

Music Imagery Information Retrieval

Sebastian Stober, Avital Sternin,
Jessica A. Grahn and Adrian M. Owen

NEMISIG 2015 – Ithaca, NY
January 31, 2015



Music Imagery:

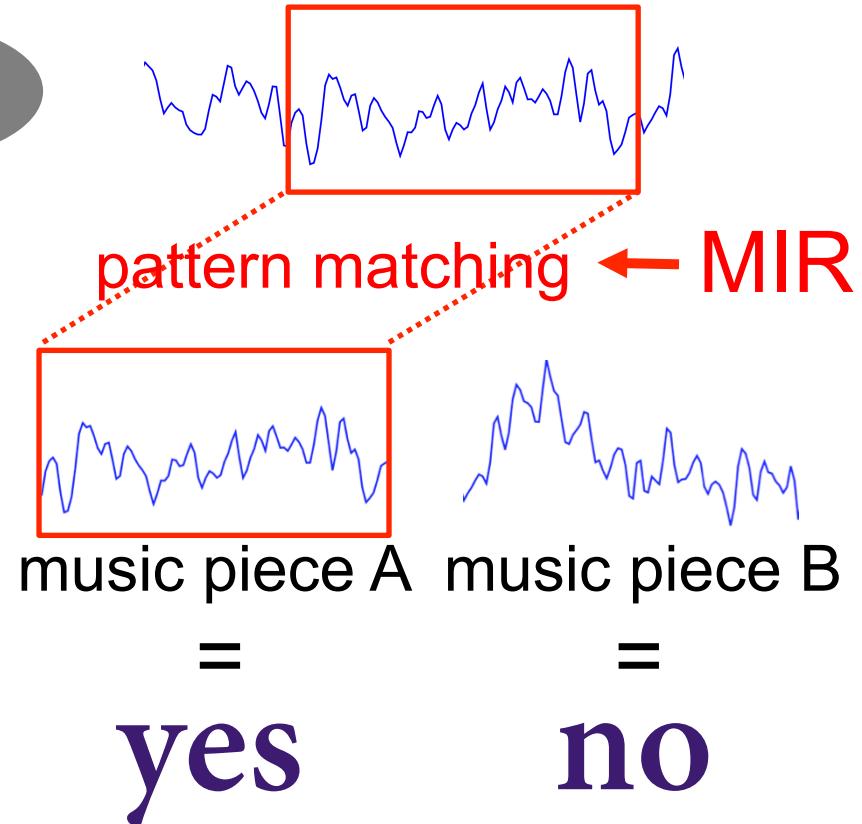
*“the process of deliberately imagining well-known music,
by recreating the perceptual experience internally”*

(R. Schaefer)

Music Imagery Information Retrieval (MIIR)

= *“retrieving music from brain signals”*

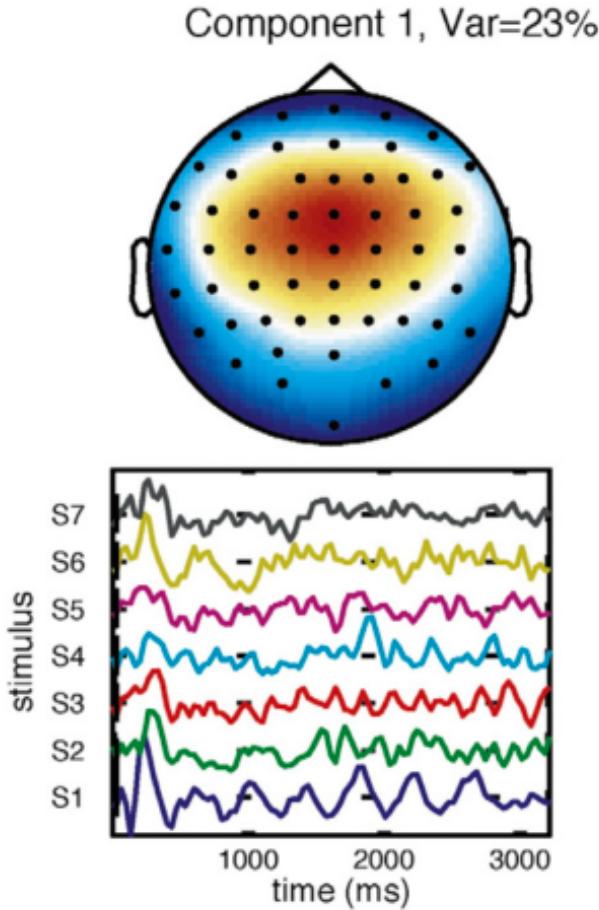
Brain-Computer Interaction through Music Imagery Information Retrieval



So Far:

