Rhetorical Figure Annotation with XML

Sebastian Ruan,* Chrysanne Di Marco,* and Randy Allen Harris**
*Cheriton School of Computer Science
**Department of English Language and Literature
University of Waterloo, Waterloo Ontario
saruan@uwaterloo.ca, cdimarco@uwaterloo.ca, raha@uwaterloo.ca

Abstract

There is a driving need to interrogate large bodies of text for pragmatic meaning, e.g., to detect sentiment, diagnose genre, plot chains of reasoning, and so forth. But this type of meaning is often implicit, 'hidden' meaning, evoked by linguistic cues, stylistic arrangement, or argumentation structure—features that have hitherto been difficult for Natural Language Processing (NLP) systems to recognize and use. Pragmatic concerns were historically the province of rhetorical studies, and we have turned to rhetoric in order to find new solutions to computational pragmatics. This paper highlights a form of rhetorical device that encodes deep levels of pragmatic meaning and yet lends itself to automated detection. These devices are the linguistic configurations known as *rhetorical figures*, which have been poorly understood and vastly underutilized in Computational Linguistics and Computational Argumentation. We present an annotation scheme using XML for rhetorical figures to make figuration more tractable for NLP, enhancing applications for argument mining, as well as such tasks as authorship attribution, genre detection, and sentiment detection. We also discuss the intellectual and technical challenges involved in figure annotation and implications for Machine Learning.

1 Introduction

Rhetorical figures are cognitively governed linguistic devices that serve functional, mnemonic, and aesthetic purposes. Take the famous maxim from Kennedy's inaugural address:

1. Ask not what your country can do for you. Ask what you can do for your country.

This expression quickly became proverbial in the American consciousness for the way it captures the spirit of a particular historical moment, the ethos of a particular administration, and the aspirations of a particular generation. Countless more prosaic formulations, by Kennedy and others, expressed that confluence too, but they left a distinctly less memorable impression. Why? Two reasons. Firstly, the formal structure and the functional structure are virtually isomorphic: Kennedy (and speechwriter Ted Sorensen) expressed the rejection of one civic attitude and its replacement by the opposite one, in the iconicity of reversing the terms of reference. Secondly, that very snug form/function coupling inhabits a material structure that is, on its own, cognitively very sticky. The Kennedy-Sorensen phrase has become so widely known, that is, so easily shared, so frequently invoked and quoted and recited because of (1) the schematic congruence with which the form matches the Rejection-Replacement function its arrangement serves, and (2) the cognitive affinities humans have for its structural properties (opposition, repetition, and symmetry).

The cognitive affinities explain its mnemonic and aesthetic effects, but, as computational argumentation scholars, we are more concerned with its tight form-functional correlation. The *form* makes it tractable for automated detection. The *function* gives us its semantic purpose. In terms of argument mining, an application that accessed this correlation could epitomize Kennedy's inaugural address (which argued for the rejection of an ethos of entitlement and its replacement by an ethos of duty) virtually on the basis of this expression alone.

We are developing an approach to computational pragmatics that combines the insights for argumentation that rhetorical figures provide together with argument mining, corpus linguistics, and machine learning, with payoffs for both computer science and for rhetoric. There has to this point been success at detecting some rhetorical figures, but little sense of what to do with them once they have been detected.

Our approach is a more sophisticated use of rhetorical figures than has been attempted, operating at layers of formal and functional abstraction. It depends fundamentally on an annotation format for rhetorical figures, which present unique challenges but also rich opportunities.

2 Opportunities and Challenges

Computationally, figures are important for three central reasons. First, they are endemic to human language. This is very well-established for a few tropes, such as metaphor, which is the central focus of Cognitive Linguistics and deeply entrenched in ontologies like FrameNet and WordNet. But it is equally true of literally (a word we don't use lightly) hundreds of other figures. If we want language-perceptive algorithms, they must have knowledge of figure structure. Secondly, many figures (especially the ones called *schemes*) work in terms of formal patterns that algorithms can detect through surface analysis; our Sentence 1 illustrates that clearly. Thirdly, they correlate with rhetorical meaning (pragmatic and argumentative functions). We will illustrate this further on. For now, the rejection-replacement function of Sentence 1 will have to stand.

