Subjectivity in the Annotation of Bridging Anaphora

Lauren Levine and Amir Zeldes
July 31, 2025

GEORGETOWN UNIVERSITY

Talk Outline

- **Introduction to bridging anaphora**
  - Current landscape of bridging resources in English

- Intro to GUMBridge, a new resource for bridging
  - Annotation Schema & Subtypes

- GUMBridge Annotation Pilot & Results
  - Agreement Study
  - Analysis of Subjectivity in the Annotation Process

What is Bridging?

- Non-identity anaphoric phenomenon

- Anaphor is in some way dependent on an associative, non-identical antecedent:

There is <u>a big yellow house</u> on the hill. **The door** is red.

associative antecedent   bridging anaphor

Bridging Varieties

There are many different kinds of associative relationships that can manifest in a discourse…

Comparative-markers

A woman → Another woman

Meronomy

A house → The door

Sense anaphora

A Chinese restaurant → The Italian one

Implicit arguments

A murder → The victim

Prototypical associations

A wedding→ The ceremony

Relational nouns

A child → The parent

Bridging Resource Landscape

- Various bridging resources, each with their own definition for bridging and annotation scheme

- Fractured landscape of resources, not easy to compare

- Difficult to establish/standardize benchmark for bridging resolution
  - Task of identifying all instances of bridging anaphora and their corresponding associative antecedents

Problems with Existing English Reference Corpora

ARRAU (Uryupina et al., 2019)
- Relying on the concept of "cohesion" & identifying instances from predefined semantic relations can lead to false positives in annotations

ISNotes (Markert et al., 2012) & BASHI (Rösiger, 2018)
- Identifies bridging instances based on information status
- Lacks granular bridging subtype annotation

All lack genre diversity
- Wall Street Journal (WSJ) news data (>30 years ago)

GUMBridge

- Goal: Reconcile the formalisms bridging resources for English with new bridging resource
  - Broad reach in its scope of bridging phenomena
  - Various levels of granularity annotated

- Construction:
  - Built on top of GUM (Zeldes, 2017)
    - Existing multi-genre corpus of English
  - IS approach to identifying bridging (as in ISNotes & BASHI)
  - Apply schema of bridging varieties for sub-categorization (as in ARRAU)

# Identifying Instances of Bridging

- Information status (IS) informed perspective

- **IS**: Extent an entity is accessible to reader/hearer of a discourse (Nissim et al., 2004)

- **Accessible**: Entity has not been mentioned, but its referent is inferable to a reader/hearer

- **Bridging**: Entity **Accessible** via inference from previous non-identical entity

Identifying Instances of Bridging

When evaluating an entity as a candidate bridging anaphor:

- Do you judge this entity to be to **Accessible** (inferable) in the discourse?

- Does that accessibility **rely on the understanding of a previous entity** in the discourse?
  - If so, it is a **bridging anaphor** and that previous entity is the **associative antecedent**.

