Lecture Notes: Scientific Writing

ENG 2020 Intermediate Composition: Science and Technology

Joshua Bowles
Utah Valley University

October 6, 2010

Contents

List of Figures

List of Tables

Part I
Preliminary Comments

This class is unconventional in many ways. The most obvious unconventionality in this class will be the disonance between what you expect to see in this class and what you actually see. That is, you may expect to focus on writing in an explicit and direct way—writing in class everyday, receiving comments from me about how to make a sentence better, and being introduced to models of writing as types to put into practice (or genres, e.g., expository essay, technical report, *et cetera...*) and styles to be emulated. And while these things are certainly discussed and introduced, we do not spend the bulk of our time with them. These more conventional tools function as the necessary background support for our prime objective: analysis. In other words, the conventional tools of the rhetorician are necessary but NOT alltogether sufficient to get us where we want to be.[1]

What you **do** see in this class is an explicit focus on the way we think, how we think, how we process information, how we validate information, how we come to accept the warrant for claims, how we evaluate types of evidence, and so on. The bulk of these topics are covered in a fairly informal way, which means you may not actually be aware of what we are doing until later. The reasoning for this 'unconventional' way of doing things actually comes from following the example of ancient Greek scholars:

$$\text{Speech is an outward expression of thought, and if your} \atop \text{speech is disorganized, then so is your thought} \tag{0.1}$$

We can turn this around by saying that if your thoughts are disorganized, then only by a rare stroke of luck will your speech be organized. Extending this to the modern day where writing is a more widely distributed technology for spreading thoughts and ideas, we can say that good writing only follows good thinking. Therefore, the following stands to reason:

$$\text{If your thoughts are well-organized, then your writing} \atop \text{may also } \mathbf{become} \text{ well-organized.} \tag{0.2}$$

I have let everything hang on the notion of 'organized' or 'well-organized' thought: this is the domain of logic.[2] If you make assumptions in your thinking, you will make them in your writing. If you make generaliztions in your thinking, you will make them in your writing.

The last crucial piece to understanding is this: it takes a very long time to become a good writer—16 weeks is not sufficient time to make you a good writer. But in 16 weeks we can evaluate the way we think and connect this to how we express our thinking in our writing. Also, the best way to becoming a good writer is by doing it over and over again year after year; and even then, you must rely on peer-review and editors.

To sum up, learning how to write is learning how to evaluate thinking. You will become a better writer by becoming a better thinker. But the obverse does not hold: you can employ

---

[1]Much of the Rhetorician's toolkit is introduced in the textbooks. Class lectures presuppose your being introduced to them—without reading the texts you will not grasp why we are doing what we are doing in class.

[2]We will bypass the task of defining '(well-)organized' and simply get down to business.

models of good writing and emulate the sentences, templates, and vocaublary suggested by all the books and still not be a very good writer. The reason for this is simple: by the doing the latter you become a clone. Instead, we begin at the source of writing: the structure of thought. If you cannot think well, no turn-of-phrase, no syllogism, no well-crafted sentence will save you from real mediocrity.

———————————————

Part II
Induction

I have provided a brief excerpt from my own personal notes on mathematics because induction is a very crucial concept. It has a formalized description that has been widely used in just every about every formal system and it is also one of the major types of informal logic used in making arguments (the others being Deduction and Abduction). The concept of induction, as well as deduction and abduction, will inform our use of the concept of INFERENCE throughout the semeseter. In other words, when we say inference, we refer to either induction, deduction, or abduction. I provide this excerpt here to demostrate that types of inference can be expressed in very rigorous ways.

A general definition for mathematical induction is based on the successor function. This is a now classic method made standard by Giuseppe Peano around the beginning of the 20th century; in which the arithmetic of cardinal numbers of axiomatized. Typically, we define a primitive and some kind of binary combinatoric operation (i.e., addition and/or multiplication), and then simply iterate the primtive and combining operation.[3]

A very common model for successor functions and induction is the domain of the whole numbers. The primitive is zero and the combinatory operation is addition (or multiplication). To this model we apply the axioms of Peano. Begin with three primitive (undefined and unmotivated) terms: **number**, **zero**, and **immediate successor of**. The following axioms can be given:

Peano Axiom 1. Zero is a number

Peano Axiom 2. The immediate successor of a number is a number.

Peano Axiom 3. Zero is not the immediate successor of a number.

Peano Axiom 4. No two numbers have the same immediate successor.

Peano Axiom 5. Any property belonging to zero, and also to the immediate successor for every number that has the property, belongs to all numbers.

The last axiom is referred to as the Principle of Mathematical Induction; this reflects that understanding that the last 'axiom' is not technically an axiom; it may also be a called an Induction Axiom Scheme. Peano's axioms can be formally represented as follows.

---

[3]It is hard to overemphasize the importance of the use of induction to contemporary formalization.

0.1   A Formalized Induction Scheme for Peano's Axioms

$$\text{Peano Axiom 3} \quad \neg 0 = sx \tag{0.3}$$

$$\text{Peano Axiom 4} \quad sx = sy \rightarrow x = y \tag{0.4}$$

0.3 says that the negation of zero is a number with a successor, or conversely, if a number has no successor, then it must be zero, or that zero is not the successor of any number. 0.4 says that if the successor of x equals the successor of y, then x equals y, or in other words, given any number, its predecessor is unique. Now we can provide representations for axioms of the binary operations addition, $+$, and multiplication, $*$,

$$x + 0 = x \tag{0.5}$$

$$x + sy = s(x + y) \tag{0.6}$$

$$x * 0 = 0 \tag{0.7}$$

$$x * sy = x * y + x \tag{0.8}$$

0.5 is the identity of addition (any number x added to zero sums to that number x); 0.6 says that any number x plus the successor of y is equivalent to the successor of $x + y$; 0.7 is straightforward, and 0.8 says that any number x times the sucessor of the number y is equivalent to $x * y$ plus the number x.

To the representations in 0.3–0.8 we can add the following instance of the induction scheme, or principle of mathematical induction, from 5.

Given any formula with a free variable x, $P(x)$, we can get the following instance of induction

$$(P(0)\&\forall x(P(x)) \rightarrow P(s(x))) \longrightarrow \forall x P(x) \tag{0.9}$$

What this says is that for the property P that belongs to 0, and for all x with that property, then the successor of x, sx, also has that property, $P(s(x))$; If the latter is the case, then this will hold for all x.

0.1.1   Writing numerals inductively

The common way of writing the Arabic numerals in the set of whole numbers, $0, 1, \ldots, n$, implies the inductive principle and is an efficient way to compact the information in that principle. A more explicit way to write the set whole numbers is the following (I also show the correlation with Arabic numerals):

$$\begin{aligned} 0 &= 0 \\ 0s &= 1 \\ 0ss &= 2 \\ 0sss &= 3 \end{aligned} \tag{0.10}$$

Addition of $2 + 2 = 4$ looks like this (the second representation is in Polish notation and is equivalent with the first):

$$s(s(0)) + s(s(0)) = s(s(s(s(0)))).$$
$$\text{or,}$$
$$+(s(s(0)),\ s(s(0)) = s(s(s(s(0))))).$$

Division in a formalized induction scheme   Induction is defined for binary combinatoric operations multiplication and addition. But we can also define division as (assume the definitions 0.5–0.8):

$$\exists z(x * z = y) \qquad\qquad (0.11)$$

Notice that we can draw a relationship with induction and observation: given n observations of the phenomena x, we may inductively conclude that x is indeed the case for $n+m$ events. Induction is also useful for understanding how we get statistical data: we take a sample of n people and iductively generalize those results to the actual (or larger) population $n * m$; given that sample n is representative. It also works for classification heirarchies: if $a \in A$, a belongs in A, and $\beta \in a$, then we might expect that properties belonging to a also belong to β, or vice versa. These three cases are very complicated and require a lot of analysis that we won't bother to go into—it is enough to know that induction relates to real world thinking and that methods using induction are much more complicated than they first appear to be.

