A Critical Role (and Dimension 20) Discourse Analysis

Ashley Bakaitus

Ling 2340

April 16, 2025

D&D Basics:

DM - Dungeon Master (narrator)
    - knows the world & story, plays out scenes for players to engage in
Players - wizards, rogues, fighters, barbarians, elves, halflings, dwarves…
    - uncover plots, fight bad guys & monsters, solve problems, cause chaos
Campaign - The story/adventure, start to finish

Three main elements: Description, imagination, rolling dice

Research Question & Goals

Analyze speech data from two D&D liveplay podcasts, Critical Role and Dimension 20. One of which has a man as the DM, and the other with a woman DM.

Using our familiar metrics, token counts, sentence counts and length, as well as looking for hedging, jargon, and "readability", investigate if there is any significant difference between our two DMs or if they both share a relatively equal balance in participation with their players.

Data Basics

Critical Role
1 season
141 episodes (147 files)
1 ep ~4 hrs long
1 DM (Matt)
7 players (3W 4M)

Dimension 20
4 seasons + 3 chat seasons
54 episodes & files
1 ep ~2-2.5 hrs, 4hrs
1 DM (Aabria)
18 players (8W 8M 2NB)

The Data

Data treatment and dataframe building

-HTML files converted to JSON files
-JSON normalized until in decent shape

	name	text	ts.h	ts.m	ts.s	episode
0	MATT	We've got some cool stuff to talk about. First...	0	0	15	1
1	LAURA	As in nnnn.	0	1	21	1
2	MATT	Nnnn. But yeah, so they're going to be a long-...	0	1	24	1
3	TRAVIS	(vomiting noises) Would you like some?	0	2	4	1
4	MATT	So yeah. I'm super excited to have that, guys....	0	2	7	1

For D20, this was a lot less decent than CR…

	index	variable	value	episode	season
9	9	Aabria: Hello, friends, and welcome. We're her...	None	1	Misfits and Magic
10	10	a new game called "Misfits and Magic"! [cheeri...	None	1	Misfits and Magic
11	11	Cast cheers and laughs.	None	1	Misfits and Magic
12	12	Aabria: I am your game master, Aabria Iyengar,...	None	1	Misfits and Magic
13	13	Magical Misfits. Say hi, Magical Misfits!	None	1	Misfits and Magic

Critical Role data was much easier…

- Data speech lines contained non-speech information
    - Get rid of it? NO
    - Spend hours and hours of weekends using regex to move to another column? YES

```
8783                    (whizzing) (explosion) (screaming)
15411        (breathing heavily) (whooshes) (surprised laug...
17487                  (whooshing) (whooshing) (flapping)
25846                (as Marius) (as Orly) (as Marius)
41274        (Luc struggle grunts) (as Pumat) (poof)
                              ...
417661                  (grunting) (gasping) (laughter)
417680      (yelling) (digital bleeping) (dial up tones)
426365                  (gasping) (laughter) (laughter)
427855        (exhale) (heavy wind gust) (rolling)
432836                  (laughs) (explosion) (laughter)
```

	index	name	text	episode	timestamp	inaudible_speech	nonspeech
0	0	MATT	We've got some cool stuff to talk about. First...	1	00:00:15		
1	1	LAURA	As in nnnn.	1	00:01:21		
2	2	MATT	Nnnn. But yeah, so they're going to be a long-...	1	00:01:24		
3	3	TRAVIS	Would you like some?	1	00:02:04		(vomiting noises)
4	4	MATT	So yeah. I'm super excited to have that, guys....	1	00:02:07		

7

Dimension 20 data had a much less tidy starting point

- Names in same column with speech
- Non speech information in same column with speech
- Single speech units split out over multiple df rows
- Character names instead of player names

	index	name	player	text	episode	inaudible_speech	nonspeech	season
0	10	Aabria	Aabria	To the churlish Captain K.P. Hob, venerated s...	10			ACOFF
1	11	Aabria	Aabria	Goblin Court. To the master of ceremonies, Del...	10			ACOFF
2	12	Aabria	Aabria	the Court of Wonder. To BINX Choppley, sole su...	10			ACOFF
3	13	Aabria	Aabria	Craft. To the tenebrous Prince Andhera, scion ...	10			ACOFF
4	14	Aabria	Aabria	to the notorious Lords of the Wing, Lady Chirp...	10			ACOFF

More and more hours of regex later, 3/4 of these problems were solved!
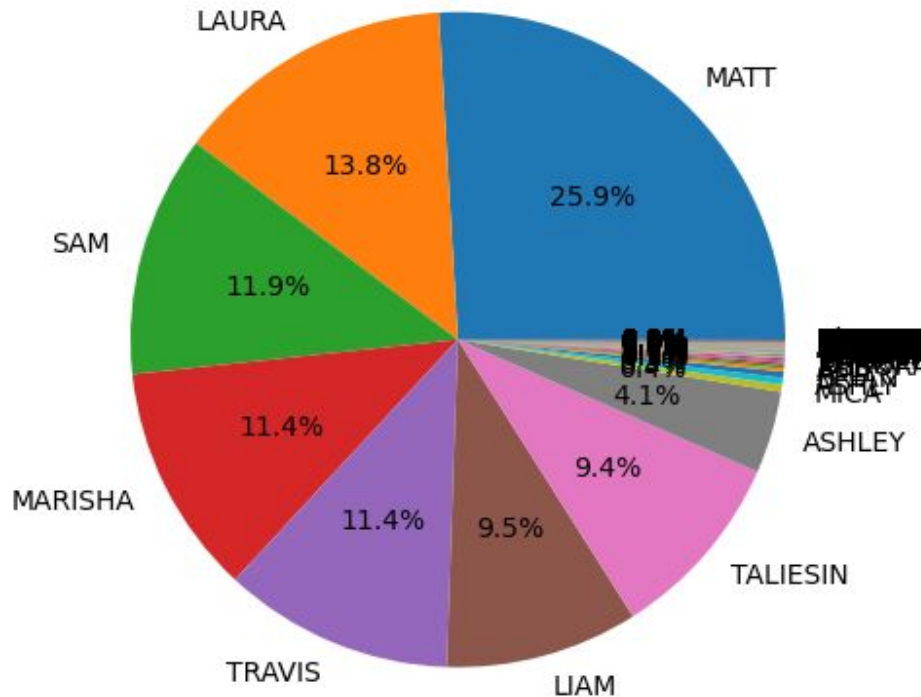
Analysis

