
The Daily W

Sunday, August 30, 2006

Martians invade earth

Incredible as it may seem, it has been confirmed that a large martian invasion fleet has landed on earth tonight.

First vessels were sighted over Great Britain, Denmark and Norway already in the late evening from where, as further reports indicate, the fleet headed towards the North Pole and Santa Claus was taken hostage by the invaders.

Afterwards they split apart in order to approach most major cities around the earth. The streets filled as thousands fled their homes, many only wearing their pajamas...

Empirical approaches to discourse

day 4

PDTB, Implicit-Explicit

ESSLLI 2012

Jennifer Spenader

Outline

1. Penn Discourse Treebank

2. Implicit vs. Explicit
   - When is something really implicit?
     - are there more subtle clues than “the usual suspects”
   - Do genres differ on how explicit or implicit they are?
   - What’s the difference between the same relationship when marked vs. unmarked?
   - quantitative or qualitiative?
3. Experiments looking at the diffrence
   - Marcu & Echihabi, Spoorleder & Lascarides, Lin et al.

Why do people want to create annotated corpora?

- To get accurate distributional information
- To use as input to supervised machine learning in order to eventually automatically reocognize the annotated categories

But annotation is difficult

- How do you know if you are really annotating what you think you are annotating?
- People’s intuitions are vague here
- SOLUTION: Let people annotate what they already know

Penn Discourse Treebank (PDTB 2.0)

- Annotation of explicit and implicit relations and their arguments in the Wall Street Journal corpus of the Penn Treebank

- Connective-based annotation
  - Lexically-based = theory neutral
  - For each connective, its **sense** is identified, disambiguating different usages
  - When no connective is present, annotators are asked to **add** the most appropriate connective
  - Local coherence relations only
  - Based on idea of connectives as discourse structural projectors

Verbs and argument structure

verbs "project" their predicate-argument structure

intransitive verbs

sleep: John sleeps.

(takes one agent argument)

transitive verbs

meet: John met Mary.

buy: John bought some gum.

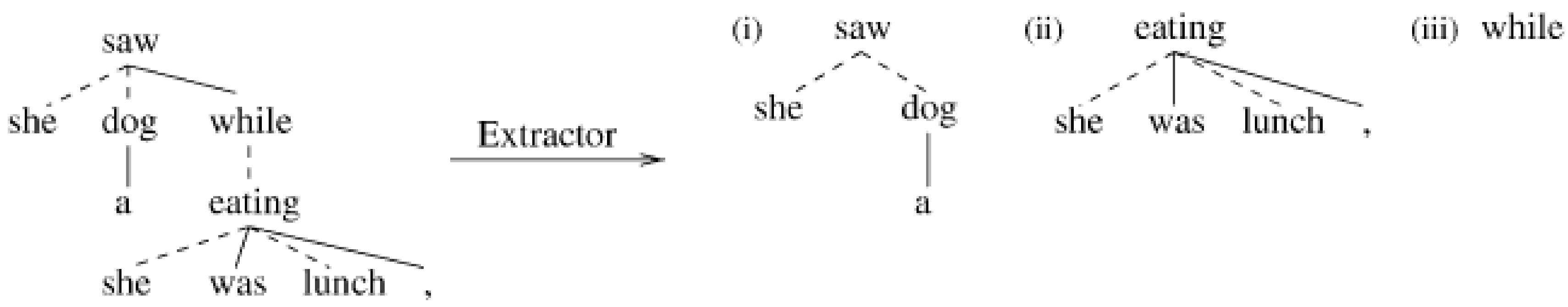
(takes two arguments, one agent and one patient)

ditransitive verbs

offer: John offered Mary some gum.

(takes 3 arguments, agent, patient and goal)

She saw a dog **while** she was eating lunch.



“Connectives are discourse level predicates which project predicate-argument structure on a par with verbs at the sentence level”

Webber and Joshi (1998; DLTAG), Webber et al. (1999b) and Webber et al. (2003)

Connective based annotation

- Connective take two abstract objects as arguments:
  - events
  - states
  - propositions
- Each annotation relates a connective with two arguments (**`Arg1`** and **`Arg2`**)
- **`Arg2`** is the clause syntactically bound to the connective.

