

# Recognising enthymemes in real-world texts: a feasibility study

Olesya RAZUVAYEVSKAYA and Simone TEUFEL

Computer Laboratory,  
University of Cambridge,  
15 JJ Thomson Avenue, Cambridge, UK  
e-mail: *or264@cam.ac.uk*, *sht25@cam.ac.uk*

**Abstract.** We present a feasibility study for the task of finding and expanding enthymemes, i.e. arguments with missing propositions, in real-world texts. We split the task into three subtasks: 1. finding the beginning and the end of the text span describing a minimal argument block, 2. deciding whether that span really represents an *enthymematic* minimal argument block, and 3. describing the missing premise in natural language. We argue that an objective ground truth for these tasks must be found before we can address automatic annotation. In our opinion, such a ground truth can only come from agreement on independent human annotation. In other words, we require that two or more annotators agree on the existence and textual span of an enthymematic minimal argument block, and on the actual paraphrase of the missing premise. We present a case study using the two authors of this paper as annotators, where we test three cue phrases for their suitability to this task, *because*, *therefore* and *let alone*. We find that minimal argument blocks centred around the cue phrase *let alone* are of a particularly high quality. We also discuss pragmatic effects of *let alone* and how they relate to argumentation theory.

**Keywords.** argumentation, enthymemes, annotation experiment, pragmatics

## 1. Introduction

In this paper, we investigate the feasibility of finding arguments with missing premises or conclusions, also referred to as *enthymemes*<sup>1</sup>, in unrestricted texts.

Many of the arguments expressed in a form of natural language are incomplete [6]. The majority of implicit propositions represent trivial facts that the speaker tends to avoid to say in order not to bore the listener [7]. This complies with Grice's Maxim of Quantity - *Do not make your contribution more informative than is required* [8]. Sometimes, the speakers also reduce the amount of required information in order to avoid potentially fallacious propositions [3]. Jackson and Jacobs [9] notice that enthymemes also play an important role in maximizing a listener's agreement, because additional information always increases the possibility of disagreement.

Automatic reconstruction of enthymemes is an interesting task from a text understanding perspective. As enthymemes are known to establish textual entailment relations between two propositions [10], the ability to reconstruct missing premises would provide a clear demonstration that the argument was understood. For the systems that aim at verifying the validity of arguments, the task of enthymeme reconstruction is important from a practical viewpoint, because such systems need to have access to the omitted parts as well as the explicitly stated ones. Moreover, universally quantified premises express general truths about our world, and could therefore be theoretically used as a source of information that is objective enough to be put into general knowledge databases.

All types of arguments can be expressed in a truncated form, but truncated syllogisms are the most studied types of these [2]. Standard-form syllogisms are convenient for the analysis of en-

---

<sup>1</sup>*Enthymemes* were first defined by Aristotle [1] as a standard-form syllogisms with one missing proposition. However, in the scope of modern usage of this term in argumentation theory [2] [3] [4] [5], it is widely accepted to refer to any type of an argument where one of the principal inferences is missing as enthymematic.

enthymemes, because such arguments consist of only three terms and satisfy a set of strict requirements [11]. Therefore, given any two syllogistic propositions, the goal of restoring the missing premise or conclusion can be achieved by following a set of deterministic steps. For instance, consider the enthymeme below:

*Pugs are mammals,*                      [minor premise]  
*because they are dogs.*                [conclusion]

As can be inferred from the formal definition of a syllogism, this is a categorical syllogism with the *minor premise* and the *conclusion* expressed. There exists a general procedure [2] that allows us to conclude that the missing proposition is:

*All dogs are mammals.*                [major premise, universally quantified]

Syllogisms allow us a straight-forward and objective definition of the missing proposition. Enthymematic syllogisms already contain all three terms required for the reconstruction, and no external information is necessary.

