

# Computational Investigations of Pragmatic Effects in Natural Language

**Jad Kabbara**

School of Computer Science, McGill University, Montreal, QC, Canada  
Montreal Institute for Learning Algorithms (MILA), Montreal, QC, Canada  
jad@cs.mcgill.ca

## Abstract

Semantics and pragmatics are two complimentary and intertwined aspects of meaning in language. The former is concerned with the literal (context-free) meaning of words and sentences, the latter focuses on the intended meaning, one that is context-dependent. While NLP research has focused in the past mostly on semantics, the goal of this thesis is to develop computational models that leverage this pragmatic knowledge in language that is crucial to performing many NLP tasks correctly. In this proposal, we begin by reviewing the current progress in this thesis, namely, on the tasks of definiteness prediction and adverbial presupposition triggering. Then we discuss the proposed research for the remainder of the thesis which builds on this progress towards the goal of building better and more pragmatically-aware natural language generation and understanding systems.

## 1 Introduction

The past several years have seen growing trends in relying on intelligent systems to carry out day-to-day tasks. From interactions with your virtual personal assistant to carrying out a simple conversation with a chatbot to asking for directives or getting restaurant recommendations, the ability of said intelligent systems to properly understand its user is becoming increasingly important and crucial to their effective functioning.

For a proper *understanding*, these systems ought to rely on two different but complementary aspects of language: semantics and pragmatics. At a very high level, semantics is concerned with the literal meaning of words and sentences that is context-free, while pragmatics is concerned with the intended meaning, one that is context-dependent.

On the frontier of semantics, the past few years have witnessed a remarkable progress in language

modeling and other tasks. Research on neural vector representations (word embeddings) has particularly exploded starting with the Word2vec model (Mikolov et al., 2013) and GloVe (Pennington et al., 2014). These neural models learn high-quality vector representations of words which are empirically shown to have state-of-the-art performance on semantic similarity tasks. Many variations have been proposed to account for varying textual structures (words vs sentences vs paragraphs vs documents), multiple languages (Luong et al., 2015), varying topics (Liu et al., 2015), etc.

Pragmatic reasoning, on the other hand, is arguably one of the major milestones on the road to general AI simply because for a machine to fully understand individuals, it has to understand the nuances and subtleties of the language that is being conveyed and which often goes beyond the literal meaning of the utterances being made. And while correctly performing pragmatic reasoning is at the core of many NLP tasks such as information extraction, automatic summarization, and machine translation, there has been not much focus on it in NLP research – at least not as much as its semantic counterpart.

The goal of this thesis is to develop computational models where pragmatics is a first-class citizen both in terms of natural language understanding and generation. We have already made progress toward this goal by (1) developing a neural model for definiteness prediction the task of determining whether a noun phrase should be definite or indefinite in contrast to prior work relying on heavily-engineered linguistic features and (2) by introducing the new task of adverbial presupposition triggering detection which focuses on detecting contexts where adverbs (e.g. “again”) trigger presuppositions (“John came again” presupposes “he came before”).

Moving forward, we propose to examine the

role of pragmatics, particularly presuppositions, in language understanding and generation. We will develop computational models and corpora that incorporate this understanding to improve: (1) summarization systems e.g. in a text rewriting step to learn how to appropriately allocate adverbs in generated sentences to make them more coherent, and (2) reading comprehension systems where pragmatic effects are crucial for the proper understanding of texts. By the end, this thesis would present the first study on presuppositional effects in language to enable pragmatically-empowered natural language understanding and generation systems.

In what follows, we will give a summary of our research effort thus far, briefly discussing the two tasks mentioned above and that can be seen as testbeds for pragmatic reasoning followed by exciting current and future research avenues for further exploration.

## 2 Definiteness Prediction

In (Kabbara et al., 2016), we focus on definiteness prediction, the task of determining whether a noun phrase should be definite or indefinite. In English, one instantiation of this task is to predict whether to use a definite article (the), indefinite article (a(n)), or no article at all. This task has applications in machine translation, and in L2 grammatical error detection and correction.

Definiteness prediction is an interesting testbed for pragmatic reasoning, because both contextual and local cues are crucial to determining the acceptability of a particular choice of article. Consider the following example: A/#the man entered the room. The/#a man turned on the TV. Factors such as discourse context, familiarity, and information status play a role in determining the choice of articles. Here, man is introduced into the discourse context by an indefinite article, then subsequently referred to by a definite article. On the other hand, non-context-dependent factors such as local syntactic and semantic restrictions may block the presence of an article. For example, demonstratives (e.g., this, that), certain quantifiers (e.g., no), and mass nouns (e.g., money) do not permit articles.