Is there a significant difference
(in these metrics)
between a man and woman DM?
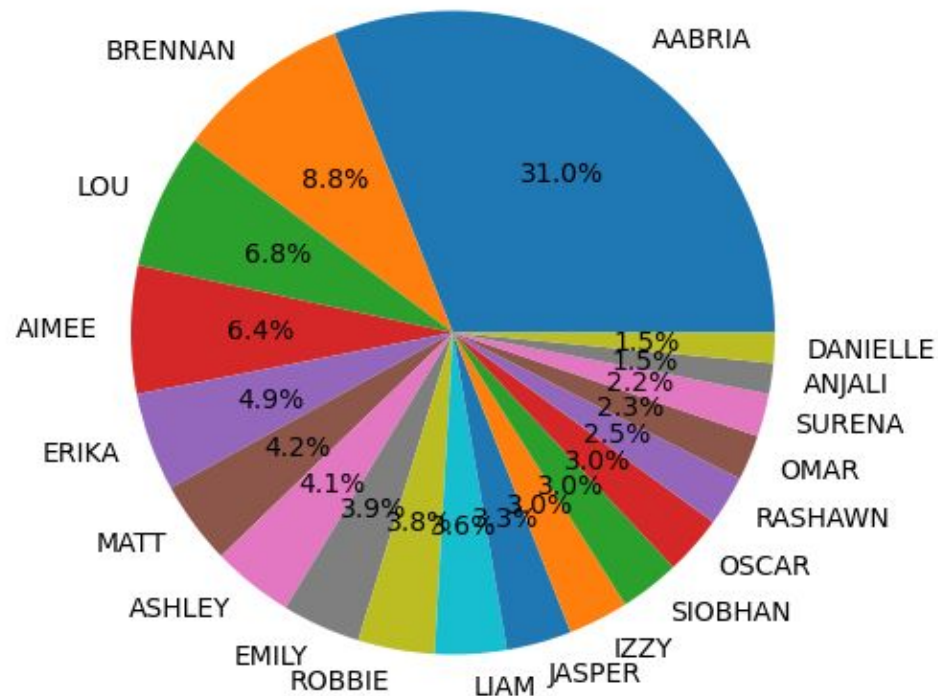
Information to consider

1. CR is only one season over 3.5 yrs, not everyone was present for every episode. Out of 141 episodes, Ashley was only there for 86!