Part III
'Expressives' and 'Argumentation' in Genres of Writing

I this part we will look at 'expressive' words (or content) as well as 'argumentation' words. I will develop some techniques to look at how language is used in large collections of text that belong to three basic genres.[4] These are what I will refer to as (i) `expressdom`: EXPRESSIVE DOMAIN = e-mail, blog, chatroom, and other digital media related text; we will assume that this genre is informal and does not consist of things like e-mail between you and your boss, or between client and lawyer, *et cetera*; (ii) `acadom`: ACADEMIC DOMAIN = transcriptions of university lectures, PhD dissertations, and academic journal articles; and (iii) `rhetdom`: RHETORICAL DOMAIN = a random sample of excerpts from your Analysis and Synthesis papers.

1   What are Expressives and Why Should we Care?

Expressive words are lexical items that contain some emotive content. They do not, however, impact the truth or falsity of a statement (usually called propositional content). For more on expressives in the context of formal pragmatics and semantics, as well as corpus linguistics, see Constant et al. (2009), Potts (2007), Potts (2003), or Potts (To appear). Potts' faculty page has most of these as downloadable pdfs: http://www.stanford.edu/~cgpotts/.

Expressives are relevant to us because we typically want to present our research through writing that is both neutral and negotiates its emotional investment. Through using expressive terms we reveal bias by revealing emotional investment. This is certainly one way

---

[4]A genre of writing is a style, domain, or type. Genres are typically defined in relation to each other. Notice this implies that more diverse your reading experiences, the richer your view of genres that exist. That is, if all you read are newspaper articles, then you have no reason to expect that writing outside of newspapers is any different. And when we write, we emulate what we read. That is, if all you read is newspapers, chances are you are going to write like newspaper writing becuase that is your only model.

to argue against using expressive terms in scientific writing. But there is, I argue, an even stronger reason: as we will see in the examples below, expressive terms do not appear to be compatible with making precise statements. Use of expressive content seems to mask precision and encourge vague reference to the world and data.

2   Some Examples and Data

2.1   Examples

(1)   Expressive words used in this study:

    a. like, damn, soo, wow, crap, super, hell, awesome, totally, really, literally, idiot, stupid, dumb, amazing

(2)   Common expressives

    a. swears, exclamatives, intensives

(3)   The cheetah can run **so** fast.

    a. The cheetah can run **so** 60 miles per hour.

(4)   Pluto is **really** far away.

    a. Pluto is **really** 5,913,520,000 kilometers away.

(5)   Toyota gas-pedal problems are **freakin'** rare.

    a. Toyota gas-pedal problems are **freakin'** 0.00001%.

Notice also that here we are addressing a problem that has come up before: vague language. We replace the terms *fast, far, rare* with measurable quantities. Notice what happens to the expressives – in some cases, problems 3a and 4a, they appear to act like an appeal to/for credibility.

Table 1: Individual counts for 'expressives'

Word	Academic	Ac-mod	Chat
like	5047	(5410)	17010
damn	9	(167)	64
soo	1	(151)	13
wow	15	(30)	224
crap	10	(10)	63
super	20	(21)	10
hell	45	(45)	162
totally	114	(118)	206
really	2666	(2734)	3305
literally	48	(51)	28
idiot	8	(16)	13
stupid	32	(50)	166
dumb	2	(2)	22
amazing	33	(33)	25
awesome	0	(0)	39
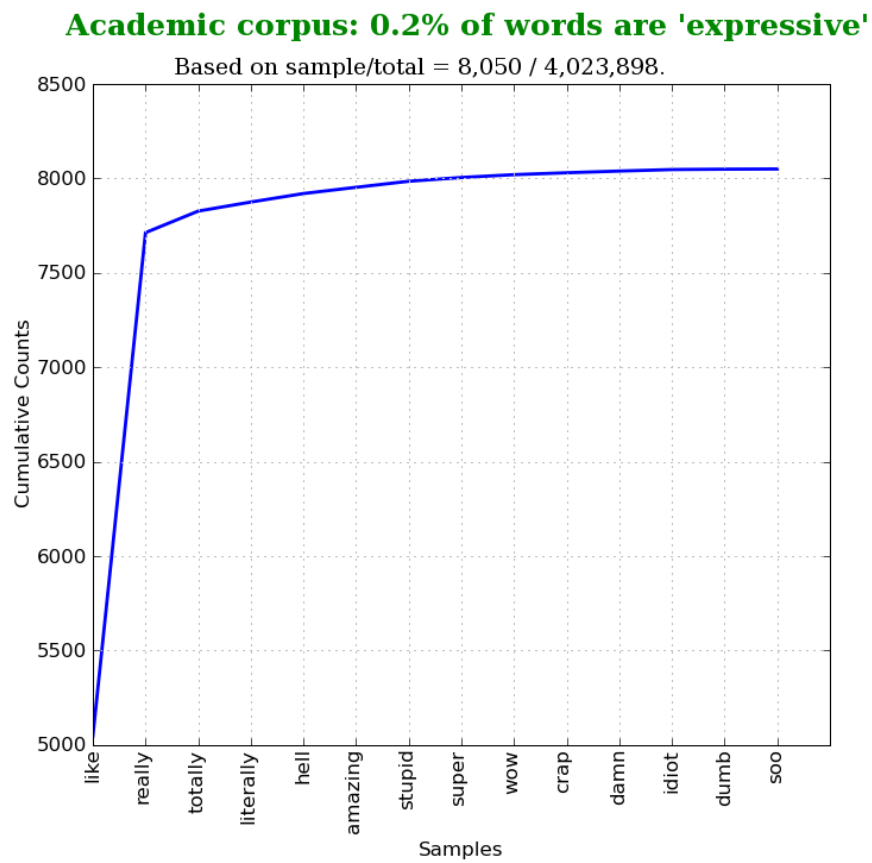
Figure 1: Cumulative plot for acadom

Chat, email, blog, reviews corpus: 1.2% are 'expressive'. Based on sampe/total = 20115 / 1627211.
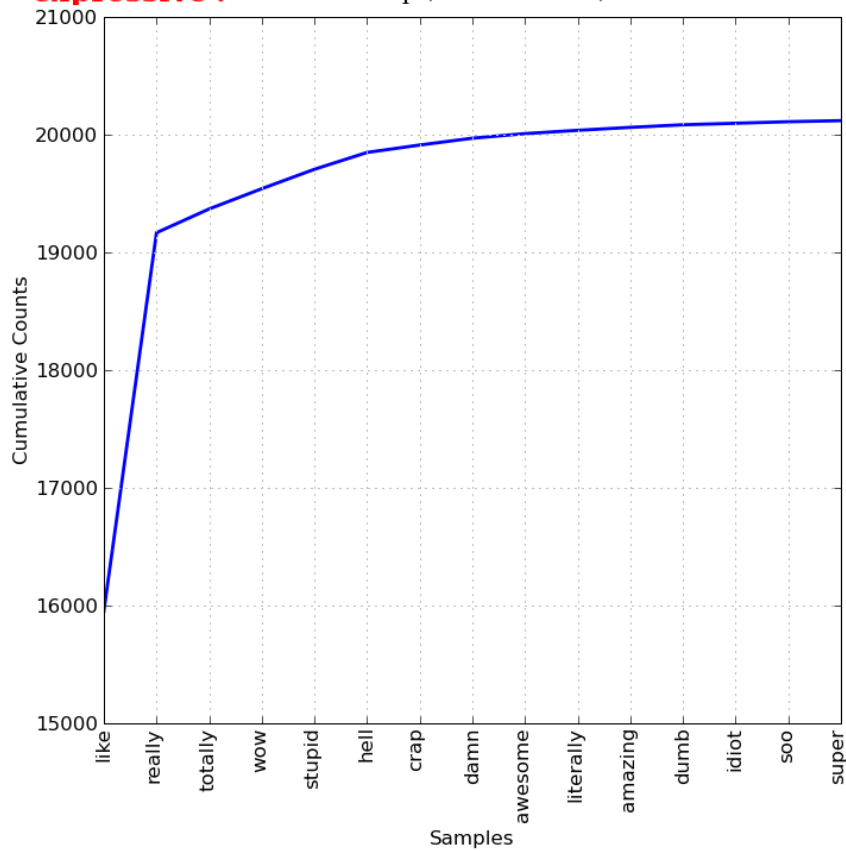
Figure 2: Cumulative plot for expressdom

