# **Turnitin Writing Project**

## Week1-2

Reading RST:

Question1: Does it only apply to letter genre?

Answer1: No, it analyzed more than 400 texts, in types of [administrative memos, personal letters, letters to the editor, advertisements, scientific American articles and abstracts, newspaper articles and editorials, organizational newsletter articles and appeals, public notices in magazines, travel brochures, and recipes.]

Question 2: What is relational patterns?

Question 3: What it means by constraints on the nucleus(p11)?

Answer: P12

- a. Constraints on the Nucleus: The reader might not believe the nucleus to a degree satisfactory to the writer.
- b. Constraints on the satellite: The reader believes the satellite or will find it credible
- c. Constraints on the combination of Nucleus and Satellite: The reader's comprehending the satellite increases his or her belief of the nucleus.

Question 4: What is locus of effect (p12)

Question 5: It is said in the paper, "the purpose relation is defined without reference to 'in order to', this intentional lack of explicit hook makes such judgements essential." Why?

Question 6: If we actually label according to RST, will the model perform well, since it detect sentence-level sequence and content, as well as the relationship in between? Are these sentence level information hard to build model and predict?

Question 7: How are we going to utilize this RST?

Question 8: What exactly does it mean by "non-volitional result?" (p30), why are 14,16 non-volitional result?

Question 9: Where are the definition of these RST relations?

In appendix.

Question 10: Circumstance: constraints on S: S presents a situation (not unrealized)- meaning? Also, circumstance doesn't seem like a transparent relation to detect, it may confuse with Elaboration? Logic may be a little bit different

Question 11: p55 Non-volitional result, constraints on S: presents a situation that is not a volitional action, volitional action meaning?

Question 12: So joint relation can just be anything? Insight:

1. the RST structure stress its use for investigating relational propositions, unstated but inferred propositions that arise from the text structure in the process of interpreting texts. Also useful in text coherence

so it should be useful in our project, because coherence should be one of our focus?

- 2. The RST is indeed detailed and complex enough for investigating student writing I think
- 3. But human coding the writing according to RST can be problematic due to its detailed and

complex nature, and may have conflict between one another.

- 4. To do NLP, it is very important to clear out structure, sometimes in very fine-grained way, about language.
- 5. "(p35) For each relation linking nucleus to satellite, if the relation is defined to have a locus of effect that is the nucleus only, remove the satellite." Like decision tree pruning, now I see the connection between machine learning and linguistics, -- computational linguistics.
- 6. In p45, status of RST, it explicitly pointed out that RST is applicable to many kinds of texts, enabling a unified description of text structure regardless of genre. It helps to factor the genuinely genre-specific aspects of text from the more genre-independent ones.

This "genre-independent" nature makes it very useful in our project.

7. Is this schema going to be too abstract to detect? And may be so hard for annotators to annotate? Because they need to understand rst first.

## Week 3

9 10

Through meeting with Shiyan, I get a better understanding of the following steps after understanding the RST. I know now that we are to write a documentation about how RST framework is used in the specific field of this project- student writing. The document should be able to provide a clear structure for those who are not so proficient in linguistics, and enable them to annotate student writing.

## 9.13

We have discussed a pair of debatable relationship, background - causal.

Shiyan used two sentence about Mars to have the relationship of "background" in RST, while I hold that these sentences should be in non-volitional causal effect relationship.

The pair of sentence is like this:

- 1. The Mars have fragile atmosphere and is also far from the sun,
- 2. It is experiencing frigid weather.

Through further reviewing back to the RST paper's definition and content, I found that one way that can settle on a relationship is to look at writer's intention. Since this pair of sentence make sense to be "background" and "causal effect" relationship, we can look at which should more closely describe the writer's intention. Do they want to stress the "background" or the causal effect? That seems to be one way to reach an agreement in terms of argument.

Through discussion, Shiyan and I also agreed on a methodology of annotation:

- 1. See if there are relationship
- 2. NN or NS relationship?
- 3. If NS: presentational or subject matter relationship?

Also, I knew now that when we annotate the relationship of sentences/ span pairs, we should not be influenced by the genre of the article. The intention of RST is to be genre-independent.

Annotating the student writing: training data.csv line 249

Some agreements that Shiyan and I reached after the discussion in the afternoon:

- 1. After ensuring that there are relationship between two sentences, and maybe divide it into either N-N/ N-S relationship, there is no need to further classify it into presentational or subject matter N-S relationship
- 2. The difference between background and circumstances:

  Background information is general information of any sort that is likely to help the reader

to understand the nucleus. This differs from Circumstance, in which the nucleus and circumstance both refer to a single situation.

3. The difference between antithesis and concession:

For both of them: the W's intention is R's positive regard for N is increased.

The constraints are both are: W hold positive regard for N

The difference is- in concession: the W is not saying that S does not hold.

4. Justify and background:

The example of justify:

1. There is a event 2. I will post more details later, but feel free to save this to your calendar

This example: 2 serve as satellite for 1, and without N, S does not make sense. But in background, the S increase R's ability of understanding N, but does not necessarily need to depend on N to make sense.

Thinking about not including justify in the documentation?

## 5. Evidence:

Regardless of the definition in the RST, it should make sense to redefine, in our particular context, evidence to be when students support N with information in the prompt sources, but not actually evaluating, interpreting or giving their own opinion, just presenting the information in the source matter-of-factly.

6. Sequence:

We debate about whether sequence have two levels of dimensions. 1. The sequence in which student write about the sources 2. The sequence in which reviewer view students' points.

But later we find these two sequences are not necessarily different

- 7. We will not distinguish evaluation and interpretation
- 8. Difference between summary and interpretation?

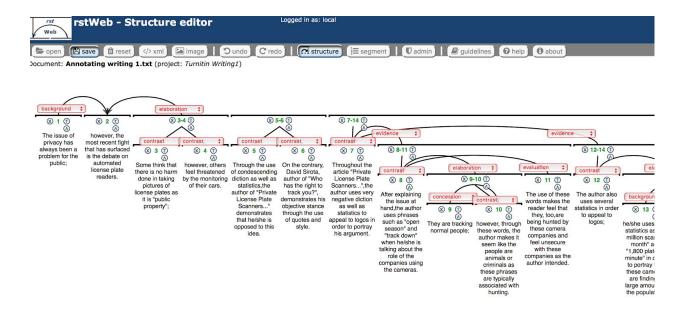
Summary are when no new points are brought out.

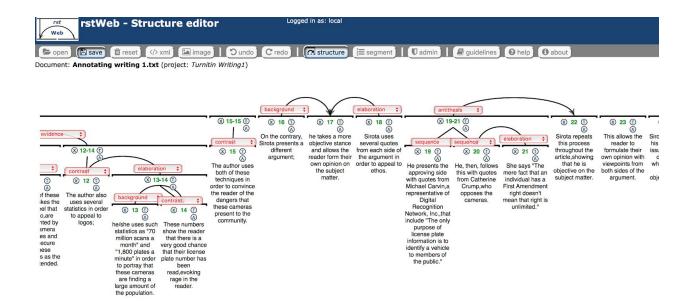
Interpretation includes new points.