However, it is almost impossible to find well-formed standard syllogisms in everyday language [2]. Even once spotted, the transformation of such a naturally stated argument into the syllogistic form requires much effort and general knowledge. It therefore does not seem feasible to automate this step to collect a large corpus of well-formed syllogisms in this way. Additionally, many of the naturally occurring arguments cannot be translated into syllogisms because they are not logically valid, but only *defeasible*, i.e. some exceptions to their application exist. It therefore does not seem feasible to use real-world arguments directly as a source of well-formed syllogisms.

The argumentation theory literature is generally pessimistic about the feasibility of insertion of missing premises or conclusions: it has been called a challenging and subjective task even for human experts, often resulting in a wrong interpretation of the arguments [12] [13] [14]. Hitchcock [14] mentions two problems in particular associated with enthymemes, a) the difficulty of distinguishing enthymemes from deductively valid arguments and from the arguments that are to be rejected ("demarcation problem") and b) the difficulty of evaluating whether the inferred expression was the one that was originally intended. Therefore, successful automation of enthymeme reconstruction is not possible without first establishing the ground truth about the premise insertion goal. We plan to investigate the feasibility of several subtasks associated with this task:

- detection of the beginning and the end of the text span describing a minimal argument block;
- agreement on whether that span really represents an enthymematic minimal argument block;
- agreement on the missing premise explicitly stated in natural language.

Ground truth for these tasks can only be established by means of human annotation, because the interpretation of enthymemes is known to be inherently subjective. If two or more annotators, working independently, can agree on the existence and identity of the missing premise, we can take it as proof that the enthymeme objectively "exists". This turns the problem of human subjectivity of enthymeme reconstruction into an empirical question – it can now be answered simply by measuring agreement of annotators.

We believe that the first step towards making the task of premise insertion objective is to detect *enthymematic minimal argumentation blocks* (EMAP for short). We define them to be text pieces containing a minimal enthymematic argument that can be interpreted without additional context, in particular, without general knowledge. This will make the task of inserting missing premises more objective. In what follows, we are developing a corpus of minimal argumentation blocks which are enthymematic, using an input stream of unrestricted text. We will assume that most of these arguments are defeasible. We need to employ an automatic procedure to find the text spans expressing the arguments; in particular, we need to avoid including non-relevant material in these text spans.

Under our scheme, annotators have to decide for each candidate EMAP, whether the given text contains a stand-alone argument. One of the possible questions that can arise is: *What if these stand-alone arguments are not really enthymematic, i.e., what if they already contain all of the required propositions?* We make the assumption, also made by Dijk and Kintch [15], that we can always

insert an implicit premise between two propositions, for the pragmatic reasons already sketched out above – humans tend not to spell out links between statements. Therefore, for any EMAP, there should always be the logical "space" to insert a new statement.

## 2. Discourse markers for enthymeme detection

The result of the human annotation studies will be a corpus of EMAPs. To conduct these experiments, we first need to build an initial corpus of candidate EMAPs for annotation. The use of discourse markers for automated recognition of arguments is standard practice [16] [17] [18] [19]. By using argumentative discourse markers, i.e., explicit signals used by the speaker, we can assume that the argumentative moves are indeed present.

However, not all of the discourse markers are equally suitable for our goal of building a corpus of EMAPs. Plausible discourse markers should have a high likelihood of indicating an EMAP in their context. We will refer to this property as the *locality* of a discourse marker. The relevant context around the discourse marker which forms part of the candidate EMAP should ideally be as small as possible.

We define three acceptable locality ranges. Here, we refer to the sentence where the discourse marker is localized as a *target sentence*.