2. D20 is 4 seasons with mostly unique players. (Brennan is in 3/4 seasons, Lou and Erika are in 2)

3. Character types: excitable and chatty or stoic and serious?
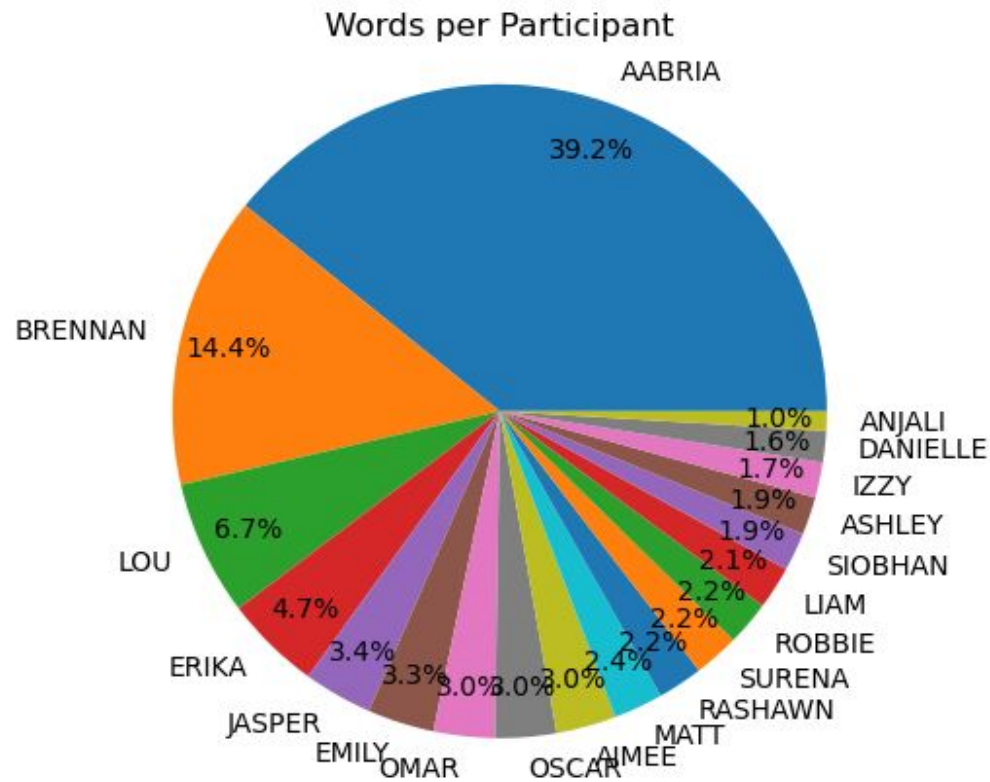
4. Role among the group?
