



Aims

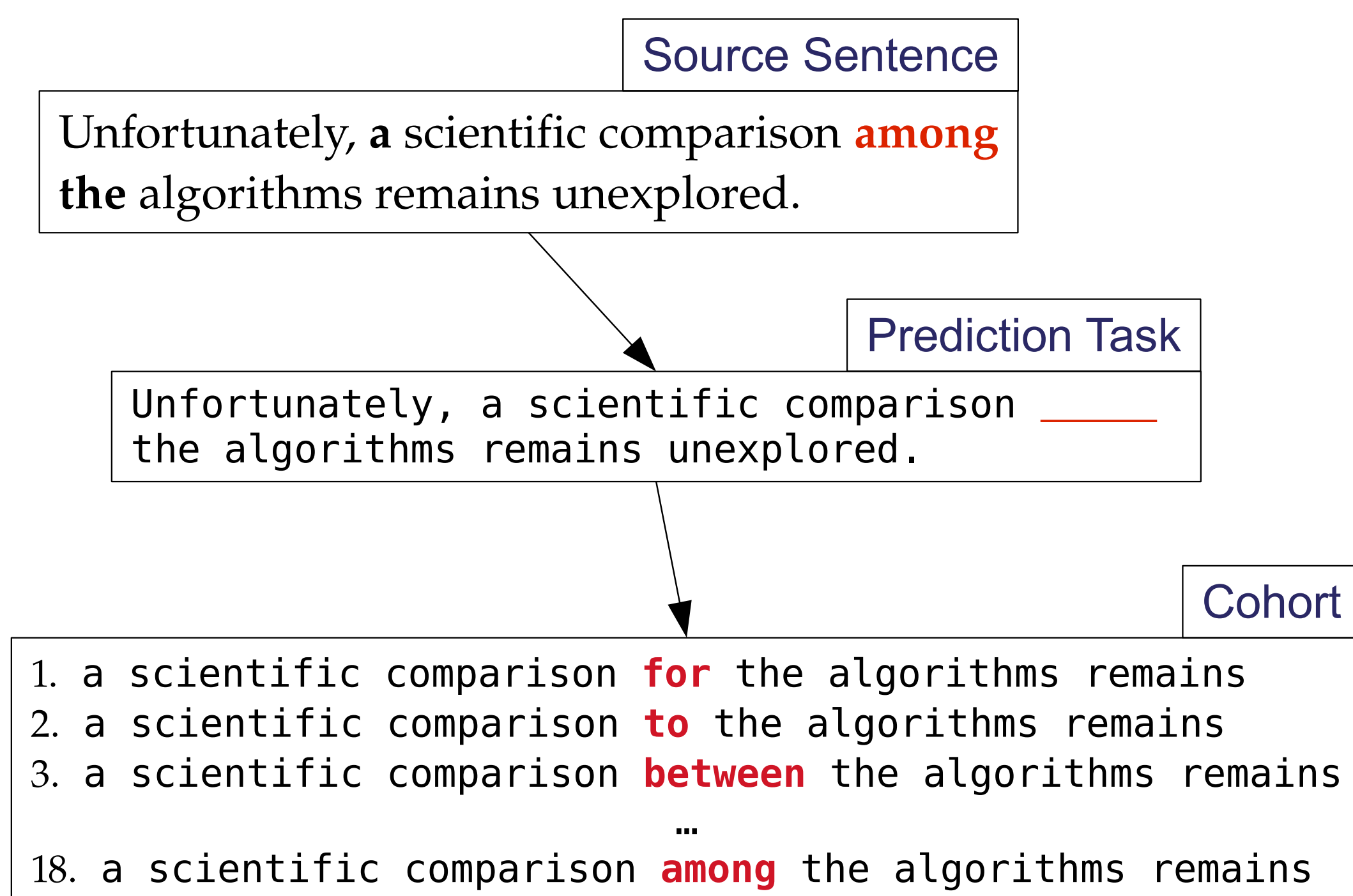
- Detect and correct **function word substitution errors** in non-native English text using the the n-gram prediction approach from Elghafari, Meurers & Wunsch (2010)
- Focus on substitution errors for prepositions, determiners, conjunctions, and quantifiers
→ **12% of the errors** annotated in the HOO training data
- Determine the tradeoff between
 - Informativeness of context (→ the size of the context n-gram)
 - Data-sparseness (→ can we find the n-gram in a reference corpus?)