But figures are not without their challenges for Natural Language Processing. Metaphor remains elusive, for instance, despite all the attention it has attracted in cognitive science, AI, and linguistics, including computational linguistics, in the last two decades. Metaphor is a type of figure known as a trope, which depends on semantic deviation. We are not yet successful enough with straight-laced semantics to support forays into semantic distortions. Some tropes (such as oxymoron, which is a juxtaposition of antonymic terms, such as *square circle* or *deafening silence*) can be reliably detected (Gawryjolek 2009). We believe antithesis (juxtaposed opposite propositions, as in Sentence 2) has a similar potential for reliable detection.

2. The young would choose an exciting life; the old a happy death.

But most semantic distortions—tropes—are far from tractable computationally. Nor do many of them provide the tight form/function coupling that has such a promising payoff for computational argumentation.

Another type of figure, schemes, are formal deviations, shifts of expected structure, as in Sentence 1, an antimetabole (reverse lexical repetition; in this case "you" and "your country"). The computational detection of antimetaboles is finding success (Gawryjolek 2009, Hromada 2011, Dubremetz and Nivre 2015).

The work of these researchers only concerned detection, an essential first step but it doesn't get us to argument mining. None of this work attempted to find *meaning* in the antimetaboles they detected. The researchers were unfamiliar with the rhetorical functions antimetabole serves.

Antimetabole has a small set of rhetorical functions, keyed to the iconicity of its formal structure (which evokes balance and opposition, as well as sequence or priority). We have very limited space in this paper to demonstrate these rhetorical functions, so a few examples will have to suffice.

One function of antimetabole is to convey Reciprocal Force, illustrated by Sentence 3, Newton's third law of motion. (We adopt the convention of identifying the defining figurative elements parenthetically.)

3. If you press a stone with your finger, the finger is also pressed by the stone. (stone / finger)

Newton's third law is often expressed as "for every action, there is an equal and opposite reaction," but Newton's own argument favored the antimetabole whose very structure suggests "equal and opposite" (We give the example in English, but Newton's original Latin is also antimetabolic.)

A very similar rhetorical function of antimetabole is to convey Reciprocal Specification, a kind of mutual definition, illustrated by Sentence 4:

4. Gay rights are human rights, and human rights are gay rights. (human rights / gay rights)

In this phrase the notions of human rights and gay rights are reciprocally identified with each other. You can't have one unless you have the other.

Another rhetorical function of the antimetabole is to convey Comprehensiveness, illustrated by the ordinary-language example, Sentence 5:

5. A place for everything, and everything in its place. (place / everything)

The reverse repetition in Sentence 5 shifts from reciprocal force to a reciprocal *coverage*, largely because it has prepositional predication rather than the transitive predication of Newton's Sentence 3. We call this function comprehensiveness because the sequential iconicity *means* a back-and-forth, alpha-to-omega, omega-to-alpha coverage of some domain—in this case, the domain of tidiness. All things have assigned places; all places have their assigned things.

A fourth rhetorical function of the antimetabole is to convey Irrelevance-Of-Order, well known from algebra and predicate calculus:

6. $m + n = n + m$ (m / n)

There are other ways to express the principle of commutation, but none as natural and iconic as formulae like 6. Opposite sequences of the same variables, on either side of the same operator, pivoted by a predication of identity, equivalence, or equality inescapably *means* that neither sequence has priority. Order doesn't matter to addition (multiplication, union, etc.).

We have built a curated list of over 400 antimetaboles illustrating these functions, but only have space for a few more representative examples:

Reciprocal Force

7. A corollary of PHC [the Principle of Hierarchical Coincidence] is that resources flow toward political power, and political power flows toward resources; or, the power of state and of capital typically appear in conjunction and are mutually reinforcing. (resources / political power)

8. Women are changing the universities and the universities are changing women. (women / universities)

Reciprocal Specification

9. The negation of a conjunction is the disjunction of the negations and the negation of a disjunction is the conjunction of the negations. (negation of a conjunction / disjunction of the negations)
10. Anger and depression, the pop-psych books tell us, are two sides of the same coin: depression is anger suppressed, anger is depression liberated. (depression / anger)

