

**„So, like, at the beginning it was like, what’s new Buenos Aires?“**

Discourse markers in unscripted  
Canadian Podcasts

Nicole Marie Benker

Nicole.Benker@campus.lmu.de

# Overview

1. Definitions: Discourse markers and Canadian English
2. Material and method
3. Results
  1. Discourse marker inventory
  2. Functions of *like*, *well*, *so*
4. Comparison to other varieties
5. Outlook



# What are discourse markers?

- Highly debated, incredibly many definitions
- Inconsistent terminology (discourse connectives, discourse particles, pragmatic markers, semantic conjuncts etc. (Fraser 2009: 294))
- Terms sometimes used interchangeably and sometimes used to distinguish between different semantic or pragmatic properties (ibid.)

# What are discourse markers?

Working definition for today:

“[Discourse markers] are, as a group, difficult to place within a traditional word class. [They are elements] [...] which are **syntactically optional**, which **may occur at the beginning, middle, or end of a discourse unit** or **form a discourse unit of their own**, which have **little or no semantic meaning in themselves**, which are **multifunctional**, [and] which **occur in oral rather than written discourse** [...].”

(Müller 2005: 27)<sup>1</sup>

<sup>1</sup> Müller adapted this definition from Stenström and Andersen 1996 and Svartvik 1980:169.

# Canadian English (CE)

- North American variety of English with a lot of British influence (part of the Commonwealth)
- Though often regarded as a mixture of American English (AE) and British English (BE), it is a variety in its own right (Brinton & Fee 2001: 426)
- Rather homogenous variety with little to no diatopic variation (Brinton & Fee 2001: 423)

# Research questions

1. What is the discourse marker inventory of CE (or at least the data)?
2. How often and in which function do the most-studied discourse markers in the English language *like*, *well*, and *so* appear in the data?

# Material

- About 2h, comprising six 20min segments taken from three unscripted Toronto-based podcasts
- The podcasts: *I Hate It But I Love It* (IHIBILI), *The Villain Was Right* (VWR), and *Talk From Superheroes* (TFSH), conversational pop-culture podcasts
- Six speakers (two per podcast), early- to mid-thirties, native-speakers of CE
- 23202 words

# Method

## Discourse marker inventory

- Criteria based on Müller 2005:
  - Syntactically optional
  - Reduced semantic meaning
  - Semantically optional
- Classified in terms of part-of-speech membership

## Functional analysis of *like*, *well*, and *so*

- *Like* (adapted from D'Arcy 2017:14)
  - Extra-clausal: example, explanation
  - Intra-clausal: approximation, hesitation, focus
  - *be like*: quotative
- *So* (extra-clausal, based on Müller 2005: 71-86):
  - Explanation, opinion, topic change, result, turn-taking, summary
- *Well* (based on Aijmer 2013: 32-41 and Müller 2005: 107)
  - Disagreement, turn-taking, hesitation, explanation, quotative



# Example analysis

KA: **Yeah**<sup>particle</sup>.

JG: **I mean**<sup>clausal</sup> *it's almost*<sup>adverbial</sup> *unfair in certain ways*<sup>prepositional</sup> () *to compare this with 'Armageddon' because, yes*<sup>particle</sup>, *the subject matter is so*<sup>adverbial</sup> *similar, but it feels*<sup>discontinuous</sup> *kind of*<sup>nominal</sup> like<sup>focus</sup> *to me* *like*<sup>discontinuous</sup>  
(clausal) D- *'Deep Impact' is like*<sup>1</sup> *a very*<sup>adverb</sup> *passably good* () *salmon dinner? Like*<sup>explanation</sup> *very*<sup>adverbial</sup> *sort of*<sup>nominal</sup>  
*well-rounded [salmon dinner]*

KA: [**sure**<sup>adjectival</sup>]

JG: **And then**<sup>conjunction</sup> *'Armageddon' is like*<sup>1</sup> *a version of the Ludovico treatment instead of like*<sup>example</sup> *strapping your eyes open and making you watch scary things, it's just*<sup>adverbial</sup> *Michael Bay pouring Pop Rocks in your* *mouth.*

(IHIBILI 180)

<sup>1</sup>*Like* is a comparative particle in this instance and will thus not be classified as a discourse marker.