- *1-span locality (complete locality)* - at least two propositions, the premise and the conclusion, are usually stored in the target sentence itself
- *2-span locality*
  - \* *backward-looking locality* - the conclusion and its preceding premise are stored in the following span:  
[1 sentence before the target sentence]+[target sentence]
  - \* *forward-looking locality* - the premise and its preceding conclusion are stored in the following span:  
[target sentence]+[1 sentence after the target sentence]
- *3-span locality* - the premise can be stored both before and after the conclusion, therefore, we need a larger span to make sure that both propositions are captured:  
[1 sentence before the target sentence]+[target sentence]+[1 sentence after the target sentence]

Our expectation is that smaller locality ranges should make objective enthymeme detection and reconstruction easier. However, we also predict that successful reconstruction of an enthymeme depends not only on the locality. Consider the following two arguments:

*A must be true, because A has never observed to be false*  
*A must be true, because X has told so*

While the same *because* discourse marker is used in both cases, these enthymematic arguments are based on entirely different reasoning types. The first argument is based on the lack of evidence about any controversial case, whereas the second argument is based on the reliability of the claim's source. These arguments also correspond to different argument schemes, *argument from expert opinion* and *argument from ignorance*; therefore, different reconstruction techniques must be applied to these enthymemes. In order to apply these techniques, discourse markers must uniquely identify their associated argument schemes. We further refer to this characteristic of a discourse marker as *scheme specificity*. With respect to scheme specificity, *because* can be used with almost any argument scheme; therefore, people are likely to disagree on the missing inference type even more than with a scheme-specific discourse marker.

## 3. Methodology and Results

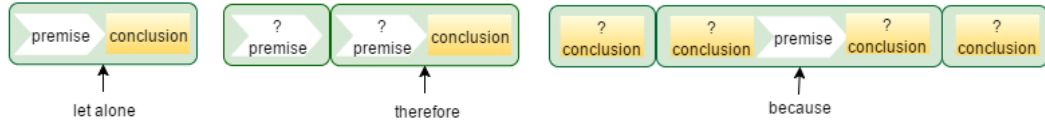
We report an experimental human annotation study that supports the assumption about the importance of both the locality and the scheme specificity of discourse markers for reaching good agree-

ment between annotators on the inserted inference. We performed a case study using two annotators, the authors of this paper, where we test three discourse markers for their suitability to this task: *because* and *therefore*, both of which are scheme-non-specific, and a scheme-specific *let alone*.

### 3.1. Annotation data

100 random sentences containing each discourse marker (300 in total) of the average length of 25 tokens were extracted from the British National Corpus (BNC).

*Let alone* constructions were limited to the target sentence only (The assumption is that both the premise and the conclusion are usually located inside a 1-span text). For each *therefore* cue phrase, we extracted the target sentence along with the immediately preceding one, because this discourse marker is of a 2-span type (The assumption is that the preceding sentence may contain the premise for the conclusion). For each *because* marker, target sentences were extracted with one directly preceding and one directly following sentence (The assumption is that both the preceding and following sentences may contain the conclusion for the premise).



**Figure 1.** Spans occupied by *let alone*, *therefore*, and *because* EMAPs

#### 3.1.1. Boundary detection task

Two annotators independently decided whether there was enough context to detect an EMAP. For the cases where they could detect an EMAP, they were also asked to indicate the EMAP-irrelevant information. This would allow us to study the locality of the discourse marker.

Table 1 shows the distribution of the total number of positive choices, i.e., the cases where annotators agreed on the presence of an EMAP, per marker. As can be seen, the number of positive choices was high for *let alone* (183/200), in contrast to *therefore* (67/200) and *because* (83/200).

The inter-rater agreement was calculated by means of Cohen's *Kappa* statistics [20].

As far as the locality property is concerned, for 2-span *therefore*, the annotators only specified EMAP-irrelevant information (17/67 cases) to the left of a discourse marker. This demonstrates that our assumption that preceding sentence often contains a premise for the conclusion was correct. For 3-span *because*, the amount of EMAP-irrelevant information was higher (25/83 to the left and 30/83 to the right of a discourse marker). This shows that the conclusion for the premise is often located either to the left or to the right of the target sentence. For 1-span *let alone*, there were only 2/183 cases where either of annotators specified EMAP-irrelevant information to the left of a discourse marker.