We present in this work a recurrent neural network model that employs an attention mechanism. The primary motivation for the use of an attention mechanism is to investigate whether LSTM models focus on certain parts of the sentence when

making predictions, and if so, to gain more insight into what parts of the sentence affect the model’s prediction.

Our model achieves state-of-the-art performance on definiteness prediction, outperforming a previous logistic regression classifier (De Felice, 2008) that uses 10 types of hand-crafted linguistic features. Our best model achieves 96.63% accuracy on the WSJ/PTB corpus, representing a relative error reduction of 51% compared to the previous state of the art. Each of the factors we examined (initializing with pre-trained vectors, giving more context, giving POS tags, attention) contributes to the performance of the model, though in different degrees. We perform a number of analyses to understand the behavior of the models, and show in particular how the attention mechanism can be useful for interpreting the model predictions.

The most interesting contribution of this work is highlighting the suitability of LSTMs for tackling complex cases of article usage where there is no obvious local cue for prediction. We find evidence that LSTMs given an extended context window can resolve cases of article usage that seem to require reasoning about coreferent entities involving synonymy. Our results suggest that recurrent neural network models such as LSTMs are a promising approach to capturing pragmatic knowledge.

## 3 Adverbial Presupposition Triggering

In (Cianflone, Feng, Kabbara et al., 2018), we introduce the task of predicting adverbial presupposition triggers such as also and again. Solving such a task requires detecting recurring or similar events in the discourse context, and has applications in natural language generation tasks such as summarization and dialogue systems.

Presuppositions have been extensively studied in linguistics and philosophy of language with the earliest work dating back to (Frege, 1892). They can be viewed as assumptions or beliefs in the common ground between discourse participants when an utterance is made. The importance of presuppositions is that they underlie spoken statements and written sentences and understanding them facilitates smooth communication. We refer to expressions that indicate the presence of presuppositions as presupposition triggers. These include, among others, definite descriptions, factive verbs and certain adverbs.

Our focus in this work is on adverbial presupposition triggers such as *again*, *also* and *still*. These triggers indicate the recurrence, continuation, or termination of an event in the discourse context, or the presence of a similar event, and are the most commonly occurring presupposition triggers after existential triggers (Khaleel, 2010).

As a first step towards language technology systems capable of understanding and using presuppositions, we propose to investigate the detection of contexts in which these triggers can be used. This task constitutes an interesting testing ground for pragmatic reasoning, because the cues that are indicative of contexts containing recurring or similar events are complex and often span more than one sentence. Moreover, such a task has immediate practical consequences. For example, in language generation applications such as summarization and dialogue systems, adding presuppositional triggers in contextually appropriate locations can improve the readability and coherence of the generated output.

We create two datasets for the task based on the Penn Treebank corpus (Marcus et al., 1993) and the English Gigaword corpus (Graff et al., 2007), extracting contexts that include presupposition triggers as well as other similar contexts that do not, in order to form a binary classification task. In creating our datasets, we consider a set of five target adverbs: *too*, *again*, *also*, *still*, and *yet*. We focus on these adverbs in our investigation because these triggers are well known in the existing linguistic literature and commonly triggering presuppositions. We control for a number of potential confounding factors, such as class balance, and the syntactic governor of the triggering adverb, so that models cannot exploit these correlating factors without any actual understanding of the presuppositional properties of the context.

In addition, we investigate the potential of attention-based deep learning models for detecting adverbial triggers and introduce a new weighted pooling attention mechanism designed for predicting adverbial presupposition triggers. Our attention mechanism allows for a weighted averaging of our RNN hidden states where the weights are informed by the inputs, as opposed to a simple unweighted averaging. Our model uses a form of self-attention (Paulus et al., 2018; Vaswani et al., 2017), where the input sequence acts as both the attention mechanism’s query and key/value. Un-

like other attention models, instead of simply averaging the scores to be weighted, our approach aggregates (learned) attention scores by learning a reweighting scheme of those scores through another level (dimension) of attention. Additionally, our mechanism does not introduce any new parameters when compared to our LSTM baseline, reducing its computational impact.

We compare our model using the novel attention mechanism against strong baseline classifiers, including a logistic regression model and RNN- and CNN-based deep learning models, in terms of prediction accuracy and show that it outperforms these baselines for most of the triggers on the two datasets without introducing additional parameters – achieving 82.42% accuracy on predicting the adverb “also” on the Gigaword dataset.

