eye-tracking data analysis procedure

Xinxin Yan

2025-03-27

Overview

Below is a pilot analysis carried out on the collected eye-movement data (n=43) for the first two trials of the stimuli (e.g., the wilting violet; the creampuff). The goal is to evaluate a set of reading-time measures that capture the cognitive effort associated with processing the critical region-—that is, either the predicate in predicative expressions or the subject/object in referential expressions—-as well as more global discourse integration processes.

Reading Measures

First-Run Dwell Time (`IA_FIRST_RUN_DWELL_TIME`)

In particular, the early/initial processing of the critical region can be investigated by the First-run dwell time which measures how long the reader stays in the region before leaving it the first time. It is calculated as the sum of all fixation durations made during the reader's first pass through the interest area. This would include the initial fixation on the first word, any refixations on that word and any fixations on other words within the same region, as long as the gaze hasn't exited the region yet. It captures all early fixations within the region, but not include any fixations made on the region after the eyes have moved out of it. For example, a reader might fixate on the phrase "a/the wilting violet" with an initial fixation on "wilting" (186 ms), followed by a fixation on "violet" (253 ms). The first-run dwell time would then be 186 + 253 = 439 ms.

As such, first-run dwell time reflects the ease or difficulty of early lexical access, syntactic parsing, and initial semantic integration. Longer durations suggest greater early processing effort.

Total Dwell Time (`IA_DWELL_TIME`)

In contrast, the total dwell time reflects the overall processing effort allocated to region. It is defined as the sum of all fixation durations within a given Interest Area (IA) during a trial, including both first-pass and any subsequent revisits (i.e., regressions). So if a reader first reads a region and later returns to re-read part or all of it, those later fixations are included. Compared with the first pass reading time and first fixation duration, it is more of a late-stage or holistic measure, indexing rereading behavior, reanalysis, or delayed integration. It is especially useful for identifying delayed comprehension difficulties or post-lexical discourse-level processing.

Number of distinct runs of fixations `IA_RUN_COUNT`

This variable captures the number of runs made to the critical region. A "run" is defined as a series of consecutive fixations within the same interest area, uninterrupted by fixations outside of it. Each time the eyes re-enter the interest area and fixate again, a new run is counted. Thus, it is a valuable indicator of re-reading behavior which may signal post-lexical processing load (pragmatic integration), discourse intergration difficulties. —

Predictions and Processing Hypotheses

Based on theoretical expectations, several exploratory predictions are proposed:

1. if the referential expression occurs at the subject position, people may have a shorter first run dwell time since they may want to continue reading the rest of the information and may tend to revisit the subject again.
2. Referential expressions in subject position may elicit shorter first-run dwell times, as readers may prioritize progressing through the sentence and postpone full interpretation. The difference in first-run dwell time between literal (L_ref_s) and metaphoric (M_ref_s) expressions may not be obvious, as both encourage forward reading. However, M_ref_s is predicted to yield longer total dwell times, since readers may feel inclined to reread the metaphoric reference when finish reading the entire sentence in the first run to engage in discourse-level integration.
3. The harder processing of M_ref_s may also disrupt processing of the subsequent sentence material, if readers try to resolve who the metaphorical phrase refers to while processing and integrating the meaning encoded in the rest of the sentence. In contrast to the literal condition, where the subject is semantically consistent with the rest of the sentence (e.g., The weak fighter is bound to lose again), the metaphoric subject creates an initial semantic mismatch (e.g., The creampuff is bound to lose again). To resolve this literal inconsistency, readers need to infer the intended referent of the creampuff in this context. This interpretive process likely unfolds during the reading of the subsequent sentence material, potentially resulting in prolonged first-run or total dwell times in that region. To investigate the potential spillover effects on later sentence regions, we need to compare the first-run and total dwell times in the later regions (e.g., is bound to lose again). This component of the analysis is not yet included in the R script below.
4. Based on what I observed from monitoring participants' gaze when reading the stimuli, it seems that metaphorical referential expressions (particularly when presented in subject position) appear to elicit multiple reading passes.

other things to be noted:

Importantly, the critical regions across conditions are presented in identical discourse, syntactic, and semantic environments. This careful control ensures that observed differences are not confounded by positional effects, syntactic framing, or semantic context. Moreover, parafoveal preview effects—-a known influence on fixation durations—are held constant across comparisons by design.

```
# Load necessary libraries
library(ggplot2)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:dplyr':
##
```

```
##     combine
```

```
library(lme4)
```

```
## Loading required package: Matrix
```

```
library(lmerTest)
```

```
##
## Attaching package: 'lmerTest'
```

```
## The following object is masked from 'package:lme4':
##
##     lmer
```

```
## The following object is masked from 'package:stats':
##
##     step
```

```
# Set a theme for plots
theme_set(theme_minimal())

# Load the pilot data CSV
pilot_data <- read.csv("critical region_pilot.csv")

# Convert relevant columns to numeric if needed
pilot_data$IA_FIRST_RUN_DWELL_TIME <- as.numeric(pilot_data$IA_FIRST_RUN_DWELL_TIME)
pilot_data$IA_DWELL_TIME <- as.numeric(pilot_data$IA_DWELL_TIME)
pilot_data$condition <- as.factor(pilot_data$condition)

# Reorder condition levels for grouped display
pilot_data$condition <- factor(pilot_data$condition, levels = c("L_pred", "M_pred", "L_ref_s", "M_ref_s

# Inspect summary of the data
summary(pilot_data)
```