A Brief Overview of the State of the Art

Differentiating Perception

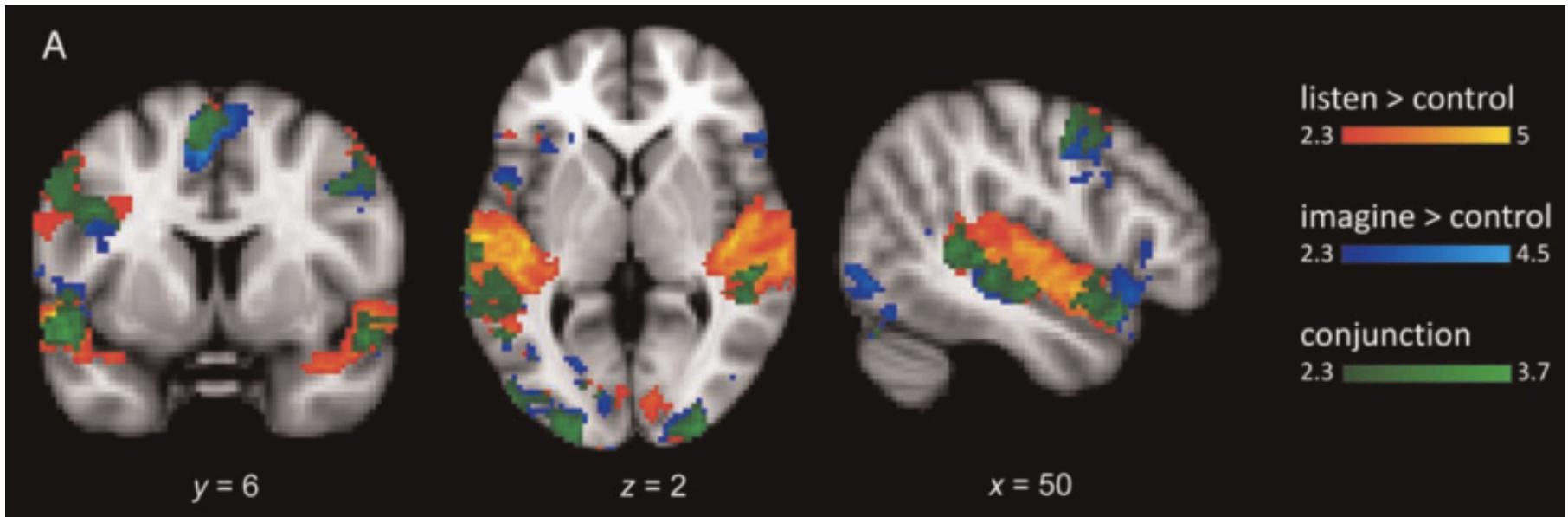


- Perceived melodies can be accurately differentiated using EEG
 - 70% accuracy - single-trial, single participant
- Common representation for music in the grand average (cross participant classification - 53%)

Schaefer et al., 2011

Music Imagination / Perception

- Imagination and perception of music uses similar brain structures



Herholz et al., 2012

Using EEG to detect imagination

- Music unfolds temporally
- EEG has high temporal resolution
- EEG used to investigate shared brain activation between perception and imagination of music

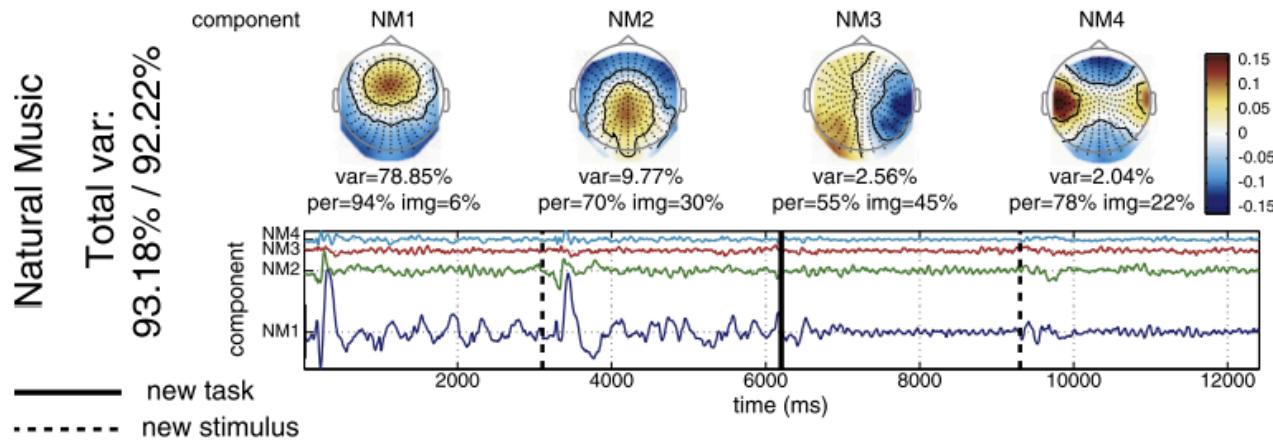


Fig. 3. Here, the distributions of the PCA components are shown for each dataset, showing all the relevant components (named for each experiment), the average explained variance per dataset and the split of the explained variance over task. In the time-courses, perception is always on the left, separated from imagery by a solid line, while the dotted lines denote different stimuli.

Schaefer et al., 2013

Rhythm Perception Classification

Dataset: Rhythm Perception Study

- recorded 2013 in Kigali, Rwanda
- 13 participants (mean age 21)
- 14 EEG channels @ 400Hz
- 12 East African/12 Western rhythms
 - constructed from 4 groups of 3 sequences
 - 3 different pairs per group (12 pairs in total)
 - for each pair within each group, one played at 375Hz and the other at 500Hz (perfect fourth)
 - 24 stimuli = 4 groups * 3 pairs * 2 pitches
- presented for 32s each, 4s break
- same tempo within participants
- block design (randomized within)

East African Sequences

①	· · · · · ·
②	· · · ·· · ··· · · ··
③	· ·· · ·· · · · ·· ··
④	· · · · · ·
⑤	· · · · · · · · · · · ·
⑥	· · · · · · · · · · · ·

