

## Abstract

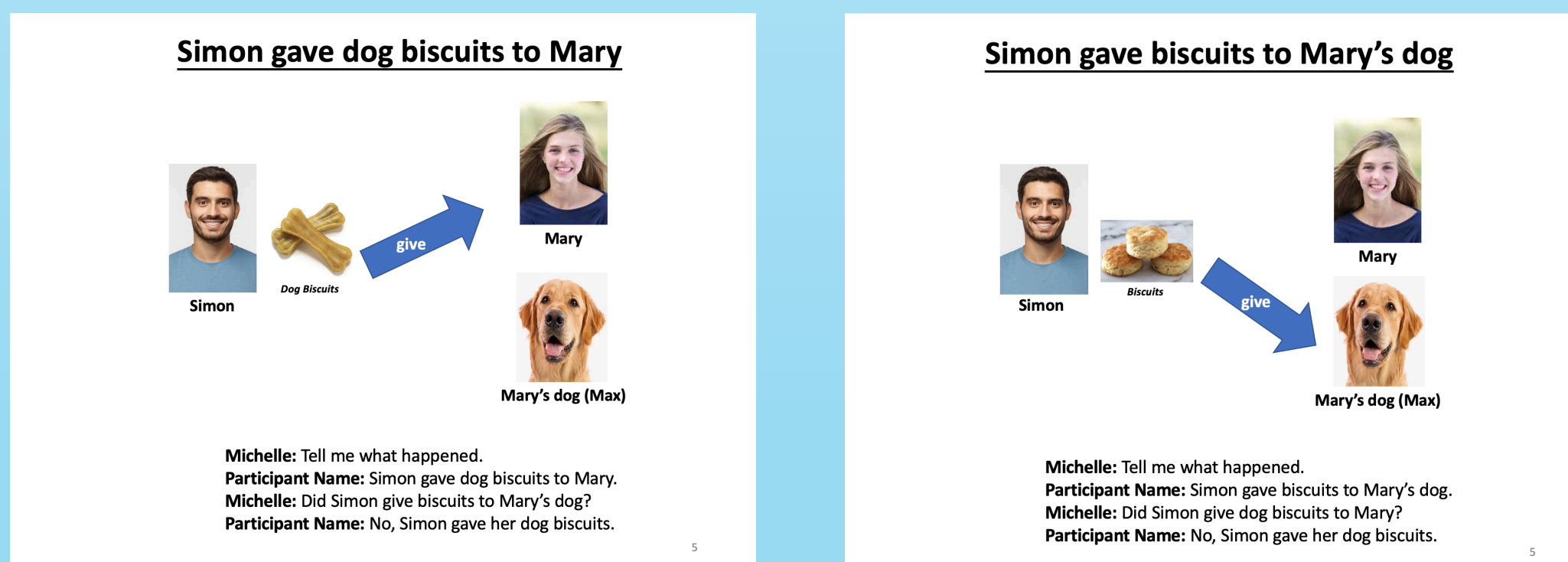
**Prosody** is "the linguistic structure expressed in the suprasegmental properties of utterances" (Cutler, 2012). Suprasegmental properties include pitch, tempo, loudness and timing patterns. Suprasegmental cues are also referred to as prosodic cues. Prosodic cues are crucial to parse sentences and resolve syntactic ambiguity. There is limited literature on prosody research because prosody is poorly integrated into speech processing models. In our study, we asked participants to mimic human conversation while going through PowerPoint presentation via BlueJeans. Audacity was used to record the conversations and PRAAT was used to annotate and analyze the recordings. In this study, we focused on comparing duration in early and late juncture sentence pairings. We completed a qualitative analysis of the data.

## Participants

- 20 participants were included in the study
- Mean Age: 26 years old
- To be eligible to participate, subjects must:
  - have been born in the United States
  - English must be their dominant language

## Methods and Materials

- Due to COVID-19, the experiment was conducted online via BlueJeans
- Participants were instructed to download Audacity and record themselves during the entire session
- I shared my screen and showed participants a PowerPoint presentation that contained forty slides
- Each slide included a title, graphics, and script like the ones pictured below



