



ULB

Creating your DMP as a project guide

Mathieu Bourguignon

Laboratoire de Cartographie fonctionnelle du Cerveau, UNI – ULB
Neuroscience Institute, Université libre de Bruxelles (ULB), 1070
Bruxelles, Belgique.

Overview of the project

Dyslexia: reading disorder affecting 5–12 % of the children

Life-time problem

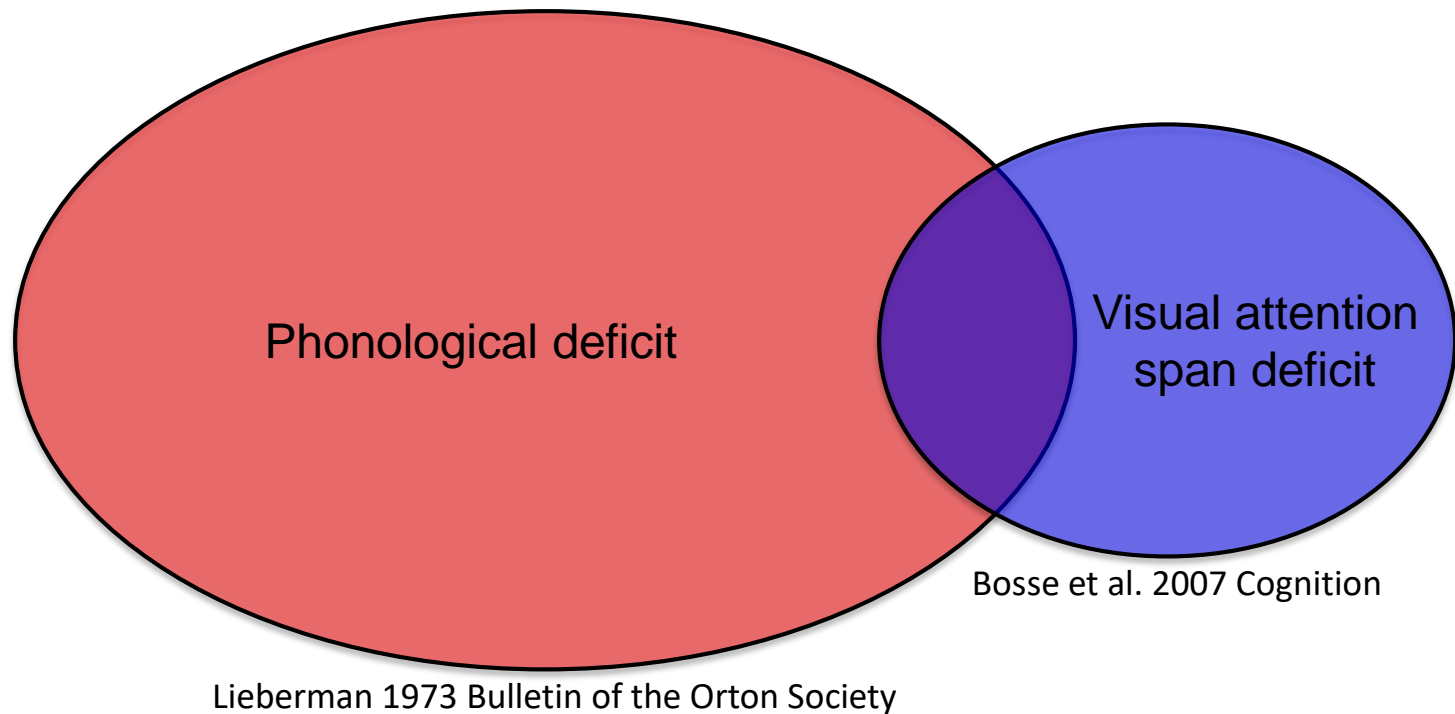


Dyslexia is recognized as a major contributing factor to school failure

Need for improved identification and remediation methods for dyslexia

Overview of the project

Cause of dyslexia?



Overview of the project

Dyslexia comes with:

→ **atypical entrainment to syllabic stress and speech rhythm** (Goswami et al. 2013 Lab Phono; Lallier et al. 2013 front human neurosci)