Western Sequences

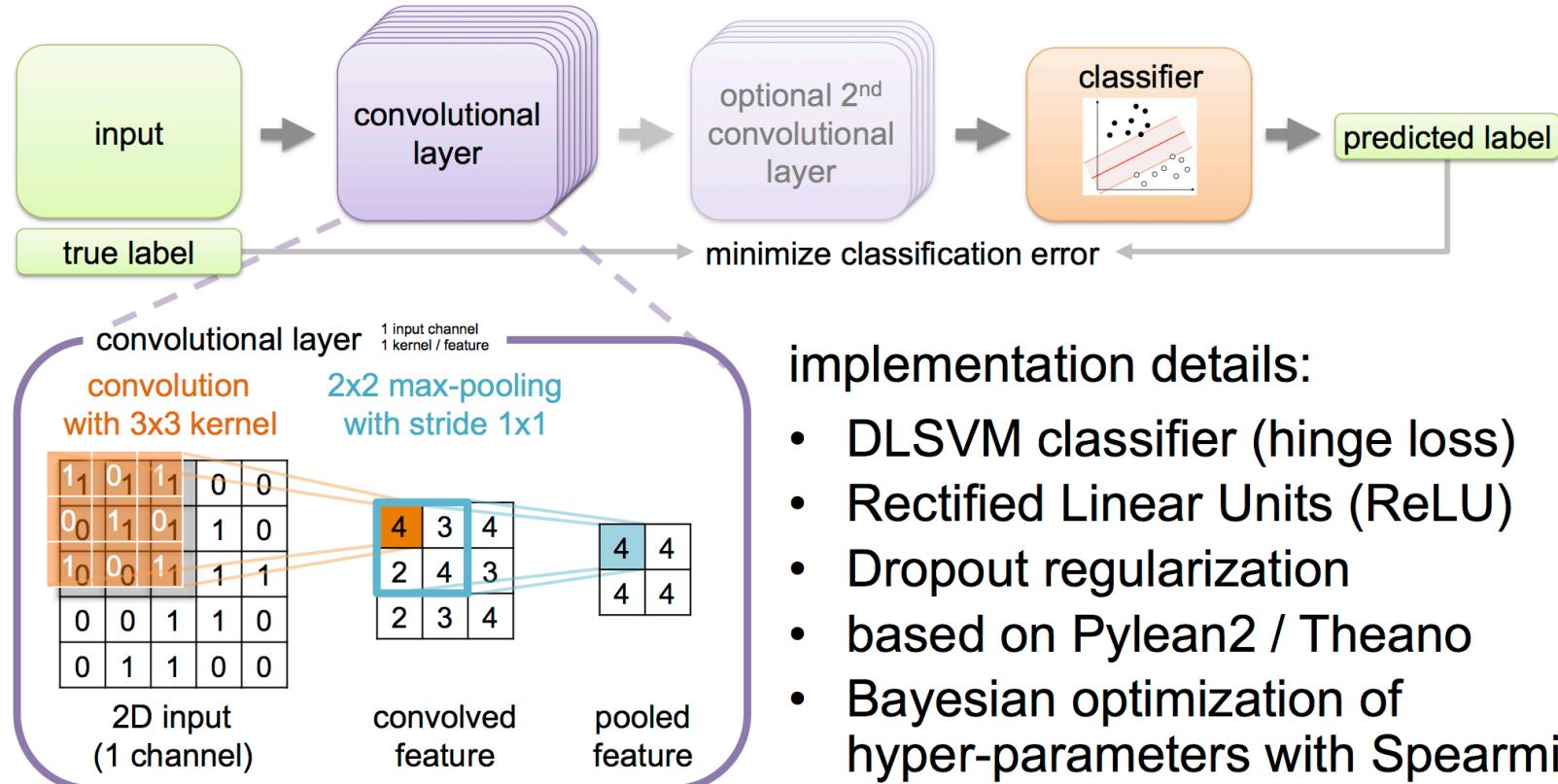
①	· · · · · ·
②	·· · ·· · · ·· · ·· · ··
③	· · · · · · · · · · · ·
④	· · · ·· · · · · ·· ·
⑤	· ·· · · ·· ·
⑥	· · · · · · · · · · · ·

Example: East African rhythm stimulus [1,2,'a']



Rhythm Perception Classification

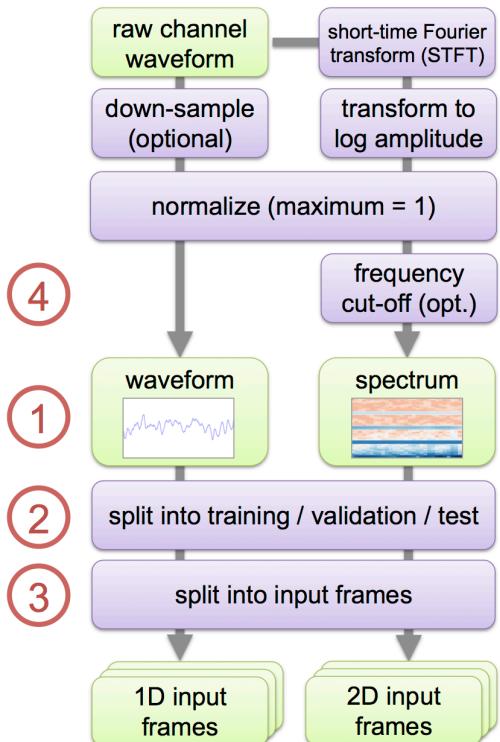
Convolutional Neural Networks (CNNs)



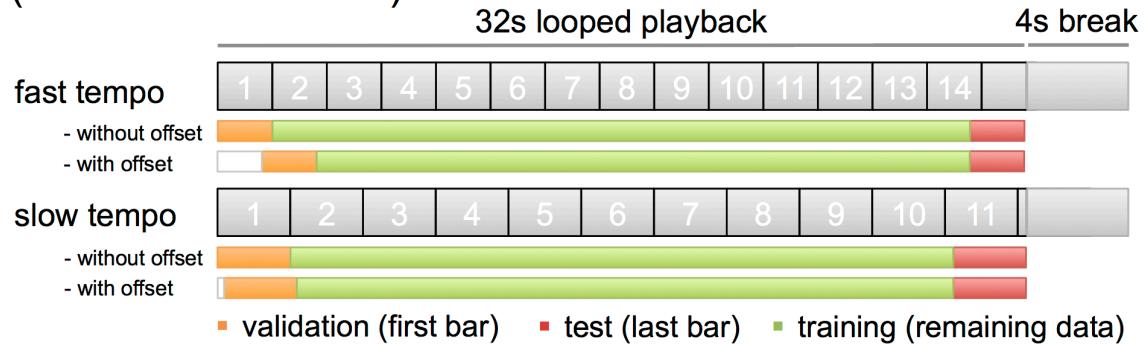
Rhythm Perception Classification

Impact of Pre-Processing

pre-processing pipelines



splitting each trial into training, validation and test data
(with or without offset)



splitting the (training) data into input frames

- with overlap (hop size 60ms)

