

TalkBankDB Workshop

- 1) Install R.

First, get R (choose link for your OS):

<https://cran.r-project.org/>

Then install Rstudio (choose the "free" version):

<https://www.rstudio.com/products/rstudio/download>

- 2) Get R script for today's workshop:

- Go to <https://sla.talkbank.org/Rdata/>

- On "demo.R":

- **Right**-click, and select "save file".

- 3) Use TalkBankDB to get data from MacWhinney corpus.

- Go to TalkBankDB: <https://talkbank.org/DB/>

- Once on TalkBankDB page, query by "Corpus" and select the path to the MacWhinney corpus:
Childes-data -> Eng-NA -> MacWhinney.

- Click blue "Add to query→" button.

- Click green "Search TalkBankDB" button (and wait for results to return).

- Click the "Participants" tab to see a table of participant info across the MacWhinney corpus.

- Click blue "Save" button to save the participant data file (**TBDB_participants.xls**) to your machine.

- Click the "Tokens" tab to see a table of all tokens (words) spoken in the MacWhinney corpus.

- Click blue "Save" button to save the token data file (**TalkBankDB.xls**) to your machine.

- 4) Open R studio and load demo script:

- Then go to menu Code->Source file... and select the demo.R file you downloaded in step 2.

- 5) Load data:

- At the R prompt, type:

```
loadLocalFiles('path/to/TBDB_participants.xls', 'path/to/TalkBankDB.xls')
```

- (Where "path/to/" is the path to where you saved these files in step 3.)

- Alternatively, if you're having trouble loading the files locally, get them directly from the web by entering at the R prompt:

```
loadFromWeb()
```

- 6) Plot some word frequencies.

- At the R prompt, type:

```
freqByAge('targetWord')
```

- (Where `targetWord` is the word you wish to see its frequency plotted against `monthAge` within the corpus.)

Some useful R commands:

- Load script:

```
source(file = "demo.R")
```

- Get working directory:

```
getwd()
```

- Set working directory:

```
setwd('path/to/set')
```

- List files in working directory:

```
list.files()
```