**Table 1.** Inter-annotator agreement results for the boundary detection task

Marker	Number of positive choices	Kappa-statistics	Number of times left boundary was moved	Number of times right boundary was moved
<i>Therefore</i>	67	K=0.451 (k=2, n=2, N=100)	17	0
<i>Because</i>	83	K=0.577 (k=2, n=2, N=100)	25	30
<i>Let alone</i>	183	K=0.729 (k=2, n=2, N=100)	2	0

### 3.2. Experimental Study on Proposition Insertion Tasks

Looking at the desired properties of discourse markers for the task of EMAP detection, we can see, based on the enthymeme detection task (Section 3.1.1), that *let alone* is always concentrated in a single sentence, i.e., its locality range is small. We will further demonstrate that it also has the other desired property, scheme-specificity.

### 3.2.1. Linguistic Background on let alone

*Let alone* represents an unusual linguistic phenomenon. According to [21], *let alone* can be syntactically treated as a coordinating conjunction, where the interpretation of the second part of the construction depends on information provided in its first part. Consider the following examples:

*This task is difficult for an **adult**, let alone a **child**.*

*The baby cannot **sit** yet, let alone **walk**.*

The phrases linked by *let alone* display parallelism with respect to grammatical functions; e.g. direct object (*adult* and *child*) or verbal complement to auxiliary (*sit* and *walk*). In the terminology of [22], the stressed elements in the first and second part of the sentence are called *correlate* and *remnant* respectively. *let alone* is similar to comparatives in this respect.

Unlike comparatives, *let alone* is frequently treated as a negative polarity construction [21]. However, [23] refined this treatment by distinguishing the following three main cases: explicit negative (a), implicit negative (b) explicit positive (c):

(a) *He cannot solve simple, let alone starred tasks.*

(b) *The task is too difficult to be understood, let alone solved.*

(c) *He could find the exact answer, let alone an approximate solution.*

From a pragmatic point of view, [24] observe that the remnant is usually more relevant to the context, and the correlate is in some respect "more likely". The pragmatic negation of *let alone* results in the negation of the remnant by first negating the more likely correlate.

[22] observes that some form of scaling information is always present in *let alone* sentences, and that a pragmatic entailment relation holds between the remnant and the correlate. The hearer can pragmatically infer, based on common knowledge [8, 25], that these stressed elements are ordered based on some scale. This hidden scalar reasoning is highly relevant to our task of finding arguments with hidden premises.

We argue that *let alone* sentences are by definition enthymemes, and that the missing scalar relation is an important part of the missing proposition.

### 3.3. Argument Scheme for let alone Sentences

The pragmatic effects observed in the *let alone* construction have an obvious similarity with *a fortiori* arguments, which were first mentioned in Aristotle's Rhetoric [1]. The main principle behind *a fortiori* arguments is that if there is a case where a certain quality is more likely to exist than in some other, less likely, but that the quality does not even exist in this case, it certainly cannot exist in the less likely case. The speakers therefore use the more likely case to reject the less likely one. [26] designed the following argumentation scheme for such arguments:

If even X does not have property P and it is a less likely case that Y has property P than that X has property P, then Y does not have P

(Even) X does not have P

(Therefore) Y does not have P

*Let alone* seems to be one of the signals of such *a fortiori* inferences in natural language. We extend a scheme suggested by Kienpointer to the scheme consisting of four terms. As an additional term, we introduce is the scaling relation holding between X and Y. In the next section, we will present such scaling relations.

### 3.4. Annotation Scheme for "let alone" scale

We categorized the scaling relations based on about 250 cases of *let alone* sentences observed, resulting in the scheme in Table 2.

The categories are as follows.

1. *Smaller than*: The scale concerned is the one with standard measurements or cardinalities of the remnant and the correlate.