Explicit and implicit connectives

EXPLICIT: because, Contingency.Cause.Reason

Arg1: In addition to the extra privacy of these trades, the transactions can often be less expensive to execute

because

Arg2: the parties don't have to pay a floor brokerage fee or a specialist's fee

IMPLICIT: because, Contingency.Cause.Reason

Arg1: Using small electrical shocks applied to her feet, they were able to monitor sensory nerves

(because)

Arg2: The shocks generated nerve impulses that traveled via spine to brain and showed up clearly on a brain-wave monitor, indicating no damage to the delicate spinal tissue

Three levels of Sense Tags

- Sense tagset:
- CLASS
  - Four major classes
  - Comparison, Contingency, Temporal & Expansion
- TYPE
  - 16 types
  - Only 10 of these occur more than 200 times in sections 2-22
- SUBTYPE
  - TYPE Cause contains SUBTYPES Reason & Result
  - Marks the type of ARG2, which is linearly after ARG1

Three levels

- Sense tagset:
- CLASS
  - Four major classes
  - Comparison, Contingency, Temporal & Expansion
- TYPE
  - 16 types
  - Only 10 of these occur more than 200 times in sections 2-22
- SUBTYPE
  - TYPE Cause contains SUBTYPES Reason & Result
  - Marks the type of ARG2, which is linearly after ARG1
  - However, this is often predictable from the connective (though not always)

CONTINGENCY:Cause:reason

Use of dispersants was approved *<u>when</u>* *a test on the third day showed some positive results, officials said.*

(CONTINGENCY:Cause:result)

In addition, its machines are typically easier to operate, *<u>so</u>* *customers require less assistance from software.*

Classes in the PDTB 2.0

- **CONTINGENCY**
  - The situations described in Arg1 and Arg2 are causally influenced
- **COMPARISON**
  - The situations described in Arg1 and Arg2 are compared and *differences* between them are identified *(similar situations do not fall under this CLASS)*
- **TEMPORAL**
  - The situations described in Arg1 and Arg2 are temporally related
- **EXPANSION**
  - The relevant to the situation described situation described in Arg2 provides information deemed in Arg1

Level 2 and 3: Types, Subtypes and senses (1)

- TEMPORAL: Asynchronous
  - Precedence
  - Succession

- TEMPORAL:
  Synchronous
  No subtypes

- CONTINGENCY: Cause
  - reason
  - Result

- CONTINGENCY: Condition
  - hypothetical
  - general
  - factual present
  - factual past
  - unreal present
  - unreal past

Level 2 and 3: Types Subtypes or senses (2)

- COMPARISON: Contrast
  - Juxtaposition
  - Opposition
- COMPARISON: Concession
  - expectation
  - contra-expectation
- EXPANSION: Restatement
  - Specification
  - Equivalence
  - Generalization
- EXPANSION: Alternative
  - Conjunctive
  - Disjunctive
  - Chosen alternative

Linear order?

- **Arg2 is the sentence/clause with which connective is syntactically associated.**
- ***Arg1*** **is the other argument.**

- **No constraints on relative order. Discontinuous annotation is allowed.**
  - **Linear:**

    The federal government suspended sales of U.S. savings bonds **because Congress hasn't lifted the ceiling on government debt.**

  - **Interposed:**

    Most oil companies**, when they set exploration and production budgets for this year,** ***forecast revenue of $15 for each barrel of crude produced*****.**

Explicit connectives

- Subordinating conjunctions (e.g., *when*, *because*, *although,* etc.)

  The federal government suspended sales of U.S. savings bonds **because** **Congress hasn't lifted the ceiling on government debt**.

- Coordinating conjunctions (e.g., *and*, *or*, *so*, *nor*, etc.)

  The subject will be written into the plots of prime-time shows, **and** **viewers will be given a 900 number to call**.

- ***Arg1*** and **Arg2**

- Discourse adverbials (e.g., *then*, *however*, *as a result*, etc.)

In the past, the socialist policies of the government strictly limited the size of … industrial concerns to conserve resources and restrict the profits businessmen could make. **<u>As a result</u>**, **industry operated out of small, expensive, highly inefficient industrial units**.

Implicit Connectives

No Explicit connective? Infer a relation based on adjacency.

> ***Some have raised their cash positions to record levels*. <u>Implicit=because (causal)</u> High cash positions help buffer a fund when the market falls.**
>
> ***The projects already under construction will increase Las Vegas's supply of hotel rooms by 11,795, or nearly 20%, to 75,500*. <u>Implicit=so (consequence)</u> By a rule of thumb of 1.5 new jobs for each new hotel room, Clark County will have nearly 18,000 new jobs.**

Insert the connective that “best” captures the relation.