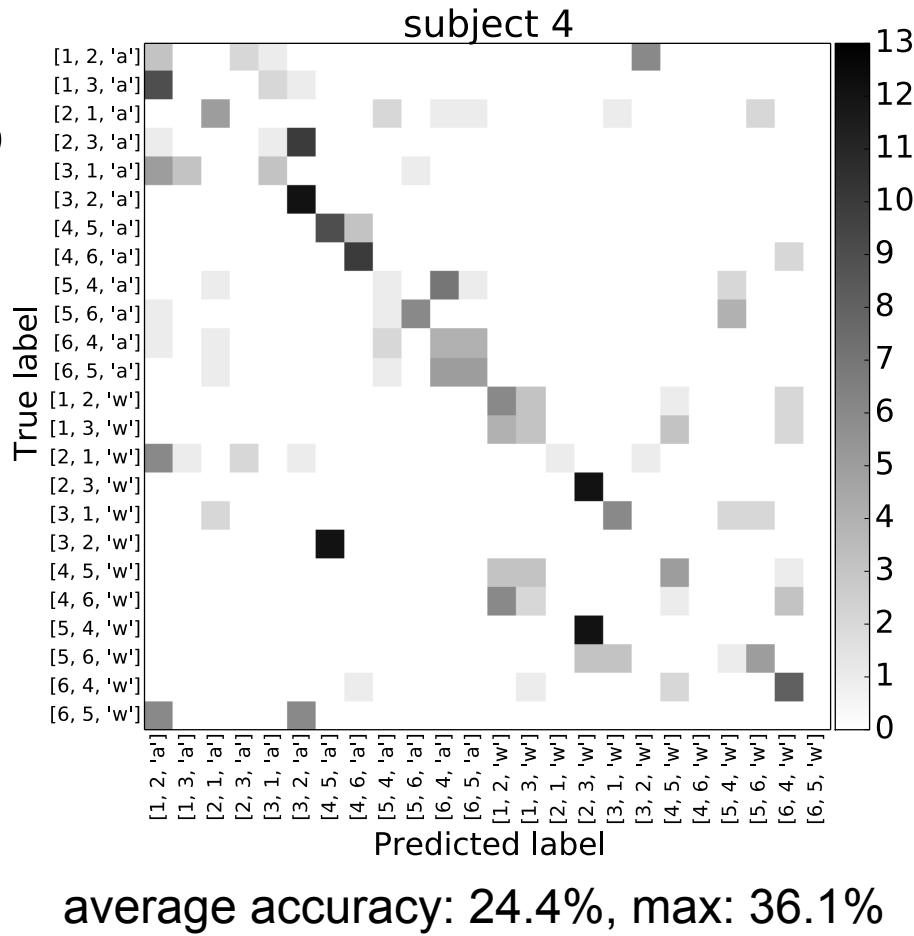


- without overlap (hop size 1 bar)



Rhythm Perception Classification

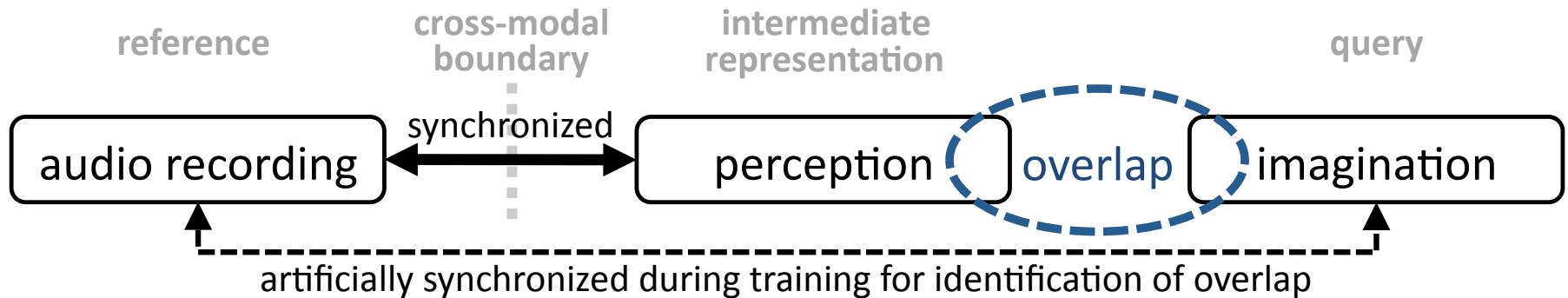
- challenging task – even for humans (rhythms not easy to distinguish)
- no improvement by second convolutional layer (may need more training data)
- problems caused by
 - limited amount of data
 - rather poor data quality (non-lab conditions, portable equipment)
 - single trial / block study design (vulnerability to block artifacts)



Next:

An Open Dataset of Music Perception and Imagination Recordings

General Approach



- primary goal: recognize individual pieces
- further possibilities:
 - cross-domain sync of EEG and audio
 - audio re-synthesis from EEG

Considerations

- explore stimulus space dimensions
 - as many trial repetitions as possible
 - maximum 2h per session
-
- support multiple MI(I)R tasks
 - make dataset access as easy as possible
 - allow easy replication / extension

Stimulus Space Dimensions

- vocals vs. instrumental
 - rhythm
 - measure (3/4, 4/4, 5/8, 6/8, ...)
 - tempo
 - emotion (e.g. VA model)
 - instrumentation (timbre)
 - ...
-
- ideally: balanced classes
 - How many trials to train useful models?

