

# Anaphora Project Update

Jasmine Falk

September 10, 2020

## 1 Brief Review

### Hypotheses

1. Children improve in anaphora resolution ability as they age.
2. Children exhibit consistent mechanisms to resolve anaphora.
3. Parent speech the parent's perception of/assumptions about their child's language comprehension abilities.

### Follow-up from 7/23

All speech transcripts were coded again to account for **child-driven anaphora** and **visually-cued anaphora**

#### 1.0.1 Child-driven anaphora

**Child-driven anaphora** refers to cases when a parent uses an anaphor to refer to an object that the child is already attending to. For example:

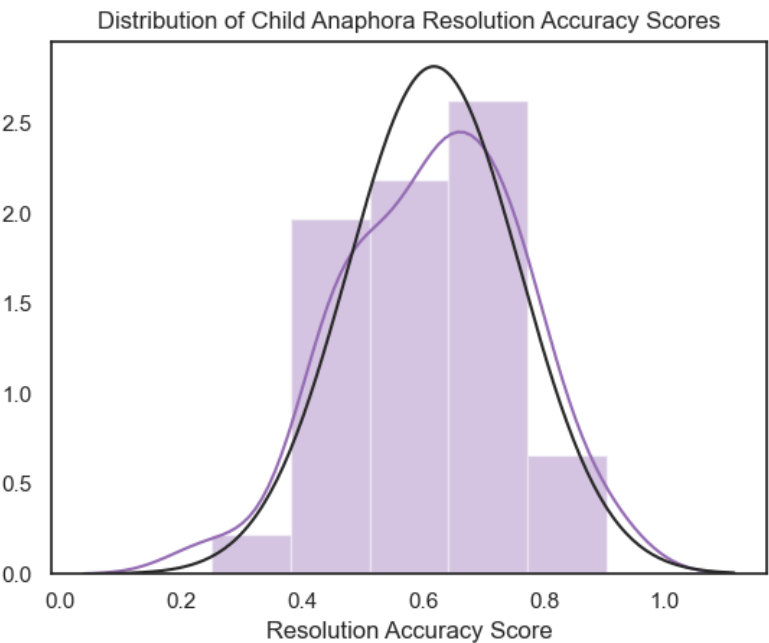
```
what did you find?  
rake  
car  
oh it has got buttons
```

#### 1.0.2 Visually-cued anaphora

**Visually-cued anaphora** refers to cases when a parent pairs an anaphor with a visual cue (e.g. pointing, picking up, etc.), whether or not that visual cue is required to resolve the anaphor. Previously, I had only coded for anaphora that required a visual cue to resolve it (could not be resolved from text alone).

## 2 Results from 7/23 Update

### 2.1 Distribution of resolution accuracy scores



mean	0.618
stdev	0.144
median	0.630
max	0.903
min	0.250

## 2.2 Child age vs. MCDI scores

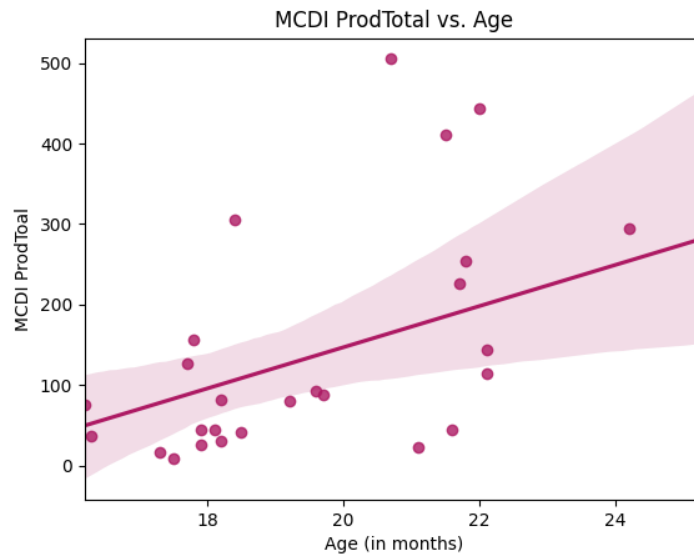


Figure 1:  $r(25) = .431$ ,  $p = .025$ .

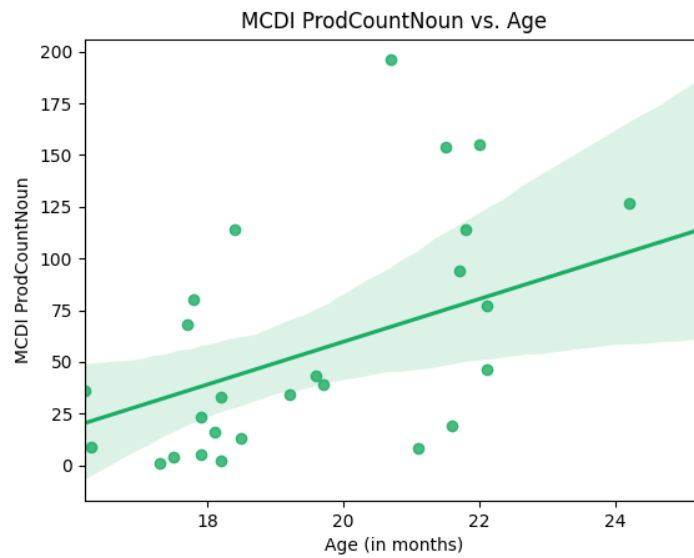


Figure 2:  $r(25) = .448$ ,  $p = .019$ .

2.3 Percentage of total utterances containing anaphora

mean	28.75
stdev	11.22
median	29.80
max	53.54
min	0.00

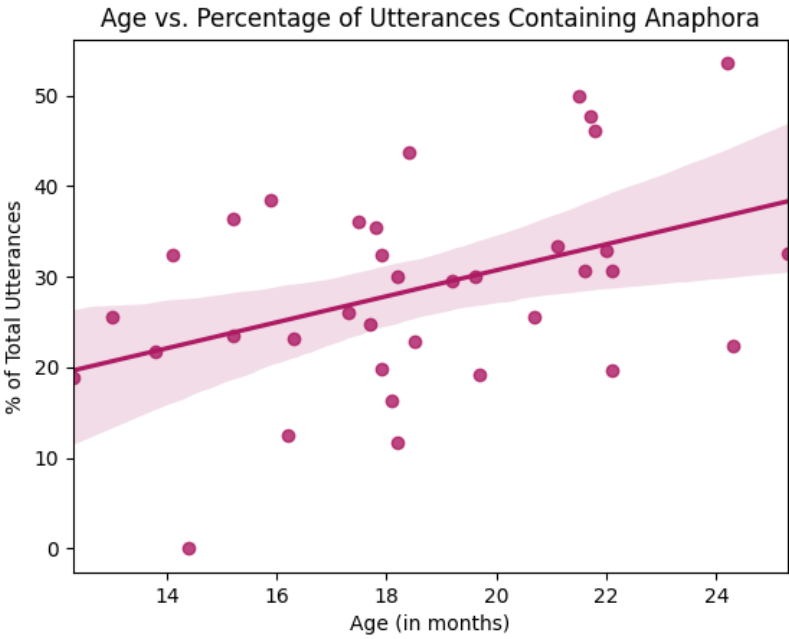


Figure 3: Compared to child age.  $r(33) = .367, p = .030$ .

