HOW ARE SPEECH SOUNDS REPRESENTED IN LONG TERM MEMORY?

Diogo Almeida

diogo@nyu.edu | *www.diogoalmeida.net*

جامعة نيويورك أبوظبي

NYU | ABU DHABI

#myNYUAD

students from over 80 countries, speaking 80-100 languages

01
THE METHODS

GOAL:

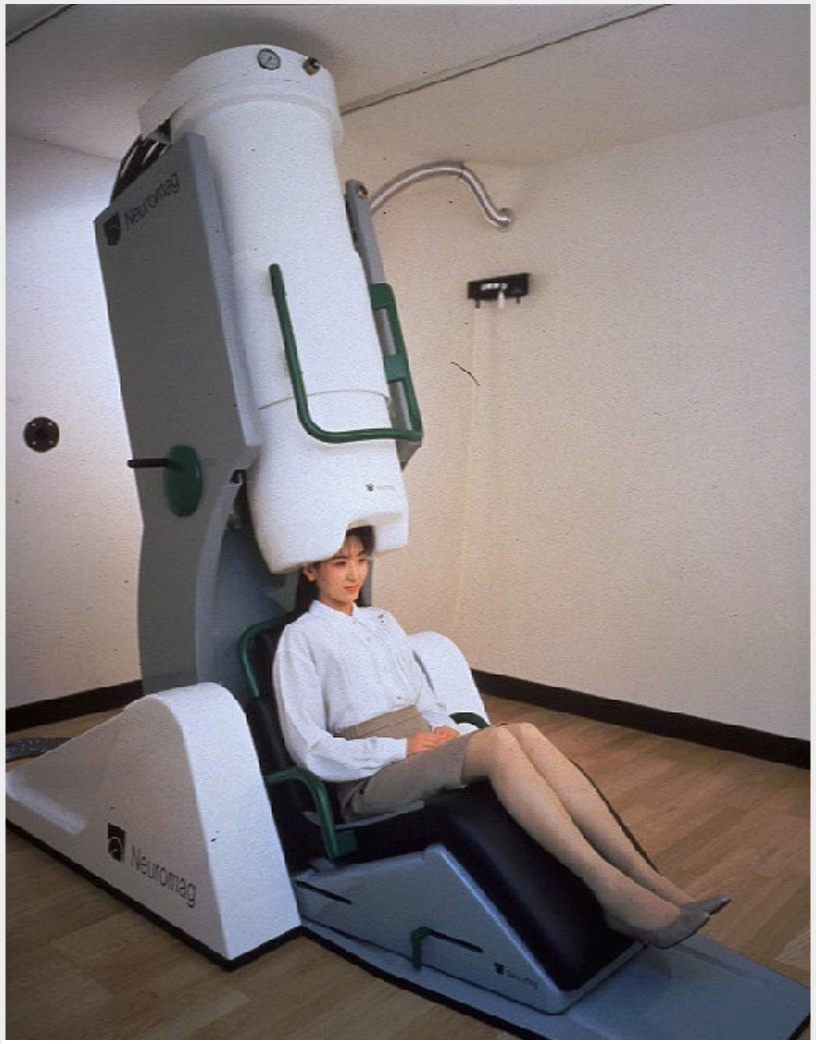
BRAIN DATA → STRUCTURE OF THE MIND

TODAY: SPEECH SOUNDS IN LONG TERM MEMORY

ELECTROMAGNETISM:

ELECTROENCEPHALOGRAPHY (EEG)
MAGNETOENCEPHALOGRAPHY (MEG)

MEG



EEG

ELECTROMAGNETISM

Millisecond temporal resolution.

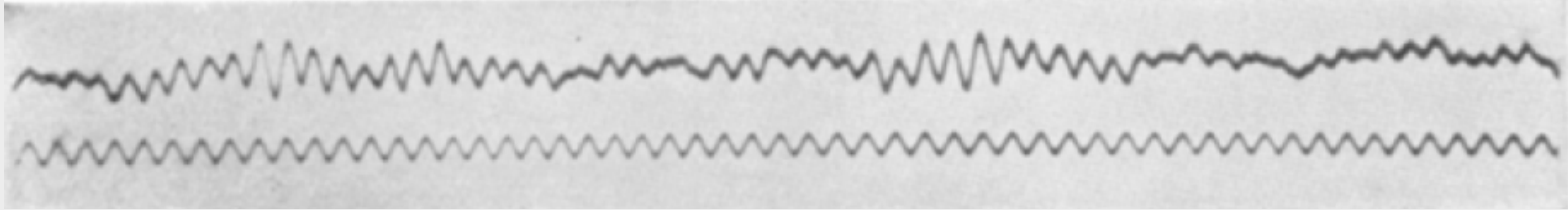
Neurons communicate thousands of times per second by sending each other tiny electrical impulses