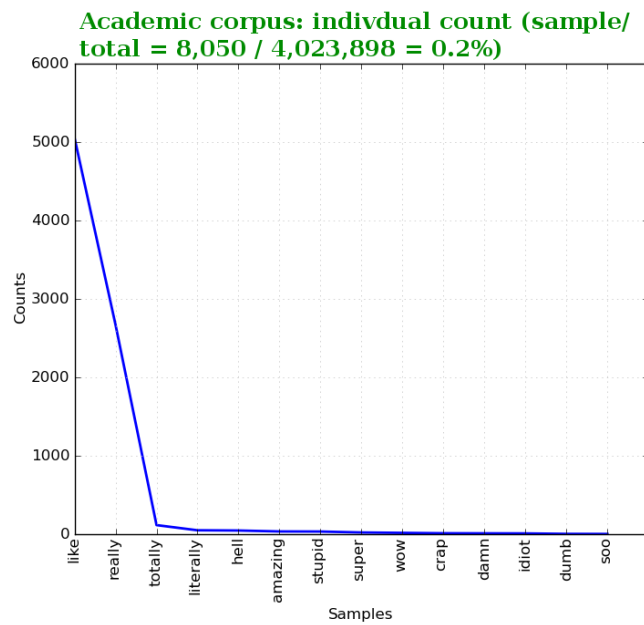**Academic corpus: indivdual count (sample/ total = 8,050 / 4,023,898 = 0.2%)**

Figure 3: Individual counts for acadom

Chat, email, blog, reviews corpus (sample/ total = 20,115 / 1,627,211
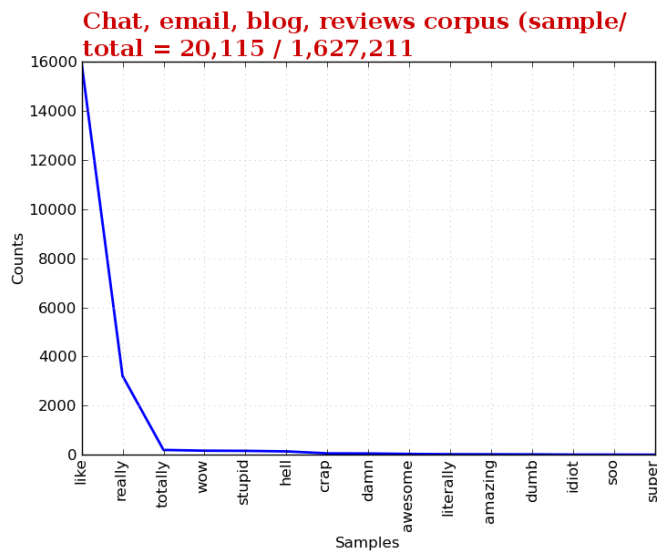
Figure 4: Individual counts for expressdom

3   Description of Method

Frequency counts of the words in example (1a) are made in two different collections of texts. One collection (or corpus) contains lectures by academics and runs a total of about 4 million words; the other corpus is based on collections of emails, blogs, chatrooms, and reviews (i.e., movie, product, restaurant reviews) and runs about 1 million. I wrote some Python scripts to search through each corpora and count the total number of instances of the set of 'expressives' from (1a) (which were lowercased beforehand). The total sample was divided by the total number of word instances for each corpora.

3.1   Corpora

The academic and chat corpora were built by me using pre-existing corpora.

3.2   Academic Corpus

For the academic corpus I used the British Academic Spoken English (BASE) http://www2.warwick.ac.uk/fac/soc/al/research/collect/base/. According the BASE conditions the following must be stated: "The recordings and transcriptions used in this study come from the British Academic Spoken English (BASE) corpus. The corpus was developed at the Universities of Warwick and Reading under the directorship of Hilary Nesi and Paul Thompson. Corpus development was assisted by funding from BALEAP, EURALEX, the British Academy and the Arts and Humanities Research Council." Some additional information about BASE comes the corpus' own description, which I quote in full:

> The BASE corpus consists of 160 lectures and 39 seminars recorded in a variety of university departments. Holdings are distributed across four broad disciplinary groups, each represented by 40 lectures and 10 seminars. These groups are: Arts and Humanities, Social Studies and Sciences, Physical Sciences, and Life and Medical Sciences. The lectures and seminars have been transcribed and annotated using a system devised in accordance with the TEI Guidelines. There is a DTD file which must be kept in the same folder as the corpus files, named 'base.dtd'. The transcription and mark-up conventions are described in the 'BASE manual' document which is in PDF format, and the holdings are described in the Excel spreadsheet, 'BASE corpus holdings.xls'. The token count for the entire corpus is 1.6 million, and the files contain the transcripts of nearly 200 hours of recording.

3.3   Chat Corpus

The e-mail, blog, chat, and reviews corpus was built by me out of 3 pre-existing corpora.

3.3.1   Loyola CMC

The first is Loyola Computer-Mediated Communication (CMC) Corpus, http://cmccorpus.cs.loyola.edu/, which describes itself as

> This site provides access to a corpus of over 800 text samples gathered from test subjects at Loyola College, Baltimore, Maryland, in 2006 and 2007. Twenty-one subjects provide a completely correlated corpus in which each

subject provided their opinion in each of six predetermined topics in each of six genres: blog, chat, discussion, email, essay, and interview.

3.3.2   NPS Chat, Brown, and NLTK

The second collection is the NPS Chat Corpus, http://faculty.nps.edu/cmartell/NPSChat. htm, which was accessed through the Natural Language Toolkit (NLTK), http://www. nltk.org/Home, which is an open source project that uses the programming and scripting language Python. It consists of many modules designed to educate, facilitate, and support development of natural language processing for linguistic data for Windows, Mac OSX, and Linux operating systems. The third collection comes from the 'Reviews' category in the Brown Corpus, http://icame.uib.no/brown/bcm.html. While there are many places to get the Brown Corpus, I accessed it through use of the NLTK.

4   Problems and Caveats

There are a number of caveats and problems that accompany this short demonstration; I will name some here.

Problem 1. *The academic corpus consists of sampling from one pre-existing corpus which is composed entirely of academic lectures.*

This is problematic becuase we can assume with some confidence that one is more likely to use expressives in discussion versus writing—even in academic lectures. This means the percentage of 'expressive' words may actually be lower than 0.2%. Generally, the chat corpus has a better sampling, consisting of three pre-existing corpora I put together.

Problem 2. *The structure of the BASE is such that each xml file consists of the same introduction.*

This is only a problem becuase I did not filter out the BASE introductions to each xml; there a quite a number of xml files. The result is that is the 4 million plus total sample count is not accurate — though I did check to make sure that the 'expressive' words I used did not occur in the BASE introductions.

Problem 3. *The words I use as the set of 'expressives' was compiled quickly by myself with little forethought*

Though it is not reasonable to compile a a list of every single expressive term, we should expect a great deal of thought to go into selection for the set. The choices I made were intuitive and were I pushed to justify some choices over others I probably could not provide satisfactory argument for the choices I made.

5   The Moral

These problems mean that my 4 million total count of the academic corpus is not accurate; 8,050 divided by 1.6 million is 0.5%. The numbers given in class and above are mere sketches of what we expect to find, and, should not be used to make actual arguments about the nature of expressive content in academic language (also given the assumption that discourse is more likely to yield a higher content of expressives, even in university lectures — and the

chat corpus consists only of written material. The more accurate number, 0.5% is probably lower given the discursive nature of hte acadmeic corpus—all things equal, the 0.2% number I give in the graphs, then, is probably not too far off. Nonetheless, it is still not accurate and serves at best as a demonstration.)

What should be taken from this demostration is that we have pretty good circumstatial evidence that expressives are not used a lot in academic language. The reason for this is twofold: the obvious assertion that expressives are 'revealing,' AND, more interestingly, expressives don't seem to modfiy precise statments very well (though they do license vague statements), and consequently, scientists don't use expressives much because scientists are more concerned with precise statements than vague ones.