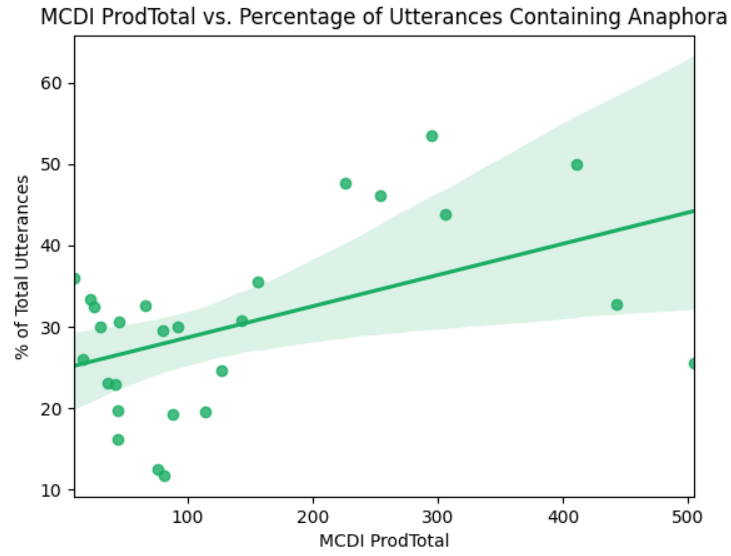


Figure 4: Compared to MCDI ProdTotal scores.  $r(25) = .490, p = .009$

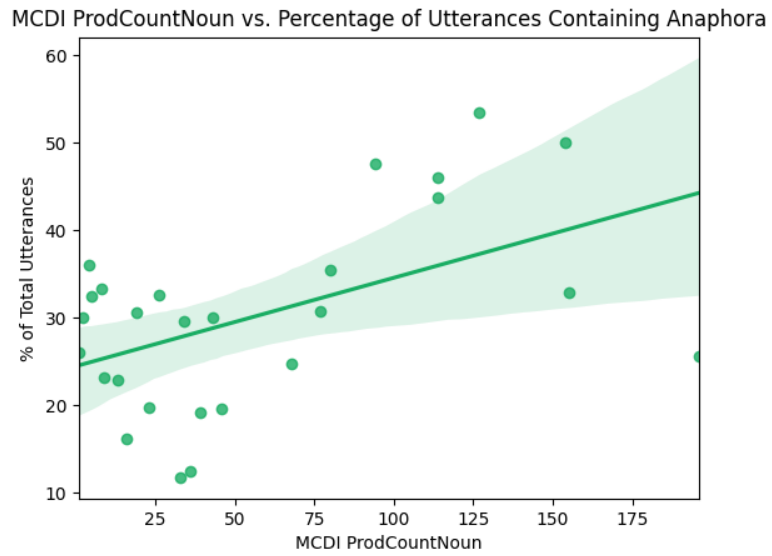


Figure 5: Compared to MCDI ProdCountNoun scores.  $r(25) = .504, p = .007$

2.4 Percentage of anaphoric utterances containing split anaphora

mean	3.82
stdev	6.98
median	0.00
max	27.12
min	0.00

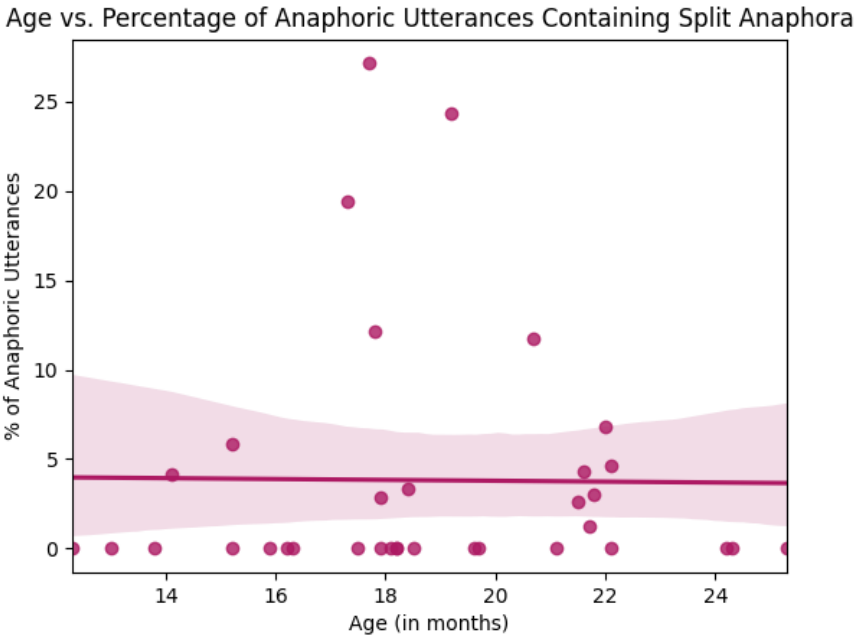


Figure 6: Compared to child age.  $r(33) = -.011$ ,  $p = .949$

MCDI ProdTotal vs. Percentage of Anaphoric Utterances Containing Split Anaphora

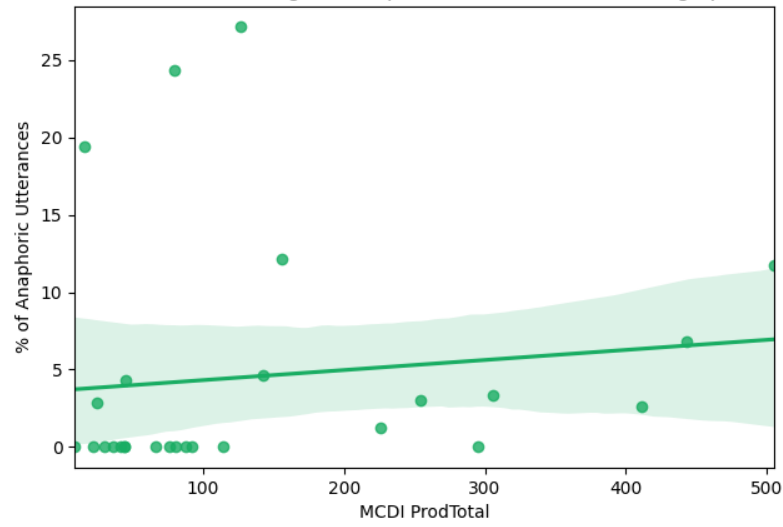


Figure 7: Compared to MCDI ProdTotal scores.  $r(25) = .118$ ,  $p = .556$ .

MCDI ProdCountNoun vs. Percentage of Anaphoric Utterances Containing Split Anaphora

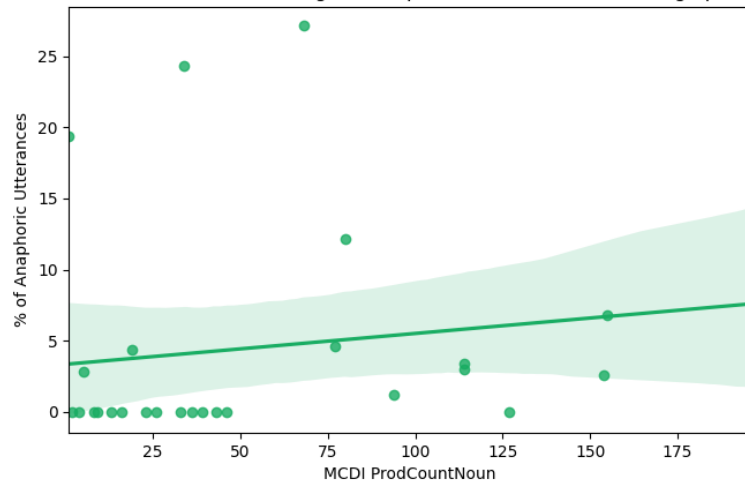


Figure 8: Compared to MCDI ProdCountNoun scores.  $r(25) = .154$ ,  $p = .444$ .

2.5 Percentage of anaphoric utterances containing one anaphora

mean	4.99
stdev	5.16
median	4.35
max	18.18
min	0.00

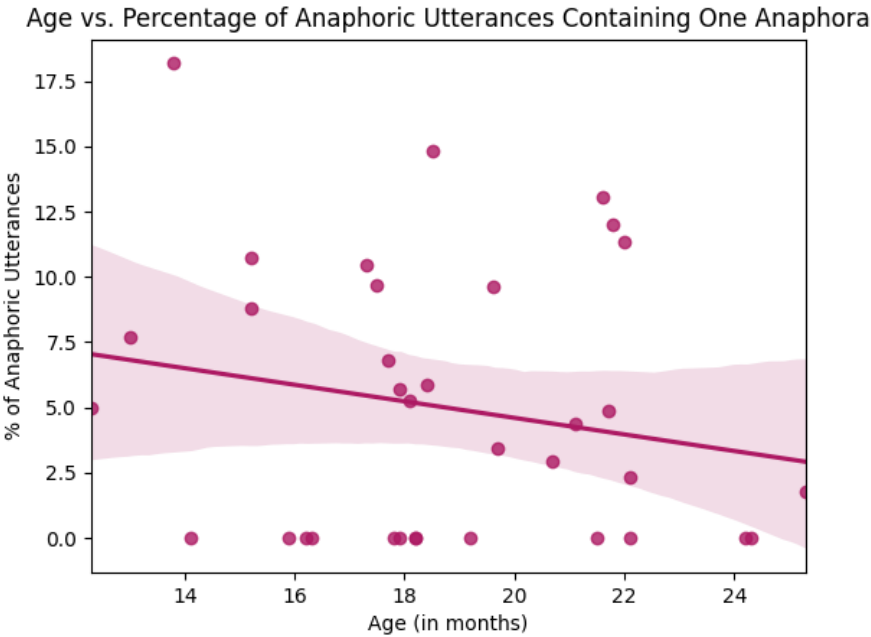


Figure 9: Compared to child age.  $r(33) = -.199$ ,  $p = .252$



MCDI ProdTotal vs. Percentage of Anaphoric Utterances Containing One Anaphora

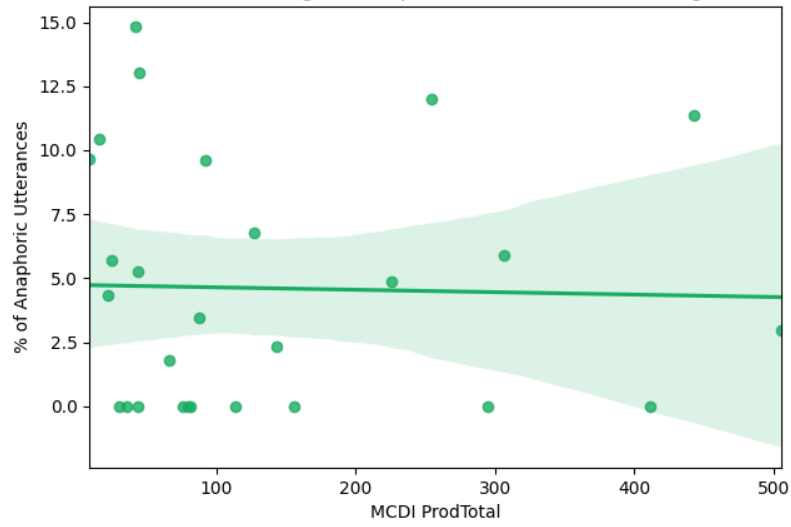


Figure 10: Compared to MCDI ProdTotal scores.  $r(25) = -.028$ ,  $p = .891$ .

MCDI ProdCountNoun vs. Percentage of Anaphoric Utterances Containing One Anaphora

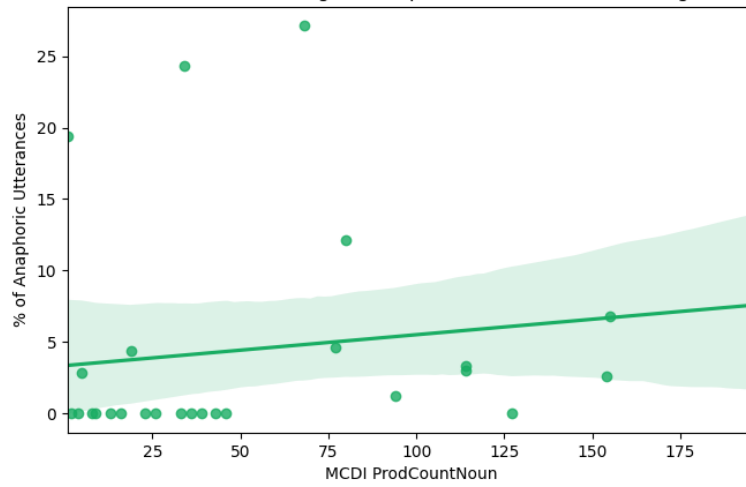


Figure 11: Compared to MCDI ProdCountNoun scores.  $r(25) = -.053$ ,  $p = .790$ .

2.6 Percentage of anaphoric utterances containing anaphora paired with exclusively verbal cues

mean	57.86
stdev	18.19
median	60.00
max	93.55
min	22.22

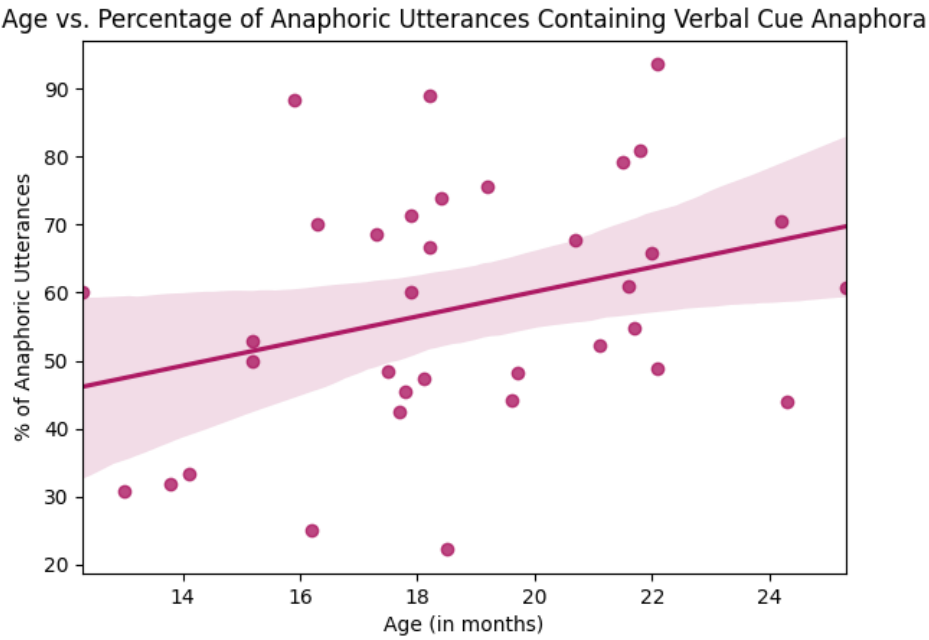


Figure 12: Compared to child age.  $r(33) = .322, p = .0589$ .