**Table 2.** Annotation scheme for *let alone* scales

	Relation type	Example
1	Smaller than	<i>You wouldn't make to New York, let alone the West Coast</i>
2	Part of	<i>This does not apply to Germany, let alone all of Europe.</i>
3	Precondition for	<i>Your talent isn't enough to participate, let alone win.</i>
4	Other lexical entailment	<i>He doesn't even sleep, let alone snore.</i>
5	Earlier date	<i>They didn't have electricity in 1923, let alone 1909.</i>
6	Additional constraint	<i>You don't know what a middle-aged person feels like, let alone a middle-aged prince.</i>
7	Additional referent	<i>The company does not even insure their employees, let alone their families.</i>
8	Cumulative/independent	<i>This Easter-egg packaging does not even protect its contents, let alone have anything to do with Easter.</i>
9	More extreme case than;	<i>They refused to refer to Kursk, let alone Moscow.</i>
10	Easier than;	<i>I could not solve the first, let alone the last tasks.</i>
11	Less likely than	<i>I have not even seen Mary, let alone Rose there.</i>

2. *Part of*: The relationship between the referents in remnant and correlate is that of superset.

3. *Precondition*: The action in the remnant necessarily requires an action in the correlate having taken place earlier (In example 3 from Table 2, *participating* is a precondition for *winning*).

4. *Other lexical entailment*: Covers the larger class of actions where the proposition expressed in the remnant lexically entails the proposition expressed in the correlate. There can be temporal overlap between the two propositions (In example 4 from Table 2, *snoring* entails *sleeping*.)

5. *Earlier date*: A special case of *Smaller than*: if a state first becomes true at time  $t_0$ , it by definition does not hold at any earlier time  $t < t_0$ . It can only be applied to irreversible events.

6. *Additional constraint*: In this category, a number of semantic predicates (constraints) applied to the correlate are compared with that same predicates plus additional predicates in the remnant.

7. *Additional referents*: Covers cases where the set of referents in the remnant logically includes the referent from the correlate.

8. *Cumulative/independent*: Here, the remnant of *let alone* is not directly comparable to the correlate, but becomes interpretable if we read it as an additive constraint.

The three categories *Easier than*, *Less likely than*, and *More extreme case than* are fallback options, which are only to be used if none of the more specific categories applies. A special rule in the annotation is that annotators should always choose the most specific case.

There is a clear connection between the scale class and the linguistic form of premise that can be constructed from that class:

- 1 X is smaller than Y
- 2 X is a part of Y
- 3 X is a precondition for Y
- 4 Y lexically entails X
- 5 Y is earlier than X
- 6 Y poses additional constraint on X
- 7 Y in addition contains X
- 8 X and Y are cumulatively more important than X
- 9 Y is a more extreme case than X
- 10 X is easier than Y
- 11 Y is less likely than X

Annotators were asked to extract the appropriate parts from the given sentence and fill them into the template of their chosen category. For instance, for the first example in Table 2 they were expected to write:

*The distance to New York is smaller than the distance to the West coast.*

### 3.5. Enthymeme Reconstruction Task

In 3.1.1, *let alone* contexts were found to have the highest positive agreement and the only substantial agreement regarding the presence of an EMAP. We next conducted more fine-grained annotation tasks only on those *let alone* sentences that were positively classified by both annotators as enthymematic (90 cases).

In this experiment, the annotators performed two tasks:

- They classified EMAPs according to the annotation scheme from Table 2 in order to describe the scale of *let alone* presented.
- They provided a paraphrase of the missing premise by composing a new sentence using material extracted from the given sentence together with the template given in Section 3.4.

#### 3.5.1. Results of Scale Classification Task

The class distribution per category presented in Table 3 is skewed.

The most frequently chosen category was the "precondition for" relation - 61 (33.9% of cases). The "additional constraint" and "additional referent" classes were also frequent (12.7% and 10% respectively). The fact that the fallback categories were chosen relatively frequently suggests that better formulation of the main categories is needed for future experiments.

**Table 3.** Class distribution (absolute and relative) per relation category

Relation type	Number of choices
Smaller than	9 (5%)
Part of	8 (4.4%)
Precondition for	61 (33.9%)
Other lexical entailment with	11 (6%)
Earlier date	5 (2.7%)
Additional constraint on	23 (12.7%)
Additional referent	18 (10%)
Cumulative/independent	9 (5%)
Less likely than	9 (5%)
More extreme case than	12 (6.7%)
Easier than	15 (8.3%)

The agreement between annotators was  $K=0.75$  ( $k=2$ ,  $n=2$ ,  $N=90$ ). This indicates that annotators were overall well able to distinguish relation categories (scales). The small numbers prevent us from analysing problematic categories systematically, but we noticed that *earlier date* and *precondition for* seem to be relatively harder for annotators to distinguish reliably.