6   Comparing Academic Terms in Three Corpora

For this demonstration I have taken the corpora and methods used to examine 'expressive' words; see the beginning of part III of the lectures for discussion of expressives. I have extended these methods to allow our own writing to be compared against the other two corpora. Specifically, a number of words from Graff & Birkenstein's (2006: 174-175) have been selected and termed 'argumentation', or `arg`, words. Additionally, the `rhetdom` corpus is brought—remember that this is a collection of randmomly selected excerpts from your Analysis and Synthesis papers.

6.0.3   A Note on Selection Criteria

In selecting excerpts from your papers I followed the following method: before grading or reading you papers I flipped to teh 2nd or 3rd page and picked the first paragraph I saw. I hand-typed this paragraph into the text editor Emacs, which produced a common `.txt` file that I named `excerpts.txt`. I then went through the papers a second round and selected a paragraph from the first page. Next, while I was grading the papers I selected another paragraph: my selection was of both good and bad paragraphs. Section titles were included but author names were not.

I then took the `excerpts.txt` file and sorted it by paragraph using the UNIX command `sort`.

(6)   `sort-R -o OUTPUTFILE INPUTFILE`

I ran the file through this command multiple times and wrote it to an external file called `excerptRandom.txt`.

6.1   Tables, graphs, and images

Table 2: Individual counts for 'argumentation'

arg Word	expressdom	acadom	rhetdom
accordingly	10	16	0
as a result	0	0	0
consequently	2 (**0.0001**%)	18 (**0.0009**%)	◇ **3 (0.0068**%)
hence	21	134	0
it follows	0	0	0
since	482	372	12
so	6174	17546	85
then	2992	5475	86
therefore	102 (**0.0062**%)	◇ **560 (0.0282**%)	3 (**0.0068**%)
thus	103	293	0
as a result	0	0	0
in conclusion	0	0	0
in short	0	0	0
in sum	0	0	0
to summarize	0	0	0
likewise	10	30	0
similarly	11	116	0
although	177 (**0.0108**%)	715 (**0.0360**%)	◇ **22 (0.0502**%)
by contrast	0	0	0
conversely	2	21	0
despite	22	71	0
nevertheless	4	73	0
whereas	99	287	0
TOTALS	10211	25727	211

Table 3: Totals for corpora

acadom	expressdom	rhetdom
BASE = 1.6 million journal = 169,236 books = 123,383 diss = 88,883	Loyola = 1,541,497 NPS = 45,010 Brown reviews = 40,704	Rhetcomp = 43,778
TOTAL = 1,981,502	TOTAL = 1,627,211	TOTAL = 43,778

6.2   Description

For this study I am using the same basic corpora as the expressives study above. I will be a bit more specific about the corpora as we move on. For one, the Academic corpus is called

Table 4: Averages for `arg`

acadom	expressdom	rhetdom
25727 ÷	10211 ÷	211 ÷
1,981,502	1,627,211	43,778
≈ 1.29%	≈ 0.58%	≈ 0.48%
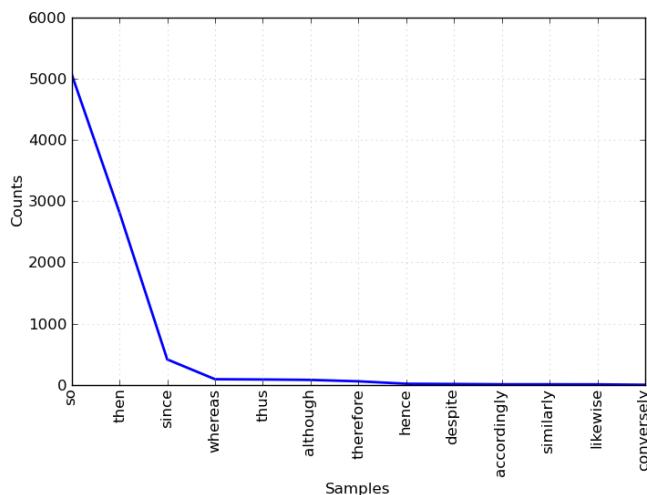
Figure 5: Individual counts for `arg` in expressdom

`acadom`: `aca`(demic) `dom`(ain), the Chat corpus is called `expressdom`: `express`(ive) `dom`(ain), and the corpus compiled from student writing is called `rhetdom`: `rhet`(oric) `dom`(ain).

Here we are looking at what I will call 'argumentation' words. I have taken these from Graff & Birkenstein's (2006: 174-175) index on templates, namely the Commonly Used Transitions. I specifically took small phrases or single words for ease of search and presentation (but there may more to come in the future that looks for distributional evidence of the templates themselves).

Totals for corpora can be seen in table 3; the set of words in `arg` is in table 2.

6.2.1   A Last Note

For this look at 'argumentation' words I have added to the Academic corpus, `acadom`, but the Chat corpus, `expressdom`, has stayed the same. Although the Academic corpus has the same name as it does in the analysis of expressive, the actual content is different. Also, I still have not filtered the BASE portion of acadom — that is, filtered out the intro on every xml file in BASE — and so total counts will not be accurate. For this reason I give the token counts in table 3 and have calculated all arithmetic means according to the totals in that table.

When we account for the redundancy in BASE (see sections 3.2, 4, and 5; also problem
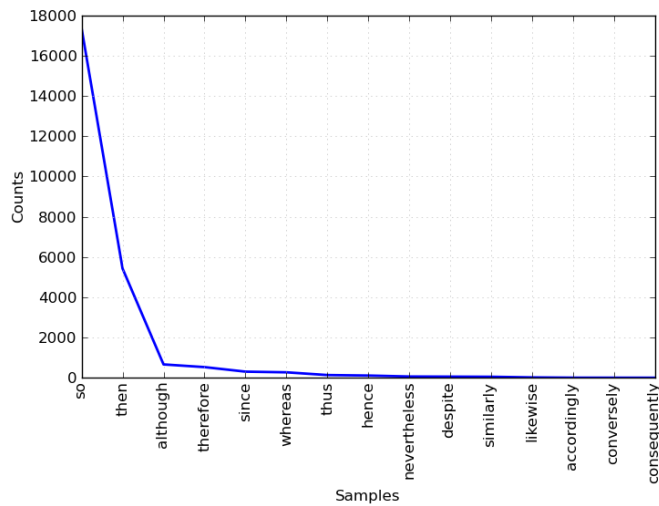
Figure 6: Individual counts for `arg` in acadom
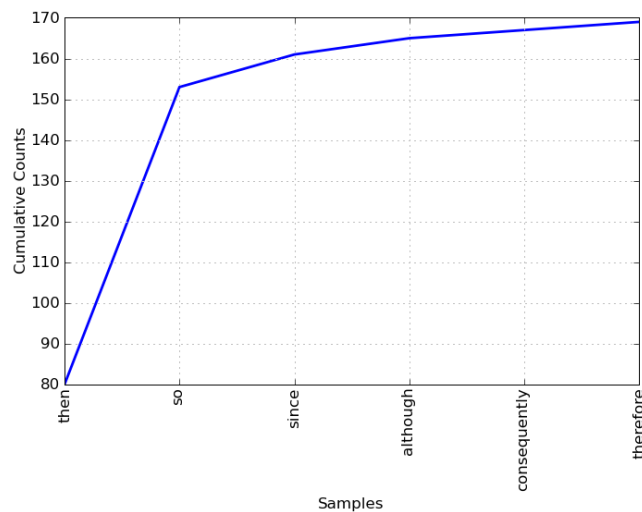
Figure 7: Cumulative counts for `arg` in rhetdom

2 for discussion), taking the BASE count at 1.6 million instead of 4.4 million, we see that the sample of 'argumentation' words for the total count of `acadom` is actually around 1.2% (25,144/1,981,502; the expressdom sample/total is 8,691/1,627,221).