Bridging Subtype Schema
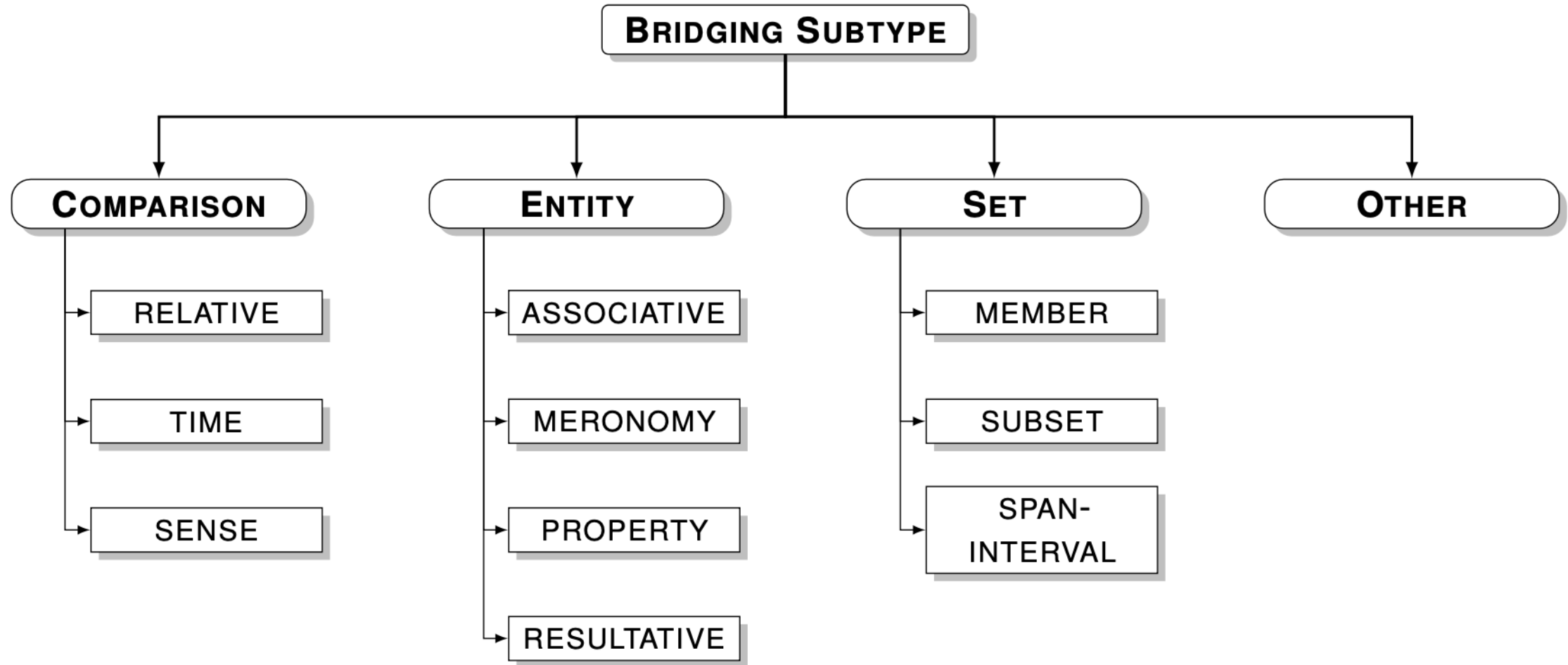
Relation Types of Bridging Subtypes

Comparison: The anaphor is interpretable via a comparison to the antecedent.

- <u>Several women</u> walked into the room. **Other women** soon followed.

Entity: The anaphor is an attribute or associated entity of the antecedent.

- There was <u>a murder</u> last night. **The victim** has yet to be identified.

Set: There is a set relation between the bridging anaphor and antecedent.

- The dealership sells <u>several cars</u>. **The sedan** is the most popular.

Annotation Pilot

- Conducted on the test set of GUM V10
  - 26 document (26k tokens) across 16 genres

- All documents double annotated
  - Annotator A: Me!
  - Annotator B: Linguistics graduate students

- Provided concise annotation guidelines (5 pages)
  - Identifying instances of bridging (anaphor-antecedent pairs)
  - Classifying bridging subtype

Annotator Agreement Results

	Precision	Recall	F1 Score
Anaphor Recognition	0.44	0.34	0.38
Anaphor+Antecedent Recognition	0.32	0.25	0.28
			Accuracy
Antecedent Resolution			0.72
			Cohen's κ
Bridging Subtype			0.58

Stages of annotation

GEORGETOWN UNIVERSITY

Test Set Adjudication

- Manual adjudication of annotations
  - Produce a single set of bridging annotations for the test set (V0.1)

- Adjudicated test set available online:
  - https://github.com/lauren-lizzy-levine/GUMbridge