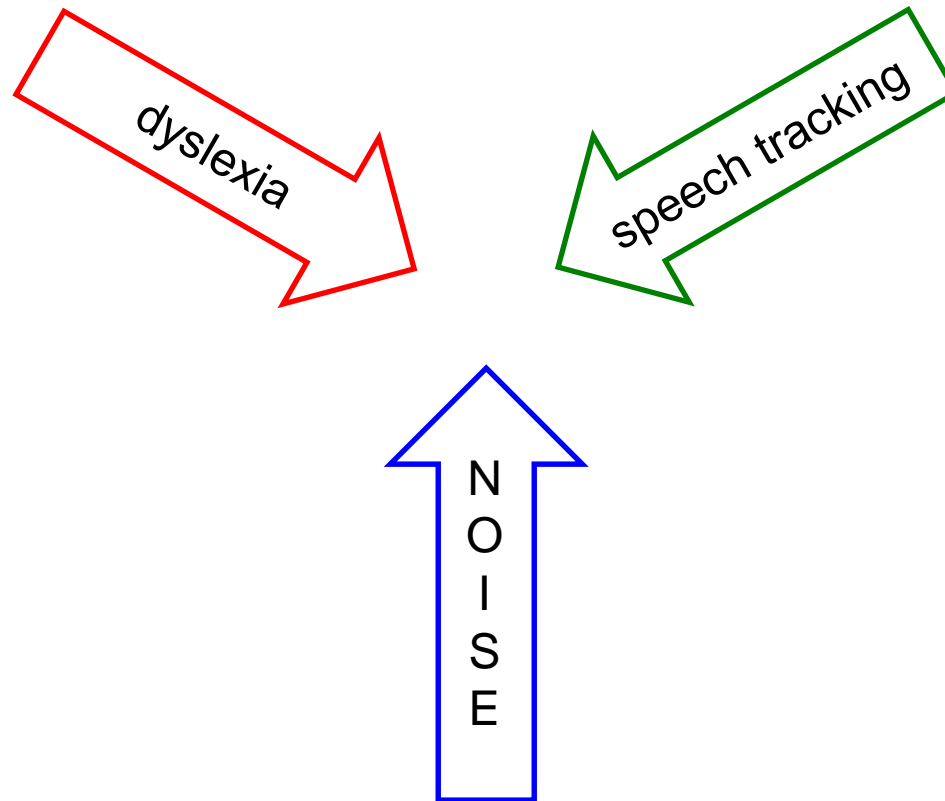
→ **exacerbated phonological deficit in the presence of noise** (Ziegler, et al. 2009 Dev Sci)



Overview of the project

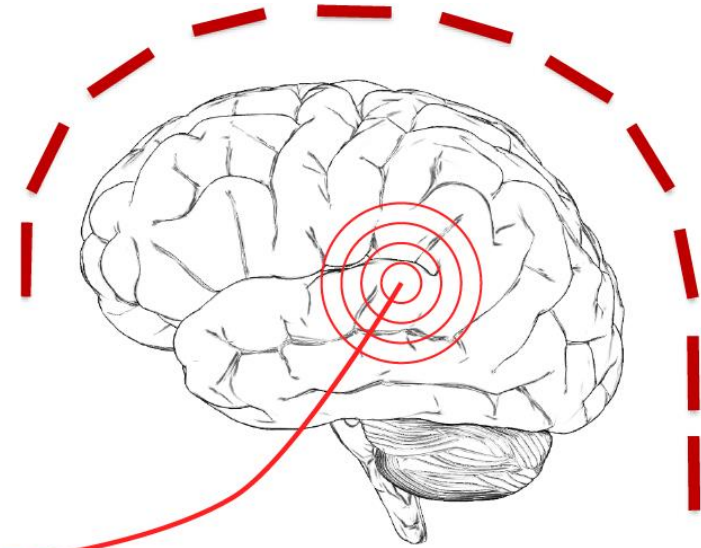
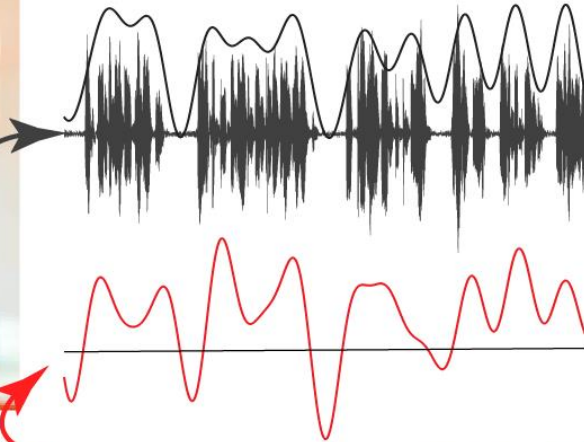
Overarching goal of the project:

To better understand cortical speech processing in the presence of noise, in order to provide novel identification and remediation methods for dyslexia



Overview of the project

Speech brain tracking



Overview of the project

Data collected from children to test some hypotheses

- Speech brain tracking in noise develops with age
- The developmental trajectory depends on noise properties
- Speech tracking in noise is altered in dyslexia