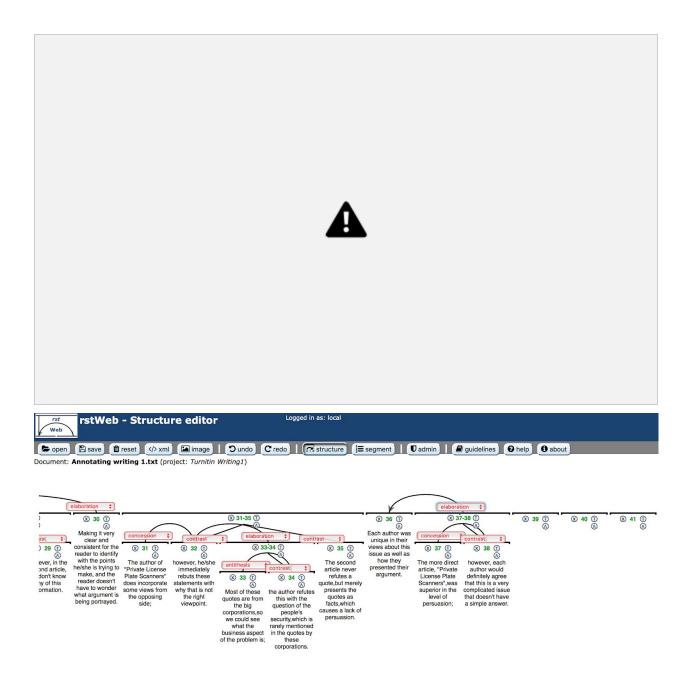
## 9.16 night

Some doubts in annotating the same writing for the second time after discussion:

- 1. Span 5-6, do they serve as a satellite to 2? Do they have relationship, such as elaboration, for 2?
- 2. Between 9-10, is it concession?
- 3. The 11-8 relationship is interpretation instead of evaluation
- 4. 8-7 elaboration or evidence?
- 5. Elaboration and evidence are still easily confused, because when elaborate, it is common to cite from the prompt source, which is also the definition of evidence
- 6. 15 should be the summary of 7-14
- 7. The contrast tag at 8 and 12 should be ignored, it is the software's bug.
- 8. 23-22, 25-24 is interpretation
- 9. The relationship between 37.38-36 is hard to decide for me, is 36 providing background for 37,38, or 37-38 is elaboration of 36?







## Week 4--9.23

On the online meeting, we talked about whether there should be sections. Advantage for having sections:

1. We can see the bigger relationships on a higher level than spans, between paragraphs, that convey the more primary opinion of the writing

Problems for having sections:

1. It is hard to define on what level should we make a section

We then, using the flowchart, tried to annotate in a sequential order, that moves from left to right, to try to annotate the same writing we did twice.

We have found that, though looking for the strongest relationship between each span, on a span level, we could sometimes lose sight of bigger idea.

But we can also look back, to see if there are bigger ideas emerging, once we finished, say maybe, after a paragraph of span annotation is done.

Through annotating, we found that since this article is quite structured, and our annotation on this writing may not be applicable to other not-as-good student writing. Specifically, we are interested in some "messy" relationship between different paragraphs, that may be beneficial for giving structure feedback.

## 9.25

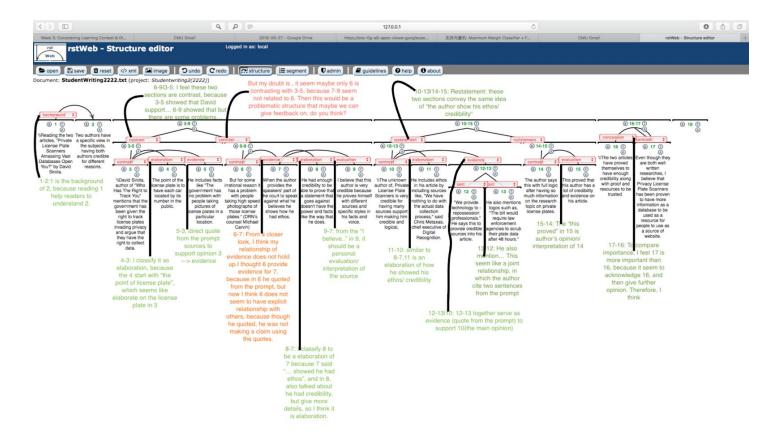
Through annotating the not-so-good student writing, I have found that the limitation of student's language greatly hinder us to see his structure and make sense of his claim. Some specific issues/questions:

- 1. Span 7-8 are confusing in language, also cannot decide if 8 is the elaboration of 7
- 2. The last spans (span16-17) is summary of , span3-9 and span 10-15, the tool did not allow me to put summary on it as I wanted to.
- 3. If we have found it very difficult to extract the higher level of structure, is it that we did not look hard enough, or is it that the students did not make it explicit enough? Should expect their writing to be ideally transparent, and categorizing them as not doing a great job in structure if we cannot see it after we read it for one or two times. If we have to consciously digging and making connections and assumptions about the deeper structure, should we still categorizing that as the writers' intended structure, or just our interpretation?

Surprisingly, since the structure of this article is 2, I expect it to have very loose and poor structure that would benefit from a lot of room of improvement. But I found this reading to be just as structured as the .

This is not to say that this writing has the same complexity of structure as the 4-4-4-4 one, but in this (2-2-2-2) one, it did not seem to have messy structure or sentences all over the place, the structure is still complete and quite coherent. (I know we were looking for relationship of different spans between paragraphs, but I did not seem to find much in this writing.)

But it is true that in terms of the layer of logic it may not be as rich, so maybe the feedback could be increasing the variety and complexity of the structure, and the richness of evidence, elaboration, background and interpretation?



## 9.26 Evening

Shiyan have fixed two technical problem for me.

1. If I want to export image:

Don't use: bash rstWeb\_local.sh

Use: ./ rstWeb\_local.sh

2. If I want to add relationship:

Copy and paste relationship in the current relationship file: Users/default\_rels.tab

We have, while looking at our difference in annotating student writing 2, make clear several important issues:

- 1. How to classify something as important, if people are having trouble deciding-Three rules of thumb:
  - a. Writer's intention: which one do they intend to emphasize on
  - b. Information contained: Which span contain more information
  - c. Complexity of the span structure: Is it straightforward or does it contain multiple layer of structure, such as concession, contrast using words like "even though, but" within the sentence.

The complexity of the sentence can signal in a way, how much information is contained, which echoes with b.

2. How to distinguish between elaboration and background:

- a. The tendency is that in elaboration, N appear before S
   But in background, S appear before N (This is not a constraints, because there may be exceptions)
- b. In background, the S is more necessary for understanding N, which means readers can be confounded in what some pronouns mean if they did not read S. But in elaboration, the dependency between S and N is less or non-existent. Reader can perfectly understand N if he was not provided with the extra information in S. The S merely serves as some additional detail.
- 3. How to distinguish between evaluation and interpretation:
  - a. Disjunct or Conjunct: Whether the S contains frame of reference from N Disjunct:

Interpretation contained logics outside of N, may judge N from a different aspects/ angle than already mentioned in N and contained new information not at all mentioned in N, which can not easily be predicted or deducted.

Conjunct:

Evaluation tend to use the frame of reference in N, which judge and evaluate N from information within the frame of reference of N. It tend to be more easily predicted and deducted from N, without big leap or turn in logical reasoning.

