



# Sense Anaphoric Pronouns: *Am I One?*

Marta Recasens | [recasens@google.com](mailto:recasens@google.com)

Zhichao Hu | [zhu@soe.ucsc.edu](mailto:zhu@soe.ucsc.edu)

Olivia Rhinehart | [orhinehart@google.com](mailto:orhinehart@google.com)

# Outline

1. Sense Anaphora
2. Annotation Task and Data
3. Anaphoric Classification
4. Evaluation

# Outline

1. **Sense Anaphora**
2. Annotation Task and Data
3. Anaphoric Classification
4. Evaluation

# Sense Anaphora

1. If you have **multiple loans**, you can consolidate **them** into a single loan.  
-> *them* anaphorically depends on and corefers with *multiple loans*
2. Consumers and companies demand fewer **loans** and struggle to pay back existing **ones**.  
-> *ones* anaphorically depends on *loans*, but does not corefer with it.
3. Although many **investors** were aware that a bond offering was being scheduled, **many** expected a much smaller amount of bonds to be sold.

(Grinder and Postal, 1971; Bresnan, 1971; Sag, 1976; Dalrymple et al., 1991)

# Candidate Anaphors

We focus not only on **one** but on other expressions that can similarly borrow their sense from a contextual expression:

- **One** (*one, ones*)
- **Quantifiers**
  - **Indefinites** (*all, any, few, many, more, most, much, some*)
  - **Numerals** (*two, three, ..., hundred, etc.*)
  - **Measure nouns** (*bits(s), bunch, couple, dozen(s), lot(s), pair, plenty, ton(s)*)
- **Demonstratives** (*that, those*)
- **Possessives** (*mine, yours, his, hers, ours, theirs*)
- **Other** (*other(s), another*)

# Classes of Sense Anaphora

## Partitive class

Subset relationship between the anaphor and antecedent

1. The blast killed two cameramen, one from Spanish TV, another from Reuters.

## Instantiator class

New instance created from the same sense as the antecedent

2. In both quantity and quality, the English teaching materials of today leave those of before in the dust.

(Dahl, 1985; Luperfoy, 1991; Gardiner, 2003; Payne et al., 2013)

# Outline

1. Sense Anaphora
2. **Annotation Task and Data**
3. Anaphoric Classification
4. Evaluation

# Annotation Basics

Custom GUI, strings from the target anaphor categories automatically highlighted in each passage

Annotation workflow:

1. Determine if the highlighted tokens were anaphors
2. If so, identify corresponding antecedent
3. Choose boundaries of antecedent span



# Annotation Tricky Cases

Sharing antecedents:

1. Added to this is the perennial problem of class sizes being too large, and not enough English classes scheduled – only one or two a week.

No explicit antecedent:

2. It must be advanced with a plan, the easy ones first and the tough ones last.

Non-anaphoric numerals:

3. The chamber voted 231 - 191, 51 votes short of the two-thirds majority needed to overturn the president's veto of legislation.

# SAAnaNotes Data

- **1163** English documents from OntoNotes (Weischedel et al., 2011)
- Average number of sense anaphors per document is **0.6**
  - Most common annotation error was anaphor omission
- Single-annotation on **1138** documents
- Four-way annotation on a set of **25** documents
- Used Fleiss' kappa to measure agreement on anaphor identification
  - $\kappa = .67$
- Pairwise agreement on antecedent spans of common anaphors was **63%**.
- Data released at: <https://github.com/dmorr-google/sense-anaphora>

# Outline

1. Sense Anaphora
2. Annotation Task and Data
3. **Anaphoric Classification**
4. Evaluation

# Anaphoric Classifier

- **Anaphoric**

Consumers demand fewer loans and struggle to pay back existing **ones**.

vs.

- **Generic**

While **some** think that the estimate may be inflated, the consensus is that...

- **Numeral**

Those corrections lasted **one** to **four** weeks.

- **Deictic**

I figured there was something like **that**.

- **Idioms**

They discovered that although the trees were uninjured, they died **one** by **one**.

# Anaphoric Classifier

- Target
  - **One** (*one, ones*)
  - **Quantifiers**
    - **Indefinites** (*all, any, few, many, more, most, much, some*)
    - **Numerals** (*two, three, ..., hundred, etc.*)
    - **Measure nouns** (~~*bits(s), bunch, couple, dozen(s), lot(s), pair, plenty, ton(s)*~~)
  - **Demonstratives** (*that, those*)
  - ~~**Possessives** (*mine, yours, his, hers, ours, theirs*)~~
  - **Other** (*other(s), another*)

# Previous Work: *one*

## Anaphoric classification

	<b>Gardiner (2003)</b>	<b>Ng et al. (2005)</b>
System	Rule-based	C4.5 decision tree
Features	5 heuristics	7 features
Classes	anaphoric numeric partitive generic	anaphoric numeric partitive generic idiomatic unclassifiable
Test set	773 BNC examples (71.5% anaphoric)	1577 BNC examples (15.2% anaphoric)

# Our System: beyond one

- SVM (LIBLINEAR)
- 31 lexical and syntactic features:

[...] *the only way to lose weight is cardiovascular* [...] *And that lifting the weights won't lose weight* [...] *I mean you might lose **some** but that's only* [...]