Comprehensiveness

11. I meant what I said and I said what I meant. (meant / said)
12. Whether we bring our enemies to justice or bring justice to our enemies, justice will be done. (our enemies / justice)

Irrelevance of Order

13. With a similar qualification, in the Cambridge Grammar of the English Language, a head 'plays the primary role' in 'determining the distribution of the phrase' (introductory chapter signed by Pullum and Huddleston, in Huddleston and Pullum 2002:24) (Pullum / Huddleston)
14. "Spanglish," [is] the combination of Spanish and English (or English and Spanish) (Spanish / English)

It is these functions, coupled with the relative ease of rhetorical-scheme detection, that make rhetorical figures so promising for computational tasks in which comprehension is central, like argument mining and text summarization.

Again, however, there are challenges. They are not as thorny as the challenges of most tropes because they concern surface analysis, not semantic plumbing. But they exist. In particular, figures rarely come in isolation. The Kennedy-Sorenson maxim, for instance (Sentence 1), is an antimetabole (you / your country). But it is also an antithesis (ask not X / ask X). It is, thirdly, a mesodiplosis (clause-medial repetition; here, *can do* occurs in the middle of both clauses).

We call this phenomenon, when figures pile on other figures, *stacking*. It is both a challenge and an opportunity. It is a challenge because rather than detecting a single figure or multiple independent figures, we need to detect overlapping figures. It is an opportunity because the functions are enhanced and stabilized under stacking. When two or more figures coincide in the same utterance, the functions they convey are highly consistent. Formal stacking breeds semantic conspiracy.

For instance, when antimetabole stacks with antithesis (conjoined or highly proximal opposite predications), the joint function is always to reject the negated predication utterly and replace it with the positive predication. Again, Sentence 1 is our paradigm, but here are two more:

15. We don't build services to make money; we make money to build better services.
16. Plain statement must be defined in terms of metaphor, not metaphor in terms of plain statement.

Some instances of stacking are so common and so predictable as to be entailments. Ploche, for instance, is simple lexical repetition, so it always stacks with antimetabole (reverse lexical repetition). If you find the latter, you always find the former. Semantically, ploche conveys the rhetorical function, Identity-Of-Reference, which is always embedded in the functions of antimetabole (if you have reciprocal force or reciprocal specification, for instance, you have identical entities in a reciprocal relationship). Further, mesodiplosis clause-medial lexical repetition) also entails ploche as well, conveying an identical force when the mesodiplosis is a transitive verb (e.g., Sentences 3, 7, and 8), identical specification when it is a copula verb (e.g., Sentences 4, 9, and 10).

We do not pretend to have a full and complete mapping of form to function, however. This work is still in the very early stages, but we believe it holds considerable promise, and we believe machine-learning corpus studies can be extremely helpful, especially for the challenges and opportunities of stacking.

Figural stacking, as we come to understand the functional combinatorics better, is perhaps the greatest promise of rhetorical figures for computational understanding of natural language. Our paradigm example, which stacks the schemes antimetabole, mesodiplosis (both entailing ploche), and the trope antithesis provides a pitch-perfect example of the rhetorical function, Reject-Replace. A computational analysis of Kennedy's inaugural address tuned to the workings of rhetorical figures could tell us what the address was about—namely, the rejection of an ethos of entitlement and its replacement with an ethos of responsibility—virtually on the basis of this particular stacking (along with, of course, the lexical semantics of *you*, *your country*, and so on)

We can, and should, rely on rhetoricians to tell us what the functions of certain figures and certain figure-stacks are, at least in these early stages. But the rhetorical tradition is haphazard, and sometimes conflicting. The terminology alone is forbidding. As much as computational argument studies can benefit from a better understanding of rhetorical figures, rhetorical figures can benefit from computational studies of form and meaning. (And, yes, that sentence was an antimetabole, stacked with mesodiplosis; the rhetorical function is Reciprocal Force, modulated by the possibility modality of *can*.)

The path forward is to bootstrap rhetoricians' knowledge by way of annotation, marked-up text corpora, and machine learning, so that computationally mined data can start to tell *them* what functions figures have, through

confirmation, through refinement, and through new discoveries, all of which we have good reason to anticipate.

