LIN 313 Language and Computers
9/15/2025
Inter-Annotator Agreement
Name:_____
UT EID#:_____
Partner: _____

Part 1: Experimental Design (2 mins).

With your partner, pick **one classification idea** from the board and pick **one of the example texts** below (a selection from Goosebumps or a selection of random sentences from our corpus). Come up with a one sentence annotation guideline. The annotation guideline should not offer a definition or interpretation of the category. Rather, it should explain how to apply the category to the sentence. The guideline should answer the question of *what* about the example you (the annotator) are labeling: The speaker? Your own feelings as the 'listener'? The situation being discussed?

Example: "Mark as "green-flag" when the *situation* being discussed is a green flag to the average college student and "red-flag" if it would be a red flag to the average college student."
Example: "Mark as "locked-in" when the speaker/person being discussed is locked in. Base your judgment on whether the speaker / person being discussed is more important in a given sentence. "

Text (goosebumps or our corpus): _____

Classification Scheme : _____ vs. _____

Annotation Guideline:

Part 2: Annotation (5 mins).

By yourself, go through each sentence in **the text you two chose** and label that sentence with one category or the other **according to your classification scheme** and guidelines.

Text 1: Random sentences from our corpus

doc	sentence	Sentence	Label
23	81	How could I lose to one so slow?”	

6	77	let us drink!	
32	41	I shook my head.	
25	12	None of the other kids in my entire school said a single word about the sailor that I was forced to don each day, which I doffed as soon as I exited the car at drop off.	
4	152	Oh, all right.	
9	30	He longs to return to\nhome and country.	
4	943	Don't waste time\ntrying to get t-two stars at a time in the scope field.	
8	173	I've already told you, we are just friends," protested Corentin while Anais stood silent, blushing just a bit.	
4	720	Do the men know of this?'\n 'No, of course not!'\n 'Good!	
4	702	'They are in trouble?'\n 'Not they.'	
7	50	He eyed them angrily as he\npassed.	
4	200	And\nthen, they say, things called Stars appeared, which robbed men of their\nsouls and left them unreasoning brutes, so that they destroyed the\ncivilization they themselves had built up.	
14	69	It is not merely an absence of sound, but a presence of concentrated thought.	
4	386	You will go mad, completely\nand permanently!	
14	22	These charges can be at rest, known as static electricity, or in motion, which constitutes an electric current.	
22	90	She tried to smile but couldn't.	

Text 2: Egg Monsters from Mars (R.L. Stine)

sent_id	sentence	label
1	Thump, thump, THUMP.	
2	I had to see what was happening in my dresser drawer.	
3	Had the egg hatched?	
4	Was the turtle bumping up against the sides of the drawer, trying to climb out?	
5	Was it a turtle?	
6	Or was it something weird?	
7	Suddenly I felt very afraid of it.	

8	I took a deep breath and rose to my feet.	
9	My legs felt rubbery and weak as I made my way across the room.	
10	My mouth was suddenly as dry as cotton.	
11	Thump, THUMP, thump.	
12	I clicked on the light.	
13	Blinked several times, struggling to force my eyes to focus.	
14	The steady thuds grew louder as I approached the dresser.	
15	Heartbeats, I told myself.	
16	Heartbeats of the creature inside the egg.	

Part 3: Calculating inter-annotator agreement (10 mins).

Calculate your level of agreement with your partner using Cohen's Kappa.