6.3   Code Samples

Here are a snippets of Python code used to analyze various data in the corpora.

```
>>> len(expressdom)
1627211
>>> percentage(len(arg_word),len(expressdom))
0.53410405903106606
>>> arg_word2 = [w.lower() for w in acadom if w in arg]
>>> len(acadom)
4405400
>>> percentage(len(arg_word2),len(acadom))
0.5707540745448767
>>>
```

Figure 8: Piece of Python code showing sample/total percentage

```
######################  ARGUMENTATION  #######################
#################freq and tab for argumentation words##########
arg = ['accordingly','as a result','consequently','hence','it follows','since','
so','then','therefore','thus','as a result','in conclusion','in short','in sum',
'to summarize','likewise','similarly','although','by contrast','conversely','des
pite','nevertheless','whereas']

#plot frequency for arg words in {expressdom}
def arg_plot(expressdom):
        fdist101 = FreqDist([w.lower() for w in expressdom if w in arg])
        return fdist101.plot(cumulative=True)

#print a tabulated frequency in {expressdom}
def arg_tab(expressdom):
        fdist202 = FreqDist([w.lower() for w in expressdom])
        for i in arg:
                print i + ':', fdist202[i]


#plot frequency for arg words in {acadom}
def arg_plot(acadom):
        fdist1011 = FreqDist([w.lower() for w in acadom if w in arg])
        return fdist1011.plot(cumulative=True)

#print a tabulated frequency in {acadom}
def arg_tab(acadom):
        fdist2022 = FreqDist([w.lower() for w in acadom])
        for i in arg:
                print i + ':', fdist2022[i]
#*************************************************************

#================cfd==================================
```

Figure 9: Piece of Python code for argumentation

7   Sample Introductions: what to shoot for

In this class we are defining a very strict style of sci-tech writing: one that is clear, concise, and gets to the point in an efficient manner. This kind of style is fairly counter-intuitive to what one is normally taught about writing. For example, in this style, repetition is often a good thing—it highlights the major points (and reminds the reader of the major points), and it also reflects consistency and gives the reader a sense that you know what you are doing.

The following excerpts are from academic papers published in peer-reviewed journals. The topics represent issues within my field of research, namely formal and computational lingusitics. All selections were pretty much random (whatever I had in my current reading file and my desktop). The excerpts are all from the last paragraph of the introduction (and in some cases I provide the last two paragraphs).

The purpose here is to get a sense of the essential pieces of what many academic introductions include. Notice that the introductions below are on averge 1.6 pages long (single-spaced); the papers are average 43.75 pages long (I added in one paper not below; intro = 2 pgs, paper = 78).[5] The excerpts below are average 176.25 words $\left(^{705}/_4\right)$

7.1   Etxepare (2010)

(intro = 1 pg; paper = 24)
"The paper is organized as follows: section 2 compares the Spanish main que-clauses with hearsay evidentials, taking as a basis Fallers (2002) careful study of theQuechuan reportative evidential. It is shown that both phenomena share a number of semantic and syntactic properties. Section 3 shows that, despite those commonalities, hearsay evidentials and Spanish que-clauses differ in several important respects, and cannot be lumped together. Section 4 introduces the proposed analysis of main que-clauses, and suggests a possible connection with so-called quotative constructions cross-linguistically. Section 5 develops the syntactic analysis introduced in the previous section. Section 6 puts the proposed analysis in a comparative perspective and suggests a way to accommodate syntactic variation in the domain of quotative constructions. Section 7 presents a remaining case of reporting main clause that does not immediately fall under the proposed analysis, and provides a minimal extension of the analysis. Section 8 concludes."

7.2   Paris et al. (2009)

(intro = 1 pg; paper = 38)
"This paper presents Myriad, a platform for information delivery that is based on NLG principles, in particular discourse planning principles, while being driven by engineering concerns. Myriad integrates a service-based architecture for data access, not unlike those provided by mash-up tools, with an NLG-based discourse planner that aggregates and tailors that information for the user in context. The paper starts with related work in NLG and discourse planning and then describes the Myriad platform in detail, using a specific application as a running example. Particular attention is paid to reusable mechanisms for data access and discourse planning. The paper then illustrates the use of the platform in the context of two additional and different applications and concludes with a discussion of

---

[5]5 page double-spaced paper $\approx$ 1,385 words :: $\frac{277\ words}{page}$ :: about half of the first page should be intro, which means your introduction **should** 277 to $\frac{277}{2} \approx 138$; one-half to one page. Why?

the utility of the Myriad platform and its implications for integrated delivery platforms in general."

7.3   Zesch & Gurevych (2009)

(intro = 2.5 pgs; paper = 35)

"In this article, we focus only on knowledge based methods for computing semantic relatedness, as opposed to distributional methods (Weeds 2003). However, distributional methods showed competitive performance on some evaluation datasets (the interested reader may refer to Mohammad et al. 2007). The article is structured as follows: In Section 2, we describe the state of the art in computing semantic relatedness as well as the knowledge sources used. Section 3 introduces the evaluation framework that was used for our experiments, whose results are presented in Section 4. Section 5 briefly explains the software tools that were used to conduct the experiments. The article concludes with a summary and future research directions in Section 6."

Remark 1. *Note the references below, formatted in the S&P style (Semantics and Pragmatics), and its use of DOIs: Digital Object Identifier — a unique id for digital archiving and very important for online journals.*

8   Informal Logic

9   Definition of Terms

The syllogism has been a form of logic since Aristotle (about 2,000 years ago) and has been studied instensively for most of that time[6]. Note that the use of the term **logic** here is technical. It does not have the normal meaning, roughly of "common-sensical," as expressed in the phrases

(7)   Let's be *logical* for a minute.

(8)   That is a *logical* proposal.

(9)   The board of trustees' policy is not very *logical.*

   Of course, these phrases may also refer to the logical structure of the argumentation and reasoning of the speakers, and there certainly are links between common sense and technical logic. But as we'll see, this is not always the case. Needless to say, informal logic deals with the abstract structure of argumentation and reasoning in normal, every day, language. However, informal logic does result in counter-intuitive rules, or laws. This is why "logic" cannot always have its normal, every day, meaning.

9.1   A Primer on the History of Logic

Informal logic has been around for a long time, and still is. At the end of the 19 century it started to change, mostly in the hands of George Boole. Boole was able to construct an algebraic system on the back of syllogistic reasoning, thus opening the door to the creation of what is called Mathematical Logic. As a side note, Boole's algebraic system was also the foundation of a binary digit algebra that was later to find itself at the center of the computer revolution.

   For a while Mathematical Logic (also called symbolic logic, or even first-order predicate calculus—though the latter is only a a kind of system within symbolic logic) ruled the day. Many logicians saw no real value in doing work on informal logic, especially when philosophers were also doing their own work on logic (called Philosophical Logic). Recently, there has been a reivival of informal logic, due mostly to the current pedagogical interests in rhetoric-composition studies and the near universal requirement for undergraduate students[7] to take two sememsters of writing and research. Informal logic begins with the paradigm example of traditional, classic, logic: the Aristotelian categorical syllogism.

9.2   Main Definitions

Most writing contains some form of syllogism, but it is usually a complex mixture of various forms. In other words, it's not always easy to find a syllogism in a text. Here is the definition.

Definition 1. SYLLOGISM*:*
A logical form of argument in which the conclusion follows directly from the premises.

---

[6]Barker (2003) says there are "64 possible kinds of premises that the 256 possible forms of the syllogism can have" (pg. 51). He also notes that it would take 32 Venn diagrams to show them.

[7]Mind you, near universal in American colleges and universities, I don't know what European or Asian systems require.

Subdefinition 1. *Barker (2003)*
"An argument is a categorical syllogism (or syllogism, for short) just in case it consists of three categorical sentences containing three terms in all, each term appearing in two different sentences" (pg. 44)

Subdefinition 2. *Boole (2005 [1854])*
"By syllogism is meant the deduction from two such propositions having a common term, whether subject or predicate, of some third proposition inferentially involved in the two, and forming the 'conclusion"' (pp. 237-238).