MCDI ProdTotal vs. Percentage of Anaphoric Utterances Containing Verbal Cue Anaphora

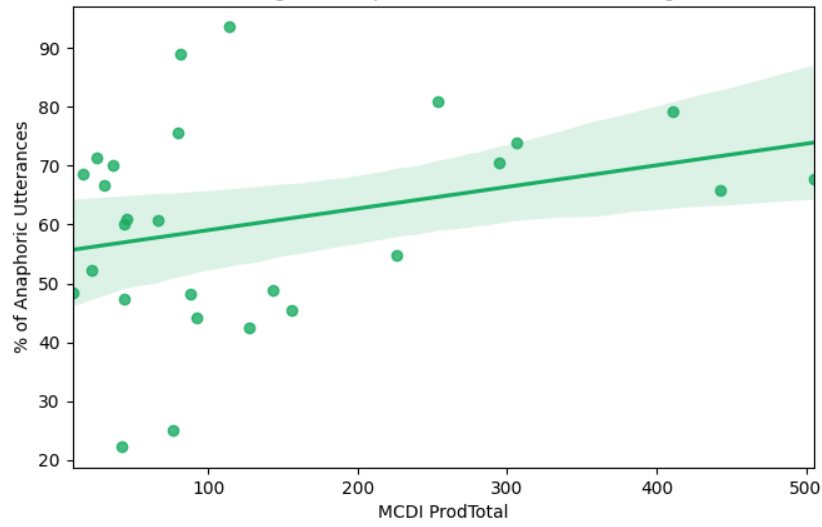


Figure 13: Compared to MCDI ProdTotal scores.  $r(25) = .297$ ,  $p = .132$ .

MCDI ProdCountNoun vs. Percentage of Anaphoric Utterances Containing Verbal Cue Anaphora

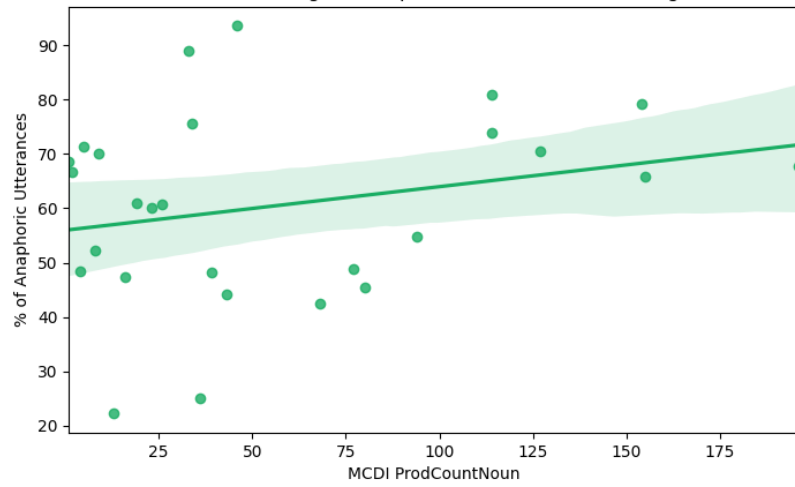


Figure 14: Compared to MCDI ProdCountNoun scores.  $r(25) = .253$ ,  $p = .203$ .

2.7 Percentage of anaphoric utterances containing anaphora paired with a required visual cue

mean	43.02
stdev	18.03
median	40.00
max	77.78
min	6.45

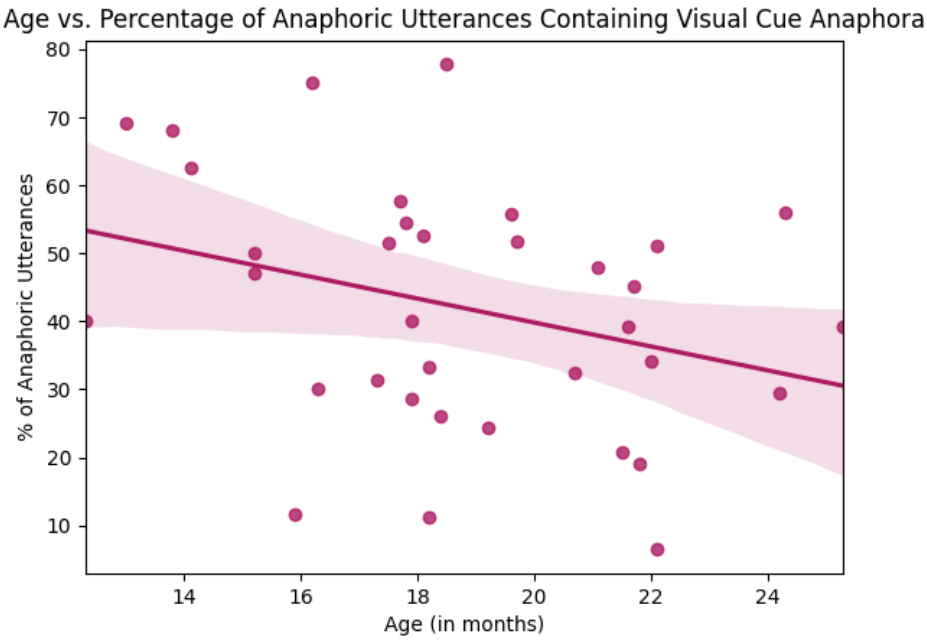


Figure 15: Compared to child age.  $r(33) = -.315$ ,  $p = .065$ .

MCDI ProdTotal vs. Percentage of Anaphoric Utterances Containing Visual Cue Anaphora

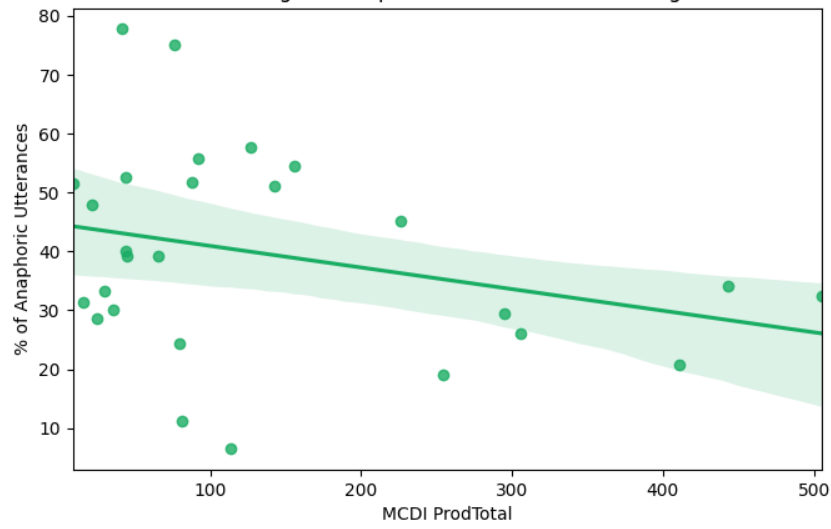


Figure 16: Compared to MCDI ProdTotal scores.  $r(25) = -.297$ ,  $p = .132$ .

MCDI ProdCountNoun vs. Percentage of Anaphoric Utterances Containing Visual Cue Anaphora

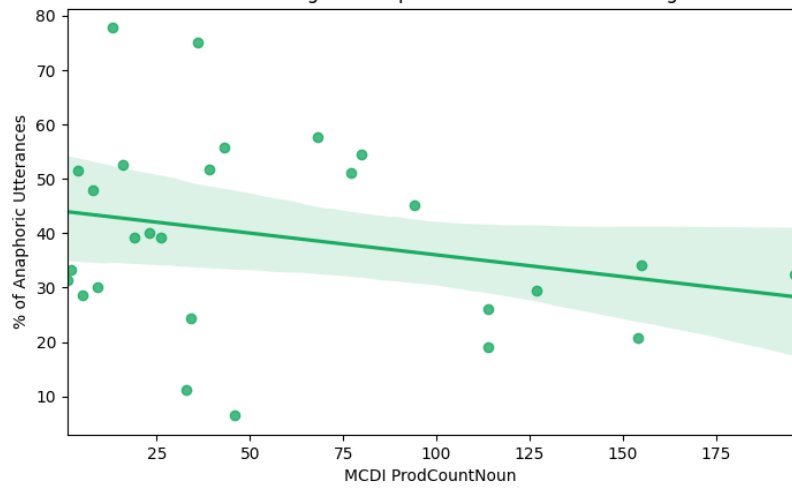


Figure 17: Compared to MCDI ProdCountNoun scores.  $r(25) = -.253$ ,  $p = .203$ .

2.8 Anaphora resolution accuracy score

mean	0.618
stdev	0.144
median	0.630
max	0.903
min	0.250

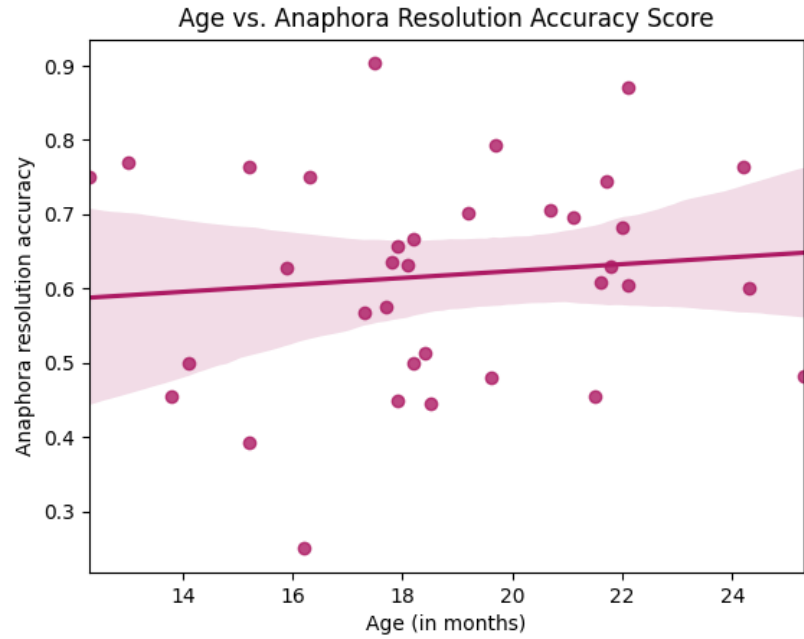


Figure 18: Compared to child age.  $r(33) = -.105, p = .549$ .