4. How to distinguish between restatement and summary:
One important difference is length:

In restatement: the length of N and S are similar

In summary: the length of N and S are drastically different- S shorter than N Information contained:

In restatement: Two spans containing similar amount of information
In summary: Two spans containing different amount of information
N is more detailed and fine-grained while S is more concise and general

# 9.28 evening:

flow-chart building

- 1. Interpretation and evaluation, shouldn't the S be the situation, and N be the evaluation/ interpretation
- 2. Conjunct, joined, flowchart not so clear
- 3. Disconjunction, contrast?

#### 9.30

We talked about how to segment spans. Our original plan was to use full sentence, signaled by beginning with capital letter, and ending with either a period or a semicolon.

But from looking at students' writing, we have found there are some students' writing that lack punctuation.

# [Joint relations, good or not?]

In this writing, I have found there are two paragraphs to have merely joint relations. We discussed would this be a problem in students writing, that need to be addressed by feedback. Our consensus is it is okay that students use relations of joint to express juxtaposed position of sentences, but they should connect the group of joint spans to other spans, so that they are not isolated in the writing.

## [Progressive relations]

We talked about we have a habit of putting elaboration(me) and background(shiyan) for spans that seem to provide additional information. We then found the that many

## [How to decide importance]

Last time during our meeting, we have agreed on three rules to decide importance, 1. Writer's intention, 2. Information contained 3. Sentence complexity

This time using the flowchart, we have found that many students' writing have a progressive pattern, that is, each sentence is used to foreshadow to the latter sentence, or elaborate for the former sentence. Shiyan and I thought that these patterns, once conceptualized, can be a clear indication of whether the students' logic is clear or not. Spiral escalation should be

## [Which span is this elaborating for?]

Once you have identify the relation to be elaboration, sometimes you cannot decide between two or several spans which one is it elaborating for. One indication is the pronoun. Look for pronouns like "it", "that", and seek what they refer to, they may help with identifying which span is the current span elaborating for.

It should also work for background.

## [Pre-summary and post-summary]

We have found in student writing 3, there are a kind of summary that exist at the beginning of the paragraph. We have discussed should this be the relations "justify", whose function is increasing readers' readiness to accept S. But we think that the latter, more detailed part, could not be S, but should be N.

We consider whether it is necessary to distinguish between a summary placed at the beginning of the paragraph or at the end of one, and whether it would make a difference in terms of giving feedback.

Then we decided on categorizing them as "pre-summary" and "post-summary". In terms of feedback giving, it would be ideal if both pre-summary and post summary exist, usually pre-summary should be more concise than post-summary.

## 10.3

Annotating student writing 4:

# [Segmentation]

I read Shiyan's annotation, she segmented 16-17, and 18-19,

I did so accordingly for later alignment in annotation

But the student clearly did not provide enough evidence in showing that there should be a segmentation there.

Therefore, how should we do the sentence segmentation is a question.

It is definitly hard to tell from the personal pronoun, such as Subject Object, because when there are clauses, there can be more than one subject or object pronoun.

But we may be able to count the pronouns, so that if student exceed certain amount of pronouns, we should give them a feedback on maybe they should segment this sentence into small parts. Maybe we can also give feedback on how to use punctuations, and whether they made mistakes.

But this may belong to the grammatical feedback, instead of structural feedback? This would still be very beneficial for their writing though.

# [Some questions in annotation]

5 also looks like evidence to 4, is our definition of evidence too narrow? A new relation **purpose** come up in my annotation, 19 and 18

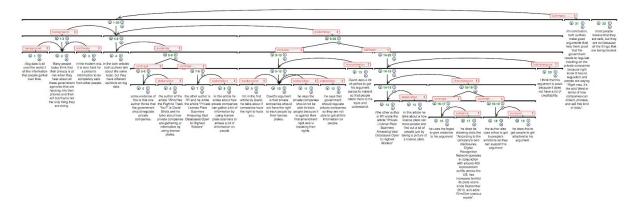
<u>Purpose</u>	on N: N is an activity;	S is to be realized through the	R recognizes that the activity in N is initiated
	on S: S is a situation that	activity in N	in order to realize S
	is unrealized		

## [rubric]

I know this is graded by experts, but this is a level three writing? Can't believe it.

[Student writing 4 annotation]

In the form before is my annotation and rationale for annotating student writing four, if requiring explanation.



N(N 1)	S( N2)	Relation s	Reasons/rationale for relation	Rationale for focus/ importance	Other possible relations/comment
2	1	bg	1 is setting background, foreshadowing/ justifying writer to talk about 2	Because 1 is the bg for 2, then 2 has to be N	
2	3	evaluati on-eval	3 can be considered as writer's personal evaluation, it is conjunct with 2 (logical prediction)	Because 3 is evaluation for 2, 2 is N	Summ ary,
8	7	bg		I feel what the author talks about is more important than who the author is, because it is information to a higher level( 递进关系?)	
5	6	evidenc e-evid	6 has prompt sources, 6 support 5	5 is the broader claim, and 6 is only facts, so 6 is supporting 5, so 5 is N	Elabor ation?
					5 also

					looks like eviden ce to 4, is our definiti on of eviden ce too narrow ?
56	78	contrast		NN	
4	5-8	evidenc e		Decided by relation (DBR*)	
10	9	backgro und	I categorize 9 to be background of 10 because 10 is more important than 9.	递进关系 DBR	
10	11	elaborati on-elab	Provide details,	其实10 和11 也像递进关系, 没把11 选成n的原因可能是, 觉得11还是说他说了啥。。虽 然更进一步,可是还是支持10 ,所以还是DBR	
10	12	elab	Provide demail	DBR	
10( 9-1 2)	13	interpret ation	Disconjunction, the information of the prediction does not come from N	DBR	
14	15	elaborati on	I initially wanted to put 14 as the background of 15, because, supposingly, this paragraph should have the same structure as the former one, shouldnt it?  But from a closer look, I think  1. 15 is not significantly more important than 14?  2. Does understanding 15 require us	This is indeed one that I have struggled about which is N, because i cant decide whether it is bg or elab relations.  I think it should still be bg, 15 provide 深一层 information for 14, so 15 N	backgr ound

			providing 14? 3. 14 is different from 9 because, 9 says something about what the author said in the article, but 14 said only the article's title and the author's name		
16	17	evidenc e	Prompt source to support claim	DBR	
18	19	purpose	Author did 18 in order to do 19	DBR	
16 17	18 19	contrast		nn	
14- 15	16- 19	interpret ation	Interpretation cannot be predicted from 14-15	dbr	
14- 19	20	interpret ation	Disjunct, jundgement	dbr	
9-1 3	14- 20	contrast		nn	
4-8	1-3	backgro und	BUT: 1. does knowing 4-8 rely on knowing 1-3 2. Is 1-3 the background for 4-8 or 4-20, or even 4-22	Dbr Since 1-3 is before 4-8, 1-3 is the s for 4-8 in relation of bg	
4-8	9-2	elaborati on	I have been hesitating to put elaboration or interpretation, because it contains both of them, but then I think some interpretation may also be considered as details, thus elaboration	Cross paragraph, harder to see which is more important, more often be decided by relations	
1-2 0	21	summar y		dbr	
	22		Really sound like a		

	stand-alone sentence.	