Populations of neurons are connected into networks

When networks fire in *synchrony*, the dynamics of the electric activity can be detected and recorded outside the skull.

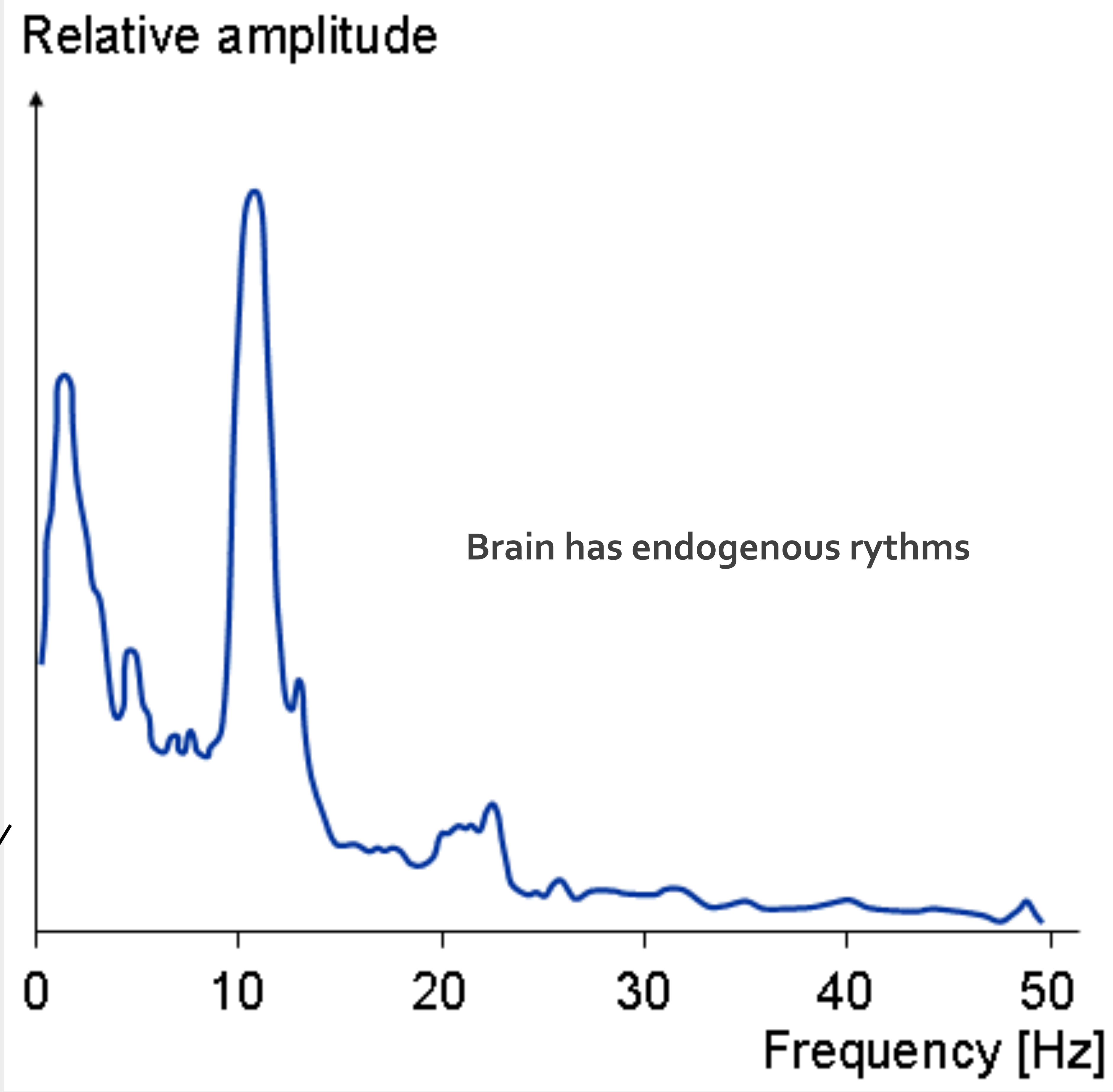
First EEG recording in a human.

