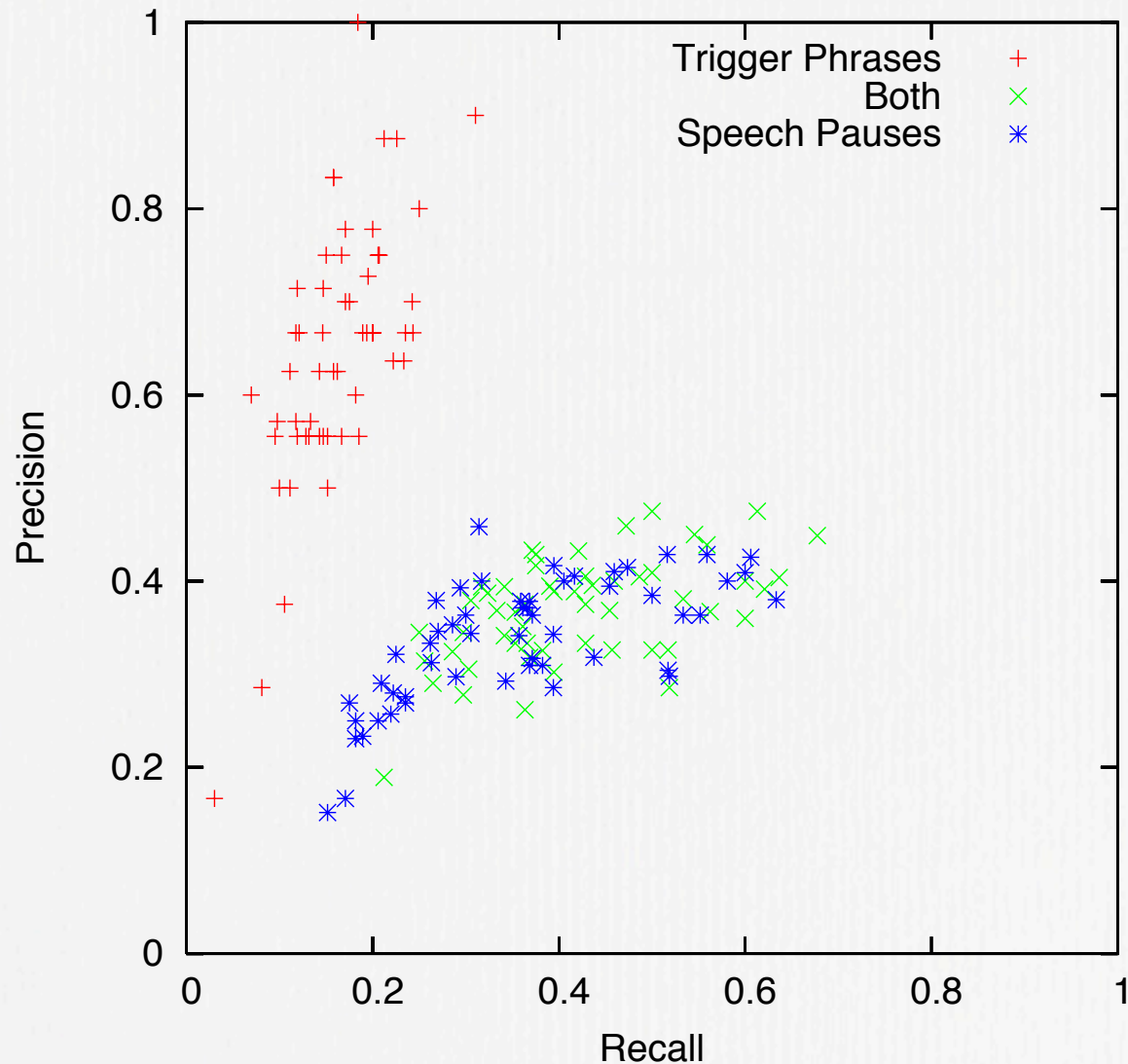
# Story Segmentation, Cond. 3, Trigger Phrases



# Story Segmentation, Cond. 3, ABC



# Story Segmentation, Cond. 3, CNN





# Conclusions

- We have some interesting performance end points with shot boundaries and trigger phrases
- Even a low-precision signal (shot boundaries) can improve both precision and recall of a signal (combined trigger phrases and speech pauses) that it's combined with
- There is a surprising distinction between and consistency within news sources(s) for our measures

# Future Work

- Explore a broader tuning range of speech pauses, particularly w.r.t. their interaction with trigger phrases
- Try separate interactions between single text measures and the video measures
- Fold in improved shot boundaries
- Improve the coverage on CNN trigger phrases, with an eye towards a generic scheme for any news source