#### 3.5.2. Results of Premise Insertion Task

In order to evaluate the annotators' agreement of inserted premises, we considered the 71 cases where the annotators agreed on the chosen scale category. Of those cases, one of the annotators judged 69 premises (97%) to be identical or near-identical. The first disagreement case was:

*Few adherents to the new classical macroeconomics trouble even to question it, let alone provide an analytical basis to justify it.*

While both annotators have specified that "questioning" is a precondition for "providing an analytical basis to justify", they specified partially different propositions:

- Questioning something is a precondition for providing an analytical basis to justify something.*
- Questioning classical macroeconomics is a precondition for providing an analytical basis to justify classical macroeconomics.*

The second EMAP on which the annotators disagreed was:

*It's not exactly as if Wimbledon are an English force – let alone European.*

Here, both of the premises were still nearly the same, but the annotators disagreed on the ellipsis resolution:

- (a) *For Wimbledon, to be an English force is easier than to be a European force.*
- (b) *To be an English force is easier than to be a European force.*

#### 4. General Discussion and Conclusion

In this work, we presented our first experiments on whether it is possible to automatically find enthymemes in unrestricted texts. The results support the hypothesis that signals, such as discourse markers of particular types, can be used as initial objective indicators of argumentative relations between two minimal units of text. For the case of *let alone*, we have observed both significant inter-annotator agreement and high positive agreements on the decision on whether the text units are argumentative. For the cases where annotators agreed, we have further investigated whether the annotators also agree on the exact type of pragmatic relation that holds between two propositions. The results show that although there exist some types of relations that are not very distinguishable from each other, the annotators are able to achieve significant agreement on this task. The fact that the exact premises specified by the annotators are identical in the overwhelming number of cases, suggests that *let alone* constructions are very good candidates for future collection, investigation and reconstruction of enthymemes. It also means that all the information needed to reconstruct an enthymeme could indeed be located in the local context of the sentence itself. As a result, human annotators can perform this task with apparent ease.

In future work it is essential to perform this task on a larger scale, with independent annotators and a larger dataset. Another aspect worth mentioning is that the amount of surrounding context for *therefore* and *because* EMAPs was larger than for *let alone* ones, which can be plausibly assumed to decrease the agreement probability. In the future, it is important to identify more precise argumentative boundaries of EMAPs before they are presented to the annotators. It is also essential to describe all types of relations that have demonstrated low agreement rate in a way that will make them more intuitive for future annotators.

Human annotation is only the first step towards automatic reconstruction of enthymemes. In particular, automatic resolution of gapping is a challenging task. Syntactic information is often not enough for intelligent reconstruction, and the semantic information is not available automatically [27] [28]. But it is important to have the ground truth on this subjective task showing that it is achievable by people.

Another insight that this study suggests is the fact that pragmatic information, which is not traditionally used in argument mining, can be beneficially incorporated as important features.