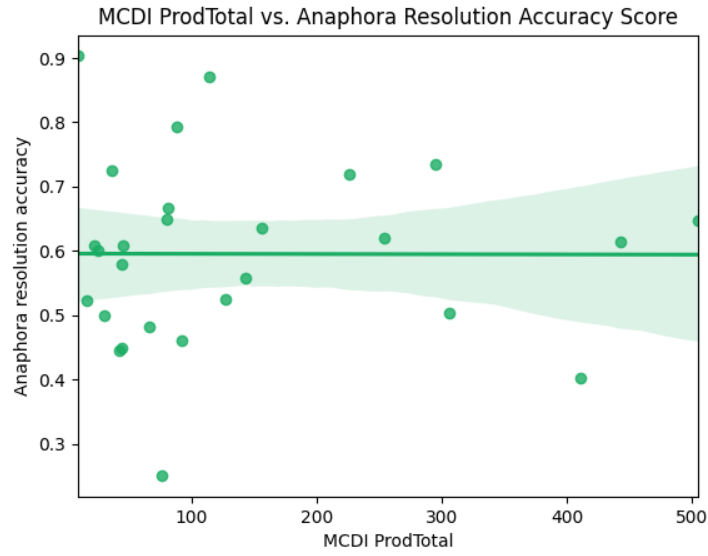


Figure 19: Compared to MCDI ProdTotal scores.  $r(25) = -.003$ ,  $p = .988$ .

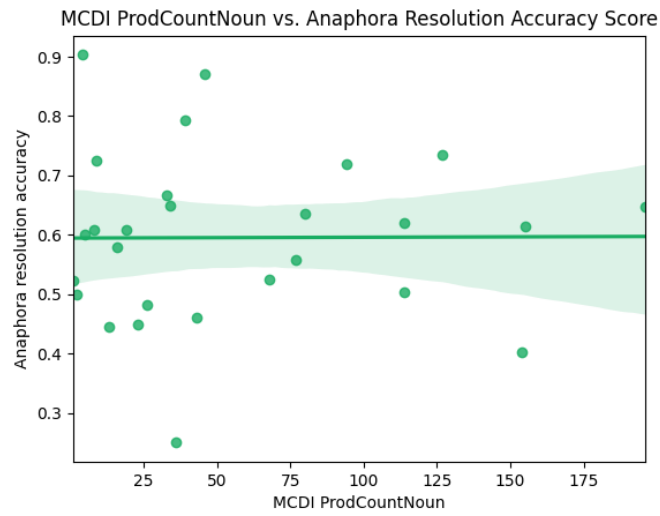


Figure 20: Compared to MCDI ProdCountNoun scores.  $r(25) = .005$ ,  $p = .980$ .

2.9 Pronominal anaphora resolution accuracy score

mean	0.623
stdev	0.156
median	0.611
max	0.893
min	0.250

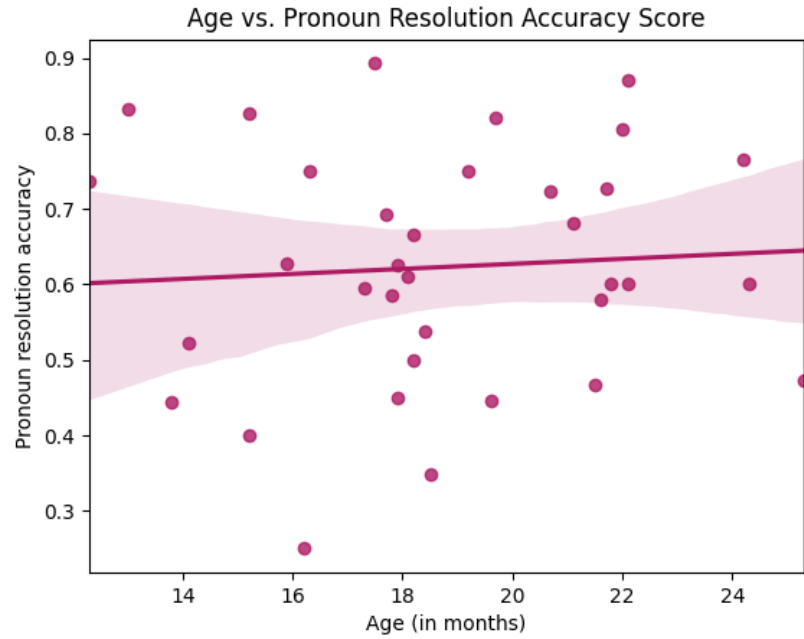


Figure 21: Compared to child age.  $r(33) = 0.069$ ,  $p = .695$ .



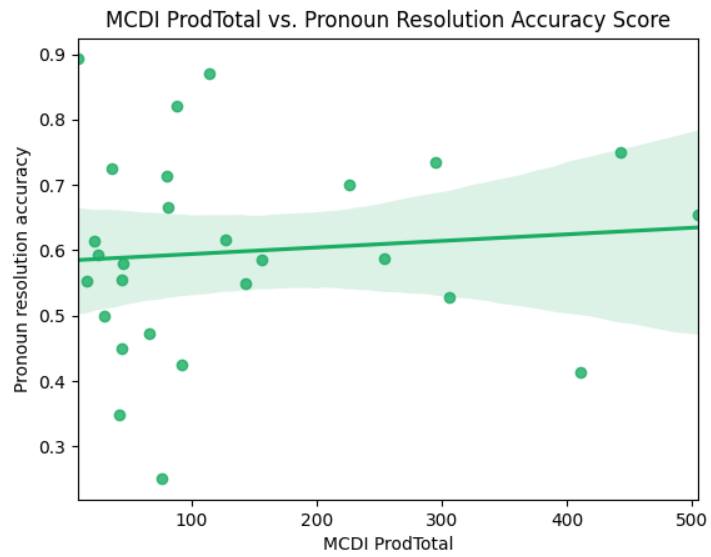


Figure 22: Compared to MCDI ProdTotal scores.  $r(25) = .092$ ,  $p = .647$ .

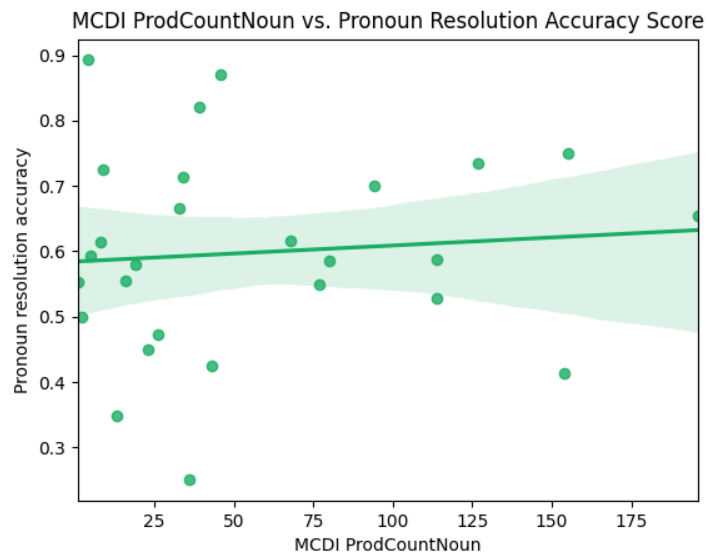


Figure 23: Compared to MCDI ProdCountNoun scores.  $r(25) = .088$ ,  $p = .663$ .

2.10 One anaphora resolution accuracy score

Note that not all subjects used one anaphora in their speech, so this data is much more limited.

mean	0.707
stdev	0.356
median	0.845
max	1.000
min	0.000

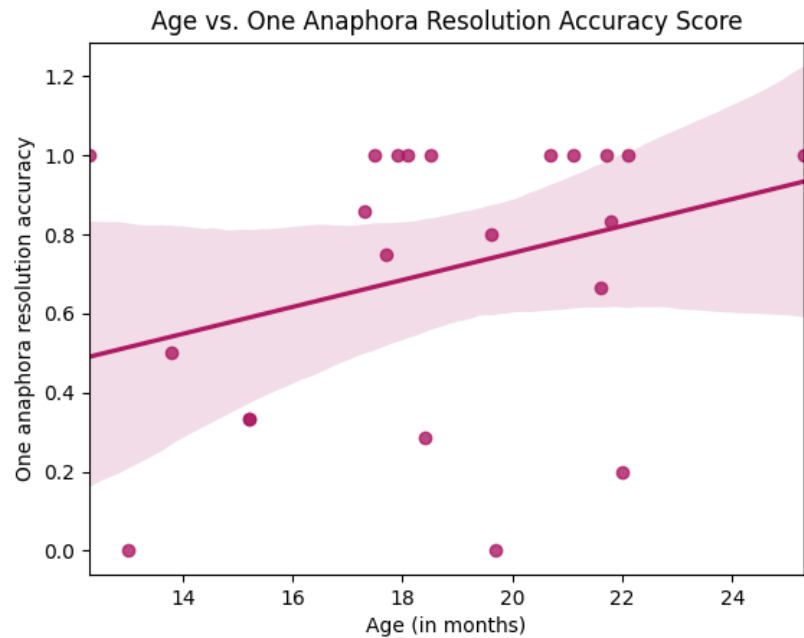


Figure 24: Compared to child age.  $r(20) = .318, p = .149$ .

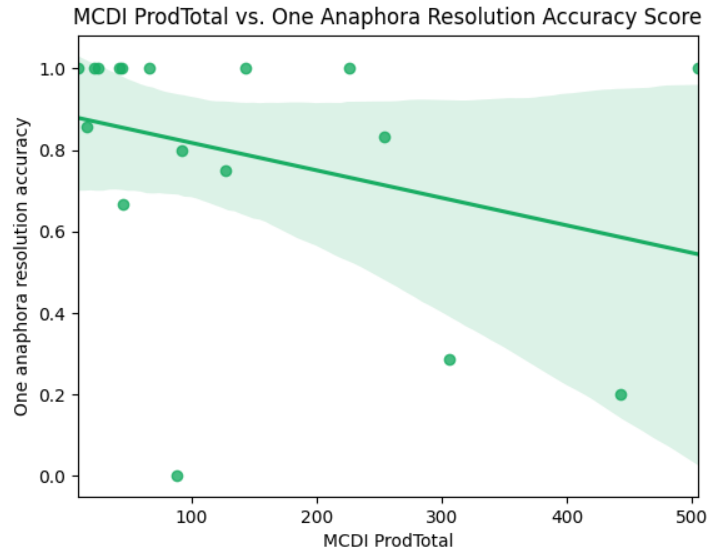


Figure 25: Compared to MCDI ProdTotal scores.  $r(15) = -.320$ ,  $p = .210$ .