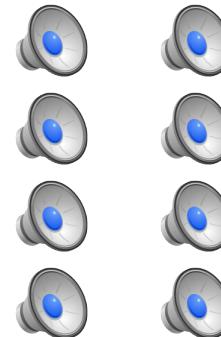
The Stimuli

- 12 audio recordings of 8 music pieces
 - 4 songs recorded with and without lyrics
 - 4 instrumental pieces
- complete musical phrases
- length between 7.6s and 16s (mean 10.5)
plus tempo cue (1 or 2 bars click)
- volume normalized
- tempo adjusted for pairs

The Stimuli

songs with / without lyrics:

- ① Chim Chim Cheree
- ② Take me out to the ballgame
- ③ Jingle Bells
- ④ Mary Had a Little Lamb



instrumental pieces:

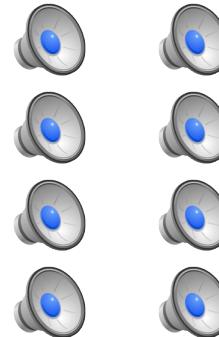
- ① Emperor Waltz
- ② Harry Potter Theme
- ③ Imperial March (Star Wars Theme)
- ④ Eine Kleine Nachtmusik



The Stimuli lyrics vs. instrumental

songs **with** / **without** lyrics:

- ① Chim Chim Cheree
- ② Take me out to the ballgame
- ③ Jingle Bells
- ④ Mary Had a Little Lamb



instrumental pieces:

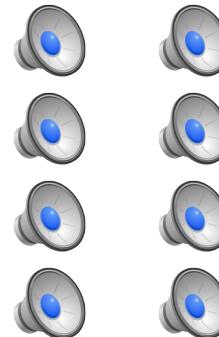
- ① Emperor Waltz
- ② Harry Potter Theme
- ③ Imperial March (Star Wars Theme)
- ④ Eine Kleine Nachtmusik



The Stimuli 3/4 vs. 4/4

songs with / without lyrics:

- ① Chim Chim Cheree
- ② Take me out to the ballgame
- ③ Jingle Bells
- ④ Mary Had a Little Lamb



instrumental pieces:

- ① Emperor Waltz
- ② Harry Potter Theme
- ③ Imperial March (Star Wars Theme)
- ④ Eine Kleine Nachtmusik



The Stimuli tempo spectrum

songs with / without lyrics:

①	Chim Chim Cheree			210
②	Take me out to the ballgame			186
③	Jingle Bells			244
④	Mary Had a Little Lamb			160

instrumental pieces:

①	Emperor Waltz		178
②	Harry Potter Theme		176
③	Imperial March (Star Wars Theme)		104
④	Eine Kleine Nachtmusik		140

The Stimuli genres

songs with / without lyrics:

- | | | | |
|-------------------------------|--|--|--------------------|
| ① Chim Chim Cheree | | | soundtrack |
| ② Take me out to the ballgame | | | soundtrack / trad. |
| ③ Jingle Bells | | | seasonal / trad. |
| ④ Mary Had a Little Lamb | | | nursery rhyme |

instrumental pieces:

- | | | |
|------------------------------------|--|--------------------|
| ① Emperor Waltz | | classic / ballroom |
| ② Harry Potter Theme | | soundtrack* |
| ③ Imperial March (Star Wars Theme) | | soundtrack* |
| ④ Eine Kleine Nachtmusik | | classic |

*same composer



The Stimuli odd ones

songs with / without lyrics:

- | | | | |
|-------------------------------|--|--|-------------------|
| ① Chim Chim Cheree | | | ugly MIDI version |
| ② Take me out to the ballgame | | | |
| ③ Jingle Bells | | | very uplifting |
| ④ Mary Had a Little Lamb | | | piano version |

instrumental pieces:

- | | | |
|------------------------------------|--|---------------------|
| ① Emperor Waltz | | very danceable |
| ② Harry Potter Theme | | timbre / instrument |
| ③ Imperial March (Star Wars Theme) | | dark / menacing |
| ④ Eine Kleine Nachtmusik | | |

Experiment Design – Part 1

- consent
- questionnaire
 - basic demographic information
 - experience with music and musical training
 - familiarity of the stimuli
- beat tapping along to stimuli
- familiarity + beat tapping required for part 2

Experiment Setup

presentation
system

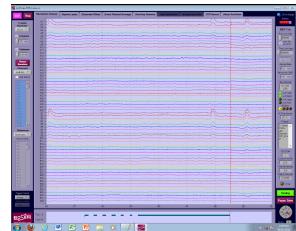
feedback

video

audio

events

recording
system



↑



markers



stimtracker

(optical)

