

Penn Discourse Treebank

Natalie Parde

UIC CS 421

Penn Discourse Treebank

- **Lexically-grounded** model of coherence relations
 - Given a list of **discourse connectives** (e.g., *because*, *although*, *when*, *since*, or *as a result*) and an unlabeled document, annotators labeled:
 - Those connectives
 - The spans of text that they connected
 - In some cases, these connectives may be implicit





PDTB Sense Hierarchy

- Four main classes:
 - Temporal
 - Contingency
 - Comparison
 - Expansion
- Numerous subtypes of each



PDTB Annotations

- Only at the span-pair level!
- No hierarchical tree structure