Paired connectives

- Take the same arguments:
  - **On the one hand**, **Mr. Front says,** ***it would be misguided to sell into "a classic panic."*** **On the other hand**, **it's not necessarily a good time to jump in and buy.**
  - **Either** ***sign new long-term commitments to buy future episodes*** **or** **risk losing "Cosby" to a competitor.**
- Treated as complex connectives – annotated discontinuously
- Listed as distinct types (no head-modifier relation)

Non-insertion: AltLex

1. AltLex

A discourse relation is inferred, but inserting a connective would be redundant.

Other lexical information (in the form of a non-connective expression) signals the same relation

New rules force thrifts to write down their junk to market value, then sell the bonds over five years.

AltLex = (result) That's why **Columbia just wrote off $130 million of its junk and reserved $227 million for future junk losses.**

Non-insertion: EntRel

- **EntRel**: Coherence is created by an entity-based relation

- ***Hale Milgrim, 41 years old, senior vice president, marketing at Elecktra Entertainment Inc., was named president of Capitol Records Inc., a unit of this entertainment concern*. <u>EntRel</u> Mr. Milgrim succeeds David Berman, who resigned last month.**

Deals with the problem of whether or not ELABORATION should be a discourse relation.

Non-insertion: NoRel

NoRel: Neither discourse nor entity-based relation is inferred.

Jacobs is an international engineering and construction concern**.** **NoRel** **Total capital investment at the site could be as much as $400 million, according to Intel.**

What can be an argument?

- Clauses or sentences (standard)
- Discourse deictic expressions (references to abstract objects)
  - Airline stocks typically sell at a discount of about one third to the stock market's price-earnings ratio – which is currently about 13 times earnings. **[That's]** *because* **[airline earnings, like those of auto makers, have been subject to the cyclical ups-and-downs of the economy].**
- Textual spans from which arguments can be derived
  - **[No price for the new shares has been set].** *Instead*, **[the companies will leave it up to the marketplace to decide].**

How much material should an argument include

- Originally: only tags **CONN, ARG1, ARG2**
- **SUP1** and **SUP2**: new tags added for information the annotator considered useful
  - *Although* **[started in 1965]**, **[Wedtech didn't really get rolling until 1975]** (when Mr. Neuberger discovered the Federal Government's Section 8 minority business program)

- **the Minimality Principle** in PDTB argument selection: be conservative in identifying **ARG1** and **ARG2**

Attribution features

- Annotated for
  - Explicit connectives
  - Implicit connectives
  - AltLex

34% of discourse relations are attributed to an agent other than the writer.

Attribution

- Relation of "ownership" between Agents and Abstract Objects
  - Abstract Objects: beliefs, facts, propositions
- Not a discourse relation
- Shows how discourse relations and their arguments can be *attributed to different individuals*:

When Mr. Green won a $240,000 verdict in a land condemnation case against the state in June 1983, [he says] *Judge O'Kicki unexpectedly awarded him an additional $100,000*.

⇒ Relation and Arg2 are attributed to the Writer.

⇒ *Arg1* is attributed to another agent.

Summary PDTB 2.0

- All WSJ sections (25 sections; 2304 texts ;1 million words)
- 100 distinct types of relations
  - Subordinating conjunctions – 31 types
  - Coordinating conjunctions – 7 types
  - Discourse Adverbials – 62 types
- About 20,000 distinct tokens

PDTB Relations	No. of tokens
Explicit	18459
Implicit	16224
AltLex	624
EntRel	5210
NoRel	254
Total	40600

PDTB style annotation

Relatively easy?

1. No embedded structures. (unlike RST!)
2. Very high level categories:
   - Almost seems as if they planned to avoid difficulties
   - Only major ambiguities seem to be when the connective itself is ambiguous
3. Connective based annotation

PDTB first annotation experiments

Miltsakaki, Prasad, Joshi and Webber(2004)

- How difficult is it to identify ARG1 and ARG2?
- 10 connectives (2717 tokens)
  - **subordinating conjunctions:** *because, although, even though, when, so that*
  - **discourse adverbials:** *nevertheless, therefore, as a result, instead, otherwise*
- Explicit tokens: 2717
- Implicit tokens: 386
- 2 annotators
- Kappa statistic requires into discrete categories, but the PDTB annotation tokens are spans and connectives, so not appropriate