sound booth

presentation system

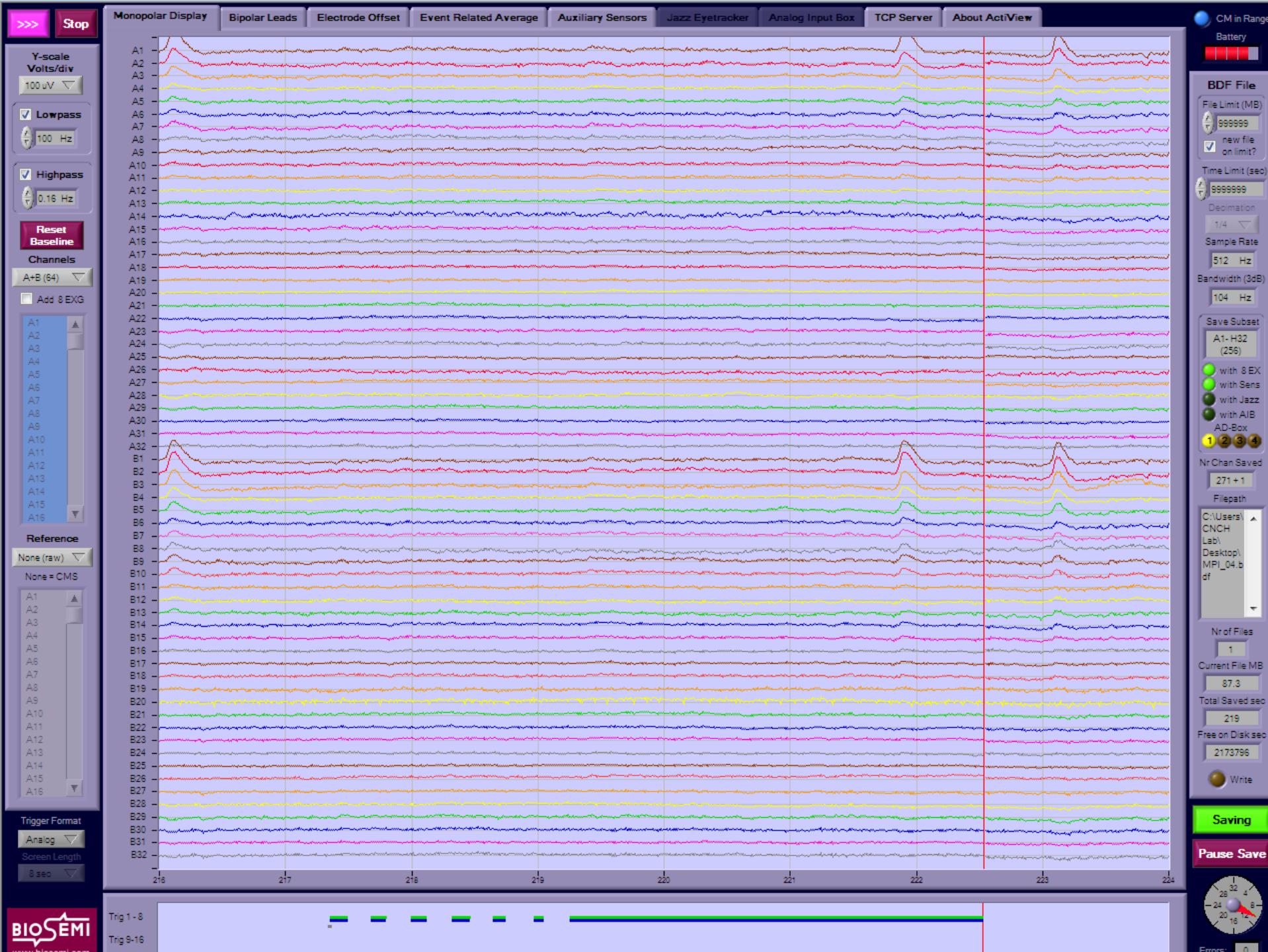
screen & speakers

feedback keyboard



EEG amp
on battery

Biosemi ActiveTwo, 64 EEG + 4 EOG channels @ 512 Hz



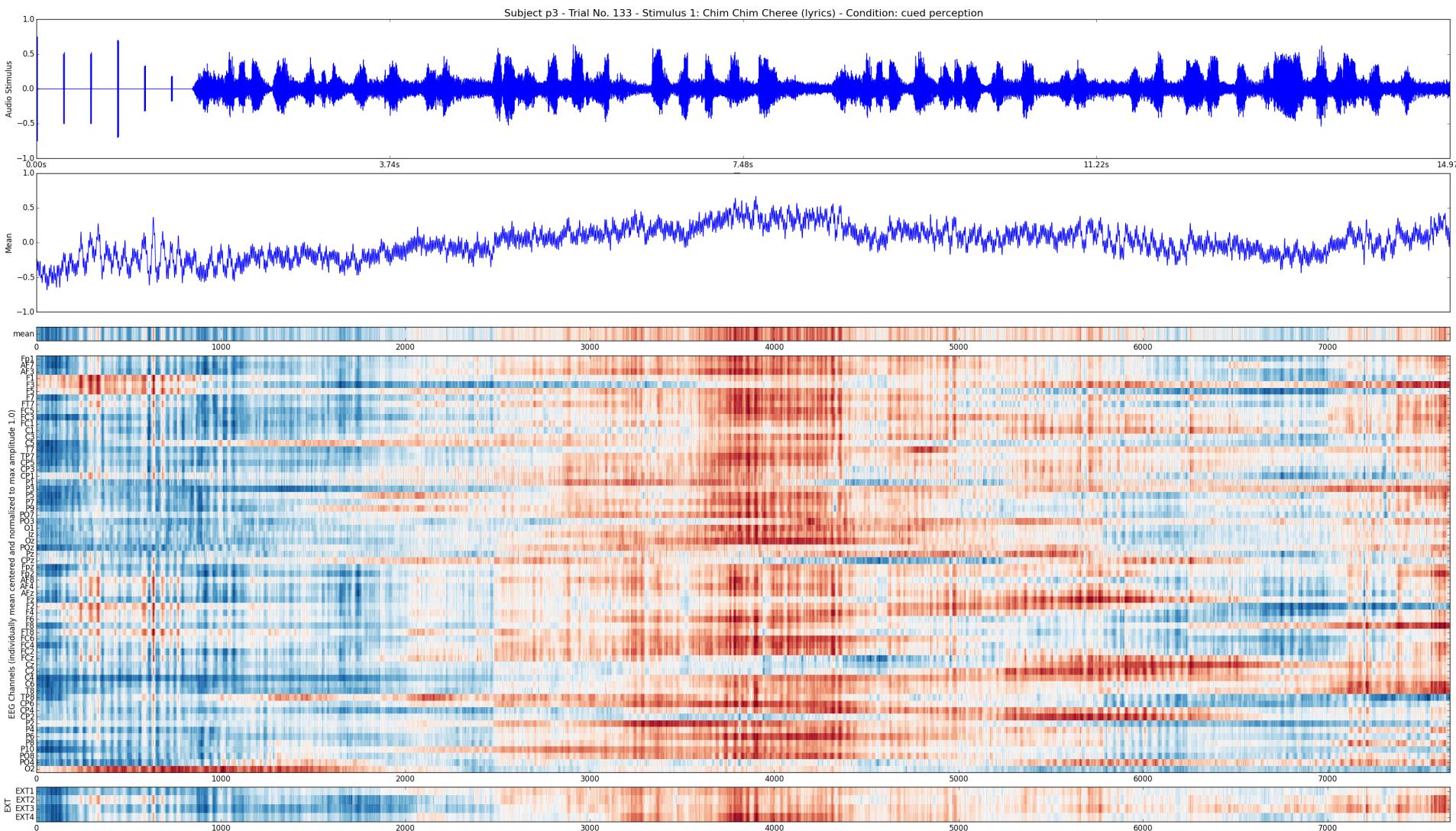
Participant Perspective



Experiment Design – Part 2

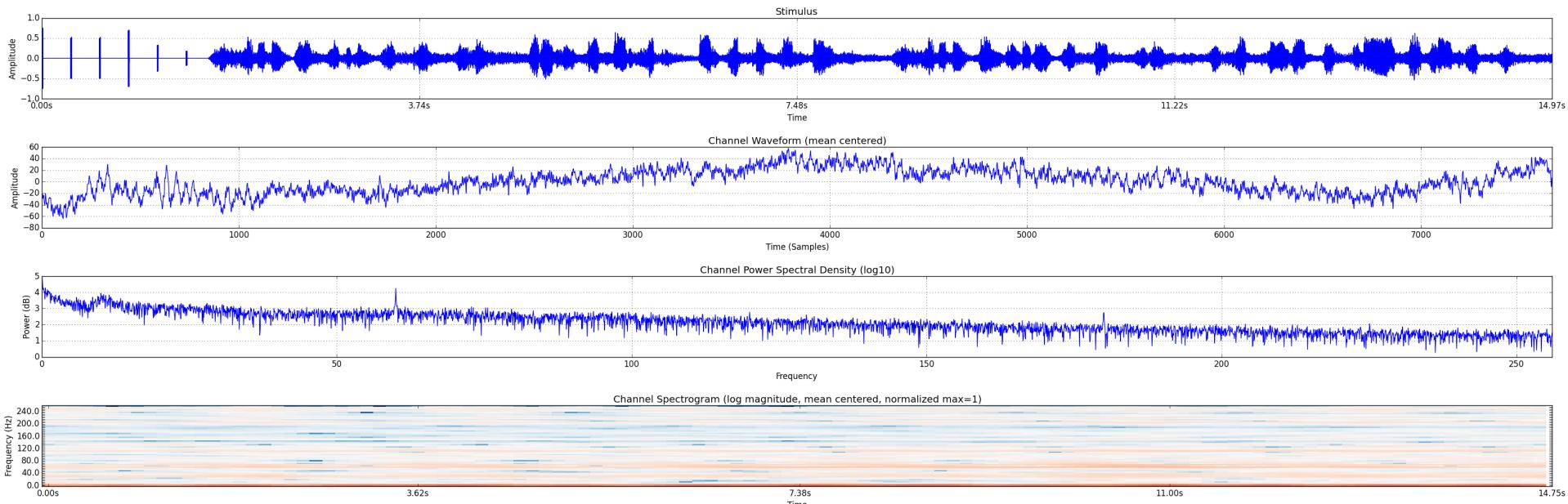
- **Block I:** 3 conditions per stimulus
(5 repetitions of all stimuli in random order)
 - 1) perception: “*Please listen to ...*”