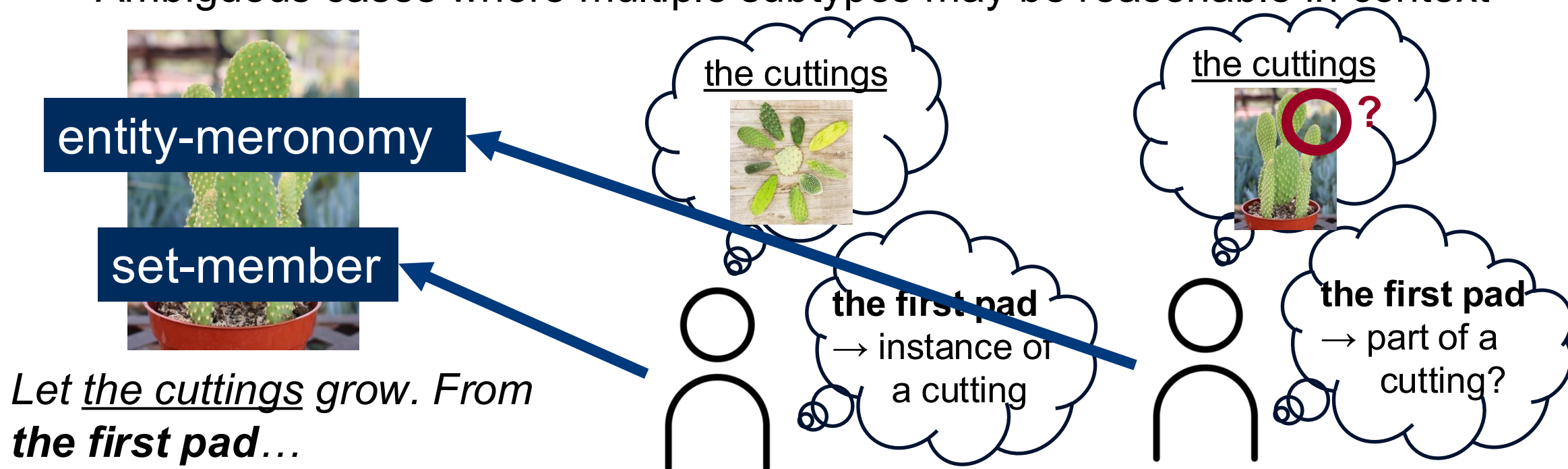| COMPARISON	
RELATIVE	59
TIME	27
SENSE	45
Subtotal	131
ENTITY	
ASSOCIATIVE	124
MERONOMY	37
PROPERTY	9
RESULTATIVE	21
Subtotal	191
SET	
MEMBER	31
SUBSET	14
SPAN-INTERVAL	18
Subtotal	63
OTHER	16
Total	401

Subjectivity in Annotator Judgement

There are 3 stages in the annotation process where annotators must make subjective judgments:

1) Recognition of the bridging anaphor

2) Resolving the anaphor back to its associated antecedent

3) Identifying the subtype category of the bridging pair

Subjectivity in Subtype Selection

- Annotators may have different conceptions of individual entities
  - A lack of familiarity with the relations between entities
  - Uncertainty in interpreting the anaphor or antecedent in context
  - Ambiguous cases where multiple subtypes may be reasonable in context

entity-meronomy

set-member

Let the cuttings grow. From **the first pad…**

the cuttings

the first pad
→ instance of a cutting

the cuttings

?