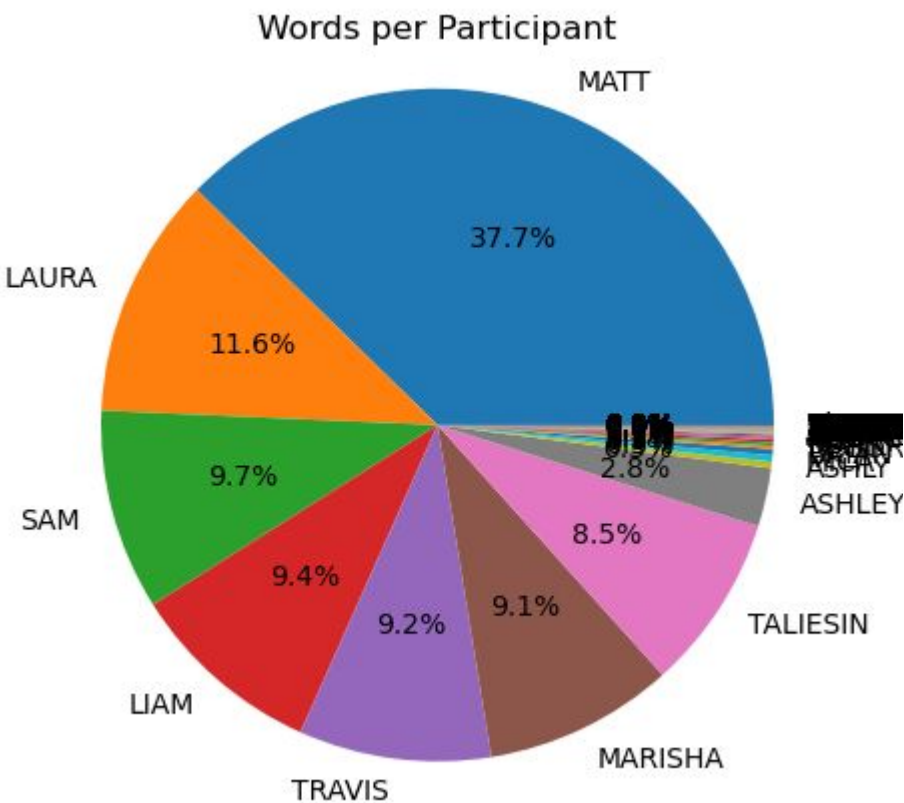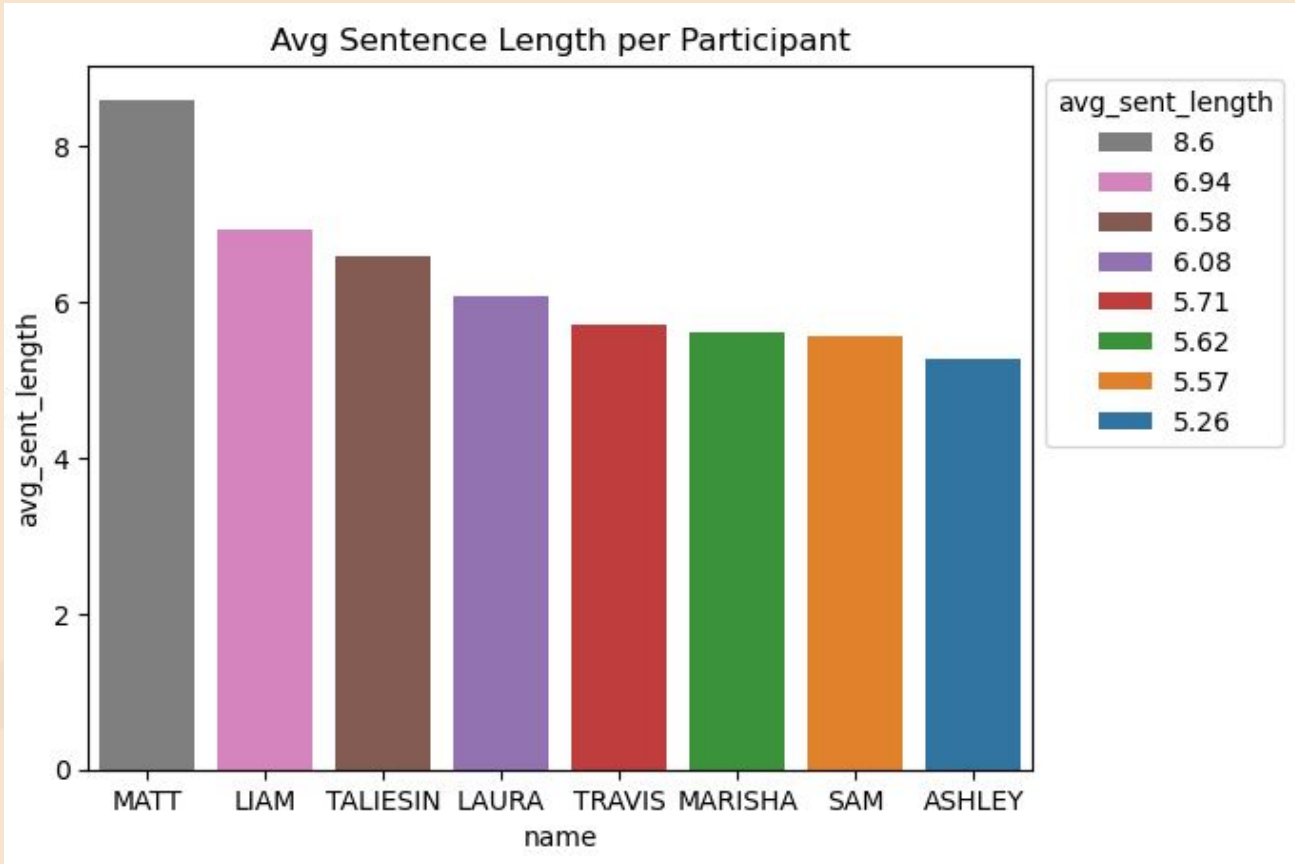
Speech turn per participant
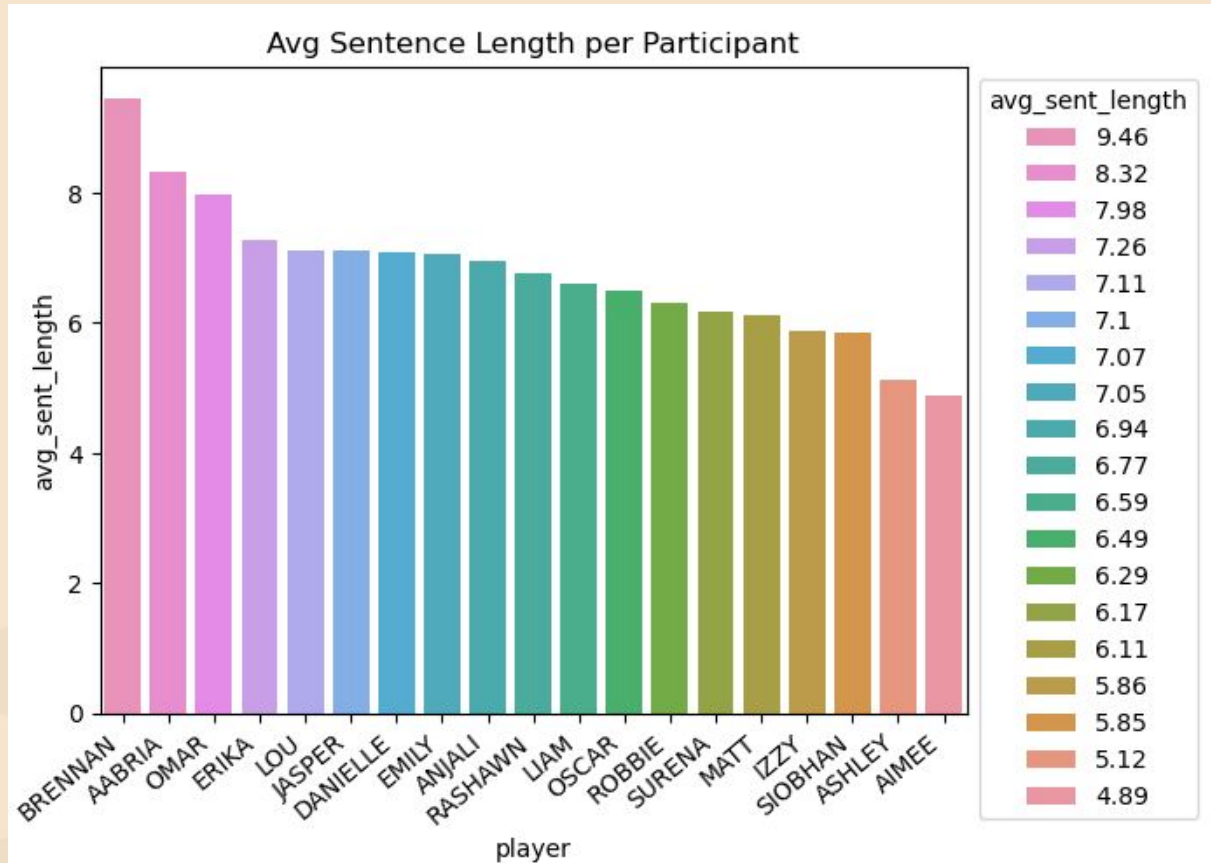
Words spoken per participant
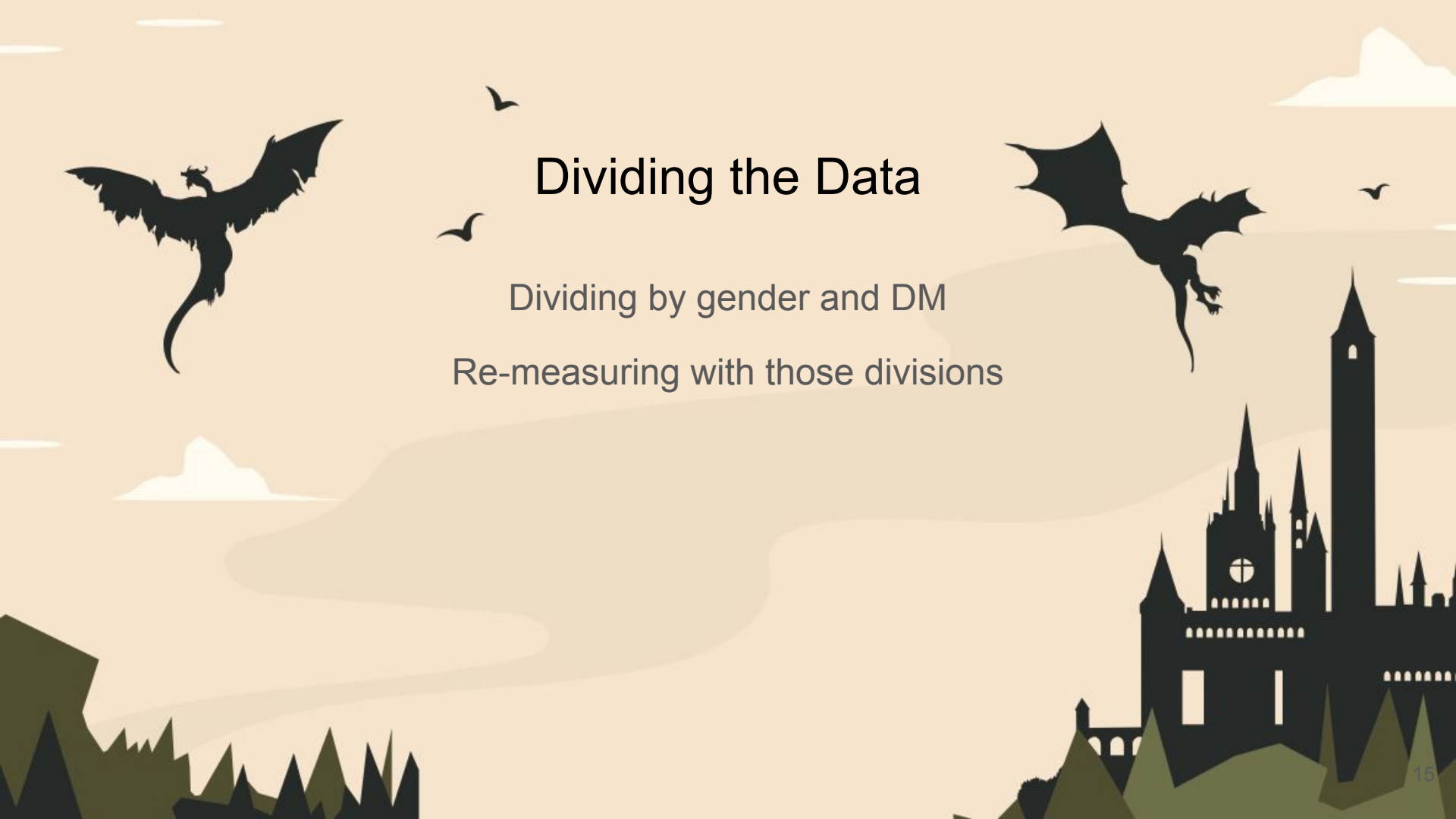
Average sentence length per participant