We can discover the proportionality of certain stackings (anecdotally, both antithesis and mesodiplosis strongly co-occur with antimetabole), the correlation of the stackings with the rhetorical functions (as specified above, on the basis of limited and anecdotal research). At its best, this work can revolutionize computational argumentation studies and rhetoric in the way corpus linguistics revolutionized lexicography and established ontologies like WordNet and Framenet. But even at its least productive, we are very confident of finding important form/function correlations that can importantly inform computational argumentation and discourse studies, in novel ways.

3 Figure Detection

There have been limited successes to figure detection over the past several years due to strict figure mappings and some unreported data (Alliheedi 2012; Alliheedi and DiMarco 2012; Dubremetz and Nivre 2015; Gawryjolek 2009; Gawryjolek, Harris, and DiMarco 2009; Hromada 2011; Strommer 2011). But it has been limited by emphasis on particular figures.

Hromada's (2011) work, for instance, was very successful at the detection of antimetabole, but he defined antimetabole in an overly specified way. Using the Waterloo Figure Representation Notation (Harris and DiMarco 2009)[1] (where *W* stands for *Word*, the subscripts indicate identity, and "…" represents other linguistic matter, extraneous to the figure, possibly null), Hromada defines antimetabole as $<W_A\ W_B\ W_C\ .\ .\ .\ W_C\ W_B\ W_A>$. That is, he searched only for antimetaboles $<W_A, W_c>$ when they stacked with mesodiplosis $<W_B\ .\ .\ .\ W_B>$, when there was no additional linguistic matter.

Most of these researchers did not look for stacked figures, except accidentally. Hromada (2011) looked for other figures (anadiplosis, epanaphora, and epiphora), but only in isolation.[2] Conversely, he 'searched' for mesodiplosis unwittingly, because of the way he defined antimetabole. He was unaware he was doing so and does not report his results. Dubremetz and Nivre (2015) found some antitheses, because they were using negation as a correlative of antimetabole (which markedly improved their success), but they were not looking for them and did not report their results. Only Gawryjolek (2009) looked for stacked figures, but that was not his focus. He did not interpret the stacking at all, nor report on the statistics. He was merely looking for multiple figures in the same corpus, many of which overlapped.

We believe this work can be strengthened by machine learning. We have developed a format for annotating rhetorical figures, in parallel to the annotation formalisms developed for part-of-speech tagging, speech-act annotation, and so on. Corpora annotated with rhetorical figures can be used to train systems on new and more sophisticated detection tasks, especially for stackings and functional correlations.

We believe that attacking the problem from another angle is required. A curated corpus approach can make additional headway, in new directions.

4 Challenges and Solutions

We want to come at the detection problem from the other end. "A serious bottleneck … is the lack of annotated data" (Dubremetz and Nivre 2015). We believe that texts curated by rhetoricians, marked up for all occurrences of certain figures, will provide rich data for machine learning, and we have developed an annotation scheme to structure the data. The labels in our figure annotation scheme are in effect features pertaining to figure identification and classification. Algorithms trained on such data will, in turn, be more fully equipped for automated figure detection.

The main challenge is developing an additional structure for language that is widely applicable when annotating figure-rich texts. The intricacies of such texts include stacking figures and interpenetrating figures. This enhancing structure needs to be easily understood by computational learning-based algorithms as well as keeping figures intact. The Extensible Markup Language (XML) syntax lends itself to these requirements and is used for the annotation method developed in this paper.

---

[1] Hromada (2011) calls this notation, Rhetoric Figure Representation Formalism or RFRF, which he adapts from Harris and DiMarco (2009). Harris and DiMarco did not label their formalism in their paper, but we use their term for it here.

[2] Anadiplosis is clause-final-clause-initial lexical repetition ($< .\ .\ .\ W_X >< W_X\ .\ .\ .\ >$). Epanaphora is clause-initial lexical repetition ($< W_X\ .\ .\ .\ >< W_X\ .\ .\ .\ >$). Epiphora is clause-final lexical repetition ($< .\ .\ .\ W_X >< .\ .\ .\ W_X >$). Note that these researchers use somewhat different terminology. Hromada uses *anaphora* for our *epanaphora*, while Dubremetz and Nivre also use *chiasmus* for our *antimetabole*. In the first case, we avoid *anaphora* (a synonym in the rhetorical tradition for *ephaphora*) because of its more prominent designation in Computational Linguistics, for pronouns. In the second, we prefer the more specialized terms. It is worth noting that the terminology of rhetorical figures, resulting from over two millennia of research, is highly inconsistent, with different labels for the same linguistic configurations, with multiple linguistic configurations corresponding to the same label, and with some linguistic activity that really isn't figurative labeled as figures. The taxonomy of figures is, in short, a mess. We have developed a much more rigorous, consistent, and principled taxonomy of figures at Waterloo. See, e.g., Harris (2013:571-575).