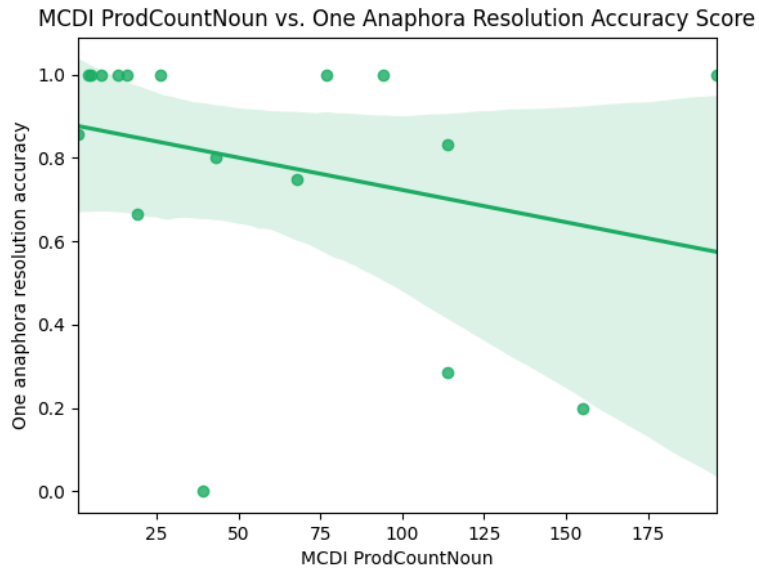


Figure 26: Compared to MCDI ProdCountNoun scores.  $r(15) = -.282$ ,  $p = .271$ .

2.11 Split anaphora resolution accuracy score

Note that not all subjects used split anaphora in their speech, so this data is much more limited.

mean	0.502
stdev	0.373
median	0.500
max	1.000
min	0.000

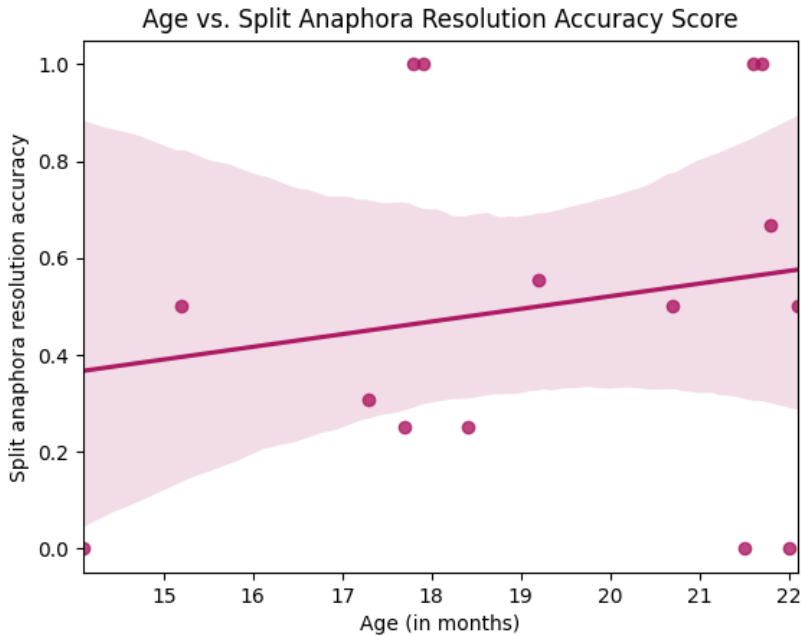


Figure 27: Compared to child age.  $r(13) = .182, p = .516$ .

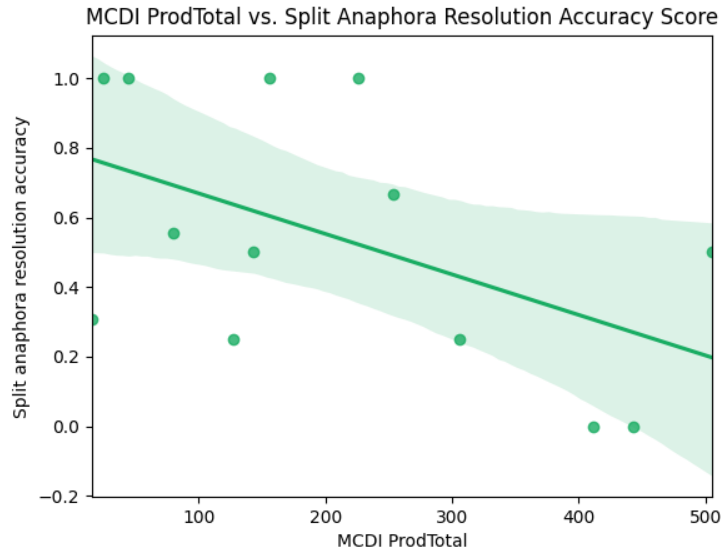


Figure 28: Compared to MCDI ProdTotal scores.  $r(11) = -.512$ ,  $p = .073$ .

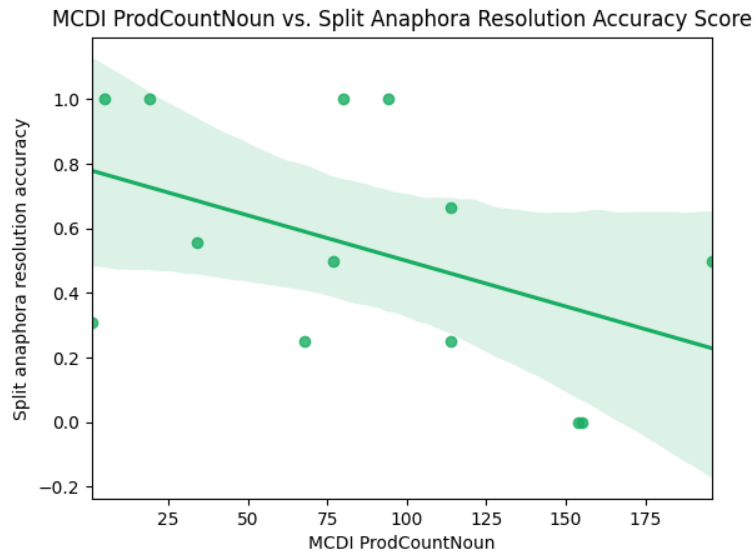


Figure 29: Compared to MCDI ProdCountNoun scores.  $r(11) = -.460$ ,  $p = .113$ .

2.12 Resolution accuracy score of anaphora paired with exclusively verbal cues

mean	0.620
stdev	0.175
median	0.625
max	1.000
min	0.261

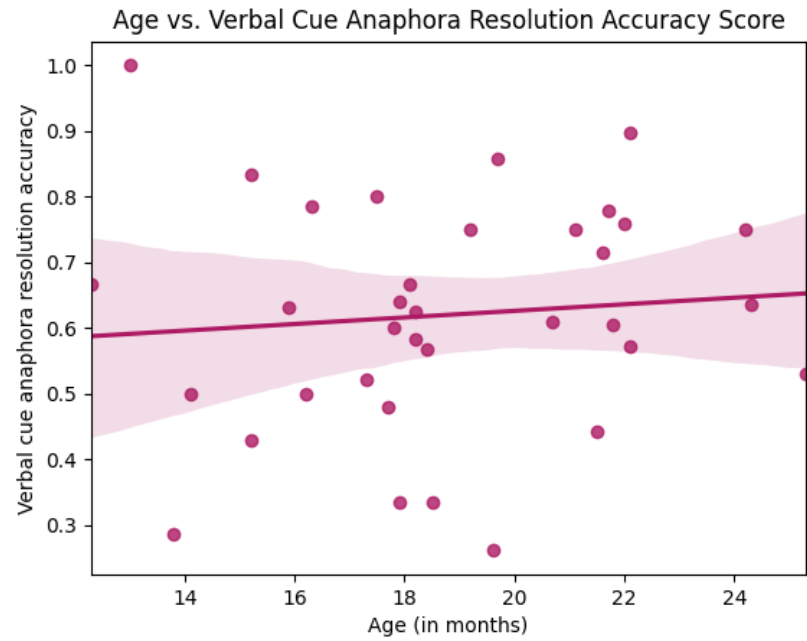


Figure 30: Compared to child age.  $r(33) = .093, p = .596$ .

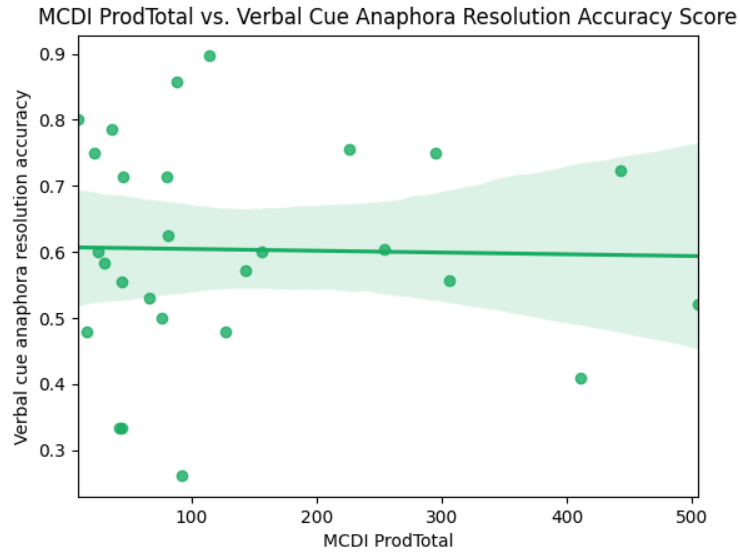


Figure 31: Compared to MCDI ProdTotal scores.  $r(25) = -.023$ ,  $p = .910$ .

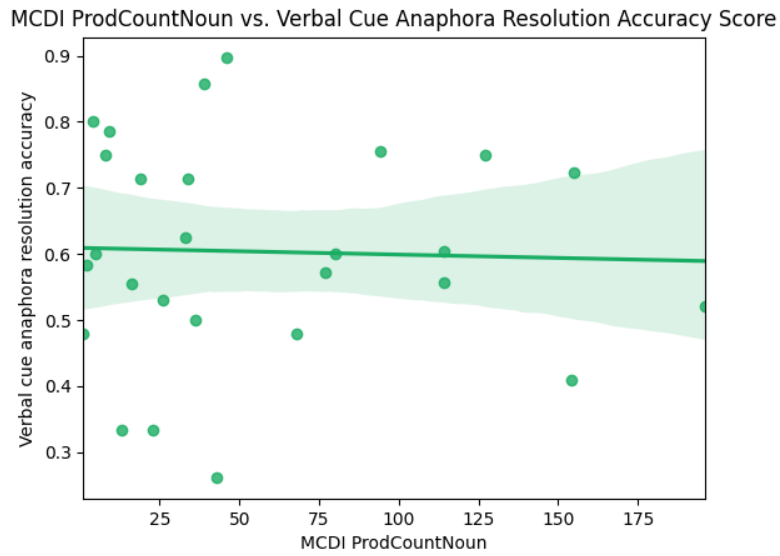


Figure 32: Compared to MCDI ProdCountNoun scores.  $r(25) = -.034$ ,  $p = .866$ .