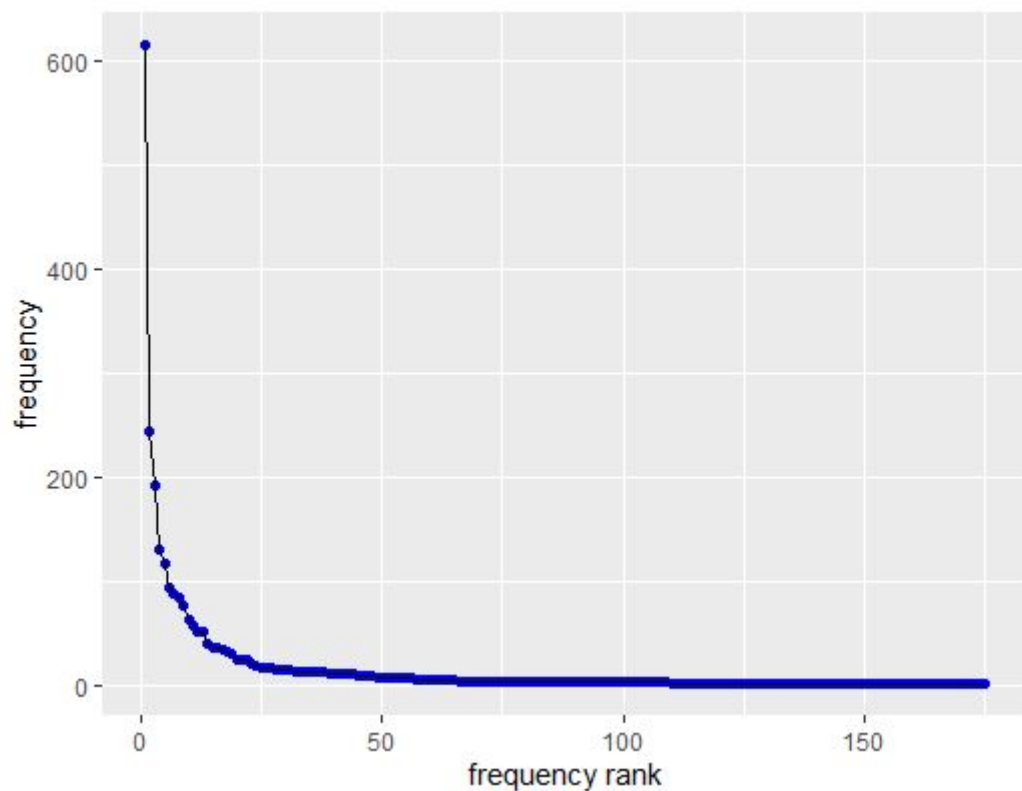
# Results: Discourse marker inventory

- 2800 tokens in total, 2019 excluding *like*, *well*, and *so*
- 12.1/100 words
- 183 types (unique discourse markers)

# Results: Discourse marker inventory

formal category	tokens	tokens/100 words	types
adverbial	806	3.5	66
particle	475	2.0	11
clausal	340	1.5	50
interjection	127	0.5	11
(pro-)nominal	103	0.4	18
adjectival	65	0.3	9
conjunction	52	0.2	7
prepositional	27	0.1	13
verbal	24	0.1	6
total	2019	8.7	183