Hans Berger (1924/1929)



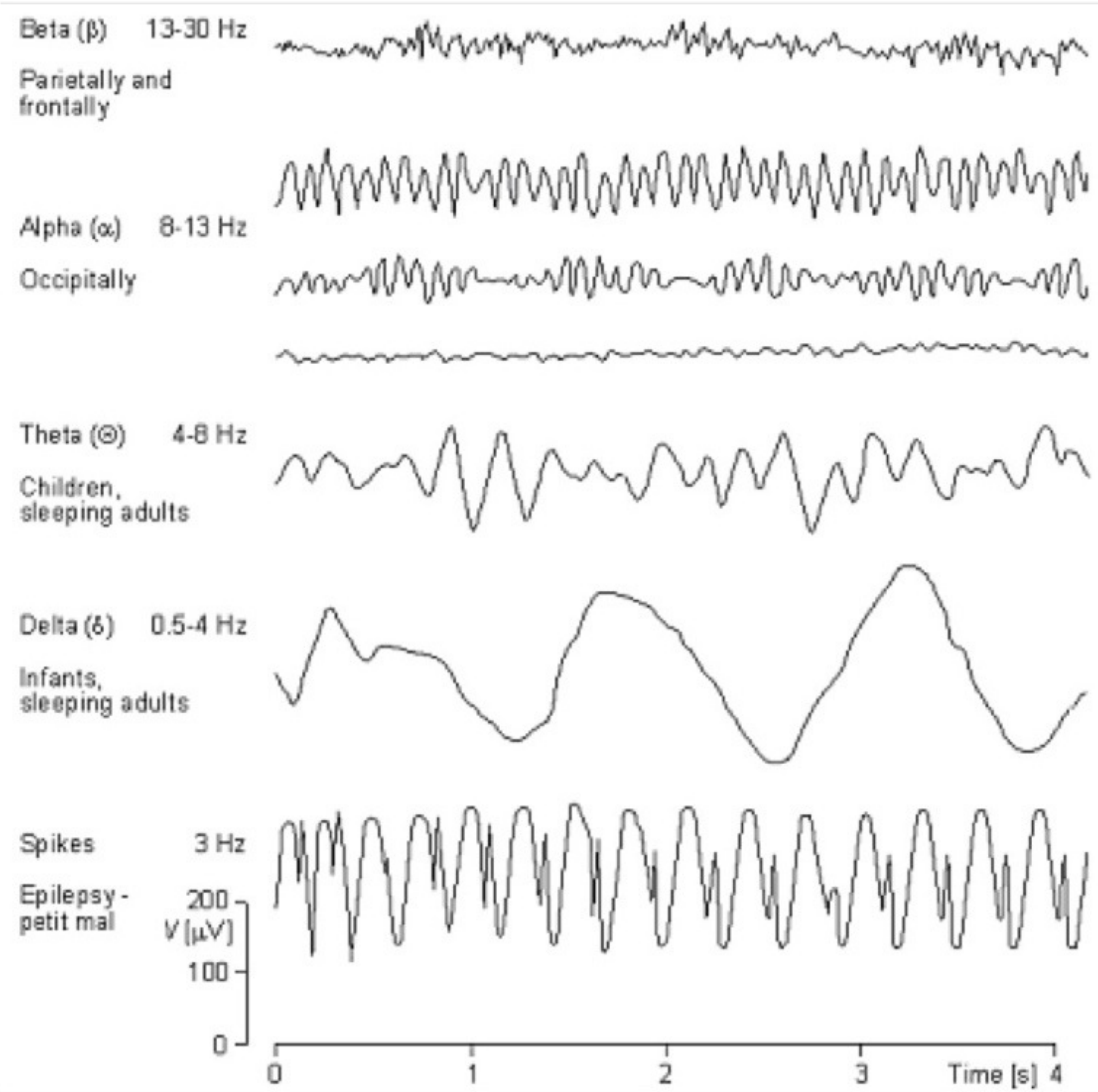
EEG

10 Hz ref signal



Brain has endogenous rythms

spectrum



- Delta 0.5-4 Hz
- Theta 4-8 Hz