Inter-annotator agreement

- All argument annotations, treating ARG1 and ARG2 as independent tokens
  - total 5434 (twice # of relations)
- Agreement *exact match* criterion
  - 90.2% agreement (4900/5434 tokens) for all
  - ARG1 Agreement: 86.3%
  - ARG2 Agreement: 94.1%

CONNECTIVES	AGR No.	Conn. Total	%AGR
when	1877	2032	92.4%
because	1703	1824	93.4%
even though	194	206	94.1%
although	635	704	90.1%
so that	66	74	89.2%
TOTAL SUBCONJ	**4469**	**4834**	**92.4%**
nevertheless	56	94	59.6%
otherwise	44	46	95.7%
instead	172	236	72.9%
as a result	110	168	65.5%
therefore	49	56	87.5%
TOTAL ADV.	**431**	**600**	**71.8%**
OVERALL TOTAL	**4900**	**5434**	**90.2%**

Table 1: Distribution of Agreement by Connective, with ARG1 and ARG2 Annotations Counted Independently

Inter-annotator agreement

For Implicit connectives:

- Connectives divided into 5 groups (based on Knotts work):
  - a) **additional information**
    - (e.g., ‘furthermore’, ‘in addition’)
  - b) **cause-effect relations**
    - (e.g., ‘because’, ‘as a result’),
  - c) **temporal relations**
    - (e.g., ‘then’, ‘simultaneously’),
  - d) **contrastive relations**
    - (e.g., ‘however’, ‘although’),
  - e) **restatement or summarization**
    - (e.g., ‘in other words’, ‘in sum’).
- 72% agreement on added connectives

In the released annotation (PDTB 2.0)

- What was the distribution of connectives like?
  - how ambiguous were connectives?

PDTB

- Explicit relations: the most frequent relation for a given connective accounts for over 90% of the discourse relations
  - most connectives are unambiguous
  - (Miltsakaki et al.,2005; Pitler et al., 2008).

Connective ambiguity
by the four sense classes:

- Discourse connectives that occur with their most common sense by connective CLASS

Comparison	93.43%
Contingency	94.72%
Temporal	**84.10%**
Expansion	**97.63%**

Pitler et al. Easily identifiable discourse relations Coling 2008

Comparison Class

connective	**#**	**% with most frequent meaning**
but	3308	97.19%
while	781	66.07%
however	485	99.59%
although	328	99.70%
though	320	100.00%
still	190	98.42%
yet	101	97.03%

(only connectives n > 50 in corpus)

Pitler et al. Easily identifiable discourse relations Coling 2008

Comparison Class

connective	**#**	**% with most frequent meaning**
but	3308	97.19%
while	781	66.07%
however	485	99.59%
although	328	99.70%
though	320	100.00%
still	190	98.42%
yet	101	97.03%

(only connectives n > 50 in corpus)

Explicit versus Implicit coherence relations

PDTB is connective based annotation:

Explicit Relation

John is very tired **because** he played tennis all morning.