# Results: Discourse marker inventory

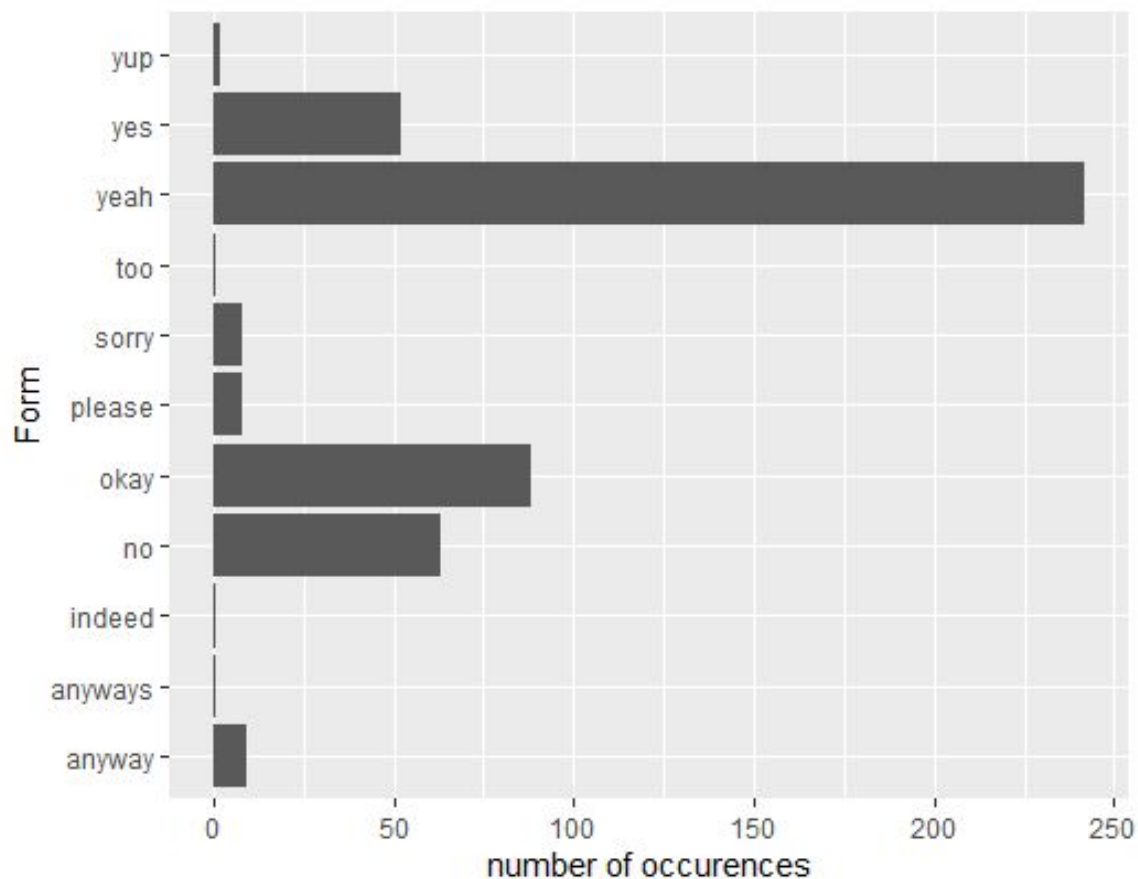


- Most frequent discourse markers:
  - *like* (616 tokens)
  - *yeah* (244 tokens)
  - *just* (192 tokens)
  - *so* (only extra-clausal uses, 130 tokens)
  - *I think* (117 tokens)
- Zipfian-like distribution

# Results: Adverbial discourse markers

- Largest group of discourse markers (806 tokens)
- Also category with the most types (66)
- Most frequent: *just* (192 tokens), *really* (84 tokens), intensifier-*so* (76 tokens), and *very* (57 tokens)
- Discourse functions:
  - Intensifiers (*very, really, literally, especially, etc.*)
  - Mitigators (*basically, apparently, supposedly, particularly, etc.*)

# Results: Particle discourse markers



- Second largest group token-wise (475 tokens)
- Very few high-frequency types (11)
- Most frequent DMs
  - *yeah* (244 tokens)
  - *okay* (88 tokens)
  - *no* (63 tokens)
  - *yes* (52 tokens)
- Discourse functions:
  - agreement/disagreement
  - “checking in”

# Results: Clausal discourse markers

- Third largest group
- 340 tokens, 48 types
- Multi-word discourse markers
- Most frequent markers:
  - *I think* (117 tokens)
  - *I mean* (31 tokens)
  - *I feel* (28 tokens)
  - *I guess* (25 tokens)
  - *you know* (25 tokens)
- Discourse functions: overall heterogenous
  - Expressing an opinion
  - Turn-yielding