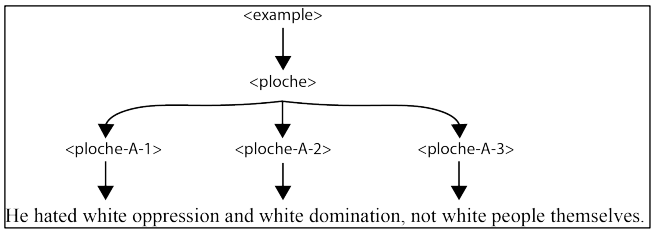

Figure 1: Hierarchical Structure of Markup (18b)

A similar markup language, Hyper Text Markup Language (HTML). has been used in the past for annotating figures. Specifically, JANTOR (Java ANnotation Tool Of Rhetoric) allowed for "manual and automated annotation of files in HTML format" (Gawryjolek 2009; Gawryjolek, Harris, and DiMarco 2009). But XML provides an advantage over HTML; creation of one's own tags and attributes. The ability to parse strings into pre-defined subsections provides the necessary data to analyze figures. XML markup provides strings with the necessary information to allow for figure analysis. Now, using XML we analyze the development process of a suitable markup.

Our original markup focused on the names of tags and did not include attributes. This was adequate for simple figures like ploche. A general markup and an example is shown below:

17. <example>
    ...text...
    <figure-name>
    ...text...
    <figure-element-number>
    ...text...
    </figure-element-number>
    ...text...
    </figure-name>
    ...text...
    </example>

18. a. He hated white oppression and white domination, not white people themselves. (white)

    b. <example>
    He hated
    <ploche>
    <ploche-A-1>white</ploche-A-1>
    oppression and
    <ploche-A-2>white</ploche-A-2>
    domination, not
    <ploche-A-3>white</ploche-A3>
    </ploche>
    people themselves.
    </example>

The container tag <example> marks off the beginning of the text while the <figure-name> tag reveals the beginning position of the figure. The vital tags of this markup are the <figure-element-number> tags which encompass the defining features of a figure. In Example 18b they are <ploche-A-[1 to 3]>. Figure 1 illustrates the hierarchical nature of the markup for 18b. These markers provide information about elements such as letter groups (A-Z) and relative positioning

```
<example>
    <antithesis>
        <antithesis-A-1>Ask not</antithesis-A-1>
        what
        <antimetabole>
            <ploche>
                <antimetabole-A-1>
                    <ploche-A-1>your country</ploche-A-1>
                </antimetabole-A-1>
                can do for
                <antimetabole-B-1>you</antimetabole-B-1>.
                <antithesis-B-1>Ask</antithesis-B-1>
    </antithesis>
                what
                <antimetabole-B-2>you</antimetabole-B-2>
                can do for
                <antimetabole-A-2>
                    <plcohe-A-2>your country</ploche-A-2>
                </antimetabole-A-2>
            </ploche>
        </antimetabole>
.</example>
```

Figure 2: The Full Hierarchical Structure of Example 19b.

(1 to 3). Issues with this markup arise quickly, but the main idea of marking defining elements still has its uses.

There are syntax and semantic issues that form when applying the markup to more figure-rich texts. By analyzing an example (1, repeated here for convenience), we demonstrate the problems.

19. a. Ask not what your country can do for you. Ask what you can do for your country. (your country / you; ask not x / ask x)

    b. <antithesis>...text...<antimetabole>...text....
    </antithesis>...text... </antimetabole>.

A syntax issue arises in Example 19b where one figure starts earlier than another but also ends earlier (a fully formatted example is given in Figure 2.)