## 4 Proposed Research

The research work thus far in this thesis has explored two classes of function words, namely articles and adverbs, and investigated the suitability of deep learning models to pick up on complex contextual cues for handling pragmatic inferences in learning tasks involving these function words. The remainder of the research work to be carried in this thesis will build on this progress to explore how pragmatic effects in language such as presuppositional effects can be leveraged to improve natural language generation and understanding systems. Accordingly, the proposed research is divided among the following three main fronts:

1. Corpus construction
2. Language generation
3. Language understanding.

### 4.1 Corpus construction

Since our work is the first to explore presuppositional effects from a computational perspective, the research community currently lacks corpora that focus on these important pragmatic effects. For this reason, we are currently in the process of constructing a new corpus that focuses specifically on adverbial presuppositional effects in English. The goal of this corpus is to provide new resources that would push further the goal of understanding presuppositional effects in specific and research on computational pragmatics more generally. The corpus construction is motivated linguistically and will involve crowdsourcing data

(e.g., through Amazon Mechanical Turk). Two approaches to the corpus construction are of interest:

1. One approach is to consider it from a generation problem perspective. Workers will be given sentences involving presuppositions and would be tasked with identifying the presupposition in context. For example, given the sentence “John went to the restaurant again”, the worker would optimally provide a simple explanation describing that the presupposition is that “John went to the restaurant before”. We envision the corpus to be useful in the context of language generation. Particularly, such corpus would be useful for designing learning models that can focus on contextual cues leading to pragmatic effects and accordingly generate what was presupposed in a sentence as a way of showing that it has a basic understanding of the relevant pragmatic effects in context.
2. The other approach is to consider it from the perspective of a presupposition-based entailment task which illustrates drawing conclusions from specific cues in text. Workers would be given a short passage involving a presupposition such as “John has been to the restaurant again”. They would be asked something along the line of: “Is it true that John has been to the restaurant before?” to which the correct answer would be “yes”. In this case, this would be an entailment setup where the presupposition is what leads to the correct conclusion.

It would be interesting to expand this to not just adverbs, but also to other kinds of presuppositions and possibly even implicature. A more general version of this crowdsourcing could involve, for example, asking the worker to qualify whether a certain statement seems to be true, or be suggested but not necessarily true or false.

We believe that such corpora are crucial for the development of learning models that can focus on the subtle pragmatic effects of language and will play an important role in improving language generation and language understanding systems.

## 4.2 Language Generation

Presuppositions are prevalent in language and they play a crucial role in shaping the conveyed meaning in a specific way. In summarization scenarios

where information needs to be acquired from different parts of the text(s), this would be particularly more important.

Two pillars of effective summarization systems are (1) the ability to pinpoint key pieces of information and crucial parts of the text and (2) appropriately rewrite the original text in order to relay the information in those parts.

On the first front, we plan to investigate links between sentences that involve presuppositional effects and sentences occurring in the previous context. That is, if a sentence includes a presupposition trigger, it would be important – from a summarization point of view – to determine whether there is a specific sentence that occurred in the previous context and that plays a crucial role in how the meaning of that (*presuppositional*) sentence is understood. We believe investigating such links between sentences is crucial for designing better summarization systems.

On the latter front, we are interested in investigating how an understanding of presuppositional effects can lead to better summarization systems. Specifically, we will investigate how a learning model, in a text rewriting step, can make informed decisions on how to properly allocate adverbs with the goal of generating a more coherent output.

Framing the problem within a summarization setting enables to not have to rely on using the original document context but instead “manipulating” the summarized version. From that version, we can select groups of sentences and ask workers to add adverbs or to remove existing ones to create manipulated versions of summarized texts to be used for designing improved summarization systems. Indeed, oftentimes, missing one presupposition trigger such as an adverb while summarizing could lead to the presupposition context (and thus the overall meaning) not holding anymore. The goal in this case would be to design summarization systems that are also trained to fill/remove adverbs such that the final summarized version is more natural and more coherent. The idea would be for the summarization system to examine each relevant sentence in the summarized text and determine whether adding/removing a specific adverb would make the phrase more informative and so whether it should be included or not in the summarized version.

### 4.3 Language Understanding

Of interest to our pursuit is the task of machine/reading comprehension also referred to as text understanding or question answering (QA), where the goal is to determine if a learning system can answer basic questions about a passage in order to show some “comprehension” of the information in that passage.

Presuppositions play a crucial role in shaping the sentence meanings and extending what is explicitly conveyed by the semantics of the words making up the sentence. We believe that understanding their role and leveraging the pragmatic knowledge resulting from that role can improve text understanding systems.

The current QA datasets are not suitable however for the task of designing pragmatically-empowered QA systems. For example, among the popular datasets for this task, is the Stanford Question Answering Dataset (SQuAD) (Rajpurkar et al., 2016). SQuAD consists of questions posed by crowdworkers on a set of Wikipedia articles, with more than 100,000 question-answer pairs on 500 articles. A key feature of this dataset is that the answer to a question is always part of the context and also always appears as a continuous span of words. One simplified way to tackle the problem is then to find the start and end of the relevant span of words (in the given passage) that corresponds to the answer. Not only the answers in such datasets are explicitly present in the passage, they are typically of fact-based nature.