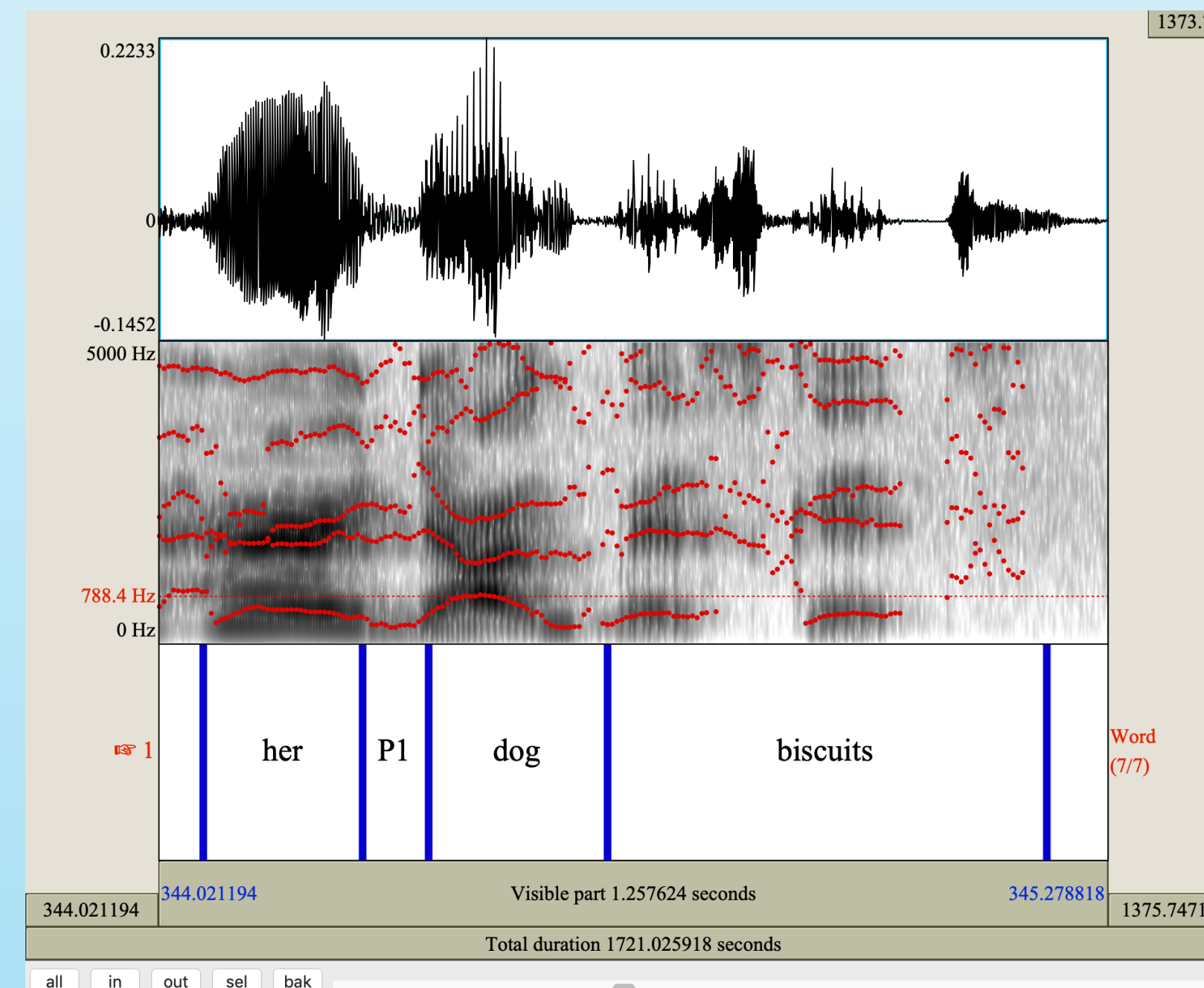
**Figure 1.** The two example slides include the same individuals, but they differ in meaning. The first slide is the early juncture sentence which suggests that Mary is receiving dog biscuits. The second slide is the late juncture sentence which suggests that Mary's dog is receiving biscuits. The images and arrow clearly display the relationship that the participant is asked to clarify in the script.

- Participants were given 20 seconds to look at each slide
- Once the 20 seconds were up, I initiated the conversation
- To best mimic a natural conversation, participants were instructed to "act out" each script instead of reading it
- The participant's goal was to clarify the situation to me
- All acoustic data was recorded from the participant's end; At the end of the session, each participant emailed his/her Audacity WAV file to me