# Results: Functions of *like*, *well*, and *so*

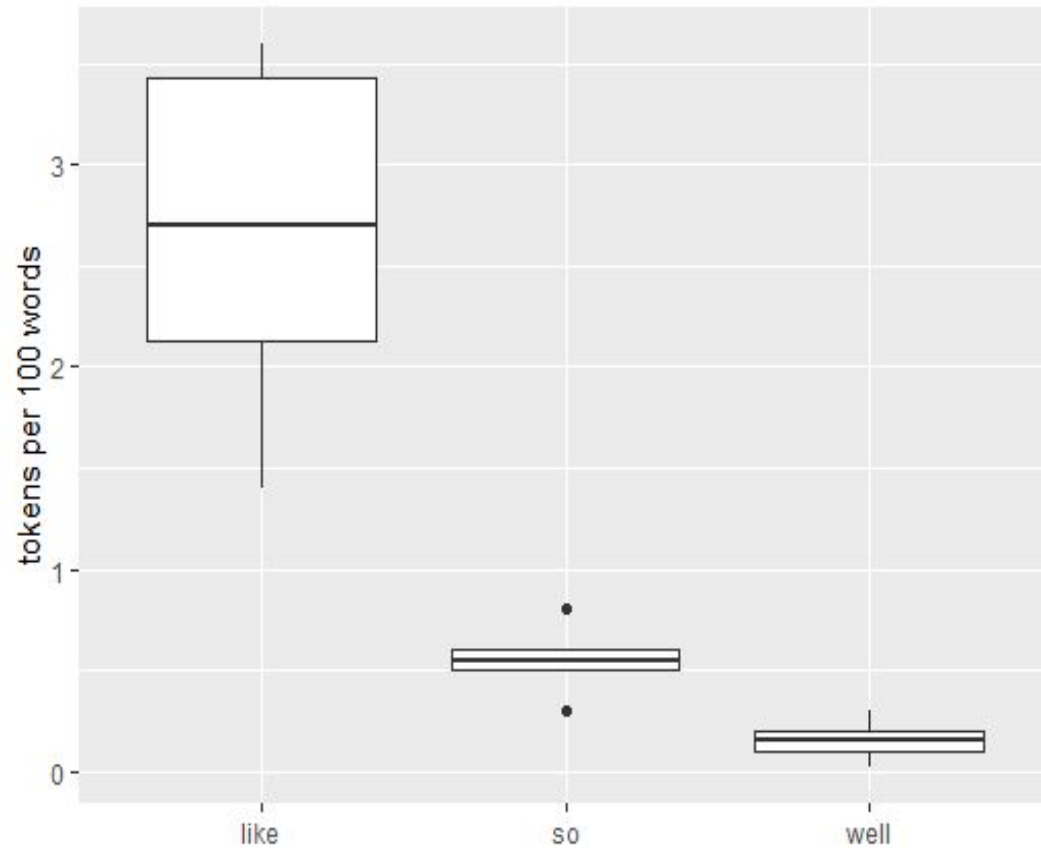
Speaker	<i>like</i>	<i>well</i>	<i>so</i>	all <sup>1</sup>
AI	1.4	0.0	0.5	10.2
CF	2.0	0.1	0.8	12.8
DM	3.6	0.1	0.6	11.8
JG	2.5	0.3	0.3	11.2
KA	3.6	0.2	0.6	12.3
RR	2.9	0.2	0.5	14.5
total	2.7	0.2	0.6	12.1

in absolute numbers: *like* (616 tokens), *so* (130 tokens), *well* (37 tokens)

