

Tieja Thomas, PhD

Vivek Venkatesh, PhD

CORPUS-ASSISTED (CRITICAL) DISCOURSE ANALYSIS

4 February 2016

@TiejaThomas



#CACDA

What we aim to accomplish today...

- 6 W's of digital research methods
- Corpus-assisted (critical) discourse analysis
- Hands-on activity

A brief introduction to

DIGITAL RESEARCH METHODS

WHY conduct digital research?

- Information, knowledge
- Learning
- Relationships, communities
- Identities

WHO (what) is my unit of analysis?

- Individual users
- Online communities
- Speech acts
- Websites
- Behaviours

WHAT are key digital research methods?

Method

- Online ethnography
- Netnography
- Computer-mediated discourse analysis
- Corpus-assisted (critical) discourse analysis
- Content analysis
- Network analysis

Unit of analysis

- Individual users
- Online communities
- Speech acts

- Websites
- Online behaviours

WHERE do I conduct digital research?

- Social media platforms
 - Facebook, Twitter, reddit, Instagram
- Online gaming platforms
 - Steam, Xbox Live, Playstation Network
- Online news outlets
 - huffingtonpost.ca, CBC.ca, theglobeandmail.com
- Digital communication
 - Email, SMS, WhatsApp

HOW do I conduct digital research?

- On-/offline divide
- Researcher role
- Privacy: contextual integrity
- Data management tools

WHEN do I stop collecting data?

- Data deluge (!)
- What data are relevant?
 - User data
 - Networks
 - Site analytics, affordances

Corpus-assisted (critical) discourse analysis

CACDA

What is CACDA?

Corpus Linguistics + Critical Discourse Analysis

- Quantitative study of language
- Analysis of large volumes of naturally occurring e-data
- Linguistic patterns
- Qualitative study of power and ideology through language
- Linguistic indicators, patterns

CACDA allows you to connect micro- and macro-level linguistic/discursive patterns.

CACDA in Practice

- Analysis of reddit conversations
- Trial and error Boolean searches
- Data management
 - Diigo, copy/paste

Preparing Data for CACDA

11. Can some of you from Quebec tell me why you think the Charter of Values is a good idea? (self.canada)
submitted 11 months ago by maybe-tomorrow

Sorry. I just do not get it. If this was happening in Ontario, I would be protesting against it. It seems based in racism to me. I am happy to be wrong though so educate me.

rienavoir 92 points 11 months ago

In principle it has merits.
Removing any and all of religion's influence on state affairs is a sound objective. Nothing new here.
Where it falls short is in its application. People who demonstrate their faith feel singled out, even attacked, by the proposed measures. In turn, the francophones are labeled as intolerant and even racist. The division it creates just sucks for everyone.
But in principle, it has merits.

wedderspoon 10 points 11 months ago*

I'm not convinced. That said it would only have merit if it applied equally to all religions including Catholicism, which it doesn't.

rienavoir 18 points 11 months ago

Yeah, that crucifix should have been taken down. That part makes us look stupid.

Look at the actual graphic produced:
<http://www.ottawacitizen.com/life/cms/binary/8988788.jpg>[1]
The good cross is a classical catholic/protestant symbol while the bad cross appears orthodox.

It's time for a

DEMO & HANDS-ON ACTIVITY

CACDA: SketchEngine

- Corpus info:
 - Provides overall statistics and information
- Word list:
 - Produces frequency counts
- Word sketch:
 - Provides information about collocations, frequency, grammatical relationships
- Wildcards:
 - Start search term with (?!) to ignore case
 - Insert .* anywhere within search term to broaden search
 - EX) (?!)qu.*be.* will return such hits as Quebec, Québec, québécois

CACDA: Preparing Data for SketchEngine

- Data need to be in .txt format
- Encode text using Unicode 5.1 – UTF8

CACDA: Collocation

- The above-chance co-occurrence of two or more words within a predetermined span – e.g., 5 words on either side of a ‘node’ or ‘key’ word.
- *logDice*: relative frequencies of XY in relation to X and Y
 - Min. value = 0 ; Max. value = 14