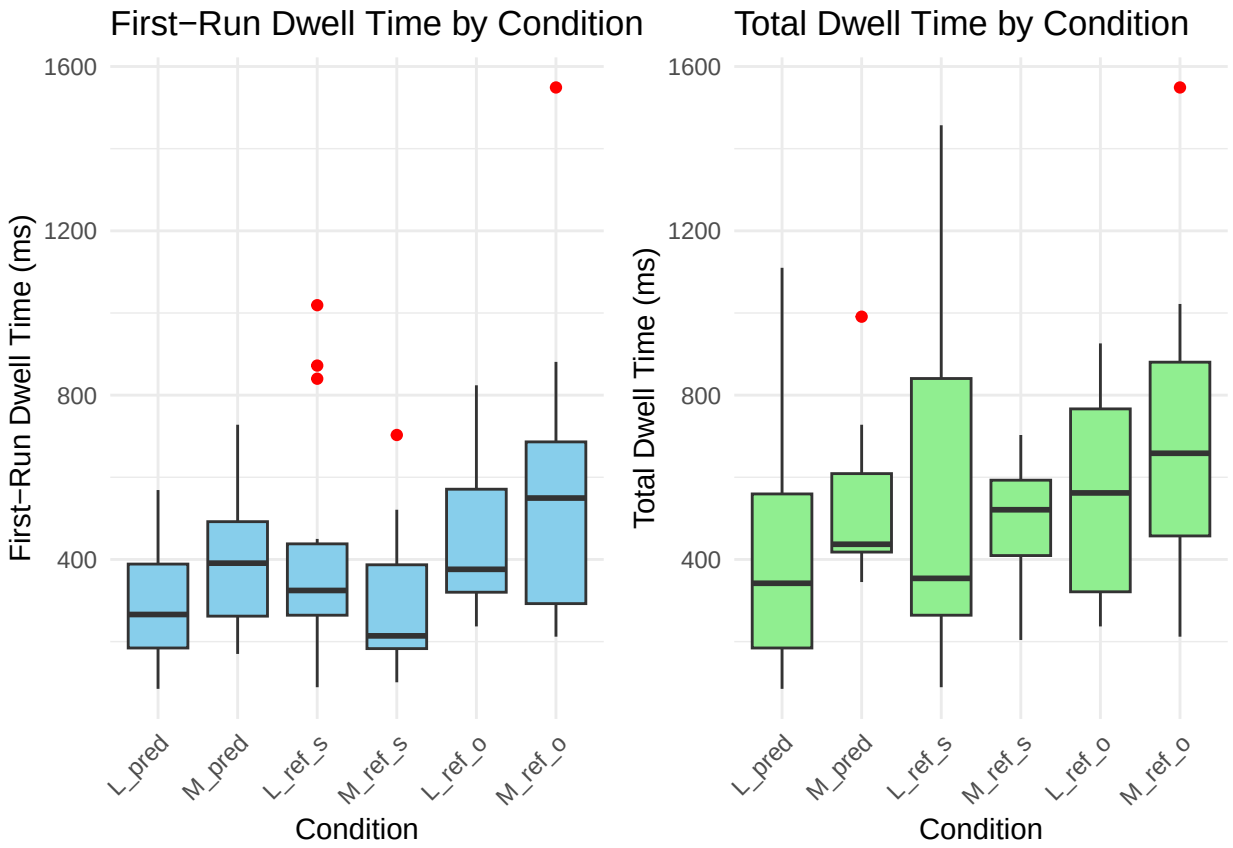
```
##   Participant        trial            list         condition
##  Min.   : 1.00   Min.   :1.000   Min.   :1.000   L_pred :14
##  1st Qu.:12.00   1st Qu.:1.000   1st Qu.:2.000   M_pred :13
##  Median :23.00   Median :2.000   Median :3.000   L_ref_s:16
##  Mean   :22.81   Mean   :1.512   Mean   :3.419   M_ref_s:15
##  3rd Qu.:33.75   3rd Qu.:2.000   3rd Qu.:5.000   L_ref_o:14
##  Max.   :44.00   Max.   :2.000   Max.   :6.000   M_ref_o:14
##    question            Accuracy         TRIAL_RT         IA_ID
##  Length:86          Min.   :-1.000   Min.   : 5533   Min.   :31.00
##  Class :character   1st Qu.:-1.000   1st Qu.: 8416   1st Qu.:35.00
##  Mode  :character   Median : 1.000   Median :11394   Median :46.00
##                     Mean   : 0.407   Mean   :12103   Mean   :42.05
##                     3rd Qu.: 1.000   3rd Qu.:14043   3rd Qu.:51.00
##                     Max.   : 1.000   Max.   :32998   Max.   :54.00
##    IA_LABEL         IA_FIRST_FIXATION_DURATION IA_FIRST_FIXATION_TIME
##  Length:86          Min.   : 85.0              Min.   : 4331
##  Class :character   1st Qu.:181.2              1st Qu.: 6482
##  Mode  :character   Median :224.5              Median : 8455
##                     Mean   :237.1              Mean   : 9189
##                     3rd Qu.:276.0              3rd Qu.:10932
##                     Max.   :518.0              Max.   :28062
```

```
##  IA_FIRST_RUN_DWELL_TIME IA_FIRST_RUN_FIXATION_COUNT IA_FIRST_FIX_PROGRESSIVE
##  Min.   :  85.0          Min.   :1.000               Min.   :0.0000
##  1st Qu.: 234.8          1st Qu.:1.000               1st Qu.:1.0000
##  Median : 342.0          Median :2.000               Median :1.0000
##  Mean   : 403.0          Mean   :1.733               Mean   :0.9651
##  3rd Qu.: 509.2          3rd Qu.:2.000               3rd Qu.:1.0000
##  Max.   :1549.0          Max.   :5.000               Max.   :1.0000
##  IA_DWELL_TIME     IA_RUN_COUNT   IA_FIXATION_COUNT IA_REGRESSION_OUT
##  Min.   :  85.0   Min.   :1.000   Min.   :1.000     Min.   :0.0000
##  1st Qu.: 312.5   1st Qu.:1.000   1st Qu.:2.000     1st Qu.:0.0000
##  Median : 486.5   Median :1.000   Median :2.000     Median :0.0000
##  Mean   : 529.3   Mean   :1.453   Mean   :2.314     Mean   :0.2442
##  3rd Qu.: 685.8   3rd Qu.:2.000   3rd Qu.:3.000     3rd Qu.:0.0000
##  Max.   :1549.0   Max.   :4.000   Max.   :7.000     Max.   :1.0000
##  IA_REGRESSION_IN IA_REGRESSION_PATH_DURATION
##  Min.   :0.0000   Min.   :  85.0
##  1st Qu.:0.0000   1st Qu.: 277.8
##  Median :0.0000   Median : 435.5
##  Mean   :0.1977   Mean   : 557.9
##  3rd Qu.:0.0000   3rd Qu.: 706.0
##  Max.   :1.0000   Max.   :2602.0
```

```r
# Box plot for First-Run Dwell Time by condition
p1 <- ggplot(pilot_data, aes(x = condition, y = IA_FIRST_RUN_DWELL_TIME)) +
  geom_boxplot(fill = "skyblue", outlier.colour = "red") +
  labs(title = "First-Run Dwell Time by Condition",
       x = "Condition",
       y = "First-Run Dwell Time (ms)") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))

# Box plot for Total Dwell Time by condition
p2 <- ggplot(pilot_data, aes(x = condition, y = IA_DWELL_TIME)) +
  geom_boxplot(fill = "lightgreen", outlier.colour = "red") +
  labs(title = "Total Dwell Time by Condition",
       x = "Condition",
       y = "Total Dwell Time (ms)") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))

# Display plots side by side
grid.arrange(p1, p2, ncol = 2)
```