Average sentence length per participant

Dividing the Data

Dividing by gender and DM

Re-measuring with those divisions
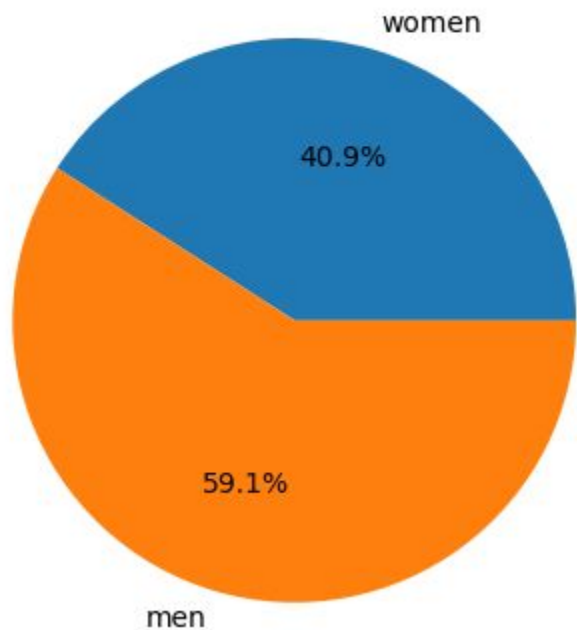
Information to remember

Critical Role has 1 more man than woman & 1 woman missed 40% of episodes

D20 has 8 men, 8 women, and 2 nonbinary players (both use she/they)

Speech Turns (Critical Role)

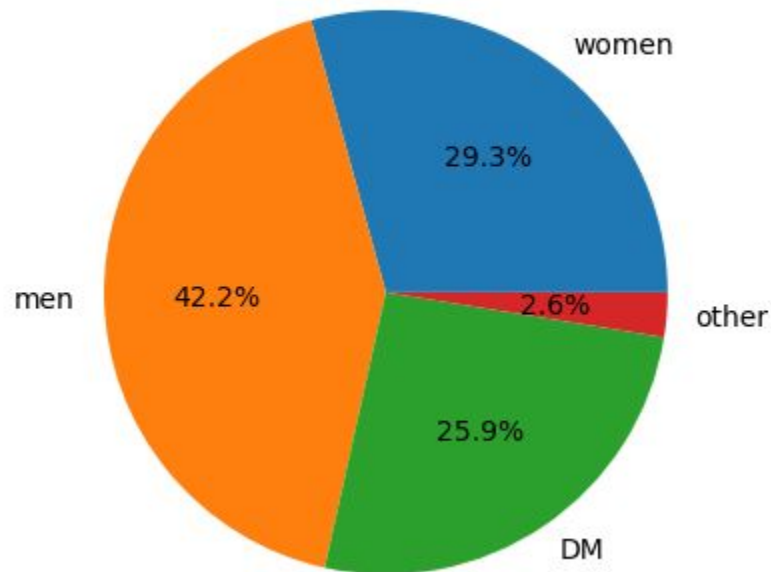# Speech Turns (D20)