 - 2) imagination with tempo cue: “*Please imagine ... following the cue*”
 - 3) imagination without tempo cue: “*Please imagine ... when the fixation cross appears*”
- **Block II:** only condition 3, with feedback on success
(5 repetitions of all stimuli in random order)

Example: Data of a Single Trial



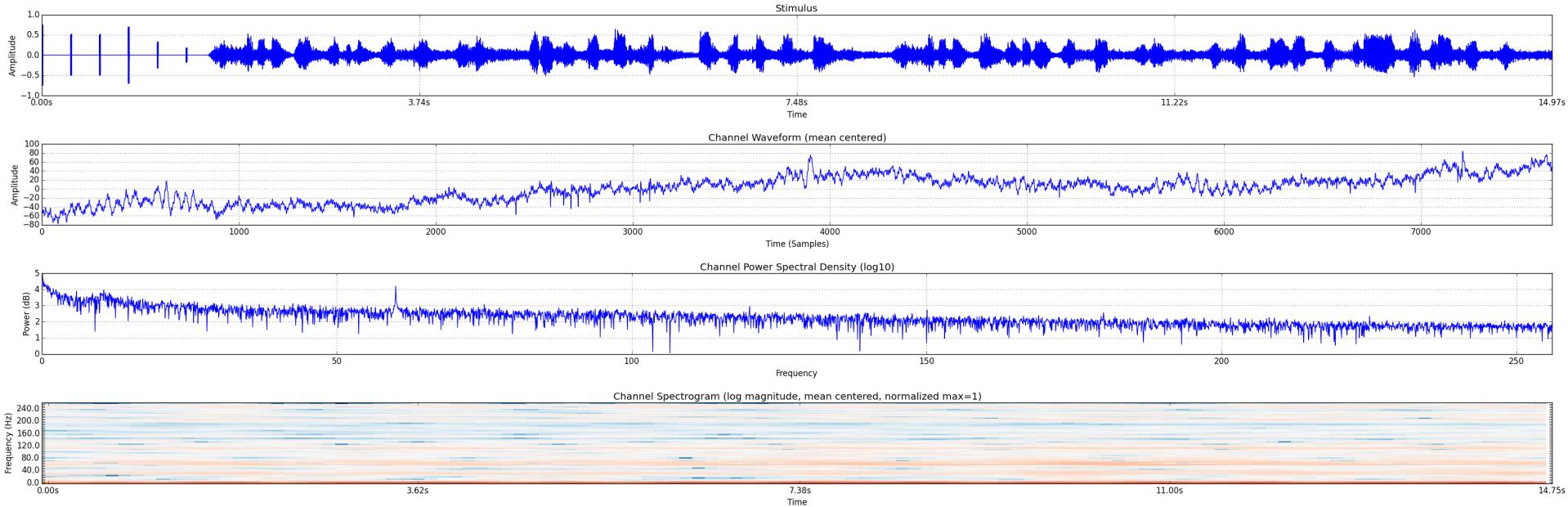
Channel Fp1

Subject p3 - Trial No. 133 - Stimulus 1: Chim Chim Cheree (lyrics) - Condition: cued perception - Channel #0 Fp1