10.5

Meeting with Carolyn, Auto-segmenting Arkensha, james, pointers

## 10.6

- 1. span1-ls the prompt asking for your conclusion?
- 2. Shiyan mentioned that we should intend the computer to do the annotation from left to right, but in this studentwriting 5 as an example, span 1 seem like a presummary, it is hard to find what it is summarying for before we find more relationship in the latter spans. In other words, after we find the relations of later spans, the relationship between 1 and other spans should have emerged. Then how can the "from left to right" scheme work here?

3. I have found another very easily confused pair: preparation and justify, and background Justify:

Justify	none	R's comprehending S increases R's readiness to accept	R's readiness to accept W's right to
		W's right to present N	present N is increased

# Preparation:

Preparation	none	S precedes N in the text; S tends to make R more ready,	R is more ready, interested or
		interested or oriented for reading N	oriented for reading N

Is the only difference between justify and preparation in that: in preparation, S has to precede N?

According to the example of preparation: it refer to titles? Can we understand preparation as "铺垫"?

The difference between preparation and background may be, in background, N depend on S to make sense?

# Two example text:

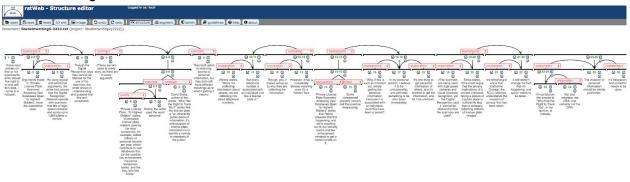
Justify	10) Let's be clear: 11) I personally favor the initiative and ardently support disarmament negotiations to reduce the risk of war. 12) But I don't think endorsing a specific nuclear freeze proposal is appropriate for CCC. 13) We should limit our involvement in defense and weaponry to matters of process, such as exposing the weapons industry's influence on the political process.	10) Let's be clear:	11) I personally favor the initiative and ardently support disarmament negotiations to reduce the risk of war. 12) But I don't think endorsing a specific nuclear freeze proposal is appropriate for CCC. 13) We should limit our involvement in defense and weaponry to matters of process, such as exposing the weapons industry's influence on the political process.	This is from the analysis of the Common Cause letter, on this website. Unit 10 is establishing the need to argue a second time toward a particular conclusion. Thus it suggests what the basis is of the writer's right to speak this item.
---------	--	------------------------------	--	---

	the <u>Bouquets</u> text analysis on this website.  mmon for titles to be preparatory (and not y) to a text.
--	--

4. What is the difference between multinuclear Restatement and NS restatement? For example: 25, 26 what are they?

High level structure usually require us to look very closely to find, should we still find them if they are not too transparent?

# Student writing 5 annotation details:



23	24	elab	Stressing, giving more details
25	26	restatement	Basically same idea
25	23-24	preparation	I think it should be bg, justify, or preparation I wanted to put background, but I think 25 does not depend on 23-24

N/N1	S/N2	relationship	reason	Rationale for focus/importance	com men ts/ alter nati ve?
23-2 6	1	background	23-26 is the conclusion that the writer get	dbr	
2	3	elab	3 provide details for 2	Quite hard, 3 is longer than 2 and also say important thing, but I guess since 2 is more	

				general like a claim, 3 support 2, 2 is N Still DBR	
4	2-3	background	2-3 provide background for 4m they .	4 shows the clear attitude of the author, so it should be more important	
7	6	justify	R's comprehending S increases R's readiness to accept W's right to present N	I feel here the writer's intention is clear, he wants to stress 7	
6-7	8	contrast	Talk about different authors' opinoin	NN	
6-8	5	pre-summary	Foreshadow about the following story	DBR	
6-8	9	post-summar y	Summarize two authors' opinion	DBR	
11	10	preparation	S precedes N in the text; S tends to make R more ready, interested or oriented for reading N	DBR	
13	12	concession	Writer admit both spans, W acknowledges a potential or apparent incompatibility between N and S; recognizing the compatibility between N and S increases R's positive regard for N	DBR	
10-1 1	12-13	evaluation	Writer's opinion	DBR	
14	15	conjunction	These two quotes seem to have no difference in importance	SHOULDn't distinguish here because it is one quote	
17	16	preparation	17 is more important, 16 prepare readers to	DBR	prep arati on?
17	14-15	background	Put background for writers, "it " in 17 refer to former parts	DBR	
17	18	elaboration	18 offer more details	17 give clearer attitude,	Just

			supporting 17	18 seem like a follow up sentence.	ify? Is ther e a relat ion for "exp laini ng reas ons ?"
20	19	preparation	It is actually so hard to put relations, and I wanted to say stand alone. But from closer look, i think 19 is saying that the way scanners are used is problematic And 20 is saying that certain way of scanner using is worse than others So I think 19 is preparation for 20	20 build upon 19 and is richer than 19  19: phenomenon A 20: there is a B, which is worse than A	
22	21	concession	Also a hard one, i feel the students' logic is not so clear. I put it as concession, because I feel that writer are acknowledge both spans, and feel that these two incompatible are actually compatible in this case.	22 seem to connect better to our core topic and argument(claim), 21 seem more like a comment or 题外话, 顺带 一提	
21-2	19-20	backgroudn	Put background to why the writer oppose it and think it should be stopped.	DBR	
23	24	elab	Stressing, giving more	Nobody in 23 include 24	

			details		
25	26	restatement	Basically same idea	This i cannot quite decide, since it is restatement, shouldnt it be the same in importance?  But the 26 has a "it" which refer to 25 I think, so does that make 25 more important?	
25	23-24	preparation	I think it should be bg, justify, or preparation I wanted to put background, but I think 25 does not depend on 23-24	DBR	

## Feedback for Shiyan's talk:

Is it possible to, based on certain performance of students' initial performace, such as their temperament, characteristics, participation, communication with others, confidence, predict students' learning trajectory maybe using model or ml? Then give proper intervention / suggestions according to these results?

## 10.7 Discussion Summary:

- 1. Shiyan: we should not separate spans in quote, whenever it is in a quote, we can just put it in one span. Because there are no structure feedback we can give for that.
- 2. We asked Prof. Rose about whether we should decide importance first or relations first
- 3. How to distinguish Conjunction/Joint/List/ Sequence

One of the point that we thought about giving structural feedback is that we can advise students to write less spans with Joint relations.

In RST, joint basically means two spans have no relations

So we have agreed on that, if the two spans are totally not related and can not be described by any relations- Joint (In which case we would advise students to strengthen their logical link between these spans)

<sup>&</sup>lt;sup>1</sup> \*DBR: decided by relation

If the two spans have some relations, equally important-- Conjunction (this is acceptable relations to write)

4. How to distinguish Preparation Background and Justify:

We think that Justify is too subtle to distinguish, so maybe we will avoid using it if possible. But Background and Preparation are easily confused,

they are both relations in which S usually appear before N, and serve as the function of foreshadowing or helping us to understand/ be more interested in N. our rationale to distinguish them is:

- a) Are there pronouns(he, she, it, this,that) in N that refer to information in S? If yes-Background, if no:
- b) Does understanding N depend on providing S? If yes: Bg, if no: Preparation
- 5. How to distinguish Sequence and List:

List	An ite	em comparable to others linked to it by the List relation	R recognizes the comparability of linked items		
Sequence		There is a succession relationship between the	R recognizes the succession relationships		
Seque	nce	situations in the nuclei	among the nuclei.		