Total Speech Turns by Gender (Among Players)

- women 37.8%
- men 51.9%
- nonbinary 10.3%
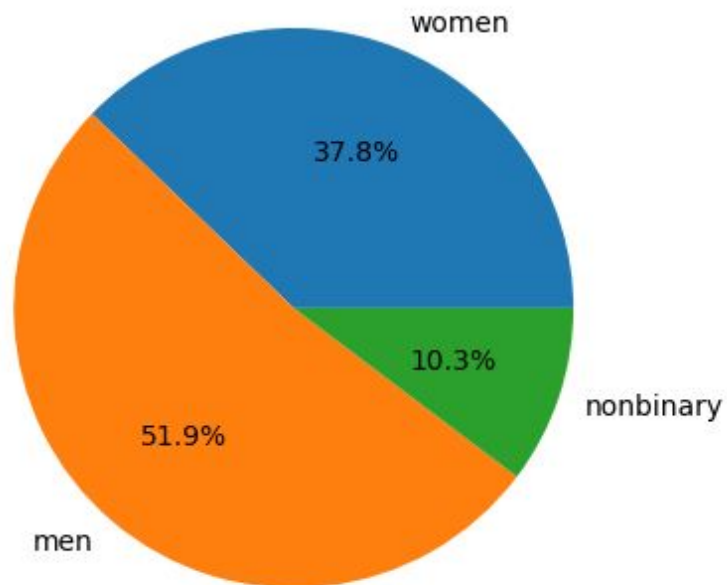
Total Speech Turns by Gender (Including DM)

- women 26.1%
- men 35.8%
- nonbinary 7.1%
- DM 31.0%
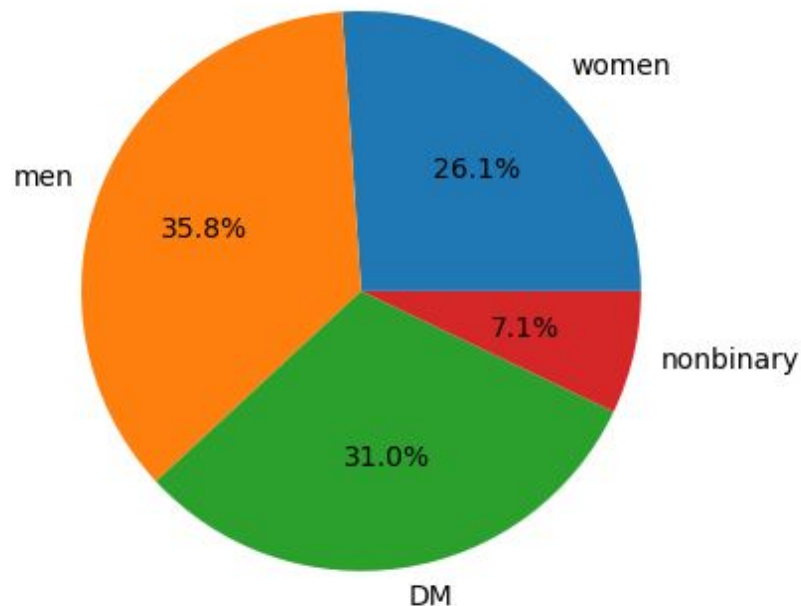
Total Words (Critical Role)

Total Words (D20)

Total Words by Gender (Excluding DM)

- women 27.2%
- nonbinary 11.5%
- men 61.4%
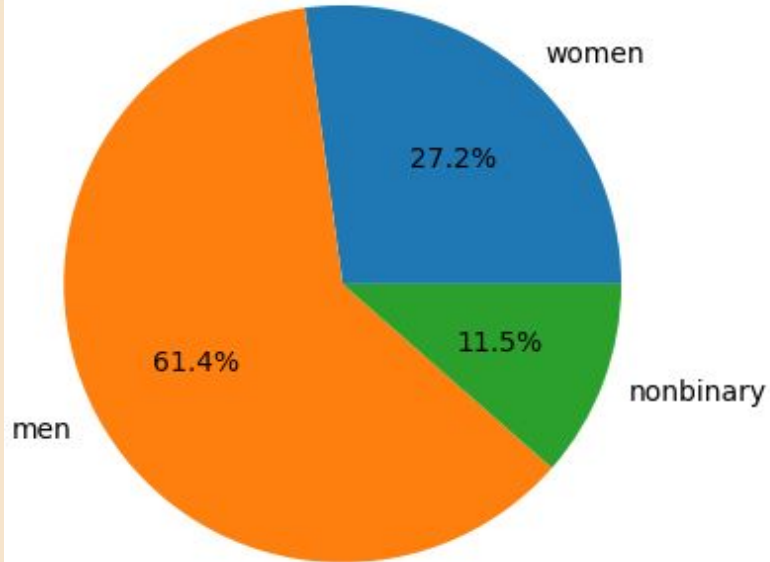
Total Words by Gender (Including DM)