Our interest, on the other hand, is to design QA systems that can answer questions that tap implicit information in the text, one that is pragmatic in nature. The answers, unlike datasets like SQuAD, would not be explicitly stated in the text which makes the task more challenging. The constructed corpus that was discussed in Section 4.1 will be crucial to designing and testing such systems. A simple example would be a small passage with a sentence involving a presupposition such as “John has been to the restaurant yesterday”. One possible question would be: “Is it true that John has been to the restaurant before?” and the answer should be true. There would be challenging *negative cases*, e.g., “Mary has been to the restaurant again, but John wants to go to the restaurant tomorrow. The answer to the same question above would be false in that case.

Answering effectively such questions would tap

into the models abilities to pick up on complex contextual cues and would be a strong hint that the model can have a basic understanding of the pragmatic effects in language.

### 4.4 Leveraging Pragmatic Knowledge in Multi-Task Scenarios

Correctly performing the pragmatic reasoning explored in this proposal, i.e. that which deals with presuppositions, is at the core of many NLP tasks such as discourse parsing, discourse segmentation and coherence modeling.

We believe that by training a model to learn to produce simple explanations of the presupposition effect in context (in the fashion described in Section 4.1), we could leverage the learned representations to “supplement” the learning of other NLP tasks as the ones mentioned above and for which such pragmatic knowledge is essential.

## 5 Conclusion

We have presented in this proposal the research progress that was accomplished thus far in this thesis, exploring two classes of function words, namely articles and adverbs, and investigating the suitability of deep learning models to pick up on complex contextual cues for handling pragmatic inferences in learning tasks involving these function words, namely, the tasks of definiteness prediction and adverbial presupposition triggering. We also discussed current and future research directions that will guide the research for the remainder of this thesis mainly in the areas of natural language generation and understanding. We believe that the research in this thesis has the potential to open up exciting research directions that are unexplored in the NLP community and that are crucial for designing improved and nuanced language generation and understanding systems that bring us closer to the vision of truly intelligent systems.

### Acknowledgments

This work was supported by the Centre de Recherche d’Informatique de Montréal (CRIM), the Fonds de Recherche du Québec Nature et Technologies (FRQNT) and the Natural Sciences and Engineering Research Council of Canada (NSERC).

## References

- Andre Cianflone, Yulan Feng, Jad Kabbara, and Jackie Chi Kit Cheung. 2018. [Let’s do it “again”: A first computational approach to detecting adverbial pre-supposition triggers](#). In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2747–2755, Melbourne, Australia. Association for Computational Linguistics.
- Rachele De Felice. 2008. *Automatic error detection in non-native English*. Ph.D. thesis, University of Oxford.
- Gottlob Frege. 1892. Über sinn und bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100:25–50.
- David Graff, Junbo Kong, Ke Chen, and Kazuaki Maeda. 2007. [English gigaword third edition](#). Technical report, Linguistic Data Consortium.
- Jad Kabbara, Yulan Feng, and Jackie Chi Kit Cheung. 2016. Capturing pragmatic knowledge in article usage prediction using lstms. In *COLING*, pages 2625–2634.
- Layth Muthana Khaleel. 2010. An analysis of pre-supposition triggers in english journalistic texts. *Of College Of Education For Women*, 21(2):523–551.
- Yang Liu, Zhiyuan Liu, Tat-Seng Chua, and Maosong Sun. 2015. [Topical word embeddings](#). In *AAAI Conference on Artificial Intelligence*.
- Thang Luong, Hieu Pham, and Christopher D Manning. 2015. Bilingual word representations with monolingual quality in mind. In *Proceedings of the 1st Workshop on Vector Space Modeling for Natural Language Processing*, pages 151–159.
- Mitchell P. Marcus, Mary A. Marcinkiewicz, and Beatrice Santorini. 1993. Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics*, 19(2):313–330.
- Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems*, pages 3111–3119.
- Romain Paulus, Caiming Xiong, and Richard Socher. 2018. [A deep reinforced model for abstractive summarization](#). In *International Conference on Learning Representations*, Vancouver, Canada.
- Jeffrey Pennington, Richard Socher, and Christopher D Manning. 2014. Glove: Global vectors for word representation. In *EMNLP*, volume 14, pages 1532–1543.
- Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev, and Percy Liang. 2016. Squad: 100,000+ questions for machine comprehension of text. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 2383–2392.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Llion Jones, Jakob Uszkoreit, Aidan N Gomez, and Łukasz Kaiser. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems 30*, pages 5994–6004.