1. Why comments?

- Online news comments are a rich resource for computational linguists
- They are an excellent source of evaluative and argumentative language; dialogic structure
- They also contain information about people’s opinion on important issues and policies
- Can be used to tackle research questions around journalism and public opinion
- Online language
- Human conversation

2. Comment thread example

**How to preserve Canada’s indigenous languages**

_Hadley_  Nov 16, 2012

How could preserving and teaching these languages that no one speaks be a benefit to anyone? Teaching young people to speak these languages “enforces their alienation” “gives no one a break” “What better way to alienate a young person then to teach them a language no one understands? More victims from the Global Free people that have probably never seen an aboriginal person in their life.”

_Sisco_  Nov 16, 2012

The better question to ask is why it is the obligation of the state to preserve these languages?

_Richelle60_  Nov 17, 2012

Now could preserving and teaching these languages that no one speaks be a benefit to anyone? Teaching young people to speak these languages “enforces their alienation” “gives no one a break” “What better way to alienate a young person then to teach them a language no one understands? More victims from the Global Free people that have probably never seen an aboriginal person in their life.”

_I'm a real person_  Nov 17, 2012

just an excuse to soak more tax dollars, they need to justify the cost of the country not hide its downturn.

_VanChadE can do it Jesus" MOVE "_"_  Nov 17, 2012

3. Raw corpus

The raw corpus comprises _opinion articles_ from the Canadian national newspaper _The Globe and Mail_ for a 5-year period (January 2012-December 2016), plus all comments posted in response to the articles.

- **Articles corpus**: 10,339 articles
- **Comments corpus**: 663,173 comments
- **Comment-threads corpus**: 303,665 threads

4. Constructiveness examples

**NDP thinks big with national daycare plan**

Simpson is right: it’s a political winner and a policy dud - just political smoke and mirrors. Mulcair is power-hungry. He wants Canada to adopt a national childcare model so he can hang on to seats in Quebec, that’s all. Years ago I worked with a political strategist working to get a Liberal candidate elected in Conservative Calgary. He actually told his client to talk about national daycare - this was in the early 90’s. The Liberal candidate said, ‘Canada can’t afford that!’ to which the strategist responded ‘Just say the words, you don’t have to actually do it. It’ll be good for votes.’ I could barely believe the cynicism, but over the years I’ve come to realize that’s what it is: vote getting and power politics. Same thing here.

**Excellent!! Maggie, I think you are the only sane person in the world!**

5. Constructiveness and toxicity annotations

Constructive comments target _specific points_; provide _appropriate evidence_; _offer a solution_ to the issues discussed in the article; share a related _personal story or experience_; and _encourage_ other readers to participate in the discussion.

**Very toxic**

- Please stop whining. Trump is a misogynist, racist buffoon and perhaps worse. Clinton is ethically challenged and craven in what she will tolerate in her lust for power. Neither of them is a stellar representative of their gender. Next time, put up a female candidate who outshines the male, not one who has sunk to his same level. Simple.

**Toxic**

- Why don’t the NDP also promise 40 acres and a mule? They will never lead this country. Panderers to socialists and unionists.

**Mildly toxic**

- Is this a joke? Marie Henein as feminist crusader, advising us what to tell our daughters?? no thanks

**Not toxic**

- In my opinion, criticizing the new generation is not going to solve any problem. If you want to produce children, you should be prepared to pay for their care.

6. Appraisal annotations

- **What is the distribution of different types of evaluation?**

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>The shallowness of this debate is depressing.</td>
<td>Affect, Negative</td>
</tr>
<tr>
<td>Thoughtful article! Thanks.</td>
<td>Appreciation, Positive</td>
</tr>
<tr>
<td>low-skilled immigrants...</td>
<td>Judgment, Negative</td>
</tr>
</tbody>
</table>

7. Negation annotations

- **What kind of opinions are negated?**
- **What is most often in the focus of negation?**
- **How can we accurately identify the item being negated?**

Clinton lost, not because <she’s a woman> but because she was perceived to be an establishment candidate.

8. Further information

All corpora are available to download at: https://github.com/sfu-discourse-lab/SOCC

Ongoing work

- Carrying out large-scale annotations for constructiveness and toxicity
- Exploring the relation between toxicity and constructiveness
- Building computational models for constructiveness