Implicit Relation w/ implicit connective

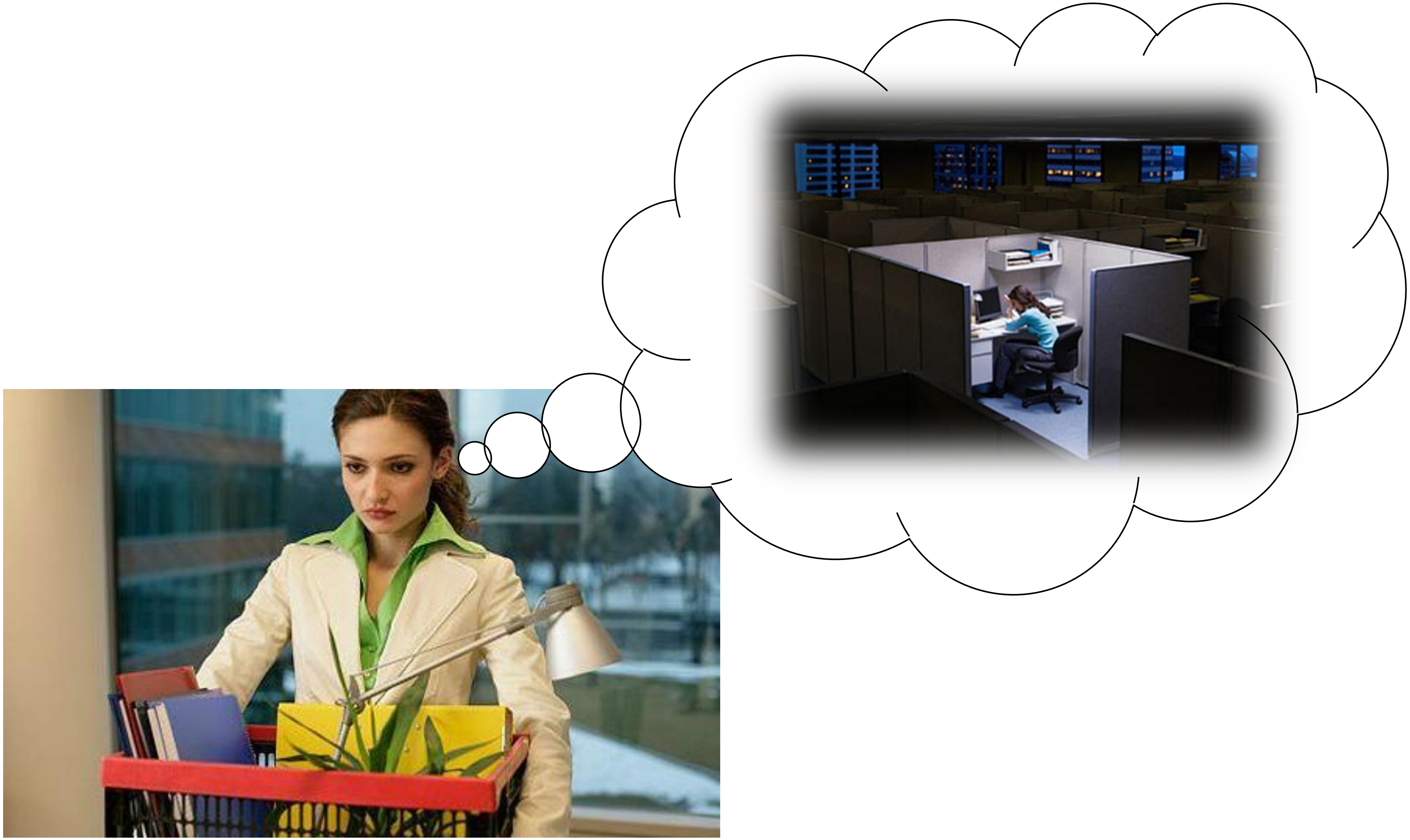
John is very tired. He played tennis all morning

Taboada, M. (2009) Implicit and Explicit Coherence Relations. In J. Renkema (ed.) *Discourse, of Course*.

- Taboda (2009) argues that all relations may be explicit.
  - because people seem to interpret relations with relative ease, so there must be signals guiding them

Taboada, M. (2009) *Implicit and Explicit Coherence Relations*. In J. Renkema (ed.) *Discourse, of Course*.

- Taboda (2009) argues that all relations may be signalled in some way
  - people seem to interpret `unmarked’ relations easily, so there must be some clue
- Two problems:
  - How do we discover which cues are signaling relations?
  - How do we test if relations are cognitively represented in the minds of hearers and readers?
    - We know from annotation experiments that identifying coherence relations is not trivial. Do we always identify these connections?



Kim

Kim quit her job.
She was tired of the long hours.

Kim quit her job ***because***
she was tired of the long hours.

Strong, causal relationship?

Kim quit her job ***because***
she was tired of the long hours.

Kim quit her job.
She was tired of the long hours.

Kim quit her job.

She was tired of the long hours ***anyway***.

No clear causal relationship.

Kim quit her job.

She was tired of the long hours ***anyway***.

Unmarked relations common

- In the PDTB (ignoring other relations)
  - 53% Explicit (18459 relations)
  - 47% Implicit (16224 relations)

- RST sources (Taboada) over 50% of relations do not have a `traditional' discourse marker
  - analyses on the RST website (Mann & Taboada, 2007), very diverse collection of 187 texts: 72% of the relations had no discourse marker
  - Taboada (2006) study (mostly discourse markers, but also mood, finiteness and punctuation)
    - In conversation; relations signaled 31% of the time
    - In newspapers: 44% had discourse markres, although a few other signals are discussed in that paper

What `signalling mechanisms’ identify the coherence relations between discourse segments? (besides discourse markers):

- morphology
- syntactic structures
- semantic and pragmatic information
- discourse particles
- real world knowledge

Taboada, M. (2009) [Implicit and Explicit Coherence Relations]. In J. Renkema (ed.) *[Discourse, of Course]*.