## Example text:

L	I am 17 years old. It is	I	It is summer,	These are the opening lines of an essay by
ist	summer, and football	am 17	and football practice	a student. They give the first of a larger set of
	practice is about to begin.	years	is about to begin.	background facts, in a list.
		old.		

Sequen	[from a recipe:] 1. Peel oranges, 2.	1.	2. and	Sequence includes both
ce	and slice crosswise. 3. Arrange in a bowl	Peel	slice	presentational sequence, e.g.
	4. and sprinkle with rum and coconut.	oranges,	crosswise.	"Secondly," and also subject
	5. Chill until ready to serve.		[The given	matter sequence, e.g. "After
			text yields 5	that," as in this case.
			nuclei.]	

## If two spans:

 a) Have a succession relationship in time between the situation in the nuclei, and the order of the spans matter

Can we change the order of the spans without distorting the original meaning of the spans?

Yes-List, no-->Sequence

5. How to distinguish evidence and elaboration?

Does the S have direct quote from the prompt source? Does the S have analysis of the claim/ prompt source? Is S supporting any claim in N?

6. How to distinguish summary and interpretation/ evaluation:

Does the S contains more information than N?-- If yes, exclude summary

Does S have analysis/ judgement for N? Y→ interpretation/ evaluation

7. How to distinguish antithesis and contrast?

Does the writer have strong preference/ favor/ emphasis of either span?-- y→ antithesis

- 8. If it is too common-sensical, then it may not be interpretation.
- 9. How to distinguish: Conjunction and contrast: see table

# Insight for Tutor building:

To organize how to decide relations in terms of production rules will make this process much clearer. And also possible easier to adapt into CTAT to build cognitive tutor, that help annotator narrow down the possible relations, and choose the same one as we would choose.

## Discussion on Oct.14th

- 1. One way to identify the nucleus is: is the span increase the information to another degree? (progressive)
- 2. One insight got from Manfred Stede and Maite Taboada's Annotation Guidelines for rhetorical structure as well as Vincent Ng's argumentation mining work, is that we can identify major claims for the whole passage and try to find claim in each paragraph while doing the first time reading. We can also ask the annotators to do so: skim and find claims that are important and mark them, as this would give insight for later finding relations. It will make it easier to find N and S, because claim and major claim would naturally fall under "N"
- 3. If it is hard to decide which one is N, we can go back to the claim of the paragraph, if it still won't come up, then go back to the article.

Meeting on Oct 17th

Will the method 2 brought up in discussion on Oct 14th be too arbitrary?

Carolyn: We will have to try and see, if it work it could be helpful. If it is hard to see the claim and major claim, or if it is not explicit enough, maybe we need to make it explicit in definition of major claim / claim.

# Oct 20th, Annotation:

N/N1	S/N2	relationship	reason	R	comments/ alternative?
	0/112	relationionip	100011	a	commente, alternative.
				ti	
				0	
				n	
				а	
				ı	
				е	
				f	
				0	
				r	
				f	
				0	
				С	
				u	
				s	
				/i	
				m	
				р	
				0	
				r	
				t	
				а	
				n	
				С	
				е	

# Student writing9-1111

First time reading and marking claim:

MC(major claim):1,13

C(claim):3,6,9,

Changed my mind 9-->10

1	2	cause	because	
3	4	joint	No particular relation	
4	5	elaboration	Give specific example showing data have	

			everything reported	
7	8	antithesis	Writer favor 7, technology than police?	
10	9	cause	That is why	
7-8	9-10	conjunctino	Joined to support 6	
6	7-19	evidence	6 seem major claim	
11	12	elaboration	12 Just give details	
13			Should be summary yet no clear trace of it.	
Studer	nt writing	g 8		
MC:1 C:4,7,9	9			
2	3	elab	defining	
4	5	elab	A deeper	
4	2-3	prep	2-3 seem to support 4, i wanted to put bg, but think that 4 does not depend on 2-3	
2-5	6	interpretatoin	Word choice is excellent personal opinion	
1	2-6	evidence	2-6 seem to support 1, some spans include quotation from the propmt	
2-6	7	post-summm ary		
8			I am not sure how 8 link to oterh spans, is the example in 8== the concept in 9?	
10	9	bg	Point of view in 10== concept in 9	
11	10	joint	No strong relations, except	

			he simply did A he simply did B	
14	13	prep		
13-1 4	15	interpret	Personal opinion	
13-1 5	16	conjunction	Want to put NN ralation here because in 16, it said "more importantly", so cannot put elab or summary.	
Studer	nt writing	<b>7</b>		
MC:1,2 C:2,4,	25 13,14,15	5,23,24		
2	1	prep	Put background but not dependent	
3	1-2	prep	Seem to be preparing for asking the questions	
5-11	4	Pre summary		
5	6-7	evidence	Quote from prompt to support 5	Or support 4?
6	7	elab	This grand refer to 6	
8	5-7	prep	5-7 talk about number, 8 introduce questions	
9	10.11	conjunction		
8	9-11	elaboration	9-11 portray what questions?	
5-11	12	post-summar y		
13	14	conjunction	I think they are both very important, even though in 14 it has "with this in mind",	

			which suggest maybe it is based on 13.	
4-12	13-14	evaluation	I struggle about evaluation or interpretation. Normally for this kind of spans i put interpretation, but in this case, I think the opinion is very well-supported by aforementioned spans.	interpretation
15	16	elab	Think 15 is more importnant	
17	18	restatement	Similar information	
19	20	elaboration		
15-1 8	19-20	interpretation		
15-2 0	21	interpretation	Not summary because information in 21 is not mentioned in nucleaus	
23	24	conjunction	Both important	
22	23-24	evidence	Support 22	
22-2 4	25	interpretation	Talk about other thing, so not summary.	
Studer	nt writing	<b>9</b> 6		
C:1,6,	11,17,22	2,27		
2	3	conjunction	Additional question	
2-3	1	preparation	Prepare for proposing questions in 2-3	
4	1-3	bg	"These" refer to 1-3	
5	1-4	bg	"Both article" refer to 1-4	
7-16	6	presummary		This seems like a very clear presummary.
8-16	7	prep		

8-12	13-16	contrast		
8-9	10	interpretation	It is hard to say if this is a interpretation or evaluation. (feel a bit like both, what do you think)	
8	9	evidence	quote	
11	12	elab		
8-10	11-12	contrast	Easy to see	
13	14	evidence	Maybe this should not be segmented?	
15	16	contrast	Easy to see	
13-1 4	15-16	evidence	15-16 support 13-14	
18	19	contrast	Should this be contrast or concession, if concession, if feel there is no span more important than the other?	
20	21	contrast		
18-1 9	20-21	contrast	The structure is so nice here.	
18-2 1	17	prep		
23-2 6	22	presummary		
24-2 6	23	prep		
24	25-26	contrast		
25	26	elaboration		
28-3 0	27	presummary		
28	29-30	evidence		
29	30	elaboration		

28-3 0	31	Post summary		
33	32	prep	I struggled about should 32 be bg for 33? Is "one article" in 33 refering to 32? But then I think not so much. Because you can still get what "one article" mean withouth 32. Then maybe we should be cautious about the distinction of whether a pronoun exist to distinguish bg and prep. This is based on the assumption that 32 is more important. (how to tell here?)	
studer	ntwring10	0(3333)		
C:1,12	2			
2-12	1	preparation	Heading should be preparation	
3-11	2	pre-summary	I am not too sure, because though 2 and 12 both stress the same thing article 1 is better than 2. But the 3-11 body part did not explicitly mention this point i think.	preparation?
4-6	3	interpretation	Quite subjective "a good point, amazing point"	
6	4-5	prep	The data and fact in 4-5 introduce author's opinion on this social issue.	
5	4	evidence		
8,9,	10-12	conjunction	Because the sentence is too long. It is hard to tell explicit relations.	