2.13 Resolution accuracy score of anaphora paired with a required visual cue

mean	0.614
stdev	0.180
median	0.636
max	1.000
min	0.167

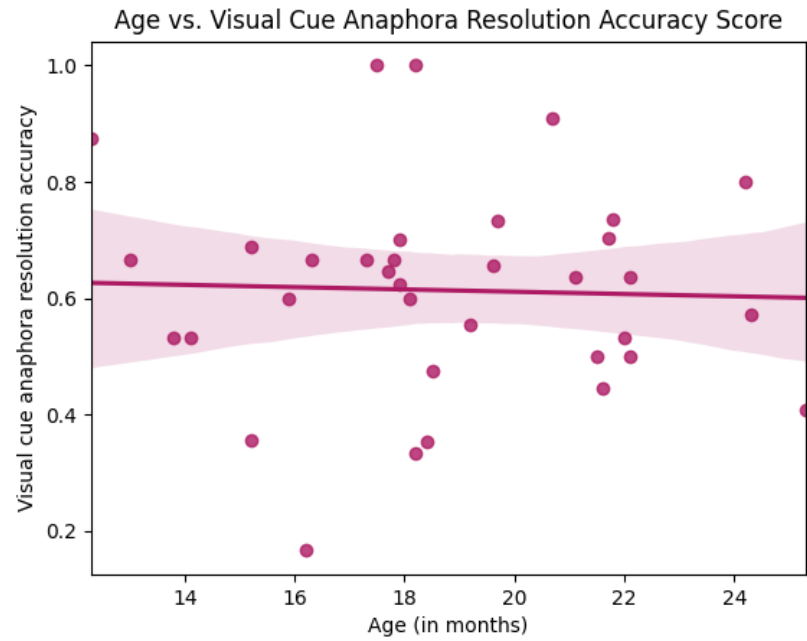


Figure 33: Compared to child age.  $r(33) = -.035, p = .840$ .



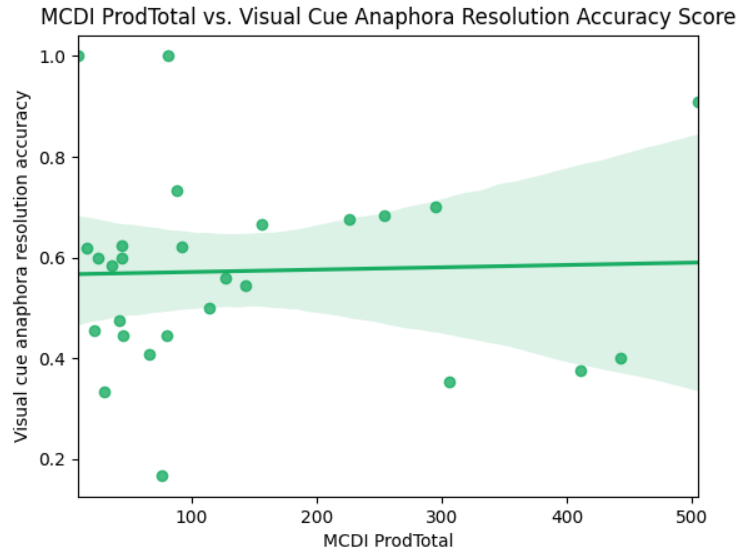


Figure 34: Compared to MCDI ProdTotal scores.  $r(25) = .033$ ,  $p = .869$ .

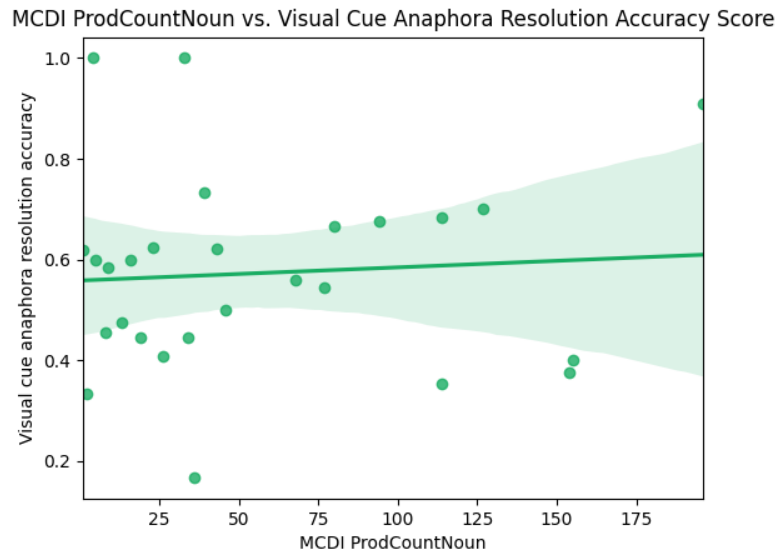


Figure 35: Compared to MCDI ProdCountNoun scores.  $r(25) = .073$ ,  $p = .717$ .

### 3 Additional Results

#### 3.1 Percentage of Anaphora Utterances Instigated by Parent

mean	52.57
stdev	23.53
median	52.24
max	91.89
min	6.06

Age vs. Percentage of Anaphoric Utterances Instigated by Parent Subject

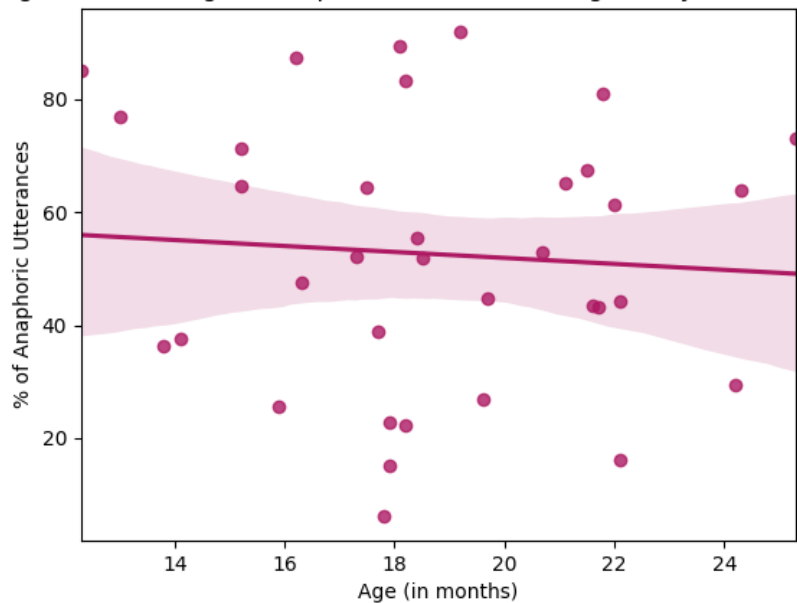


Figure 36: Compared to child age.  $r(33) = -.073$ ,  $p = .678$ .

MCDI ProdTotal vs. Percentage of Anaphoric Utterances Instigated by Parent Subject

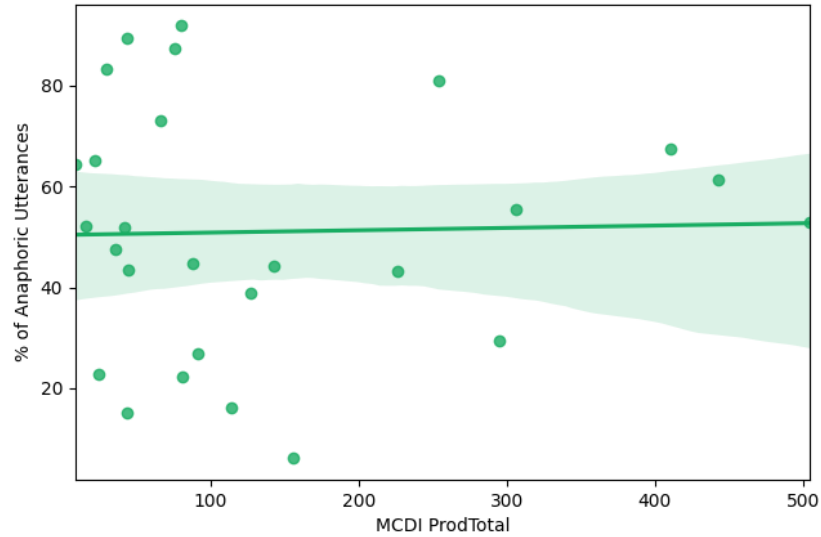


Figure 37: Compared to MCDI ProdTotal scores.  $r(25) = .027$ ,  $p = .894$ .

MCDI ProdCountNoun vs. Percentage of Anaphoric Utterances Instigated by Parent Subject

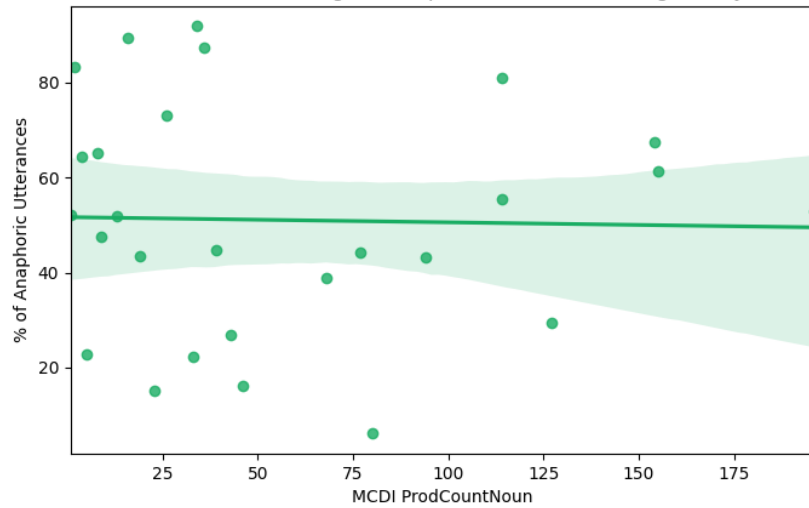


Figure 38: Compared to MCDI ProdCountNoun scores.  $r(25) = -.025$ ,  $p = .902$ .

