Dialogue Data Evaluation

Zoltán Csik

2023-09-08

Data Preparation

This R Markdown serves the purpose of evaluating acquired dialogue data. The collected data represents several linguistic dialogue features.

```
##                      Feature    HH1    HH2    HH3    HC1    HC2    HC3
## 1            Sentence Length 52.558 55.209 73.681 33.372 22.452 38.393
## 2         Words per Sentence 13.767 13.478 20.931  8.721  5.710  9.714
## 3 Unique words per Sentence 12.093 11.791 17.056  8.093  5.613  8.857
## 4          Lexical Diversity  9.287 11.006  9.816  9.618  6.915  8.277
```

Divided Dialogues:

To compare human-human data with human-robot data, the dataframe has to be split into two:

```
##                      Feature    HH1    HH2    HH3
## 1            Sentence Length 52.558 55.209 73.681
## 2         Words per Sentence 13.767 13.478 20.931
## 3 Unique words per Sentence 12.093 11.791 17.056
## 4          Lexical Diversity  9.287 11.006  9.816
```

```
##                      Feature    HC1    HC2    HC3
## 1            Sentence Length 33.372 22.452 38.393
## 2         Words per Sentence  8.721  5.710  9.714
## 3 Unique words per Sentence  8.093  5.613  8.857
## 4          Lexical Diversity  9.618  6.915  8.277
```

Human-Human Diaogue Data

Count the mean value for each linguistic feature data:

```
##                      Feature    HH1    HH2    HH3     Mean
## 1            Sentence Length 52.558 55.209 73.681 60.48267
## 2         Words per Sentence 13.767 13.478 20.931 16.05867
## 3 Unique words per Sentence 12.093 11.791 17.056 13.64667
## 4          Lexical Diversity  9.287 11.006  9.816 10.03633
```

Human-Robot Diaogue Data

Count the mean value for each linguistic feature data:

```
##                     Feature    HC1    HC2    HC3      Mean
## 1           Sentence Length 33.372 22.452 38.393 31.405667
## 2        Words per Sentence  8.721  5.710  9.714  8.048333
## 3 Unique words per Sentence  8.093  5.613  8.857  7.521000
## 4         Lexical Diversity  9.618  6.915  8.277  8.270000
```

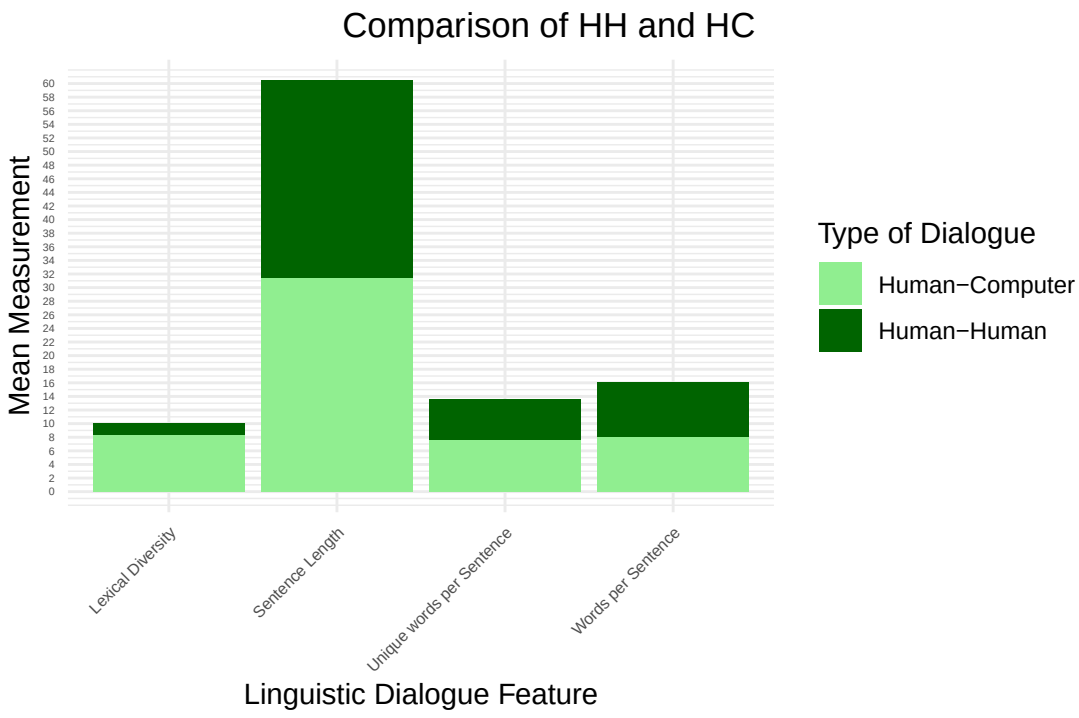
Combined Mean Values

Combine mean values from both Data Frames

```
  dialogue_summary
```

```
##                     Feature       HH        HC
## 1           Sentence Length 60.48267 31.405667
## 2        Words per Sentence 16.05867  8.048333
## 3 Unique words per Sentence 13.64667  7.521000
## 4         Lexical Diversity 10.03633  8.270000
```

Visual Comparison



Statistical Relevance

Since we have two different types of Dialogues, the Two-sample independent t-test is applied:

First application: sentence lengths compared

```
##
##  Welch Two Sample t-test
##
## data:  group_1_sentence and group_2_sentence
## t = 3.5716, df = 3.6033, p-value = 0.02778
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   5.457557 52.696443
## sample estimates:
## mean of x mean of y
##  60.48267  31.40567
```

Second application: Average words compared

```
##
##  Welch Two Sample t-test
##
## data:  group_1_word and group_2_word
## t = 2.9465, df = 2.9208, p-value = 0.06222
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.7758904 16.7965571
## sample estimates:
## mean of x mean of y
## 16.058667  8.048333
```

Third application: Average unique words compared

```
##
##  Welch Two Sample t-test
##
## data:  group_1_unique_word and group_2_unique_word
## t = 3.1129, df = 3.1877, p-value = 0.04863
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   0.0667547 12.1845786
## sample estimates:
## mean of x mean of y
##  13.64667   7.52100
```

Fourth application: Lexical Diversity

```
##
##  Welch Two Sample t-test
##
## data:  group_1_lex and group_2_lex
## t = 1.8967, df = 3.4384, p-value = 0.1423
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
```

```
##  -0.9946258  4.5272924
## sample estimates:
## mean of x mean of y
##  10.03633   8.27000
```