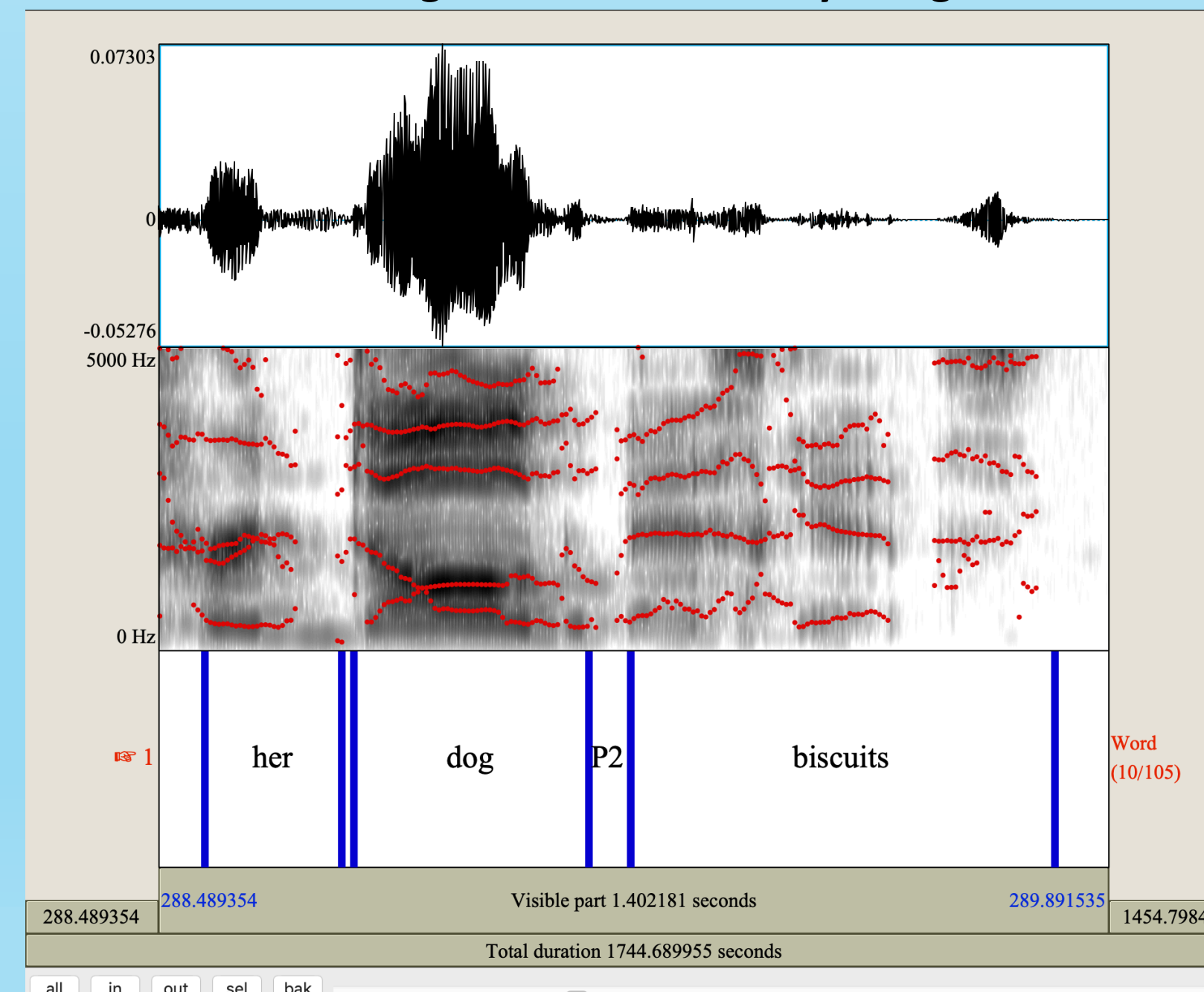
## Qualitative Analyses

- PRAAT was used to annotate each WAV file
- The file was segmented into trials, items, words, and sonorants
- Pauses are labeled in the word tier
- We extracted the word tier to compare duration between each sentence pairing

### "Simon gave dog biscuits to Mary."



### "Simon gave biscuits to Mary's dog."



**Figure 2.** The PRAAT TextGrids above represent a sentence pairing. The boundaries illustrate the start and end of the target words as well as the pauses in between them (P1 & P2). Although the participant uses the same words in each sentence, duration values can be used to better illustrate the correct meaning of the corresponding sentence.

## Qualitative Analyses (cont.)

### Duration in Milliseconds

Juncture Type	"her"	P1	"dog"	P2	"biscuits"
Early	211	87	237	0	582
Late	202	18	347	62	626

- This example illustrates a common trend in our data
- P1 and "her" is longer in the early juncture sentence than in the late juncture sentence
- P2 and "dog" is longer in the late juncture sentence than in the early juncture sentence
- P2 was zero in early juncture sentence

## Conclusions

Due to time constraints, we were unable to complete statistical analysis of the data. Nonetheless, after annotating the Audacity files on PRAAT, our preliminary qualitative analyses showed a similar pattern throughout the data. In early juncture sentences, P1 and "her" are longer. In late juncture sentences, P2 and "dog" are longer. Therefore, duration is used to parse sentences and resolve syntactic ambiguities. Prosodic cues are central to sentence comprehension. It would be interesting to look at how other prosodic cues (e.g., pitch) are used to resolve syntactic ambiguity. It would also be interesting to replicate this study in person and use eye tracking technology to see how participants make use of different visual cues. Finally, future research should also explore how listeners' use of prosodic structure differs across languages. Ultimately, our current data demonstrate that duration is an important cue in everyday discourse.

## Acknowledgements

Thank you to the Penn Undergraduate Research Mentorship Program (PURM) for funding this project. I would like to thank John Trueswell for offering guidance and access to his Language Learning Lab, Jianjing Kuang for her expertise in phonetics, and Martin Ho Kwan Ip for his diligent mentorship.