11	10	justify	Just feel this is a very classic use of justify?	
8-11	7	preparation		
3-6	7-11	conjunction		
3-11	12	post-summar y		

## Reflection Oct. 21

- 1. I think we may need a rationale for the first time reading and choose which seem more important, or it gets subjective.
- 2. Evaluation and Interpretation: Should we advocate for more evaluation than interpretation, since evaluation is more supported by logical deduction?
- 3. When sentence span is very long, then it is harder to find relations

Reading: <a href="https://corpling.uis.georgetown.edu/wiki/doku.php?id=gum:rst">https://corpling.uis.georgetown.edu/wiki/doku.php?id=gum:rst</a> Complete tree is wanted.

The corpling@GU wiki put evidence for our treating joint as relations for islands.

# 10.26 Meeting with Prof. Rosé

1. Automatic feedback is the goal, which will depend on later rules writing and structural pattern recognition. We are still at the very beginning.

# 10.27 Annotation New writing

N/N1	S/N2	relationship	reason	Ration ale for focus/i mporta nce	comments/ alternative?			
Studer	Student writing 11-1111							
MC(ma	First time reading and marking claim: MC(major claim):2,3 This writing is incomplete.							
2	1	justify	R's comprehending S increases R's readiness to accept W's right to present N		preparation			

4	5	restatement	No new information in 5	Insight: this kind of restatement should be minimize, because it is redundant.
3	4-5	elaboration	3 is more important	
3-5	6-7	restatement	Seems all be talking about why	
2	3-7	evidence	Show why he said 2.	I find now that finding nucleus is quite important.

# Student writing 12

First time reading:

Claim: 1,2,4,5,6,7,11,12

The writing is very structured, the structure is very self-explanatory, regardless of its

simplicity.

Con: only one paragraph.

1	2	evaluation		
4	3	preparation		
5,6	7	conjunction	1,2,3 where is 4?	
4	5-7	evidence	Should it be evidence or elaboration?	
9-11	8	prep		
9	10-11	conjunction		
10	11	elaboration		
8-11	5,6,7	conjunction?	I am not sure whether it should be conjunctino or elaboration	
3-11	1-2	Pre summary		
3-11	12	post-summar y		

Student writing 13

Claim:2,3,6,9,12,16,21

I believe picking out claim, major claim, even if subjectively, help a great deal in later identification, because we are not looking at local relations, but can have more global picture

## in mind.

I found that interpretation, tend to be more important nucleus i found during the first time reading. But in relation annotation, they are satellite. This is problematic, because suppose we want them to find n/s first, (deciding importance), then their later relation identification would conflict with that.

2 1 2 3		background		
2 3				
		justify		
5 4		preparation		
6 4-	-5	preparation		
1-3 4-0	-6	elab	Elab? Background? Justify? They seem both very important.	
7 8		evidence		
9 10	0	elab		
7-8 9-	-10	evaluation	Thought about it being interpretation, but I think maybe evaluation is better, because it seem not really involve a framework outside of reference in N	
11 12	2-13	interpretation	I am a bit confused with interpretation/ evaluation now	
12 13	3	elaboration		-
14 15	5	evidence		-
16 17	7	elaboration	Progressive, should progressive be elaboration relation?	_
16 18	8	restatement	Same information	-
19 20	0	evaluation		
16-1 19 8	9-20	elaboration		
22 23	3	elaboration		
21 22	2-23	evidence		-
24 25	5	conjunction		
21 24	4-25	elaboration		
21 26	6	restatement		

27	28	elaboration	
27-2 8	29	evaluation	Some doubts:  1. Is 29 evaluation, or summary?  2. Is 29 evaluating/ summarizing for 27-28, or for 21-28?
7-11, 11-1 5	16-20	conjunction	7-11 what some people said PA is 11-15 what patriot act is not 16-20 what it is actually doing So conjunction
7-20	21-26	background	
7-26	27-29	post-summar y	
7-26	1-6	Pre summary	

Studentwriting14

Claim: 10,16,17,20,22,28

How should we define the sentences we are to find here, are they claim? Or are they nucleus(just seem relatively more important)?

1	2	elab		
3	1-2	background		
4	5	evaluation		
6	4-5	bg		
7	4-6	prep		
8-9	4-7	preparation		
9	8	preparation		
11-1 5	10	pre-summary		
11-1 2	13-15	elaboration		
11	12	elaboration		
14-1 5	13	antithesis	First time i use antithesis, but i feel very proper here.	
15	14	preparation		

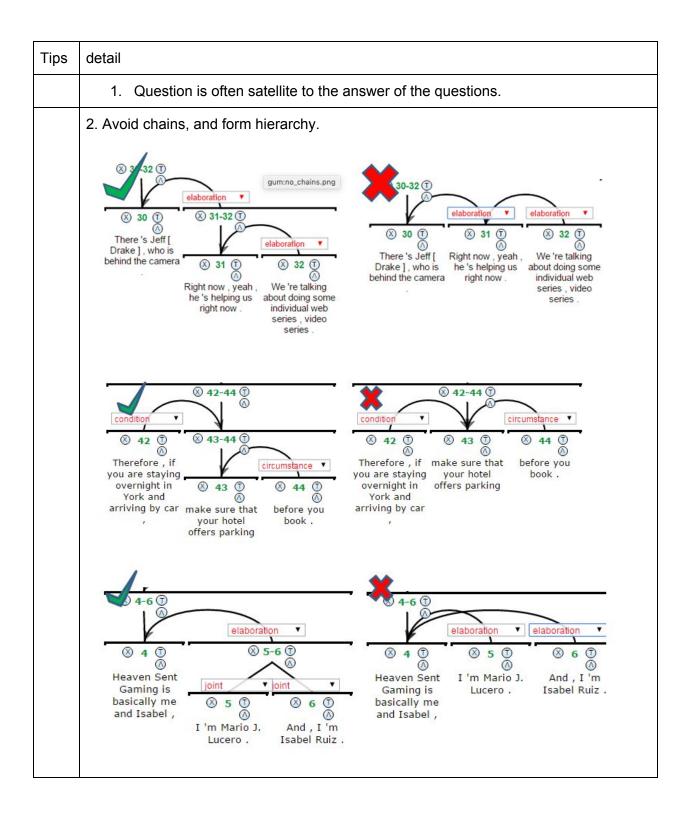
17	16	preparation		
18	16-17	motivation	18 seems to be advocating something.	
19	20	conjunction		
21	22	joint		
23	24	joint		
21-2 2	23-24	restatement		
19-2 0	21-24	conjunction		
16-1 8	19-24	evidence	Evidence or elaboration?	
25	26-7	elab		
26	27	elab		
29	28	bg		
30	31	evidence		
33	32	restatement		
10-3 1	32-33	post-summar y		
	ntwriting 4,7,14,1	15 5, 21,24,26,27		
2	1	cause	First time of using this, but feel proper	
1-2	3	solutionhood	First time of using this, but feel proper	
4	5	elaboration		
6	4-5	concession		
7	8	elaboration		
9	10	evidence		
11	12	evidence		