[S] Some entrepreneurs say the red tape they most love to hate is red tape they would also hate to lose.

[N] They **concede** that much of the government meddling that torments them is essential to the public good, and even to their own businesses.

(Concession, RST Discourse Treebank)

What do we find in the PDTB?

- How often are connectives present in the PDTB?

CLASS	**Explicit (%)**	**Implicit (%)**	**Total**
Comparison	5590 (69%)	2505 (31%)	8095
Contingency	3741 (47%)	4261 (53%)	8002
Temporal	3696 (80%)	950 (20%)	4646
Expansion	6431 (42%)	8868 (58%)	15299

- How often are connectives present?

CLASS	**Explicit (%)**	**Implicit (%)**	**Total**
Comparison	5590 (69%)	2505 (31%)	8095
Contingency	3741 (47%)	4261 (53%)	8002
Temporal	**3696 (80%)**	**950 (20%)**	**4646**
Expansion	**6431 (42%)**	**8868 (58%)**	**15299**

- What do these distributional differences tell us? Do they make sense?

CLASS	**Explicit (%)**	**Implicit (%)**	**Total**
Comparison	5590 (69%)	2505 (31%)	8095
Contingency	3741 (47%)	4261 (53%)	8002
Temporal	**3696 (80%)**	**950 (20%)**	**4646**
Expansion	**6431 (42%)**	**8868 (58%)**	**15299**

Binary classification of CLASS

- Pitler et al. Easily identifiable discourse relations
- Can you distinguish one class from the other three?
  - Connective info only, “All“ includes implicit connectives
  - Decision Tree classifier, binary classification task

CLASS	All	Explicit only
Comparison	91%	97%
Contingency	84%	94%
Temporal	95%	95%
Expansion	77%	98%

Four-way classification of CLASS

- Pitler et al. Easily identifiable discourse relations
- Can you distinguish CLASS with connective info alone?
  - Decision Tree classifier,

CLASS	Precision	Recall
Comparison	0.84 [0.84]	0.72 [0.90]
Contingency	0.66 [0.97]	0.98 [0.96]
Temporal	0.95 [0.95]	0.37 [0.844]
Expansion	0.93 [0.93]	0.67 [0.97]

- What does this tell us about the difference between implicit and explicit relations?
  - Does it look as if some implicit relations show more connective ambiguity than explicit relations of the same type?

CLASS	**Precision**	**Recall**
Comparison	0.84 [0.84]	0.72 [0.90]
Contingency	0.66 [0.97]	0.98 [0.96]
Temporal	0.95 [0.95]	0.37 [0.844]
Expansion	0.93 [0.93]	0.67 [0.97]

What does this mean?

- If connectives are present…
  - Determining CLASS easy
    - BUT is CLASS enough information?
    - CLASS is very vague

  "Level 1 classes are too general and coarse-grained for downstream applications while Level 3 subtypes are too fine-grained and are only provided for some types." (Lin et al 2009)

- If there is no connective
  - CLASS identification is not so easy
  - Implicit relation recognition for TYPE most useful

Implicit vs. Explicit relations

- Are implicit and explicit relations the same?
  - fact that Explicit relations occur with a cue phrase suggests that they might need to be signaled
  - implicit relations may have clearer features than Explicit relations
  - research that manipulates Explicit relations to try to find features for Implicit relations might be on the wrong track

Evidence that Implicit relations are qualitatively different

- different senses occur with Explicit versions and Implicit versions of the same connectives
  - suggests that certain senses may be only possible with Explicit relations.
- Anderson & Spenader (ms.)
  - suggest a qualitative difference between Implicit and Explicit PURPOSE examples, but not between Implicit and Explicit RESULT relations.
  - To even be recognizable as an Implicit PURPOSE additional features are necessary, e.g. cues like a modal auxiliary, explicit connectives,
  - Contrasts with RESULT relations. When the event pairs are RESULT, 84% to 99% of the time with or without the connective the relation will be identified as a RESULT.
  - Explicit and Implicit versions of RESULT are identified equally as well

Getting a decent set of implicit relations to study is hard.