<sup>1</sup> all items in the data classified as discourse markers

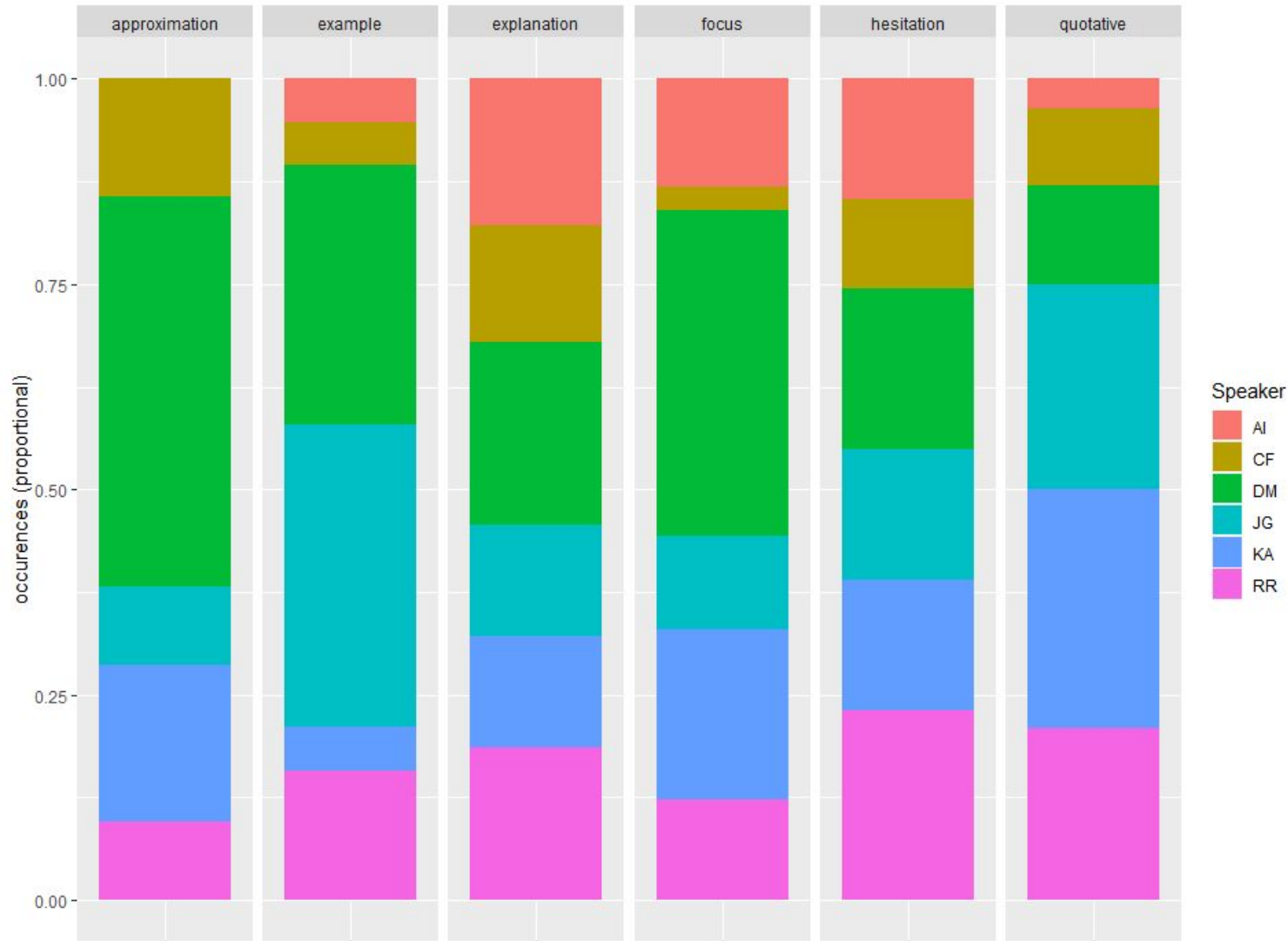


# Results: Functions of *like*, *well*, and *so*



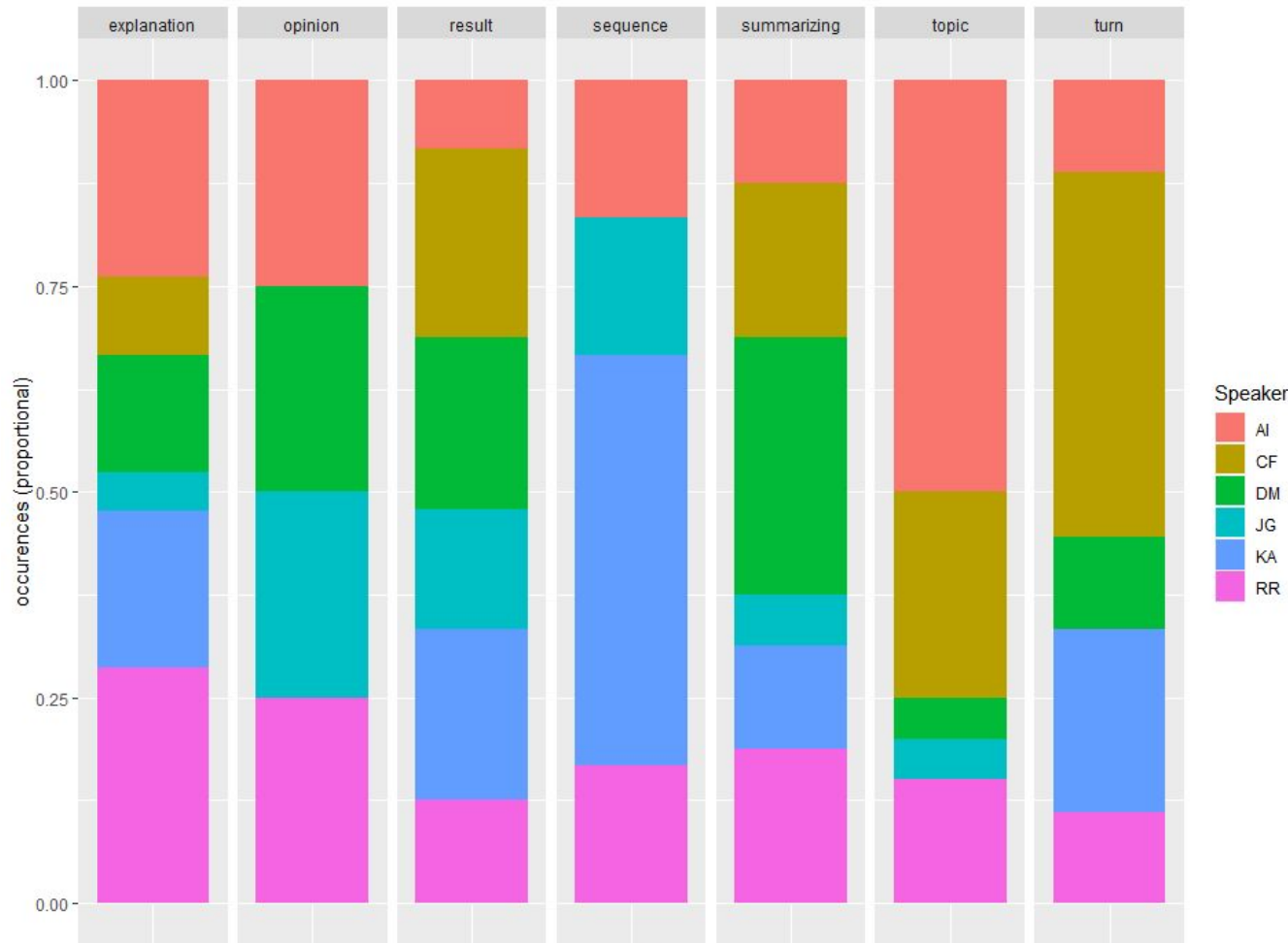
- Frequency of *like* is speaker-dependent
- *So* and *well* are both rare and not as variable