Subdefinition 3. *Aristotle (1966)*
"A syllogism is discourse in which, certain things being stated, something other than what is stated follows of necessity from thier being so. I mean by the last phrase that they produce the consequence, and by this, that no further term is required from without in order to make the consequence necessary. I call that a perfect syllogism which needs nothing other than what has been stated to make plain what necessarily follows; a syllogism is imperfect, if it needs either one or more propositions, which are indeed the necessary consequences of the terms set down, but have not been expressly stated as premises" (pp. 95-96).

Next is the Premise. Most real world syllogisms contain many premises, while the classic syllogism usually has only two.

Definition 2. PREMISE:
The propositions that come before the conclusion.

Subdefinition 4. *Aristotle (1966)*
"A premiss [sic] then is a sentence affirming or denying one thing of another. This is either universal or particular or indefinite" (pg. 95).

Finally, the conclusion. Note that in a syllogism the conclusion is completely determined by the information in the premises. There is **no new** information in the conclusion.

Definition 3. CONCLUSION:
The final proposition that is derived from the premises as a necessary consequence.

9.3   Some Other Helpful Definitions

Definition 4. PROPOSITION
We can think of propositions as sentences that state a suppossed fact, or simply state anything.

Definition 5. DERIVE
*We can think of the term **derive** as referring to a mechanistic procedure. That is, if a numerical sum is derivable from it parts, then we assume there exists a procedure for stating what that sum is; even though we are not required to say what that procedure is.*

9.4   Some Remarks on Definitions

Remark 2. *On syllogism:*
*Note that the premises do not have to correspond to any known facts or definitions in the world. A syllogism is a **closed** form in the sense that any inferences to be drawn from it depend entirely on the form itself. You may already begin to see weaknesses in the application of the syllogism.*

Remark 3. *On conclusions:*
In most cases of a syllogism it should be obvious what the conclusion is. For simple syllogisms there should be no need, hypothetically, to have to state the conclusion. This is because it is derivable from the premises, and therefore, can be done mechanistically.

Remark 4. *On the term 'derive:'*
Although the example given for a derivation above is mathematical, it works as a good analogy for the notion that a conclusion must follow from its premises. In other words, the sum 4 must follow from $2 + 2 = x$ (in the normal mathematical world).

Remark 5. *Boole (2005 [1854]) criticizes universality of syllogistic form:*
"It is maintained by most writers on Logic, that these processes [Syllogism and "Conversion"], and according to some, the single process of Syllogism, furnish the universal types of reasoning, and that it is the business of the mind, in any train of demonstration, to conform itself, whether consciously or unconsciously, to the particular models of the processes which have been classified in the writings of logicians" (pg. 238).

10 Examples of Syllogisms

Syllogism 1. *The classic example*

All humans are mortal
Socrates is a human
Therefore, Socrates is mortal.

Syllogism 2. *Cats*

All felines are spineless
All cats are felines
All cats are spineless.

10.1 More Detail

Syllogism 3. *Classic*
All humans[8] are mortal[9]—`Major Premise`
Socrates[10] is a human—`Minor Premise`
Therefore, Socrates is mortal.

Notice that the conclusion is composed **only** of elements from the premises. Namely, the minor term *Socrates* and the major term *mortal* are now in a Subject-Predicate relation.

Remark 6. *On Frege's Advance:*
One of the major advance that Gottlob Frege made in logic was the re-formulate the subject-predicate relation of propositions in terms of a function-argument relation. At the time this was a tremendous leap because it meant that propositions could find epxression in mathematicized terms as functions. This kind of re-formulation has found great success and wide-spread adoption in the modern linguistic approach of formal semantic theory, such as Chierchia & McConnell-Ginet (2000) or Heim & Kratzer (1998).

11 Problems and Excercises

Look at the following syllogisms[11] and answer the questions as either Yes or No.

(10)   Do the conclusions follow from the premises?

(11)   Are the conclusions true in the real world?

(12)   Are there any examples where the conclusions follow from the premises, but either the conclusions or the premises are not true in the real world?

    a. Which examples are these?

    b. How do you know the premises/conclusions are not true?

---

[8]Middle term.

[9]Major term.

[10]Minor term.

[11]Many of the examples are adapted from Barker (2003). Also, a syllogism, both technically and classically, only has two premises. When it has more, or an argument is composed of a series of syllogisms, it is called a **sorite**, a Greek word meaning 'pile.'

Syllogism 4. *Pakistani*
All Pakistanis are Moslems
No Sinhalese are Moslems
Therefore, no Sinhalese are Pakistani's.

Syllogism 5. *Accidents*
Accidents are frequent
Getting struck by lighteing is an accident
Therefore, getting struck by lightening is frequent.

Syllogism 6. *Mormons*
All Mormons are pious persons
No Samoans are Mormons
Therefore, no Samoans are pious persons.

Syllogism 7. *Identity*
All cells die and regenerate,
X is composed of cells,
(Therefore,) X is always dying and regenerating,
X is always dying and regenerating,
There exists an identity which is a property of X,
Therefore, The identity of X is always dying and regenerating.

Syllogism 8. *Numbers*
Some prime numbers are integers
All rational numbers are real numbers
All integers are rational numbers
Some prime numbers are real numbers.

12   Fallacies

There a many types of fallacies and there is no need to go into all the various kinds here. Instead, we will look at some of the more common ones. Before categorizing and defining the various fallacies, look at some examples.

  (13)   DRINK I SAY!
        A - You should not drink liquor.
        B - Why do you say that?
        A - Because God does not like it.
        B - How do you know that?
        A - My scriptures tell me so.
        B - But how do you know the scriptures are right?
        A - Because everything in the scriptures is right.
        B - How do you know that?
        A - Because scripture is divinely inspired.

      B - But how do you know that?
      A - Because scripture itself says it is divinely inspired.
      B - Yes, but why believe that?
      A - You have to believe the scriptures, everything in it is right!

(14) DIRECTIONS FROM NOWHERE!
How do you get to Hyde Park? There is a straight road from Blythe Park to Hyde Park. Blythe is exactly 100 miles north of Hyde. Gorky Park is on the way to Hyde and is exactly 20 miles south of Blythe. Green Park is also on the way to Hyde and is 80 miles south of Gorky. Once you get to Green park you are only fifteen miles North of Hyde Park.

(15) I'M SICK OF THIS!
If people are not deathly sick, then healthcare unnecessary.
And if people are deathly sick, then healthcare is ineffectual.
Now, people are either deathly sick or they're not deathly sick.
Therefore, healthcare is either unnecessary or ineffectual.

(16) MY FRIEND SAID SO!
My friend always speeds and never gets a ticket. That means that either police are not giving tickets for speeding or we can get away with speeding. Either way, I'm gonna start speeding.

(17) YOU SOCIALIST!
The U.S. government wants to give healthcare to everyone. Only socialistic systems give healthcare to everyone. Therefore, the U.S. is going to become a socialistic system.

(18) HE'S NOT MUSLIM, HE'S A DECENT GUY!
That guy over there is a Muslim. No he's not. I know him and he's a decent guy with morals and everything.

Definition 6. *Barker (2003)*

> *"Any logically defective argument that is capable of misleading people into thinking that it is logically correct" (pg. 260).*

Remark 7. *Note that you do not have to intentionally deceive someone. A fallacy can arise from a mistake in reasoning. That is, one can intend to to be logically correct but through defective application of reasoning principles still mislead people. Of course, one can also only pretend ignorance when caught in a fallacious deception.*

12.1   Fallacious Reasoning: How to be a Con-artist

A con-artist exploits momentary lapses of good judgment and sound reasoning. A scam is the mechanism by which the con-artist propagates their deceit. Apparently, this is a huge problem—but it should not be. I am constantly *shocked* at the lack of good judgment shown by people who have been conned. Expectedly, these same people recognize their lapse of good judgment in hindsight—but clearly too late. And as I constantly say to my wife (who is also shocked and appalled) "I hope that person (the one who got conned) is not a college graduate. If they are, then their education failed them." For this reason, I feel compelled to provide a small lecture on the "art" of scamming. This "art" consists of nothing but the ability to take advantage of fallacious reasoning.

Below are some proposals. We are going to look at them and discuss certain assumptions that, once clear, should prove without a doubt that they are "too good to be true." I also encourage you to contribute any scams you have come across—I am sure you have received some sort of e-mail scam at one time or another[12].