The syntax of XML does not allow the interpenetration of tags. When considering this problem it becomes apparent that the tags marking off the beginning of a figure are unnecessary. The key components of a figure are their defining elements such as repeating or contrasting elements (words, sounds).

The semantic complication has to do with nesting XML tags. When one tag is nested inside another, a hierarchy is created. Figures do not always have a priority associated with them, thus requiring another method to avoid forming hierarchies. This is achievable with the introduction of attributes.

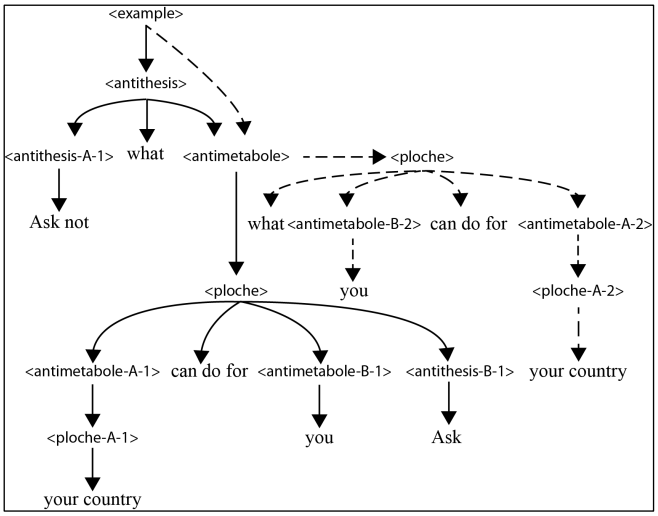

Figure 3: Problems arise from ending the antithesis tag before ending the antimetabole tag

Figure 3 displays the complexity of this version of the markup. The dashed arrows represent the effect of ending the antithesis tag before ending the antimetabole tag. Hierarchy problems also become apparent as <antimetabole-element-number> tags are sub-tags of <ploche>. Another iteration of the markup is clearly needed.

The improved markup recognizes the above problems and attempts to resolve them. It focuses on highlighting the defining elements of figures. A general markup is shown in number 20 (a fully formatted example for this markup is provided in Figure 5, given after the References for purposes of space):

20. <element figure='figure1 [figure2]' lettergroup='[figure1-(A to Z)...]' position=[figure1-(1 to n)...]'>...text...</element>

The removal of the starting and ending tags along with the introduction of attributes prevents both the hierarchy and intersection problems. As observed in Figure 5, if two figures contain the same element, the tags simply have two figure attributes; figure='ploche antimetabole'. Since XML parsers place no precedence on which attribute comes first, no hierarchy is created.

As Figure 4 reveals, the improved markup focuses on tagging parts of strings and providing them with more information. The concept of a hierarchy is also shown to be reduced. This is to minimize miscommunicating figure dependence.

Using attributes also helps to isolate information about a tag providing algorithms with easier access. The lettergroup attribute grants information on which letter grouping the tag surrounds, while the position attribute clarifies the location of the letter group in the figure. Implementing the developed markup on a string containing figures would create the required computational structure for figure analysis.

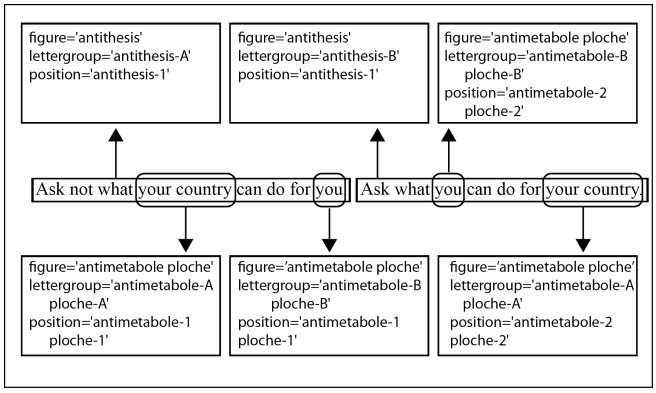

Figure 4: Improved markup tags parts of strings and provides more information

5 Conclusion

The computational uses of rhetorical figures are indisputable. We can clearly see their ability to enhance fields such as author and genre detection, NLP systems, and argumentation mining. We also know how intricate they can become. Stacking and intersecting with one another, many figures can be overlooked as observed in the previous works mentioned here. To exploit their uses, yet overcome their intricacy, a rhetorical figure markup becomes imperative and should be thought of as such.