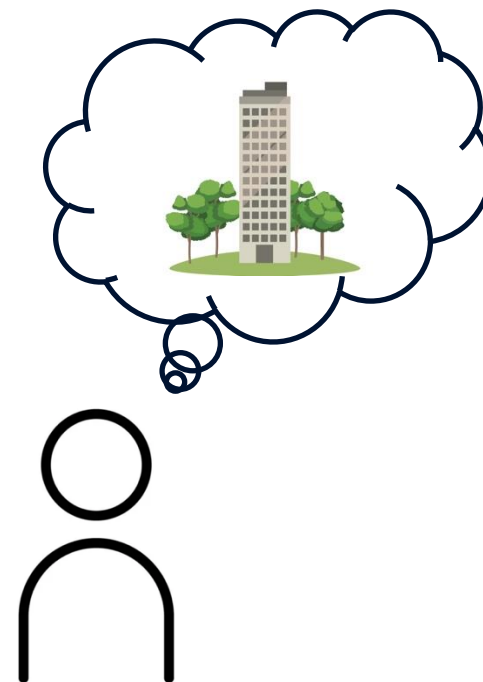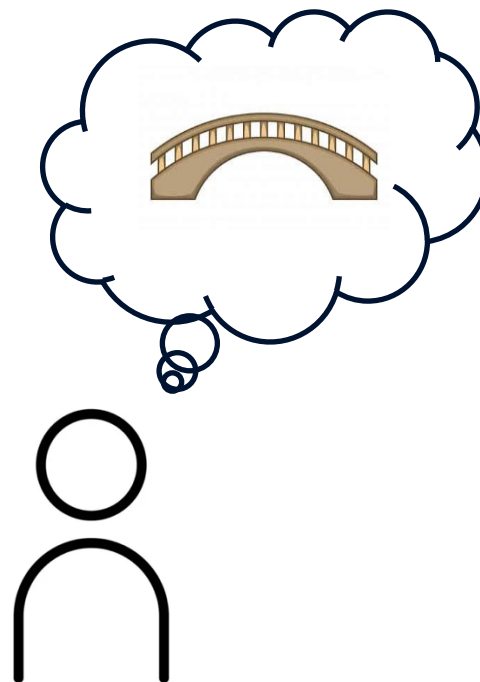**the first pad**
→ part of a cutting?

Subjectivity in Antecedent Resolution

- Multiple preceding entities which could be reasonable candidates for a bridging antecedent
- Disagreement on the meaning of the anaphor in context