(19) You receive a check in the mail for $4,000 with the following directions:

    a. Welcome. You have been selected to participate in an elite secret shopper program. The check you have received is real!! In order to get your free money simply follow three easy steps.

    b. Deposit the check and wait for it to clear.

    c. Wait for an e-mail from us to tell you where your shopping assignments are.

    d. Once you complete the shopping assignments, you will send us two processing fees (one for $1500 and one for $2300) and you are done!

(20) You receive a letter in the mail from Publisher's Clearing House informing you that you are the winner of a special random search award: your name was picked randomly from a list of millions and you have won $5,000!!! You just need to send a processing fee of $600 to receive your check.

(21) You receive an e-mail from the IRS asking you to update your information. The need you social-security number along with credite card numbers and the security codes. This information will allow the IRS to quickly process your stimulus package reward based on your credit debt.

(22) You receive an e-mail from the Prince of Congolese. His country is in turmoil and the government is collapsing. He and his extended family must escape soon. But there is one problem! The Prince's money (over 6 million U.S. dollars) is in a Swiss bank account and he cannot withdraw it unless he has a co-withdrawal agent. Would you please act as this agent, you will be rewarded!! Please respond quickly, thank you for saving the Prince's life. Yes, one more thing, there is a small processing fee of $6,000 and the Prince has no money at all!!! If you can provide this fee you will receive the fee back, plus 25% interest. Please help!!!!

(23) You receive a text message from a friend. She is major trouble. She has been mugged at gun point! Please wire $500 dollars, and make sure to spell the name right so that she is the only one who can receive it (sorry she can't call you in person, her

---

[12]Notice that con-artists can only stay in business if they can sucker people. The fact that scams are so prevalent in our society is an indication that such scams are actually working and the people proposing them are actually making money. What is so horrific is that these scams take little more than a minute of thought to reveal their clear absurdities/fallacies.

cell phone is acting weird and will only let her text message). Please help, she is stranded and scared!

(24)   An investor with an good reputation wants you to invest in his stock portfolio. The return is an amazing 12%!! This is a great opportunity and only a select group has been hand-picked to invest. Think of it, with any other investment you will see at best 2% - 3% returns—at best! With Playdoff investment you are guaranteed 10% - 12% returns over the life of you investment, perhaps even more.

(25)   You see a job advertisement: Work from home. Make Money!! Make your own hours. Tired of living up to someone else's schedule for hardly any money? Data entry for $35 an hour. Simple data entry for websites. No experience necessary, will train. Reply to `www.greenenvyjobs.com`.

13 Terrence Parsons on What an Argument Is (and is not)

Parsons (1996) discusses the basic and intuitive notion of what an argument is. This article was published in *The Journal of Philosophy*, and so, we can assume he is talking to philosophers—people who, we might also assume, already *know* what an 'argument' is. The first questions that might entertain us are below.

(26)   Why to do philosophers need to know what an argument is?

(27)   Why does Parsons need to tell philosophers what an argument is?

(28)   Are we philosophers? (If not, why should we care what an 'argument' is!)

Why do we need to know what an argument is? Surely I can argue with my parents, or girlfriend, or wife. But is this what Parsons means my 'argument?' **No**. There is a parallel here with the use of the term 'logic' discussed earlier. Here, the technical term 'argument' can be satisfied by the following:

Definition 7. ARGUMENT
An argument refers to the abstract structure of reasoning employed in the process of analyzing various concepts, ideas, opinions, ... It may be dialogical (requiring two or more speakers) but it does not have to be. In any case, it is dialectical[13].

With all this in mind I turn now to some lengthy discussion based on details and quotes in P's paper.

13.1 I see that I don't really see

Parsons (1996) begins his paper with the following bit of hypothetical (but all too real) statement.[14]

> "I see what your premises are," says the philosopher, "and I see your conclusion. But I just don't see how you get there. I don't see the *argument*."

Directly following this hypothetical piece of conversation he goes on to elaborate

> We hear such comments often. They indicate that there is a notion of "argument" in philosophy in which an argument does not consist just of premises and conclusion; it has additional structure. This distinguishes the notion of argument in philosophy from the technical notion most commonly found in logic texts, where an argument is an ordered pair consisting of the premises and the conclusion. The philosopher's argument is something with more structure, more akin to the logician's notion of *derivation*: a series of statements with intermediate steps providing the transition from premises to conclusion.

---

[13]Thesis : Antithesis : Synthesis⟶Thesis : Antithesis : Synthesis...
[14]Page numbers refer to the Word document retrieved from Parsons' UCLA faculty site: http://admin.cdh.ucla.edu/facwebpage.php?par=91

Notice here that we have another definition of *derivation* that can be compared to the one given previously. Also notice, the Parsons here catches the informal and intuitive sense that we all have when writing papers: Sure the syllogistic logic is cool, but nobody writes— let alone talks or thinks—like that; at least not naturally and intuitively with little to no training! It may be one thing to be able to put your research into a (series of) syllogism(s), but it is another thing to string those premises together in a coherent prosaic heap. And in fact, you may be thinking, if good arguments really were just **sorites**[15], then why are we not taught to write research papers that way? Clearly, there is more to argument than sound premises and valid derivations of conclusions from them.

Remark 8. *For example, in most writing the conclusion is made explicit at he beginning of the text, usually in the abstract and/or introduction. You then proceed through a series of premises and justifications (and evidence) in order to reach your conclusion in a justifiably satisfactory way.*

Here we can see that Parsons is not only concerned with the holistic form of arguments, but in the relationship between the logic and the language of arguments[16]—and he wants to get at what falls between the cracks of normal deductive reasoning. Parsons, it seems, assumes that even philosophers suffer from bad reasoning and fallacious thinking in their written and spoken texts. And in fact, this assumption can easily be verified. P is also concerned with, it seems, the status of INFORMAL LOGIC as a tool for the 'stuff between' the premises and conclusions of typical syllogistic reasoning. Obviously, he does not limit his sense of logic to the syllogism—it merely stands as a concrete example of (stereo)typical deductive reasoning and logical structure. We might assume from all of this that P sees informal logic as taking up the ground between the premises and conclusions. It forms the prosaic heap that gets us from introducing premises, explaining them, justifying them, and deriving the conclusions. In fact, informal logic is what is doing the 'derivation.' In just the same way that we might say it is number theory (or recursive definitions, or the definition of the successor function, or inductive mathematical definitions) that do the 'derivation' in the arithmetical sequence "2 + 2 = 4" in the example for derivations given earlier.

13.2   Interpreting and Assessing are like Analysis and Synthesis

Interpretation is always subjective. It cannot be otherwise. Even in such fields as mathematics or physics, the form of a proof or the observation of some phenomenon are subject to interpretation. For example,

(29)   There are many ways to solve '2 + 2 = x,' but some are more elegant, efficient, and simply more beautiful than others. You don't hear people commonly refer to the solution of addition problems in these ways, mostly because people who do solve them are not mathematicians. But in high-powered mathematics where abstraction is the norm and practical application is of little to no consequence, elegance is a powerful evaluation tool. But 'elegance' is subjective. (Or is it? If you think elegance is not subjective, then you are interpreting the nature of mathematics in specific way: i.e., constructive or intuitive versus Platonic or realist. These interpretations about the reality of mathematics are matter of personal evaluation.)

---

[15]Series of syllogisms.

[16]Not surprising, given that Parsons could be classified as a Philosopher-Logician-Linguist, a PHILOGUIST, if you will permit.

(30)   Quantum effects and phenomena have typically found two distinct interpretations that have a bearing on the nature of reality. Currently, we can see two modern interpretations in the form of the Multiple Universes model and the Single Universe model. These are extended to, or underly, arguments for the general design features of Order in Reality, call it Ω: Did GOD create Ω by design, or is Ω a product of nature. More concretely, (i) did God design beehives, or design bees to make beehives—or even, do bees make beehives because the are the most efficient way to pack lots of material in small spaces, or (ii) is the shape of beehives a natural product of packing stuff together, i.e., reality is such that when you try to pack a lot of stuff together you tend get certain shapes and designs, like the beehive.