1. First, fill in the confusion matrix below for your annotation task. It should record counts of each possible combination of agreement/disagreement between you and your partner.
2. Then, navigate to https://www.graphpad.com/quickcalcs/kappa1/ and calculate agreement
    a. Cohen's Kappa: _____
3. Finally, navigate to the Instapoll called "Experiment 1" and enter your results in this format:

> Classification scheme:
> Text:
> Annotation guideline:
> Cohen's kappa:

Example confusion matrix: Gabriella and Sooji annotate locked in vs locked out for goosebumps

		Person 1: Sooji		
		label a: "locked in"	label b: "locked out"	total:
Person 2: Gabriella	label a: "locked in"	4	1	5
	label b: "locked out"	3	8	11
	total:	7	9	total N = 16

Number of observed agreements: 12 (75.00% of the observations)
Number of agreements expected by chance: 8.4 (52.34% of the observations)

Kappa= 0.475

Your confusion matrix:

		Person 1: _____		
		label a: _____	label b: _____	total:
Person 2: _____	label a: _____			
	label b: _____			
	total:			total N =

Number of observed agreements:
Number of agreements expected by chance:

Kappa=

Part 4: Experiment 2 (10 mins).

Repeat your first experiment, changing **one thing**. You can either a) pick a *new classification scheme* and annotate the *same text*, **or** b) keep the *same classigication scheme* and annotate the *other text*.

Text (goosebumps or corpus): _____

Classification scheme: _____ vs. _____

Annotation guideline (if different):

Text 1: Random sentences from our corpus:

doc	sentence	Sentence	Label
23	81	How could I lose to one so slow?”	
6	77	let us drink!	
32	41	I shook my head.	
25	12	None of the other kids in my entire school said a single word about the sailor that I was forced to don each day, which I doffed as soon as I exited the car at drop off.	
4	152	Oh, all right.	
9	30	He longs to return to\nhome and country.	
4	943	Don't waste time\ntrying to get t-two stars at a time in the scope field.	
8	173	I've already told you, we are just friends," protested Corentin while Anais stood silent, blushing just a bit.	
4	720	Do the men know of this?'\n 'No, of course not!'\n 'Good!	
4	702	'They are in trouble?'\n 'Not they.'	
7	50	He eyed them angrily as he\npassed.	
4	200	And\nthen, they say, things called Stars appeared, which robbed men of their\nsouls and left them unreasoning brutes, so that they destroyed the\ncivilization they themselves had built up.	

14	69	It is not merely an absence of sound, but a presence of concentrated thought.	
4	386	You will go mad, completely\nand permanently!	
14	22	These charges can be at rest, known as static electricity, or in motion, which constitutes an electric current.	
22	90	She tried to smile but couldn't.	

Text 2: Egg Monsters from Mars (R.L. Stine)

sent_id	sentence	label
1	Thump, thump, THUMP.	
2	I had to see what was happening in my dresser drawer.	
3	Had the egg hatched?	
4	Was the turtle bumping up against the sides of the drawer, trying to climb out?	
5	Was it a turtle?	
6	Or was it something weird?	
7	Suddenly I felt very afraid of it.	
8	I took a deep breath and rose to my feet.	
9	My legs felt rubbery and weak as I made my way across the room.	
10	My mouth was suddenly as dry as cotton.	
11	Thump, THUMP, thump.	
12	I clicked on the light.	
13	Blinked several times, struggling to force my eyes to focus.	
14	The steady thuds grew louder as I approached the dresser.	
15	Heartbeats, I told myself.	
16	Heartbeats of the creature inside the egg.	

Part 5: Experiment 2 Agreement (10 mins).

Calculate your level of agreement with your partner using Cohen's Kappa.

4. First, fill in the confusion matrix below for your annotation task. It should record counts of each possible combination of agreement/disagreement between you and your partner.
5. Then, use the forumula for Cohen's Kappa below to calculate the interannotator agreement
    a. Cohen's Kappa: _____
6. Finally, navigate to the Canvas Assignment called "In Class: Round 1 Annotations" and enter your results from both experiments

Your confusion Matrix:

		Person 1: _____		
		label a: _____	label b: _____	total:
Person 2: _____	label a: _____			
	label b: _____			
	total:			total N =

$$\kappa = \frac{p_0 - p_e}{1 - p_e},$$

p_0 = observed agreement
p_e = agreement expected by chance
$p_e = p_e(\text{label 1}) + p_e(\text{label 2})$

Proportion of observed agreement:

Probability of agreement expected by chance (calculate for each category label separately and then add):

Kappa =