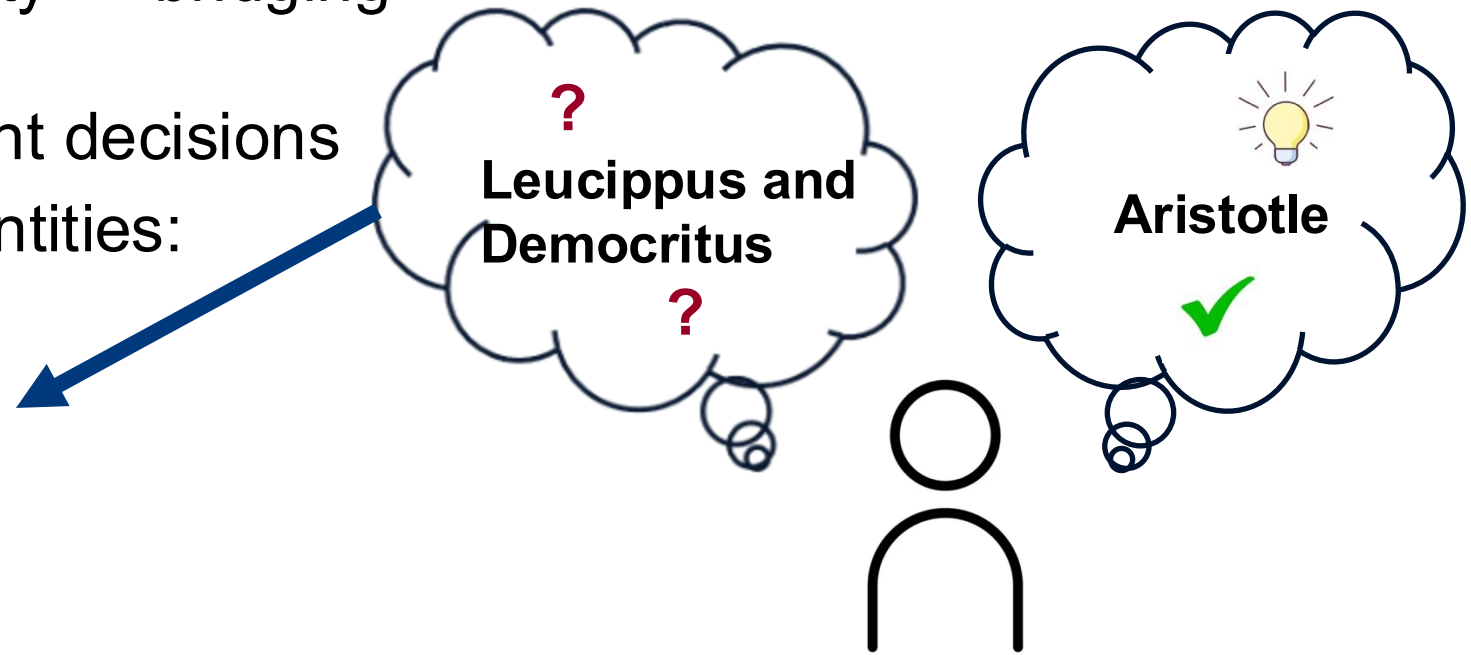*…leaning over **the edge** into…*

The edge of *what*?

Subjectivity in Anaphor Recognition

- Annotators must make subjective judgments on whether an entity is accessible:
    - Due to generics/world knowledge → **not** bridging
    - Due to an antecedent entity → bridging

- Annotators may make different decisions based on the "familiarity" of entities:

…ancient Greek philosophers…

Takeaways

- Annotator subjectivity can affect bridging annotation at 3 different levels: (1) anaphor recognition, (2) antecedent resolution, (3) subtype selection.

- We found **moderate agreement** on **antecedent resolution** and **subtype selection**, but **lower alignment** on the **recognition of bridging anaphora.**

- While subjectivity cannot be eliminated, additional guidance on the **preferred default selection** when multiple options are available will likely **improve consistency.**

References

Herbert H. Clark. Bridging. In B.L. Nash-Webber and Roger Schank, editors, Theoretical Issues in Natural Language Processing, 1975. URL https://aclanthology.org/T75-2034.

Katja Markert, Yufang Hou, and Michael Strube. Collective classification for fine-grained information status. In Haizhou Li, Chin-Ye w Lin, Miles Osborne, Gary Geunbae Lee, and Jong C. Park, editors, Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 795–804, Jeju Island, Korea, July 2012. Association for Computational Linguistics. URL https://aclanthology.org/P12-1084. (ISNotes)

Malvina Nissim. Learning information status of discourse entities. In Dan Jurafsky and Eric Gaussier, editors, Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, pages 94–102, Sydney, Australia, July 2006. Association for Computational Linguistics. URL https://aclanthology.org/W06-1612.

Ina Rösiger. BASHI: A corpus of Wall Street Journal articles annotated with bridging links. In Nicoletta Calzolari, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Koiti Hasida, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asuncion Moreno, Jan Odijk, Stelios Piperidis, and Takenobu Tokunaga, editors, *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan, May 2018. European Language Resources Association (ELRA). URL https://aclanthology.org/L18-1058.

Olga Uryupina, Ron Artstein, Antonella Bristot, Federica Cavicchio, Francesca Delogu, Kepa Joseba Rodríguez, and Massimo Poesio. Annotating a broad range of anaphoric phenomena, in a variety of genres: the arrau corpus. Natural Language Engineering, 26:95 – 128, 2019. URL https://api.semanticscholar.org/ CorpusID:164858637. (ARRAU)

Amir Zeldes. The gum corpus: Creating multilayer resources in the classroom. *Language Resources and Evaluation*, 51(3):581–612, 2017.

Questions?