### 3.2 Percentage of Anaphora Utterances Paired with Visual Cues

Note the difference between the covariate here and in Section 2.7. Here, we use the number of anaphoric utterances that are paired with a visual cue, *whether or not that visual cue is required to resolve the anaphor in question*. In Section 2.7, we use the number of anaphoric instances that are paired with a visual cue that is required to resolve the anaphor in question.

mean	55.48
stdev	20.15
median	52.24
max	100.0
min	6.06

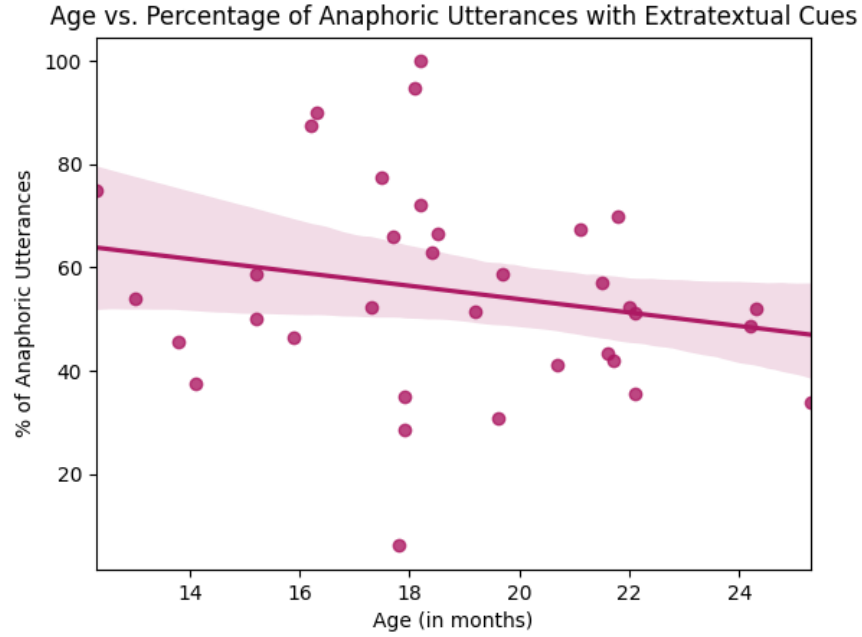


Figure 39: Compared to child age.  $r(33) = -.209$ ,  $p = .229$ .

MCDI ProdTotal vs. Percentage of Anaphoric Utterances with Extratextual Cues

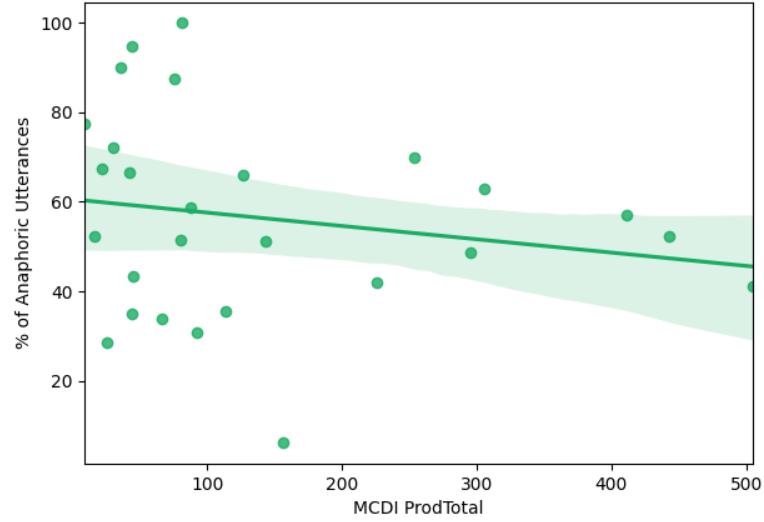


Figure 40: Compared to MCDI ProdTotal scores.  $r(25) = -.188$ ,  $p = .347$ .

MCDI ProdCountNoun vs. Percentage of Anaphoric Utterances with Extratextual Cues

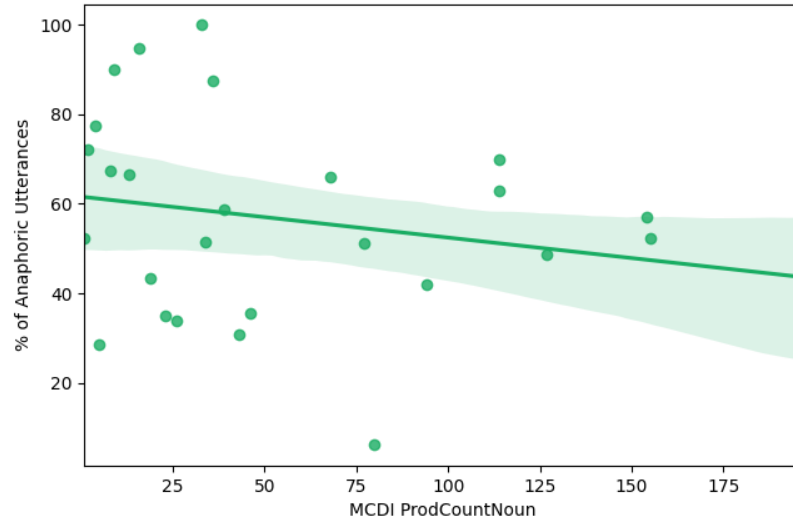
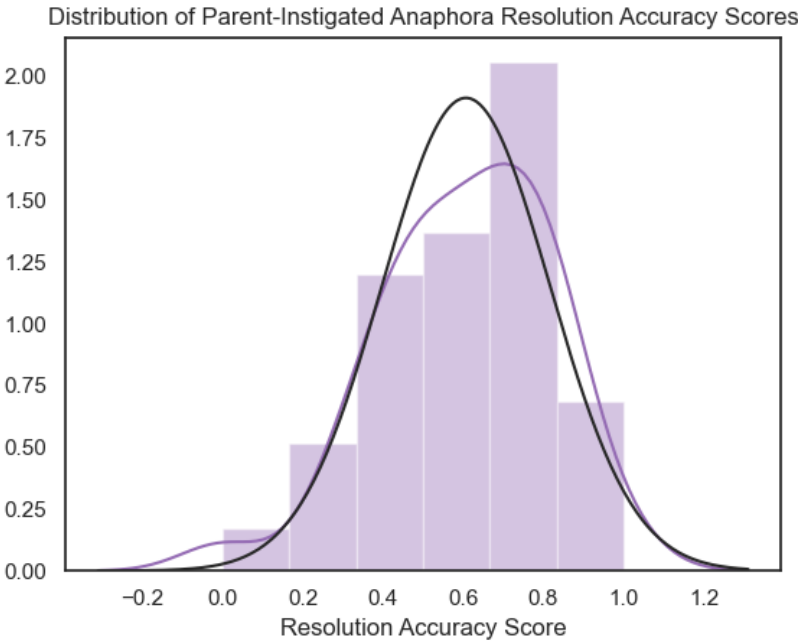


Figure 41: Compared to MCDI ProdCountNoun scores.  $r(25) = -.226$ ,  $p = .258$ .

3.3 Parent-instigated anaphora resolution accuracy score

mean	0.607
stdev	0.211
median	0.647
max	1.00
min	0.00



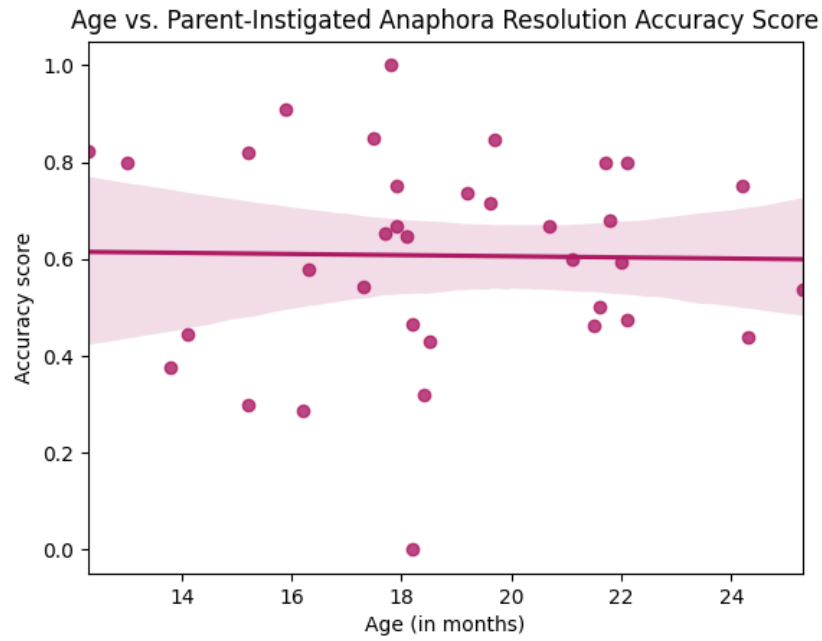


Figure 42: Compared to child age.  $r(33) = -.018$ ,  $p = .919$ .

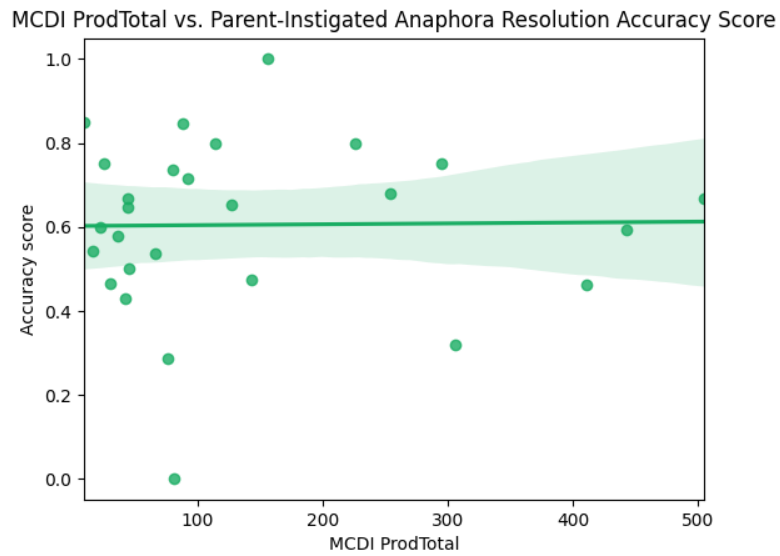


Figure 43: Compared to MCDI ProdTotal scores.  $r(25) = .014$ ,  $p = .945$ .