Our Approach

Using the **immediate distributional context** of a function word, how accurately can we detect errors and suggest corrections?

We extract one **prediction task** for each occurrence of the candidate function words.

Prediction Tasks



Prediction Algorithm

- Determine the **number of occurrences** for each 7-gram in a cohort in the genre-specific ACL Anthology Reference Corpus (Bird et al., 2008).
 - Limited POS and lemma substitutions abstract away from a purely surface-based context
- Pick the word for which the **most hits** were found.
- In case **no hits** are found, back off to **shorter n-grams**.
- If no hits are found for the **minimum n-gram length**, predict the original word.

Data

Function Word Errors in the HOO Test Data

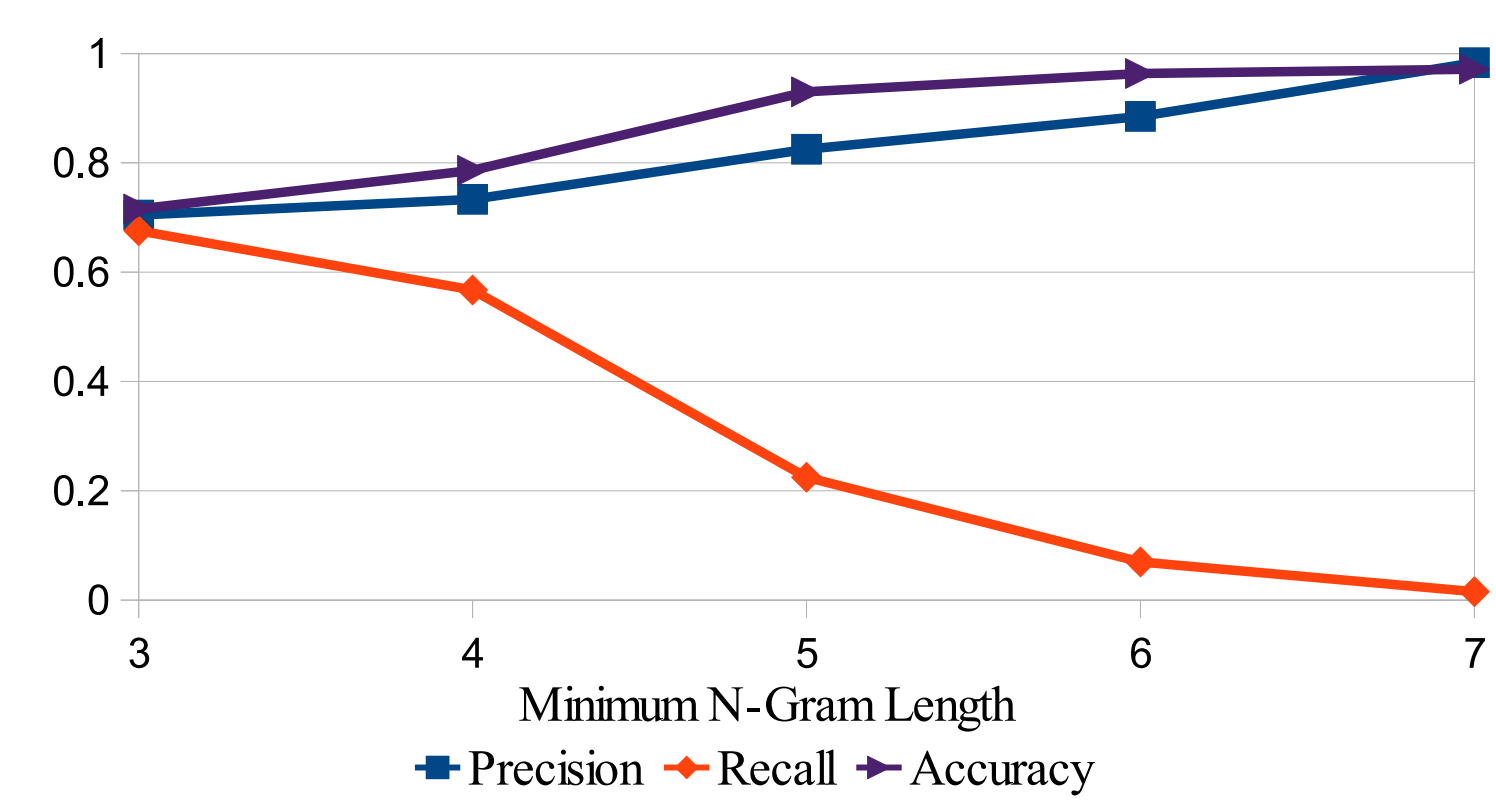
Category	Error Codes	# Errors	Candidates	# Occurrences
Conj.	RC	2	but, if, whether, whereas, however, although	80
Det.	RD, FD, DD, AGD, CD, ID	17	a, whose, their, this, an, these, the, its, those	1572
Prep.	RT, DT	86	in, on, about, over, from, onto, for, among, of, into, within, to, as, at, under, between, with, by	2126
Quant.	RQ, FQ, CQ, DQ, IQ, AGQ	4	less, many, some, fewer, much, certain	78
Total		109		3856