Neuropsychological evaluation



Neuroimaging session



Guidelines on FAIR Data Management in Horizon 2020

1. Data summary
2. FAIR Data
 1. Making data findable, including provisions for metadata
 2. Making data openly accessible
 3. Making data interoperable
 4. Increase data re-use (through clarifying licenses)
3. Allocation of resources
4. Data security
5. Ethical aspects
6. Other

Also build it online at: <https://dmponline.be/>

1. Data summary

1. State the purpose of the data collection/generation
2. Explain the relation to the objectives of the project
3. Specify the types and formats of data generated/collected
4. Specify if existing data is being re-used (if any)
5. Specify the origin of the data
6. State the expected size of the data (if known)
7. Outline the data utility: to whom will it be useful

1. Data summary

Neuropsychological data



5. Origin of the data:

- Written tests to assess IQ, reading abilities, memory,...

3. Types and formats of data generated/collected:

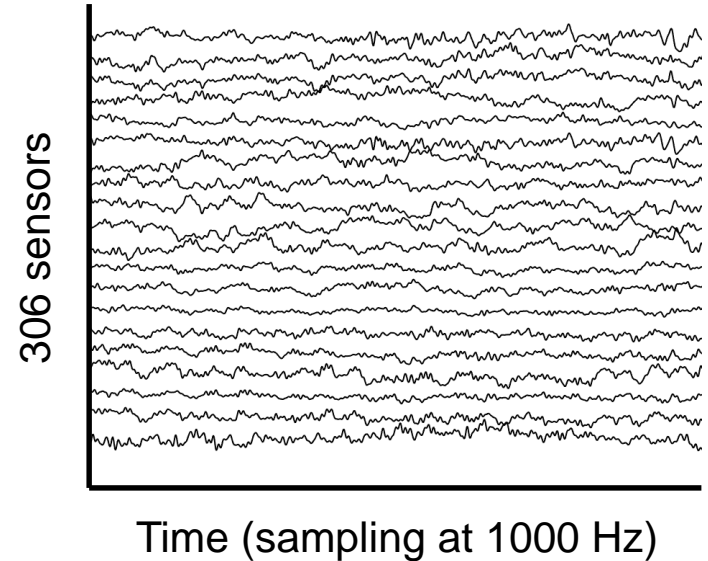
- Paper sheets saved in binders
- Data also encoded in an excel document

6. Expected size of the data:

- negligible

1. Data summary

MEG recording during connected speech listening and an MRI scan



5. Origin of the data:

- Elekta triux MEG device
- an MRI scanner

3. Types and formats of data generated/collected:

- MEG data saved in .fif files and MRI data in .dicom files

6. Expected size of the data:

- 110 participants
- 35 min of MEG (~100 Mb/min) and an MRI scan (~20 Mb) } ~400 Gb

1. Data summary

4. Re-use of the data (if any)

- I will not re-use existing data for the sake of the DysTrack project.

7. Outline the data utility: to whom will it be useful

- To the research team, and hopefully to a broader research community
- To individuals with reading difficulties

Further precisions:

- MEG data will be stored on the lab server and on 2 independent hard-drives
- MRI data will be stored on the hospital repository and on 2 independent hard-drives

2. FAIR data

1. Making data findable, including provisions for metadata

- Data shared with interested researchers upon request (mentioned in publication)
- Publications will include the following keywords: speech tracking, MEG,...
- Participant code: MEG_[x], x being a 4-digit number
- MEG file names: MEG_[x]_[vid_name].fif, vid_name being the name of the video

2. Making data openly accessible

- Upon request, code and heavy files will be sent by a yet-not-defined mean
- Deposit the code on OSF (<https://osf.io>)

3. Making data interoperable

- All data can be read with an open access toolboxes (MNE) coded in matlab and python. A text file will describe what was recorded by all channels

4. Increase data re-use (through clarifying licences)

- For now, no decision has been reached as to the type of open license that will be used for the data.

3–6. All other items

3. Allocation of resources

- There is not cost associated with the procedure I will use to share the data
- I will be responsible for sharing the data with others

4. Data security

- The data will be saved on two hard drives
- Disk integrity will be checked yearly

5. Ethical aspects

- Data collected from children that include sensitive information (IQ and reading abilities)
 - Pseudonymization
- The informed consent clearly states that pseudonymized data can be shared with third parties
- The informed consent clearly states that data will be saved electronically at the Erasme hospital

6. Other issues

- There are no other issues

Pros and Cons of the DMP

Pros

- Takes little time to build
- Forces to think about practical aspects of data management

Cons

- One more thing that takes little time
- Strategies defined in it might be already implemented
- Overlap with ethics agreement request

That's it

Thank you for your attention