# Results: Functions of *like*



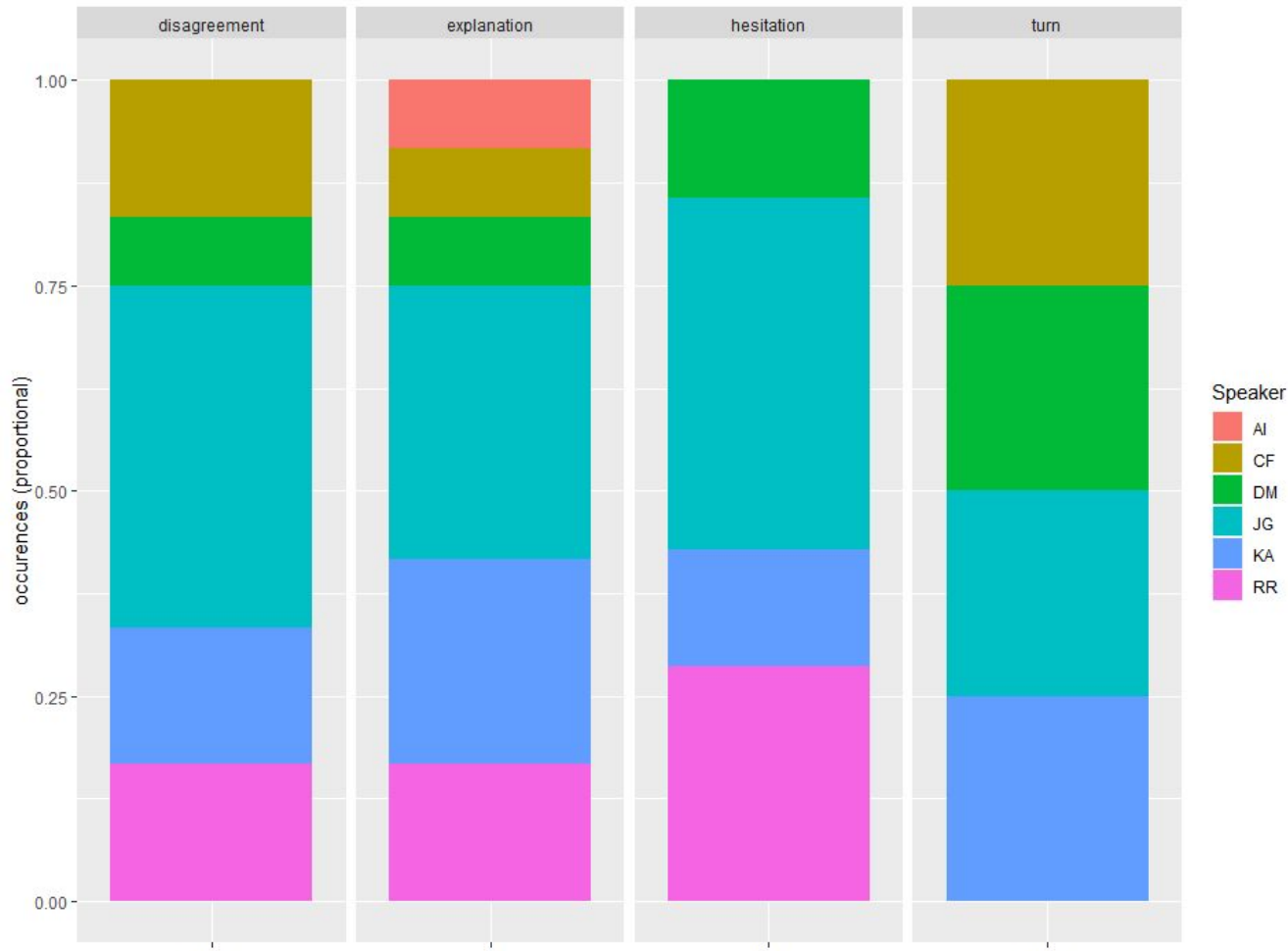
- 616 tokens overall
- Most frequent functions:
  - quotative (*be like*) (248 tokens)
  - explanation (140 tokens)
  - focus (106 tokens)