#### References

- [1] E. M. Cope and J. E. Sandys, eds., *Aristotle: Rhetoric*, vol. 1. Cambridge University Press, 2010. Cambridge Books Online.
- [2] I. Copi and C. Cohen, *Introduction to Logic*. Maxwell Macmillan international editions, Macmillan, 1990.
- [3] D. Walton and C. A. Reed, "Argumentation schemes and enthymemes," *Synthese*, vol. 145, no. 3, pp. 339–370, 2005.
- [4] E. Black and A. Hunter, *Using enthymemes in an inquiry dialogue system*, vol. 1, pp. 437 – 444. International Foundation for Autonomous Agents and Multiagent Systems, 2008.
- [5] V. W. Feng and G. Hirst, "Classifying arguments by scheme," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1*, HLT '11, (Stroudsburg, PA, USA), pp. 987–996, Association for Computational Linguistics, 2011.
- [6] M. Lippi and P. Torroni, "Context-independent claim detection for argument mining," in *Proceedings of the Twenty-Fourth International Conference on Artificial Intelligence*, pp. 185–191, 2015.
- [7] P. Hurley, *A Concise Introduction to Logic*. Cengage Learning, 2014.
- [8] H. P. Grice, "Logic and conversation," in *Syntax and Semantics: Vol. 3: Speech Acts* (P. Cole and J. L. Morgan, eds.), pp. 41–58, San Diego, CA: Academic Press, 1975.



- [9] S. Jackson and S. Jacobs, "Structure of conversational argument: Pragmatic bases for the enthymeme," *Quarterly Journal of Speech*, vol. 66, no. 3, pp. 251–265, 1980.
- [10] K. Morrell and A. Hewison, "Rhetoric in policy texts: the role of enthymeme in darzi's review of the nhs.," 2013.
- [11] J. Z. Sukkarieh, "Mind your language! controlled language for inference purposes," 2003.
- [12] M. B. Burke, "Unstated premises," *Informal Logic*, vol. 7, no. 2, 1985.
- [13] J. Gough and C. Tindale, "Hidden' or 'missing' premises," 1985.
- [14] D. Hitchcock, "Enthymematic arguments," 1985.
- [15] T. A. V. Dijk and W. Kintsch, "Strategies of discourse comprehension," 1983.
- [16] S. Somasundaran, J. Wiebe, and J. Ruppenhofer, "Discourse level opinion interpretation," in *Proceedings of the 22Nd International Conference on Computational Linguistics - Volume 1*, COLING '08, (Stroudsburg, PA, USA), pp. 801–808, Association for Computational Linguistics, 2008.
- [17] A. Tseronis, "From connectives to argumentative markers: A quest for markers of argumentative moves and of related aspects of argumentative discourse," *Argumentation*, vol. 25, no. 4, pp. 427–447, 2011.
- [18] E. Florou, S. Konstantopoulos, and A. Kukurikos, "Argument extraction for supporting public policy formulation."
- [19] J. Eckle-Kohler, R. Kluge, and I. Gurevych, "On the role of discourse markers for discriminating claims and premises in argumentative discourse," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, (Lisbon, Portugal), pp. 2249–2255, Association for Computational Linguistics, Sept. 2015.
- [20] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and Psychological Measurement*, vol. 20, no. 1, pp. 37–46, 1960.
- [21] C. J. Fillmore, P. Kay, and M. C. O'Connor, "Regularity and Idiomaticity in Grammatical Constructions: The Case of Let Alone," *Language*, vol. 64, no. 3, pp. 501–538, 1988.
- [22] M. Toosarvandani, "Letting negative polarity alone for let alone," In Tova Friedman and Satoshi Ito (eds.) *Proceedings from Semantics and Linguistic Theory XVIII*, 729- 746. Ithaca, New York: CLC Publications., 2008.
- [23] O. Sawada, "Rethinking the let alone construction: what are its construction specific characteristics.," 2003.
- [24] B. Cappelle, E. Dugas, and V. Tobin, "An afterthought on let alone," *Journal of Pragmatics*, vol. 80, pp. 70 – 85, 2015.
- [25] R. Stalnaker, "Common ground," *Linguistics and Philosophy*, vol. 25, no. 5-6, pp. 701–721, 2002.
- [26] M. Kienpointner, *Alltagslogik: Struktur und Funktion von Argumentationsmustern*. Problemata (Stuttgart), Frommann-Holzboog, 1992.
- [27] M. McShane and P. Babkin, "Automatic ellipsis resolution: Recovering covert information from text," 2015.
- [28] N. Suszczaska, J. Romaniuk, and P. Szmaj, "Automatic analysis of elliptic sentences in the thetos system1," 2005.