In general, metaphoric conditions tend to elicit both longer first pass reading times and longer total reading times. As expected, the first dwell time of M_ref_s and L_ref_s do not differ much (if getting rid of the extremely long outliers from the L_ref_s group, their avaerage first run time should be similar). However, the total dwell time difference between the L_ref_s and M_ref_s was the most pronounced compared with the other two groups, indicating that much of the effort of processing the metaphoric reference happens at the later stage (reference resolution, figurative meaning derivation).

```
# Remove missing values in IA_RUN_COUNT (if any)
pilot_data_clean <- pilot_data %>%
  filter(!is.na(IA_RUN_COUNT))

# Summary statistics
run_count_summary <- pilot_data_clean %>%
  group_by(condition) %>%
  summarise(
    mean_run_count = mean(IA_RUN_COUNT),
    sd_run_count = sd(IA_RUN_COUNT),
    n = n()
  )

print(run_count_summary)
```

```
## # A tibble: 6 x 4
##   condition mean_run_count sd_run_count     n
##   <fct>              <dbl>        <dbl> <int>
## 1 L_pred              1.29        0.469    14
## 2 M_pred              1.46        0.519    13
## 3 L_ref_s             1.31        0.704    16
```

```
## 4 M_ref_s             1.73        0.458    15
## 5 L_ref_o             1.36        0.633    14
## 6 M_ref_o             1.57        1.02     14
```

Metaphoric conditions tend to elicit more runs than their literal counterparts:

M_pred > L_pred (1.46 vs 1.29)

M_ref_s > L_ref_s (1.73 vs 1.31) ← most pronounced difference

M_ref_o > L_ref_o (1.57 vs 1.36)

M_ref_s has the highest mean number of runs (1.73) (consistent with what I observed), suggesting that metaphorical subjects are revisited more frequently, potentially due to greater difficulty in referent resolution or discourse integration.