9-10	11-12	conjunction		
9-12	13	evaluation		
7-8	9-13	evidence		
15	14	preparation		
16	17	elaboration		
18	19	evidence		
16-1 7	18-19	result	First time of using this, but feel proper	
16-1 9	14-15	solutionhood		
20	21	solutionhood		
22	23	evaluation	Or elaboration, or conjunction?	
20-2 1	22-23	evidence		
20-2 3	24	Post summary		
26	25	antithesis		
25-2 6	27	justify	Or evaluation, i guess it depend on which is more important.	
25-2 7	28	evaluation		
7-13, 14-1 9	20-23	conjunction		
7-24	25-28	concession		
7-24	1-6	preparation		

## 10.28 Discussion

- 1. evaluation/ interpretation: have logical analysis→ then evaluation; no→ interpretation
- 2. motivation/ enablement: desire or method- increase R's desire to do N  $\rightarrow$  motivation Enable R a method to do N  $\rightarrow$  enablement

# 1. In flowchart: After finding relations, go to next one.



- 3. Headings are typically seen as a 'preparation' for the following group of segments comprising the section under the heading. This is especially true if the heading does not contain information not covered again in the section. The 'preparation' should target an added span covering the entire section, and not just the head segment of the section. (see image below)
  - In some cases, the heading contains the main gist of a (usually short) section, and the section itself may be seen as an elaboration of the heading

## 4. Insertion tests

When considering two similar relations between sentences without an explicit connective like 'beacuse' or 'if', sometimes inserting a connective or phrase can help to disambiguate. Useful phrases include:

- 'because' if you can insert 'because' between clauses, often you have a cause or result relationship
- 'the reason I say this...' if you can insert this, it can indicate justify
- 'what you need to know about this...' can indicate background
- 'proof of this is...' can indicate evidence

## Some examples:

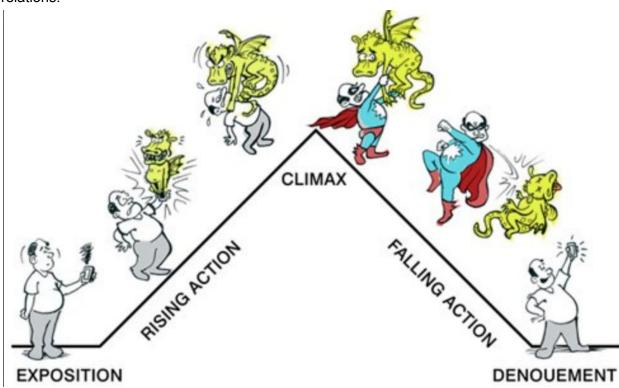
- "[IE's market share has dropped to 56%.] [Mozilla's Firefox has been actively increasing its market share] " in this example, it's easy to insert 'because', and the relationship is **cause**.
  - If this were justify, we could say "the reason I say this is that Mozilla..."
  - If it were background, we could say "what you need to know about this is that Mozilla..." which is also more forced
- "[York is a fairly small city -] [four days is enough to see the major sights] " in this example, we can add "proof of this is..." between the two units, and the relation is **evidence** 
  - If it were cause we could say "the city is small because four days is enough..." - but actually it is not small because of this fact
  - It if were background, it would be as natural or more natural to say "York is small. What you need to know about this is that four days are enough..."

I believe picking out claim, major claim, even if subjectively, help a great deal in later identification, because we are not looking at local relations, but can have more global picture in mind.

## 11.14

Meeting with Turnitin: really interesting and exciting, love it!

- 1. David mentioned about segmenting sentences into parts, instead of using the full sentence.
- 2. David also mentioned we can possibly look at relation between spans in different sections, which we ignored before. Shiyan seconded this and said that
- 3. Jill and the other girl said the narrative genre is a whole different story, its organization stress pushing the plot to the climax, therefore, maybe we can consider adding new relations.



## 11.17

Studentwriting 13 -redo					
Span relation alternative? comment					
12-13	12-13 interpretation		Think it should be elaboration, but according to flowchart, It will be interpretation,		

			Prob1: a lot of things will become interpretation.
15-16	no		Seem like having relations, but not a clear fit in N/S? But I am not fully convinced that it is totally unrelated with 15, i think high correlation.
17/19			Seems like restatement/ elaboration, Prob2: A lot of things will fall into evidence
16/17	Preparation,	conjunction	
23/24	evaluation	preparation	I feel preparation is more proper, yet evaluation is before that in the flowchart.
24/25	Cannot put preparation because 1 s can not serve 2 n		
21-25	Conjunction or restatement?		
studentwriting14			
1/2	1 is more important, but according to flowchar, it would be bg, pre		
3/2	interpretation	concession	prob1
10/11 关系	summary/ restatement?		
23/24			Prob3: anything will help understanding
25/26	Bg groups	evaluation	Prob4: Sometime you can clearly identify which is n and s, but flowchart prevent you from doing as you like.
			Prob5: section hard to see relations.

		Prob6 cause/result保留几个?
studentwriting15		
6/7关系?		
7/8		prob4
		Prob6: different level of sections find relations at the same time?

Foreshadowing: evaluation, preparation, background.

Discussion on week 13":

Evidence: must be fact/ anecdotal, autoritative

Evaluation: must be evaluating things in the other span.

Student writing 16		
1		Problem 1: crossing relation, should i mark them?
4-3		PROB2:Like result, maybe result and cause is not the same.
9-10	evaluation	The first part of the sentence is
1-3 4-5	evaluation-N, 4-5 is the evaluation of 1-3	

# Week 15

Discussion:

Evaluation: two condition, 1. judgement and analysis in one span 2. Judging on the Fact or claim in the other span

2. Section separation: if 3 point at 2, 3 also point at 4, then we compare which pair is better then decide.

## Distinguishing Similar relations:

Relation1	Other relations	Rationale
Conjunction:	Contrast	Is the writer comparing the two spans? Y→ Contrast
	joint	If the two spans are totally not related and can not be described by any relations  → Joint  If the two spans have some relations, equally important  → Conjunction
·		The example of justify:  1 There is a event 2. I will post more details later, but feel free to save this to your calendar
		This example: 2 serve as satellite for 1, and without N, S does not make sense.
		But in background, the S increase R's ability of understanding N, but does not necessarily need to depend on N to make sense.
		Thinking about not including justify in the documentation?
But in background, S appear be		The tendency is that in elaboration, N appear before S But in background, S appear before N (This is not a constraints, because there may be exceptions)
		In background, the S is more necessary for understanding N, which means readers can be confounded in what some pronouns mean if they did not read S.  But in elaboration, the dependency between S and N is less or non-existent. Reader can perfectly understand N if he was not provided with the extra information in S. The S merely serves as some additional detail.