# Results: Functions of *so*



- 130 tokens overall
- Most frequent functions:
  - resultative (48 tokens)
  - explanation (21 tokens)
  - change of topic (20 tokens)
- A lot of inter-speaker variation (however, small sample size!)

# Results: Functions of *well*



- Hardly occurs in the data at all (37 tokens)
- Most frequent use is quotative (16 tokens)

# Comparison to other varieties of English

- No analyses of discourse marker inventories using the same or similar definition used in this paper
- *Well, so, like* in CE compared to AE (Müller 2005: 244) and BE (Beeching 2015: 181)

	<i>well</i>	<i>so</i>	<i>like</i> <sup>3</sup>
Müller 2005: 244 <sup>1</sup>	0.14	1.20	1.00
Beeching 2015: 181 <sup>2</sup>	0.44		1.02
This study	0.16	0.56	2.65
		excluding quotative- <i>like</i>	1.59

# Discussion and outlook

- Discourse markers are very diverse class
- Existence of clausal discourse markers means that DMs should not be considered a formal class (i.e., a part-of-speech), but a functional class (analogous to adverbials)
- *well* does not play a huge role in CE
- Comparison to AE and BE: CE separate variety

# List of references

- Aijmer Karin. 2013. *Understanding pragmatic markers: A variational pragmatic approach*. Edinburgh: Edinburgh University Press.
- Beeching, Kate. 2015. "Variability in native and non-native use of pragmatic markers: The example of well in role-play data". *Researching sociopragmatic variability: Perspectives from variational, interlanguage and contrastive pragmatics*. London: Palgrave Macmillan. 174–197.
- Brinton, Laurel J., and Margery Fee. 2001. "Canadian English". *The Cambridge history of the English language* vol 6. Ed. John Algeo. Cambridge: Cambridge University Press. 422–440.
- Fraser, Bruce. 2009. "An account of discourse markers". *International review of pragmatics* 1: 293–320.
- Müller, Simone. 2005. *Discourse markers in native and non-native English discourse*. Amsterdam: John Benjamins.
- Stenström Anna-Brita & Gisle Andersen. 1996. "More trends in teenage talk: A corpus-based investigation of the discourse items cos and innit". *Synchronic corpus linguistics*. Ed. Carol E. Percy, Charles F. Meyer and Ian Lancaster. Amsterdam: Rodopi.
- Svartvik, Jan. 1980. "Well in conversation". *Studies in English linguistics for Randolph Quirk*. Ed. Sydney Greenbaum, Geoffrey Leech and Jan Svartvik. London: Longman. 167–177.
- Cover image: [https://upload.wikimedia.org/wikipedia/commons/b/b5/Toronto\\_August\\_2017.jpg](https://upload.wikimedia.org/wikipedia/commons/b/b5/Toronto_August_2017.jpg)