- men 37.3%
- women 16.5%
- nonbinary 7.0%
- DM 39.2%
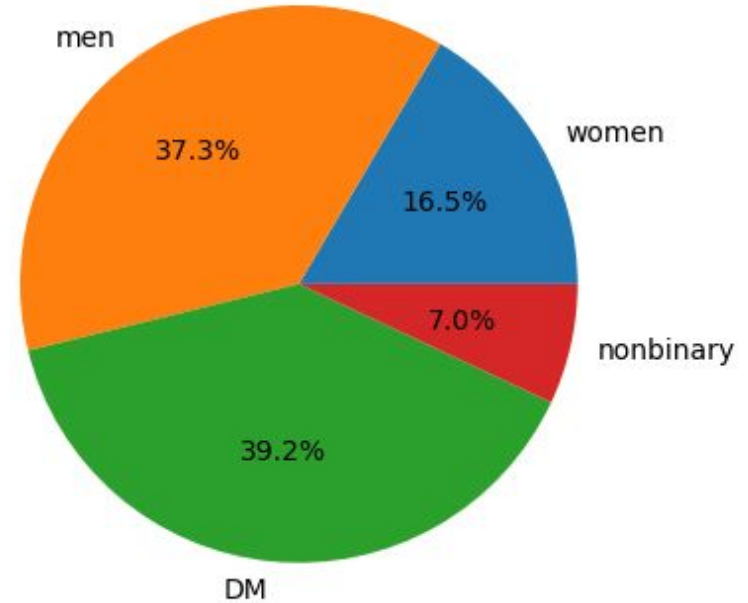
Average Sent Length (Both)
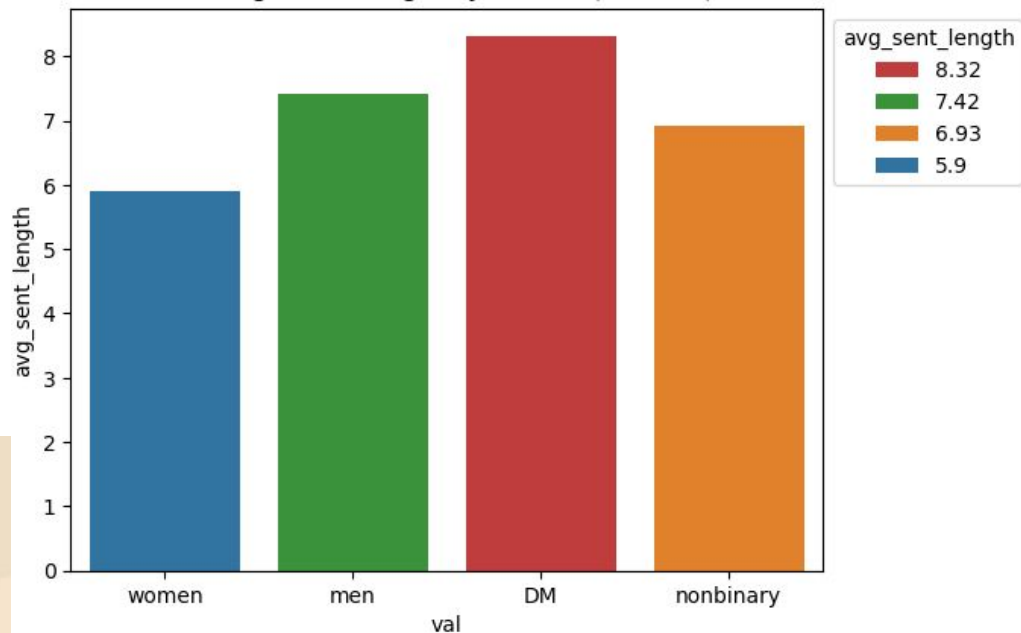
Hedging

Searching for and calculating hedging

Combination of list comprehension and regex string count for words like:
Maybe, might, suggests, possibly, appears, seems, tends, I guess, I think, could be.

In Critical Role data:
For every thousand words spoken:
5.68 words spoken by women,
5.75 words spoken by men, and
2.64 words spoken by the DM are among the scanned for hedging words and phrases.

In Dimension 20 data:
For every thousand words spoken:
6.93 words spoken by women,
6.25 words spoken by men,
5.64 words spoken by nonbinary players, and
2.35 words spoken by the DM are among the scanned for hedging words and phrases.

Jargon

Defining Jargon

Using the full text of the Player's Handbook, compare with the ~330k word Norvig Google unigram data. If not in the unigram list, it's defined as jargon.

What was left over?

1. **vocab**: subrace, clanless, cantrip, multiattack
2. **names**: thamior, breena, stumbleduck, myrkul, agathys, keyleth, strahd
3. **locations**: shadowfell, feywild, candlekeep
4. **spells**: thaumaturgy, barkskin, truesight, enfeeblement, longstrider, cloudkill, revivify, countercharm, thunderwave
5. **groups**: archfey, dragonborn, tiefling(s), lightfoots, merfolk, lizardfolk
6. **creatures**: hippogriffs, wererats, tarrasque, owlbear
7. **weapons**: greatclub, handaxes, scimitars, greataxe, longswords, shortbow, shortsword
8. **languages**: dwarvish, undercommon, draconic

Some of my personal favorites:
Sibilants, blibdoolpoolp, mordenkainen, maglubiyet

Jargon Frequency

In Critical Role:

For every million words spoken:

304.8 words spoken by the women players,

285.2 words spoken by men players, and

262.6 words spoken by the DM will be from this jargon list.