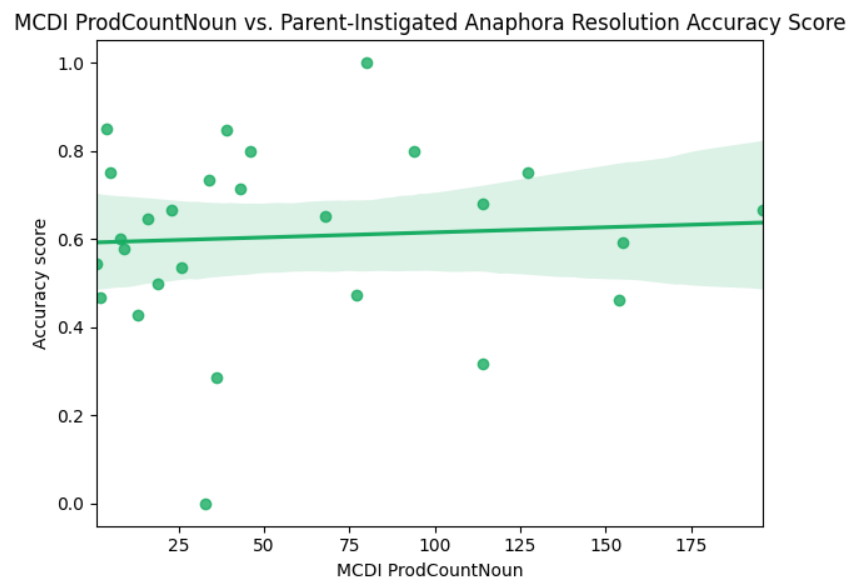
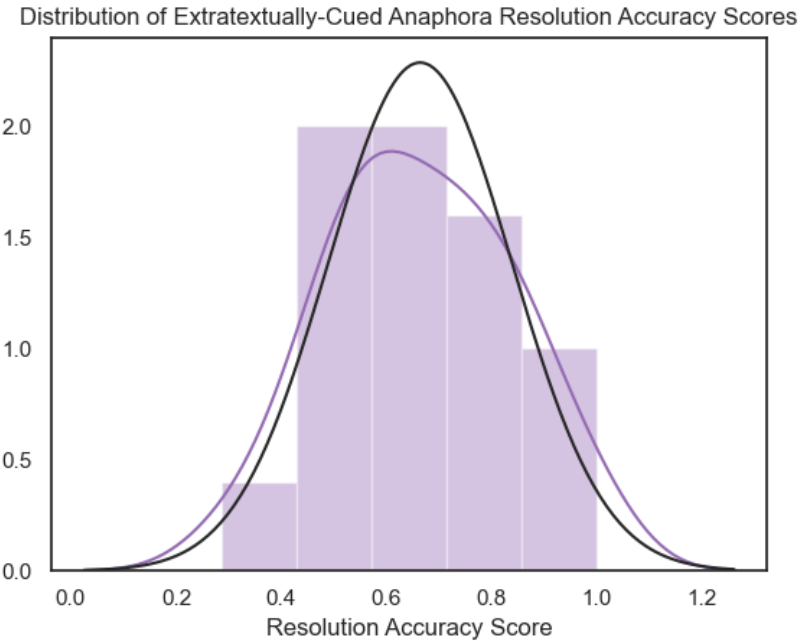


Figure 44: Compared to MCDI ProdCountNoun scores.  $r(25) = .061$ ,  $p = .761$ .



3.4 Visually-cued anaphora resolution accuracy score

mean	0.664
stdev	0.177
median	0.667
max	1.000
min	0.286



Age vs. Resolution Accuracy Score of Anaphora Paired with Extratextual Cues

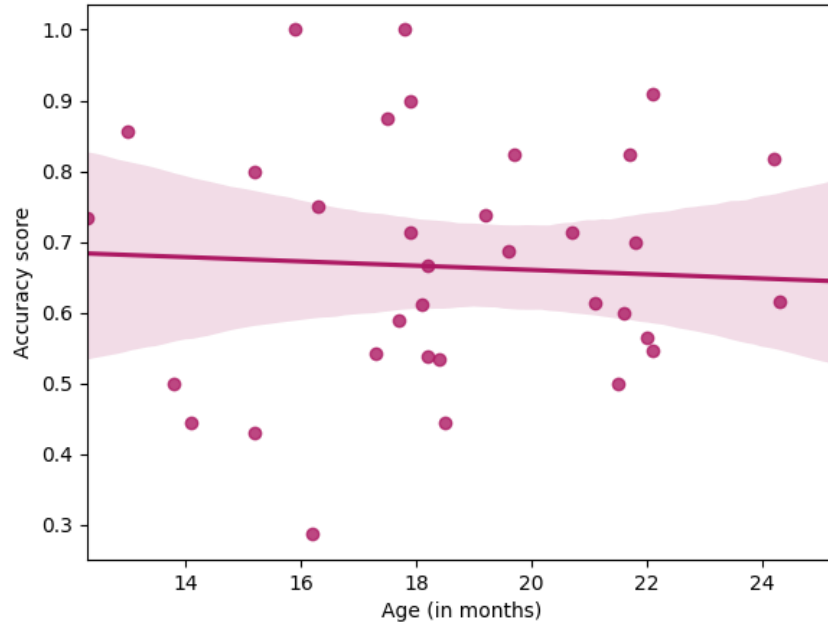


Figure 45: Compared to child age.  $r(33) = .055$ ,  $p = .752$ .

MCDI ProdTotal vs. Resolution Accuracy Score of Anaphora Paired with Extratextual Cues

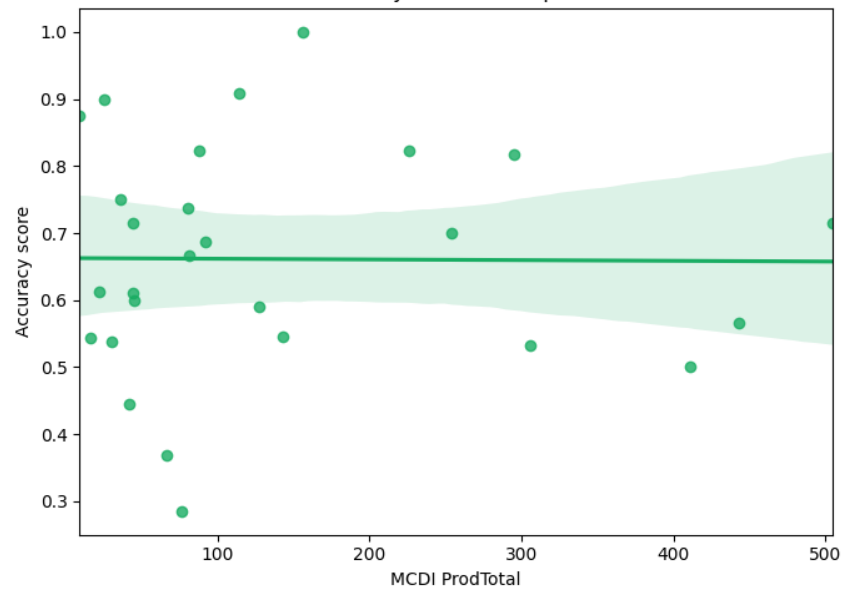


Figure 46: Compared to MCDI ProdTotal scores.  $r(25) = -.008$ ,  $p = .919$ .

MCDI ProdCountNoun vs. Resolution Accuracy Score of Anaphora Paired with Extratextual Cues

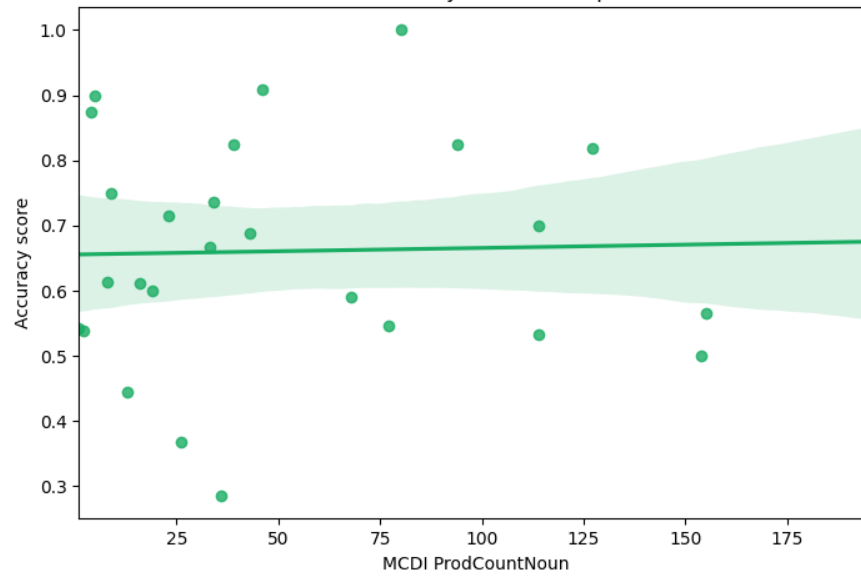


Figure 47: Compared to MCDI ProdCountNoun scores.  $r(25) = .032$ ,  $p = .873$ .

## 4 Summary

### 4.1 New coding protocol

Speech transcripts were coded again to annotate instances of both child-driven anaphora and anaphora that were paired with a visual cue.

### 4.2 Past Results & Analysis

**Figures 1, 2, 3, 4, 5, 12, 15** report statistically significant findings (or almost significant in the case of Figures 12 and 15) using a significance level of .05. However, few strong correlations beyond those were found based on the reported Pearson  $r$  values.

These findings primarily support **Hypothesis 3**, that parent speech reflects the parent’s perception of/assumptions about their child’s language comprehension abilities. Figures 3, 4, 5 most clearly indicate this, showing that parents use more anaphora in dialogue with children who are older, as well as children who are more linguistically competent based on their MCDI scores.

Additionally, Figures 12 and 15, although not quite statistically significant, show much stronger correlations than most of the other correlations measured. This finding would support Hypothesis 3 as well. Figure 15 shows that parents used more anaphora that required a visual cue to be resolved for children who were younger, and fewer anaphora requiring visual cues for older children. In other words, parents expected younger children to require greater assistance in resolving anaphora through the use of visual cues, and older children to require less. Figure 12 shows the inverse of the results in Figure 15, since verbal and visual cues were coded with binary values.

Lastly, Figures 1 and 2 indicate that there is a statistically significant positive correlation between child age and MCDI scores.

### 4.3 New Results & Analysis

None of the new results from the additional covariates coded for this time were statistically significant. The strongest correlation was between age and the percentage of anaphora that were paired with a visual cue, as shown in **Figure 39**. This supports the previous findings that parents are more likely to use visual cues to accompany anaphora when children are younger, possibly indicating that parents adjust their language and behavior to accommodate for perceived improvements in their child’s language comprehension ability.

The finding that there was a stronger correlation between age and anaphora paired with *required* visual cues (Figure 15) than between age and anaphora paired with both required and non-required visual cues (Figure 39) makes sense as parents may be more intentional about using anaphora that don’t require strong language capabilities to resolve (e.g. pointing at an object and saying “What is that?”) for younger children.

Note also that for anaphora resolution accuracy scores with child-driven instances removed (see **Section 3.3**), the mean decreased compared to accuracy scores with no cases removed (original mean=0.618, parent-instigated mean=.607), which was expected. The mean accuracy score for visually-cued anaphora (see **Section 3.4**) increased in comparison to the overall accuracy scores (original mean=0.618, visually-cued mean=0.664), which was also expected.