- Only 2.8% of the targeted function words are in error
→ **97.2% baseline**

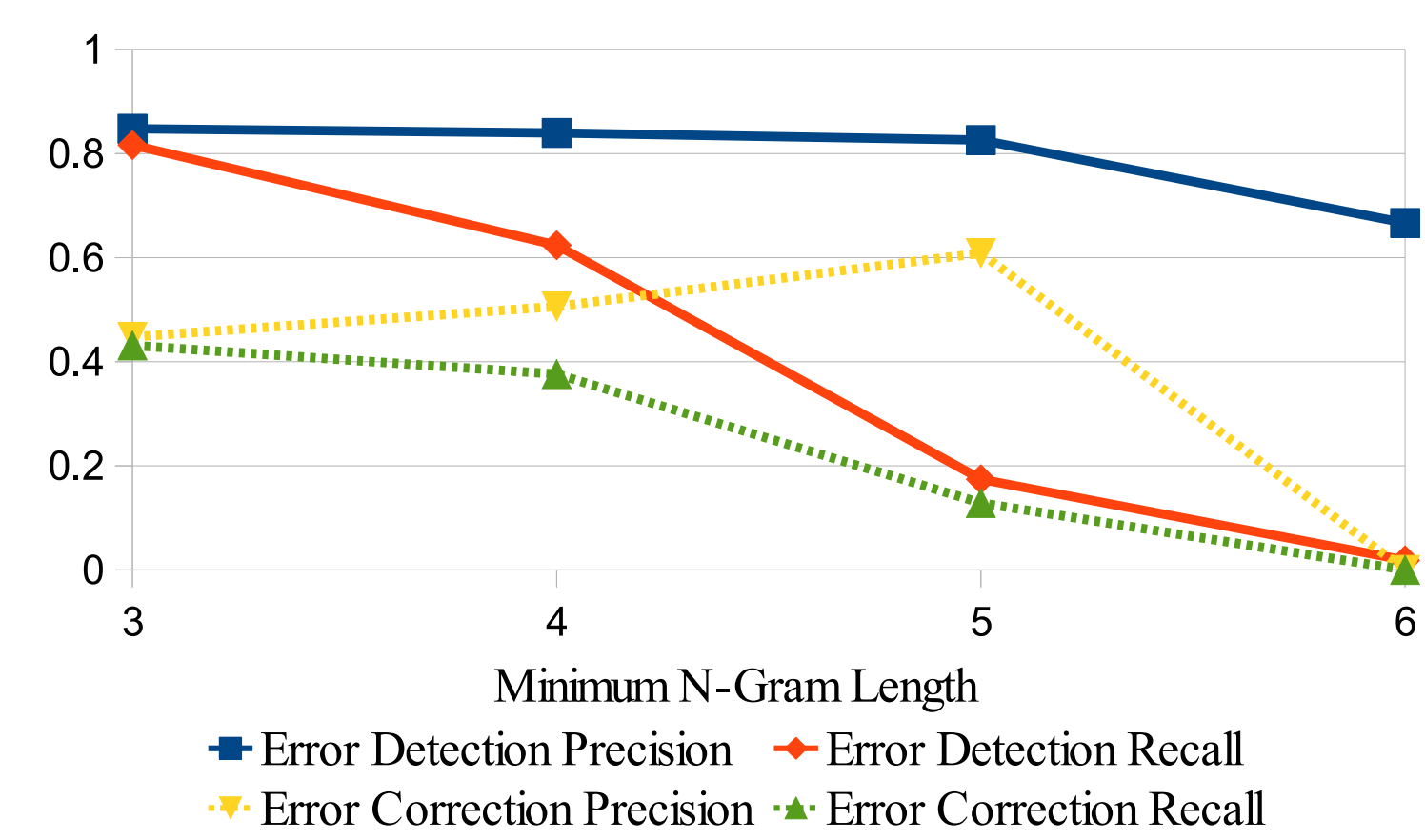
Results

Evaluation of N-Gram Prediction Approach

- **Global:** For each function word (correct or incorrect), was a correct prediction made?



- **Error detection and correction:** For each function word substitution error, was the error detected/corrected?



HOO Challenge Scores

Our best-performing submission (#2) in terms of overall detection f-score:

	Detection	Recognition	Correction
F-Score	0.126	0.080	0.039

Detection recall for targeted error types:

- 67% of preposition and determiner substitution errors
- 40% of conjunction substitution errors
- 33% of quantifier substitution errors

Accurate corrections provided for ~50% of detected errors

Discussion

- Our approach currently detects **33%–67%** of the targeted errors, but miscorrects ~10 words for each error detected.
- Current HOO annotation scheme lacks the **granularity** to identify all function word errors:
 - Our approach detects many other error types annotating multiple words (compound change, phrasal verb errors, adverbial errors).

Future Work

- Weight the words in the candidate sets to account for **global frequency**
- Error correction \neq word selection: add an **explicit bias** towards the original word
- Vary the size of the **context window** based on linguistic information
- Explore **backoff strategies** based on a greater degree of linguistic generalization

References

- Bergsma, S., D. Lin & R. Goebel (2009). Web-scale N-gram models for lexical disambiguation. In *IJCAI*.
- Bird, S., R. Dale et al. (2008). The ACL Anthology Reference Corpus. In *LREC*.
- De Felice, R. (2008). Automatic Error Detection in Non-native English. Ph.D. thesis, Oxford.
- Elghafari, A., D. Meurers & H. Wunsch (2010). Exploring the Data-Driven Prediction of Prepositions in English. In *COLING*.
- Gamon, M., J. Gao et al. (2008). Using Contextual Speller Techniques and Language Modeling for ESL Error Correction. In *IJCNLP*.
- Meurers, D. (2012). Natural Language Processing and Language Learning. In *Encyclopedia of Applied Linguistics*, Oxford: Wiley-Blackwell.
- Tetreault, J. & M. Chodorow (2008). Native Judgments of Non-Native Usage: Experiments in Preposition Error Detection. In *COLING*.