Our annotation scheme represents the first move in what we hope will be a line of research that others will find profitable to join. The outline of the annotation scheme has been developed, and now the flexibility of XML allows others to improve and customize the mechanism for their own uses. The eventual goal is to develop a markup scheme that provides computationally accessible information for all rhetorical figures.

Acknowledgements

We would like to thank our colleagues at the University of Waterloo, including Elena Afros, Adam Bradley, Ashley Kelly, Isabel Li, and Terry Stewart; our international colleagues, including Marie Dubremetz, Jelena Mitrovic, Cliff O'Reilly, Chris Reed, and James Wynn; as well as the Social Sciences and Humanities Research Council of Canada for financial assistance. Our figure-annotation research is part of an overall project of Computational Rhetoric at the University of Waterloo, organized around a comprehensive computational OWL-based ontology of rhetorical figures.

References

Alliheedi, Mohammed. 2012. Multi-document summarization system using rhetorical information. Master of Mathematics thesis, Cheriton School of Computing, University of Waterloo. [Supervised by Chrysanne DiMarco; Randy Harris, Second Reader.]

Alliheedi, Mohammed, Chrysanne DiMarco. 2014. "Rhetorical figuration as a metric in text summarization." *Proceedings, 2014 Canadian Artificial Intelligence Conference*, Montreal, QC, May 6-9, 2014.

Dubremetz, Marie, and Joakim Nivre. 2015. "Rhetorical figure detection: the case of chiasmus." *Proceedings of NAACL-HLT Fourth Workshop on Computational Linguistics for Literature*, Denver, CO, June 4, 2015.

Gawryjolek, Jakub J. 2009. Automated annotation and visualization of rhetorical figures. Master of Mathematics thesis, Cheriton School of Computing, University of Waterloo. [Supervised by Chrysanne DiMarco; Randy Harris, Second Reader.]

Gawryjolek, Jakub J., Randy Allen Harris, and Chrysanne DiMarco. 2009. "An annotation tool for automatically detecting rhetorical figures." *Proceedings, CMNA IX (Computational Models of Natural Argument)*, held with IJCAI-09, Pasadena, CA, July 13, 2009.

Harris, Randy Allen. 2013. "Figural logic in Mendel's *Experiments on plant hybrids*." 2013. *Philosophy and Rhetoric* 46.4

Harris, Randy Allen, and Chrysanne DiMarco. 2009. "Constructing a rhetorical figuration ontology." *Symposium on Persuasive Technology and Digital Behaviour Intervention, Convention of the Society for the Study of Artificial Intelligence and Simulation of Behaviour (AISB)*, Edinburgh, Scotland, April 2009.

Hromada, Daniel Devatman. 2011. "Initial experiments with multilingual extraction of rhetoric figures by means of PERL-compatible regular expressions." *Proceedings of the Second Student Research Workshop associated with RANLP 2011*, Hissar, Bulgaria.

Strommer, Claus. 2011. Using rhetorical figures and shallow attributes as a metric of intent in text. Doctoral Dissertation, Cheriton School of Computing, University of Waterloo. [Supervised by Chrysanne DiMarco; Randy Harris, Committee Member.]

```
<example>
    <element figure='antithesis' letter group='antithesis-A' position='antithesis-1'>
        Ask not what
        <element figure='antimetabole ploche' lettergroup='antimetabole-A ploche-A' position='antimetabole-1 ploche-1'>
            your country
        </element>
        can do for
        <element figure='antimetabole' lettergroup='antimetabole-B' position='antimetabole-1'>
            you
        </element>.
    </element>
    <element figure='antithesis' lettergroup='antithesis-B' position='antithesis-1'>
        Ask what
        <element figure='antimetabole' lettergroup='antimetabole-B' position='antimetabole-2'>
            you
        </element>
        can do for
        <element figure='antimetabole ploche' lettergroup='antimetabole-A ploche-  A' position='antimetabole-2 ploche-2'>
            your country
        </element>.
    </element>
</example>
```

Figure 5: The Full Hierarchical Structure of Sentence 19, in accord with the tagging specified in 20