	preparation	Background and Preparation are easily confused, they are both relations in which S usually appear before N, and serve as the function of foreshadowing or helping us to understand/ be more interested in N. Our rationale to distinguish them is: c) Are there pronouns(he, she, it, this,that) in N that refer to information in S? If yes- Background, if no: d) Does understanding N depend on providing S? If yes: Bg, if no: Preparation	
Evaluation	interpretation	<ul> <li>b. Disjunct or Conjunct: Whether the S contains frame of reference from N Disjunct: Interpretation contained logics outside of N, may judge N from a different aspects/ angle than already mentioned in N and contained new information not at all mentioned in N, which can not easily be predicted or deducted. Conjunct: Evaluation tend to use the frame of reference in N, which judge and evaluate N from information within the frame of reference of N. It tend to be more easily predicted and deducted from N, without big leap or turn in logical reasoning.</li> <li>Does this sentence pair contains writer's personal judgement?</li> <li>Yes→ Evaluation/ interpretation Is S basically a logical deduction from N, and do not involve frame of reference outside N?→&gt; Yes→ Evaluation No→ Does S contains some logical analysis of N so that the reader can make sense of the logic flow?</li> <li>Yes→ Evaluation No→ Interpretation</li> </ul>	

	summary	Does the S contains more information than N? If yes, exclude summary  Does S have analysis/ judgement for N? Y→ interpretation/ evaluation	
restatement	summary	One important difference is length:  In restatement: the length of N and S are similar  In summary: the length of N and S are drastically different- S shorter than N  Information contained:  In restatement: Two spans containing simil amount of information  In summary: Two spans containing different amount of information  N is more detailed and fine-grained while S more concise and general	
sequence	list	b) Have a succession relationship in time between the situation in the nuclei, and the order of the spans matter  Can we change the order of the spans without distorting the original meaning of the spans?  Yes-List, no>Sequence	
evidence	elaboration	Does the S have direct quote from the prompt source? Does the S have analysis of the claim/ prompt source? Is S supporting any claim in N? Evidence support the claim in the other span, While elaboration does not have this function.	

antithesis	contrast	Does the writer have strong preference/ favor/ emphasis of either span? y→ antithesis No→
	concession	
circumstance	condition	? Are they both using, if, when
motivation	enablement	desire or method- increase R's desire to do N $\to$ motivation Enable R a method to do N $\to$ enablement

# Structural feedback that we can give:

Problem	Suggestions	Revision
Too flat structure		
Too many joint relations		
One sentence paragraph		
Island sentence that have no relations with other		
Too many interpretation	Interpretation is more personal opinion without clear logic flow.	More supported relations that people can deduct logically should be evaluation.
Restatement	<ol> <li>Restatement may not seems that valid in students writing, because it may seem verbose if overuse.</li> <li>Also, if there are some spans that trespass other spans to restate a span, then it doesn't really make that much sense and will get a reader confounded. The immediate restatement is better than this kind of trespassing restatement.</li> </ol>	

	elaboration \$\displaystatement\$\$  \times 16 \times 20 17 \times 20 18 \times 20 \times	
Background, preparation, evidence are favorable relations, especially if hierarchy formed	They show that the claim is well-supported	
The most core claim should be at the center of the framework.	If it is not all pointing at the major claim, then maybe it can be improved on.	

	Task	Time	Time total (week)	Date
W ee k 1- 2	Read and understand RST and Text Analysis, thinking about how best to apply it.	20 hrs	20 hr	9.7-9.8

W	Lunch meeting with Shiyan	0.5 hr	9.5 hr	9.10 Mon
ee k 3	Meeting with Shiyan	0.5hr	-	9.13 Thur
	Review and familiar myself with the RST structures and examples on the website, and annotate a student writing	3.5hr		9.16 Sun 11:00-14: 30
	Discuss with Shiyan about the difference of our annotation result, in order to reach an agreement on the definition and classification, which will pave way for the documentation	3 hr		9.16 Sun 14:30-17: 30
	Reflect on the discussion, and document it down Re-organize my annotation result into a complete schema	2hr		9.16 Sun 21:00-23: 00
W ee k 4	Online meeting -sections, relationship between paragraphs	3 hr	3 hr	9.23 Sat 9:00-12:0 0 a.m.
W ee k 5	Annotating a not-as-good student writing, with the same prompt	1 hr	12.5 hr	9:25 Tue 19:30-20: 30
	Supplement explanation and rationale for annotation 2	2hr		9.26 Wed 10:00-12: 00
	Discuss about our difference in annotation, and definition of relationship	2.5 hr		9.26 Wed 18:40-21: 00
	Work on the clarifying the flowchart, deleted the subject matter/ presentational relationship	2 hr		9.28 Fri 10:00-12: 00
	Annotated student writing 3, discussion about flowchart.	5hr		9.30 Sun 12:00-17: 00
we ek 6	Reading the prompt source and annotate	2.5 hr	9 hr	10.3 Wed 17:30-20: 00
	Annotate student writing 5	0.5hr		10.5 Fri

				18:00-18: 30
	Annotate and write rationale	3 hr		10.6 Sat 11:30-15: 00
	Discussion about annotation	3 hr		10.7 Sun 14:30-17: 30
we ek 7	Organizing the note	1 hr	5 hr	10.9
	Revising Rationale for N and S	1 hr		10.11 0:00-0:50
	Discussion with Shiyan	3hr		10.14 14:00-17: 00
we ek 8	Read Vincent Ng argumentation mining paper	1.5hr	7.5hr	10.16 0:00-1:30
	Read the revision paper	1hr		10:17 10:00-11: 00
	Annotated 5 writing, read the website guide, document the rationale.	5 hr		10.21 10:00-12: 00, 18:00-21: 00
we ek 9	Reading the annotation paper	1hr	12.5 hr	10.24 20:00-21: 00
	Find writing, organize, annotate and document.	6.5 hr		10.2712:0 0-13:00 17:00-19: 00 20:30-24: 00
	annotate and document.	1.5 hr		10.28

				10:00-11: 30
		3.5 hr		10.28 14:00-17: 30
W ee k 10	Refining flowchart	3hr	7hr	10.31 20:00-23: 00
	Refining guideline	1hr		11.2
	Discuss with Shiyan, annotate while trying to change the guideline/ flowchart	3hr		11.4 14:00-17: 00
we ek 11	Revise guideline according to discussion	1 hr	6hr	11.4 19:00-20: 00
	Revise flowchart and guideline, (including add typical connectives)	2hr		11.7 10:00-12: 00
	discussion	3hr		11.11 17:00-20: 00
we ek 12	Turnitin meeting& later revision	3hr		11.14
	Helping zalfa with implementing	1hr		11.14
	Helping zalfa with answering questions about rst annotation	1hr		11.16
	Reannotate student writing13-15	4hr		11.17
we ek 13 -1	Annotation 2 writing	3 hr		11.25
	Discuss about the relation	1 hr		12.1

W	annotate	2hr	12.3
ee			
k			
15			