Recognition of the subjective nature of interpretation is not an allowance to BE purely subjective. Instead, it acts as a *checks-and-balances*—if we know that interpretations cannot help but be subjective, we then ask, How much subjectivity is in my interpretation? Although there is not way to quantify such an answer, asking in the first place is what is crucial here. Recognizing a subjective bias does not necessitate the reflex to eliminate it; only to negotiate it. That is, once I realize I have a subjective bias towards the notion that nature is not random—that there must be some underlying principle to the apparent probability of quantal reality—then I can continue with my argument knowing full well that I am going to require an extra stringent standard of evidence for quantal behavior. If I make my bias epxlicit, and seek to be fair in my stringent standards, then I am not in trouble. That is, there is nothing wrong with being critical about things, and thus, requiring stringent standards of proof. As long as I make clear my reasons for requiring extra-strength proof, then I can continue on. But for the quantum scientist, some degree of weakening the standards for proof is needed in order to explore various hypotheses. In this sense, then, I am actively engaging in the academic process—I am the extra strength counter-weight evidence, or the check-and-balance, to the quantum scientist's weakening standard of evidence. I am the critic. The scientist needs me. But she also needs to have a fairly weak sense of evidence in order to experiment and hypothesize. Once an experiment goes well, or a hypothesis seems to work out, she must strengthen her standard for evidence. But since she doesn't really spend her time worrying about how strong her standard must be—instead she is just trying to make her theory work—she can turn to the critic for help. Obviously, she is'nt going to adopt all the extra-strength standards of evidence the critic has set forth, but she knows if she can meet some of the criteria the critic has set forth, then she is making progress. Thus, the critic helps make progress possible bu giving an impossible standard of evidence or confirmation for the scientist to meet. But if the scientist can meet some of those standards, then we might call that progress. Of course, defining *some* is also a matter of interpretation.

For example, making quantum computers is a hard thing. No one has yet made one compute for a longish period of time (not even for 24 hours). But this (perhaps impossible) requirement does not hamper scientists from experimenting. In fact, simply making a quantum computer run for a few minutes is considered progress. Clearly, this does not meet the criterion set forth, but it satisfies if partially and this is good enough to call progress. another interesting issue here is that a scientist may be their own critic, setting standards and goals that they may never actually accomplish—like establishing lunar communities. In one sense, you never know until you try. But in an even greater way, setting high critical standards gives an interpretive measure for progress. In this sense, I would think that all successful scientists, philosophers, writers, and researchers are their own critics. But you cannot be your own critic unless you can recognize your own subjective interpretations—to

which your self-critic would offer contrary interpretations. This is what P is getting at when he refers to interpretation, and especially when says things like

> The first step, interpreting the text, is a sophisticated scholarly task. It is typically underdetermined by all available evidence, evidence must often be balanced against counterevidence, and there may be an ineliminable element of subjectivity to it. The second step is the logical task of assessing an argument. This is mostly clear and objective.

The "mostly clear and objective" *assessment* of an interpretation does not need rely so much on subjectivity. It can proceed by relying on models of logic—i.e., artificial languages that have been constructed on purpose to contain mathematical rigor and delete unfortunate structures in natural languages, like ambiguity, in order to investigate notions like 'true,' 'false,' 'contradiction,' and 'consistent.'

P introduces what he calls an 'ur-argment' (or 'source-argument'), which is the basic structural 'thing' we find when looking for an argument in a text—either spoken or written. This ur-argument may be incomplete and might require the reader/listener to fill-in gaps and steps. P states that the ur-argument is what is interpreted by the reader. It signals to the reader that there exists a more refined argument: it is an overt manifestation, he says, of a refined argument. The refined argument has a more developed and definite structure, but is mediated by our first introduction to the ur-argument, which we must interpret. Such interpretation, P remarks, goes

> beyond the ur-argument, seeing it [the ur-argument] as the overt manifestation of a refined argument, one with a more developed and more definite structure, (or sometimes seeing it as an equally good manifestation of several such developed arguments). This involves filling in steps that are not articulated in the ur-argument, and it involves clarifying meanings, e.g. so as to resolve ambiguity (pp. 2-3).

Remark 9. *For our purposes, we can think of the abstract and the introduction as the ur-argument. It asserts multiple things (claims, warrants, evidence, opinions, conclusions) without really supporting such assertions. In this way, the reader is left to interpret such assertions, waiting until the more fully developed body of the argument for your (the author's) evidence/justification/support.*

13.2.1 Refined Argument

A refined argument must meet the following conditions.

Setting

Set of assumptions about the world; assumed rules, principles, propositions. (**Do we want our setting to explicit? Why? Why not?**)

Targets

The goal (to establish some particular proposition or position). An argument is REDUNDANT if its target is assumed in the setting.

Reasoning Structures

A sequence of statements meant to reach a **Target** in a specific **Setting**. See DERIVATION, to which this is similar. Also, a SYLLOGISM is a type of reasoning structure.

Remark 10. *Technically, P requires that Reasoning Structures be annotated so that premises and conclusions are overtly identified by the annotation.*

Based on the conditions of a refined argument, we can define an argument as the following (refer the previous definition of ARGUMENT, Are these two definitions the same? What is different? Why?):

Definition 8. ARGUMENT, *(Parsons 1996: 5)*
An argument is a 'task' that employs a reasoning structure in a setting with a target.

A *successful* argument must meet the following criteria.

Criterion 1. Every premise is among the statements assumed in the setting.

Criterion 2. Every inference is in accordance with a principle of inference assumed in the setting.

Criterion 3. The conclusion is the target identified in the goal.

Criterion 4. The reasoning structure is noncircular.

Criterion 5. There is no infinite regress of justifications for any step.

Part IV
Syntax, Semantics, Pragmatics

References

Aristotle. 1966. Definition of the syllogism. In Irving M. Copi & James A. Gould (eds.) *Readings on Logic*, 95–99. New York: The Macmillan Company.

Barker, Stephen F. 2003. *The elements of logic.* Boston, MA: McGraw Hill.

Boole, George. 2005 [1854]. *An investigation of the laws of thought.* New York: Barnes and Noble Books.

Chierchia, Gennaro & Sally McConnell-Ginet. 2000. *Meaning and Grammar: An introduction to Semantics.* Cambridge, MA.: MIT Ptess, 2nd edn.

Constant, Noah, Christopher Davis, Christopher Potts & Florian Schwarz. 2009. The pragmatics of expressive content: Evidence from large corpora. *Sprache und Datenverarbeitung* 33(1–2): 5–21.

Etxepare, Ricardo. 2010. From hearsay evidentiality to samesaying relations. *Lingua* 120: 604–627. doi:10.1016/j.lingua.2008.07.009.

Graff, Gerald & Cathy Birkenstein. 2006. *They Say, I Say.* New York: Norton.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar.* Oxford: Blackwell.

Paris, Cècile, Nathalie Colineau, Andrew Lampert & Keith Vander Linden. 2009. Discourse planning for information composition and delivery: A reusable platform. *Natural Language Engineering* 16(1): 61–98. doi:10.1017/S1351324909990192.

Parsons, Terrence. 1996. What is an argument. *The Journal of Philosophy* 93: 164–185. URL http://admin.cdh.ucla.edu/facwebpage.php?par=91.

Potts, Christopher. 2003. Expressive content as conventional implicature. In Makoto Kadowaki & Shigeto Kawahara (eds.) *Proceedings of the North East Linguistic Society 33*, 303–322. Amherst, MA: GLSA.

Potts, Christopher. 2007. The expressive dimension. *Theoretical Linguistics* 33(2): 165–197.

Potts, Christopher. To appear. Conventional implicature and expressive content. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.) *Semantics: An International Handbook of Natural Language Meaning.* Berlin: Mouton de Gruyter.

Zesch, Torsten & Iryna Gurevych. 2009. Wisdom of crowds versus wisdom of linguists — measuring the semantic relatedness fo words. *Natural Language Engineering* 16(1): 25–59. doi:10.1017/S1351324909990167.