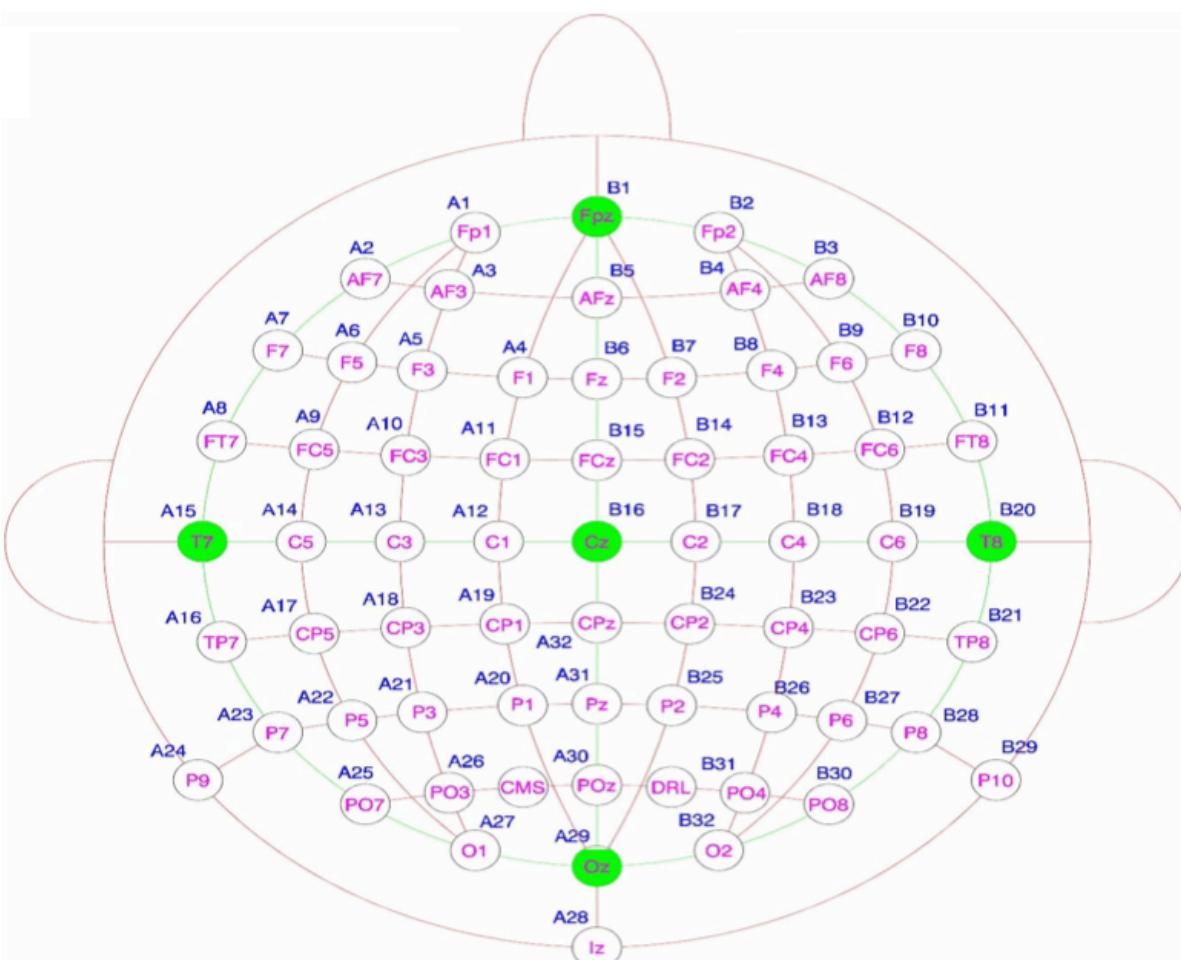


Channel T8

Subject p3 - Trial No. 133 - Stimulus 1: Chim Chim Cheree (lyrics) - Condition: cued perception - Channel #51 T8



EEG Channel Layout



64 EEG channels
+2 reference electrodes
+4 EOG channels

http://www.biosemi.com/download/Cap_coords_all.xls



Western
The Brain and
Mind Institute

Sebastian Stober - Music Imagery Information Retrieval

2015-01-31 30

Collected Data

- 5 repetitions x 4 conditions x 12 stimuli
= 240 trials per subject
- ~55min EEG recording (only trials)
with 64+4 channels @ 512 Hz
= ~400 MByte per subject
- anonymized answers from questionnaire
- success feedback in Part II

ODC Public Domain Dedication License (PDDL)

You are free:



To Share: To copy, distribute and use the database.



To Create: To produce works from the database.



To Adapt: To modify, transform and build upon the database.

As long as you:



Blank: This section is intentionally left blank. The PDDL imposes no restrictions on your use of the PDDL licensed database.

<http://opendatacommons.org/licenses/pddl/>

Tentative Schedule

- 1st participant on January 30
- 1st preview @ NEMISIG 2015
- 1st batch (~20 participants) by early March

Questions?

- How would you like your data?
 - format, download options, ...
- Who would like to try tomorrow?

Thank You!

DAAD



**Deutscher Akademischer Austausch Dienst
German Academic Exchange Service**

Canada Excellence
Research Chairs
Chaires d'excellence
en recherche du Canada



contact: sstober@uwo.ca | <http://www.owenlab.uwo.ca>
<https://github.com/sstober/deephought>



Western
The Brain and
Mind Institute

Sebastian Stober - Music Imagery Information Retrieval

2015-01-31 34