

# Practical exercise: Temporal ambiguity

## Instructions

You will find a dataset prepared for you if you modify the **boldfaced** portion of the following link:

[https://raw.githubusercontent.com/MetodesEmpirics/MetodesEmpirics.github.io/main/material/2024q3/session08/data/csvs/insert\\_your\\_student\\_ID.csv](https://raw.githubusercontent.com/MetodesEmpirics/MetodesEmpirics.github.io/main/material/2024q3/session08/data/csvs/insert_your_student_ID.csv)

Each student has their own dataset, identified by their student ID. For instance, if your student ID is **u103672**, then your link is:

<https://raw.githubusercontent.com/MetodesEmpirics/MetodesEmpirics.github.io/main/material/2024q3/session08/data/csvs/u103672.csv>

Note that your computer may insert a space in your URL because it spans two lines. Alternatively, if you are having trouble copy/pasting, you will find the link to your dataset here:

[https://docs.google.com/spreadsheets/d/1XMks1AmLyH\\_9BQVz4duZvDPz1q0pwwvKYJ7JTzBMqW2U/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1XMks1AmLyH_9BQVz4duZvDPz1q0pwwvKYJ7JTzBMqW2U/edit?usp=sharing)

You can then load your data as

```
df <- read.csv('replace_this_with_URL_to_your_data.csv')
```

The data is described below.

Submit your answers to the questions as a PDF file, through Aula Global. Round all answers to the second decimal after the comma.

## Data description

There are sentences that are syntactically ambiguous while you are processing them. That is, while you're hearing/reading the sentence, you can be misled to interpret it in a way that turns out to be wrong. This effect is called a *temporal ambiguity*. Temporal ambiguities can be so strong that it may take you a couple of times hearing/reading them before you understand them. For instance, compare the two sentences in (1) and (2).

1. While the narrator read the story **was dramatized** by the troop of skilled actors.
2. While the narrator read(,) the story **was dramatized** by the troop of skilled actors

When reading (1) for the first time, the sentence may strike you as odd when you reach the boldfaced portion of the text. This effect should be less pronounced for (2), where the comma helps you interpret the sentence as intended. When reading (1), instead, you may be misled into thinking that **the story** is the object of **the narrator read**; and not its complement. This kind of syntactic ambiguity is called NP/Z in the literature.

There are other types of syntactically ambiguous sentences. The sentences in (3) and (4) exemplify a so-called NP/S-type ambiguity.

3. The tourists saw the palace **was being restored** to its original condition.
4. The tourists saw (that) the palace **was being restored** to its original condition.

As (2) did for (1), the sentence in (4) shows a disambiguated version of (3). Both types of syntactic ambiguity, NP/Z and NP/S, are well studied in the literature. Comprehenders being confused by them gives us hints about how we process sentences and interpret sentences!

In experiments, the difficulty of a sentence is often measured by how long it takes people to read them (in milliseconds). Past literature suggests two things. First, it should take less time to read the unambiguous sentence (2) than the ambiguous one in (1); and the same for (4) over (3). Second, the difference between (3) and (4) should be smaller than that between (1) and (2). That is, processing NP/Z sentences should be harder than processing NP/S sentences, when compared to their unambiguous variants.

Your colleague has collected data from subjects that were asked to read sentences like (1) - (4). The experiment had 5 different pairs of NP/Z and NP/S sentences. Each sentence pair consisted of an ambiguous and an unambiguous version of the sentence. For instance, the sentences in (1) and (2) were one of the five pairs of sentences for the NP/Z case; and the sentences in (3) and (4) for the NP/S case. 12 different subjects read each pair. Your colleague measured how long it took them to read the sentences (in milliseconds) and recorded it in a CSV file.

Here's how the data may look like:

##	sentence.ID	status	reading.time	type
## 1	1	Amb	586.9277	npz
## 2	1	Unamb	420.7923	npz
## 3	2	Amb	582.9802	npz
## 4	2	Unamb	427.8409	npz
## 5	3	Amb	588.3345	npz
## 6	3	Unamb	423.8143	npz

The column `sentence.ID` identifies which of the 5 sentence pairs the reading times are for. The column `status` shows whether the reading time is for the ambiguous (**Amb**) or unambiguous sentence of a pair. The column `reading.time` shows the reading time recorded (in milliseconds). The column `type` specifies whether this is a NP/S ambiguity (**nps**) or an NP/Z ambiguity (**npz**)

## Questions

Help your colleague analyze the data from her experiments! Remember to round all answers to the second decimal after the comma.

### Descriptive statistics (0.5 points each)

1. What is the mean reading time of ambiguous NP/Z sentences?
2. What is the standard deviation of reading times of ambiguous NP/Z sentences?
3. What is the mean reading time of unambiguous NP/Z sentences?
4. What is the standard deviation of reading times of unambiguous NP/Z sentences?
5. What is the mean reading time of ambiguous NP/S sentences?
6. What is the standard deviation of reading times of ambiguous NP/S sentences?
7. What is the mean reading time of unambiguous NP/S sentences?
8. What is the standard deviation of reading times of unambiguous NP/S sentences?

### Inferential statistics (1 point each)

Using two regression models, with `status` as predictor and `reading.time` as outcome, estimate

1. The difference between ambiguous and unambiguous NP/Z sentences, and the standard error of this difference
2. The difference between ambiguous and unambiguous NP/S sentences, and the standard error of this difference
3. Based on your answer to (1): Do you think there is evidence to suggest a difference between ambiguous and unambiguous NP/Z sentences? Why (not)?
4. Based on your answer to (2): Do you think there is evidence to suggest a difference between ambiguous and unambiguous NP/S sentences? Why (not)?
5. Compare how well the models you used for (1) and (2) explain the variance (adjusted  $R^2$ ). Is the performance of the two models similar? Why do you think that this (not) the case?
6. Do your findings agree with the expectations from the literature? Explain why (not).

## References

If you are interested in the real experiment that this exercise is based on, see Grodner et al. (2003): Against Repair-Based Reanalysis in Sentence Comprehension.