Jargon Frequency

In Dimension 20:
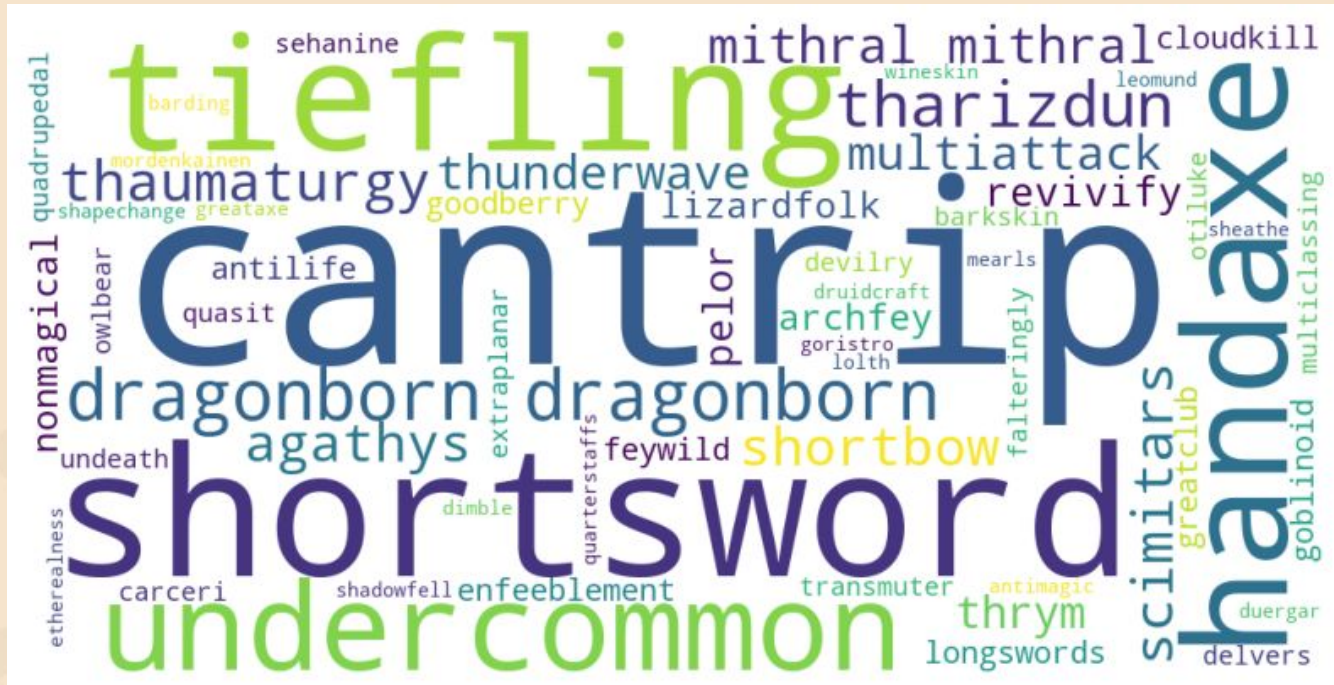
For every million words spoken:

154.3 words spoken by the women players,

211.8 words spoken by men players,

73 words spoken by nonbinary players, and

201.7 words spoken by the DM will be from this jargon list.

Why the difference?

In my opinion, genre.

CR tends to play high fantasy, classic D&D - like from the PHB

D20 are more varied, playing as stoats, magical school, regency romance… and one high fantasy story. Plus the chat shows!

Readability Scores

Getting a Baseline

The Flesch–Kincaid readability test formula:

$$206.835 - 1.015 \left(\frac{\text{total words}}{\text{total sentences}} \right) - 84.6 \left(\frac{\text{total syllables}}{\text{total words}} \right)$$

Designed to ensure things like military instructional books and medical aftercare reports are readable to a majority of people.

1-120, the higher a score, the easier it is to read. (also converts to a grade level)

Time Magazine : 52, Moby Dick : 57.9, highest score 121 is all words are 1 syllable (like Dr. Seuss!)

The Scores

For PHB: ~49, at the easier end of the "difficult" level. Grade level: 11

For Critical Role:
Women: 85
Men: 83
DM: 73
Range of easy to fairly easy, 6/7th grade.

For Dimension 20:
Women: 85
Men: 79
NB: 80
DM: 74

Problems with "readability"

Readability vs comprehensibility
- IRS tax forms have a "readability" grade score of 8th grade

NLTK syllable tokenizer is not very good, these scores are iffy

```
[['gno', 'mes'],
 ['Most'],
 ['gno', 'mes'],
 ['in'],
 ['the'],
 ['worlds'],
 ['of'],
 ['DND'],
 ['a', 're'],
 ['rock'],
 ['gno', 'mes'],
 ['in', 'clu', 'ding'],
 ['the'],
 ['tin', 'ker'],
 ['gno', 'mes'],
 ['of'],
 ['the'],
 ['Dra', 'gon', 'lan', 'ce'],
 ['set', 'ting'],
 ['Abi', 'li', 'ty'],
 ['Sco', 're'],
 ['Increa', 'se'],
 ['Yo', 'ur'],
 ['Cons', 'ti', 'tu', 'tion'],
 ['sco', 're']]
```

Thank you!