- Marcu & Echihabi (2002).
- Sporleder & Lascarides (2008)

- Create **synthetic** examples of implicit relations

It really looked like rain **but** I didn’t take my umbrella.

ARG1 but ARG2

It really looked like rain **but** I didn’t take my umbrella.

ARG1 but ARG2

but seems to be an unambiguous marker of a CONTRAST relation

It really looked like rain **but** I didn’t take my umbrella.

ARG1 but ARG2

but seems to be an unambiguous marker of a CONTRAST relation

Note: in PDTB = Comparison.Concession: contra-expectation

It really looked like rain…..I didn’t take my umbrella.

ARG1 ARG2

[Arg1, Arg2] = synthetic implicit CONTRAST relation

Marcu & Echihabi

Marcu & Echihabi

- Bag-of-words type language model
- Naïve Bayes learning
- Model of what occurs on either side of a given connective

- Such standards would preclude arms sales to states like Libya, which is also currently subject to a U.N. embargo.
- **But** states like Rwanda before its present crisis would still be able to legally buy arms.

Marcu & Echihabi

CONTRAST	CAUSE-EXPLANATION-EVIDENCE	ELABORATION	CONDITION
ANTITHESIS (M&T) CONCESSION (M&T) OTHERWISE (M&T) CONTRAST (M&T) VIOLATED EXPECTATION (Ho) (CAUSAL | ADDITIVE) - (SEMANTIC | PRAGMATIC) - NEGATIVE (K&S)	EVIDENCE (M&T) VOLITIONAL-CAUSE (M&T) NONVOLITIONAL-CAUSE (M&T) VOLITIONAL-RESULT (M&T) NONVOLITIONAL-RESULT (M&T) EXPLANATION (Ho) RESULT (A&L) EXPLANATION (A&L) CAUSAL - (SEMANTIC | PRAGMATIC) - POSITIVE (K&S)	ELABORATION (M&T) EXPANSION (Ho) EXEMPLIFICATION (Ho) ELABORATION (A&L)	CONDITION (M&T)

Table 1: Relation definitions as union of definitions proposed by other researchers (M&T – (Mann and Thompson, 1988); Ho – (Hobbs, 1990); A&L – (Lascarides and Asher, 1993); K&S – (Knott and Sanders, 1998)).

- Raw Corpus : 1 billion words, 41 million sentences
- BLIPP corpus = 1.7 million sentences
- 5000 examples of each collapsed relation class:
  - Contrast: simulated with *but, although*
  - *CEV* = cause-explanation-evidence
  - COND = condition
  - ELAB = elaboration
- Binary classification problem
- Naïve Bayes classifier

Marcu & Echibabi

	CONTRAST	CEV	COND	ELAB	NO-REL-SAME-TEXT	NO-REL-DIFF-TEXTS
CONTRAST	-	87	74	82	64	64
CEV			76	93	75	74
COND				89	69	71
ELAB					76	75
NO-REL-SAME-TEXT						64

Table 3: Performances of classifiers trained on the Raw corpus. The baseline in all cases is 50%.

	CONTRAST	CEV	COND	ELAB	NO-REL-SAME-TEXT	NO-REL-DIFF-TEXTS
CONTRAST	-	62	58	78	64	72
CEV			69	82	64	68
COND				78	63	65
ELAB					78	78
NO-REL-SAME-TEXT						66

Table 4: Performances of classifiers trained on the BLIPP corpus. The baseline in all cases is 50%.

	CONTRAST	CEV	COND	ELAB	NO-REL-SAME-TEXT	NO-REL-DIFF-TEXTS
CONTRAST	-	87	74	82	64	64
CEV			76	93	75	74
COND				89	69	71
ELAB					76	75
NO-REL-SAME-TEXT						64

Pretty amazing results!

- Maybe it shows that if we have enough data, we can build lexical models that will identify in general terms, coherence relations!

Spenader, J. & G. Stulp (2007). Antonymy in Contrast Relations.

	All Antonyms	Direct	Indirect
Antonym Source of Contrast	177 (52%)	120 (68%)	57 (32%)
Antonym not Source of Contrast	41 (12%)	21 (51%)	20 (49%)
Wrong Sense	124 (36%)	38 (31%)	86 (69%)
Total	342 (100%)	179	163

Tab. 1: *but*-marked sentences with WordNet identified adjective antonyms

We couldn't find strong evidence that WordNet antonyms would help much in identifying contrastive relations.