## ANAPHORIC

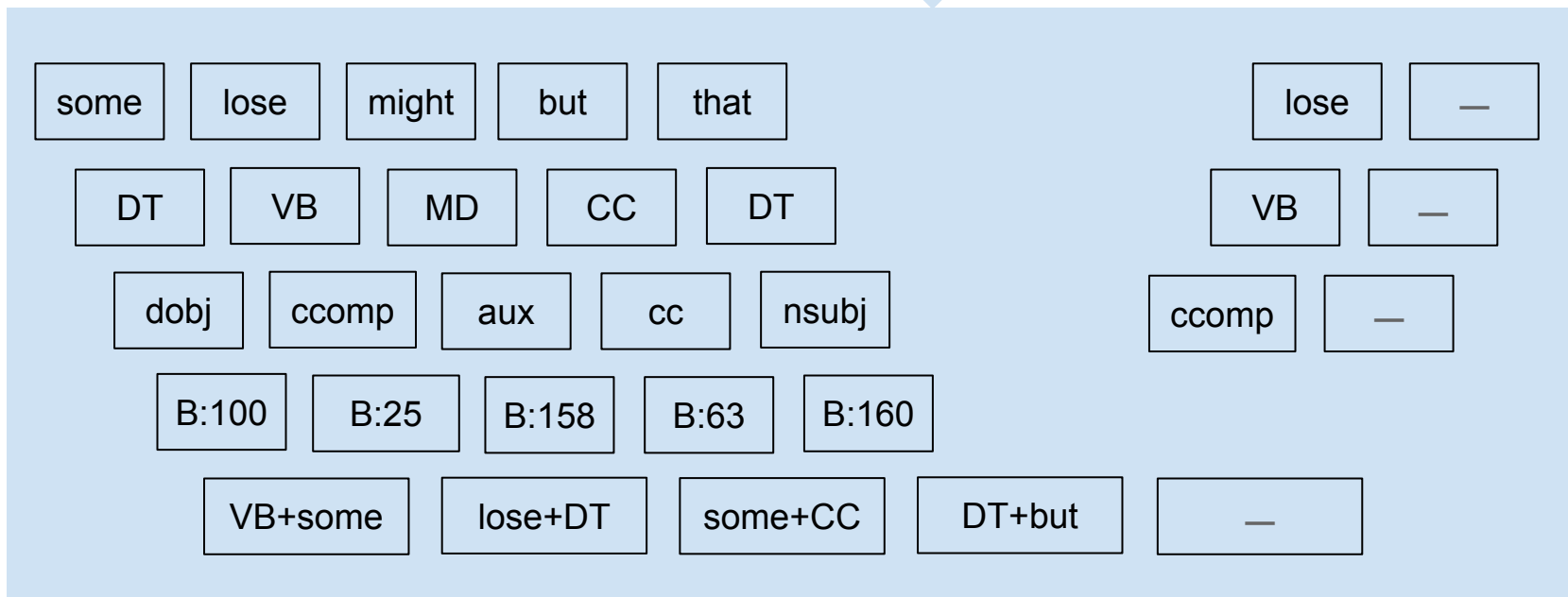
*one* + animate verb  
*the* + ADJ + *one*  
verb + *some*  
*that* + *of*  
*those* + relative clause  
...

## NON-ANAPHORIC

... + noun  
*some* + *think*  
*some* + verb  
*that* + verb  
verb + *those*  
...

# Our System: beyond one

[...] the only way to lose weight is cardiovascular [...] And that lifting the weights won't lose weight [...] I mean you might lose **some** but that's only [...]





# Outline

1. Sense Anaphora
2. Annotation Task and Data
3. Anaphoric Classification
- 4. Evaluation**

## Results: *one(s)*

<b>System</b>	<b>P</b>	<b>R</b>	<b>F1</b>
Gardiner (2003)	40.00	94.44	56.20
Ng et al. (2005)	62.50	55.60	58.80
Our system	74.42	88.89	79.01

## Results: all anaphors

Anaphor class	P	R	F1
Ng et al. (2005)	45.60	23.90	31.30
ALL	61.02	62.61	61.80
One	71.11	88.89	79.01
Indefinites	38.46	38.46	38.46
Numerals	27.78	31.25	29.41
Demonstratives	87.50	53.85	66.67
Other	61.54	66.67	64.00

# Error Analysis

- **Precision errors**

They were able to whittle down the number of missing aircrafts uh to a **few**.

Two **others**, her father apparently and another sister, have been wounded in this incident.

- **Recall errors**

Today's Tanshui residents are living their own stories [...] **Some** are active in the morning, **some** late at night.

This was an **October massacre** like **those** that occurred in 1978 and 1979.

# Conclusions and Future Work

- Tail phenomenon: identity-of-sense anaphora
- Beyond *one*
- Automatic classification with surface features at 62% F1
- SAnaNotes corpus available from GitHub  
<https://github.com/dmorr-google/sense-anaphora>
- Antecedent resolution jointly with anaphor identification
- Discourse features: parallelism, discourse relations, etc.