- Alpha 8-13 Hz
- Beta 13-30 Hz
- Gamma 30-50 Hz

Brain rythms might reflect different consciousness/mood states

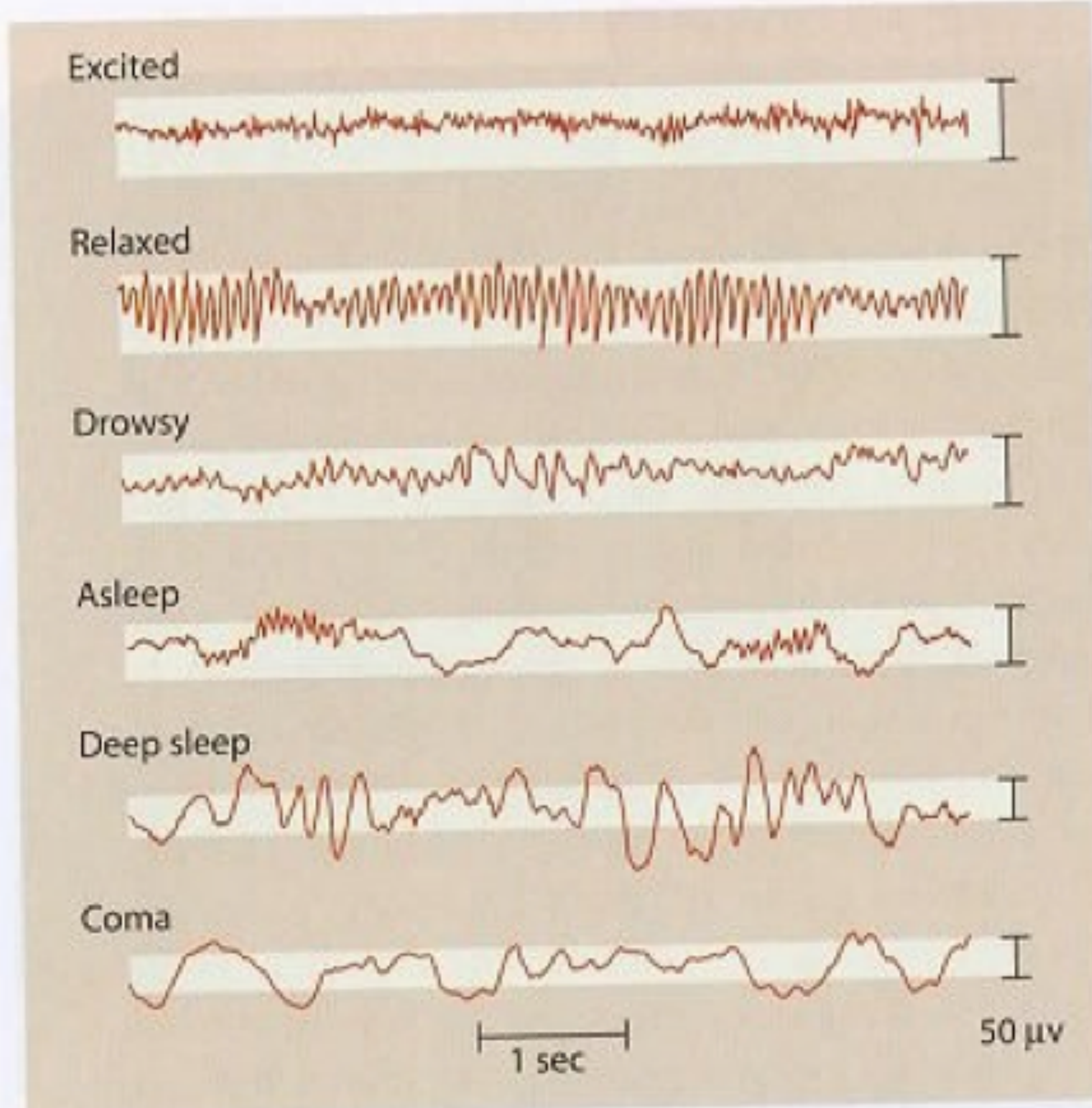
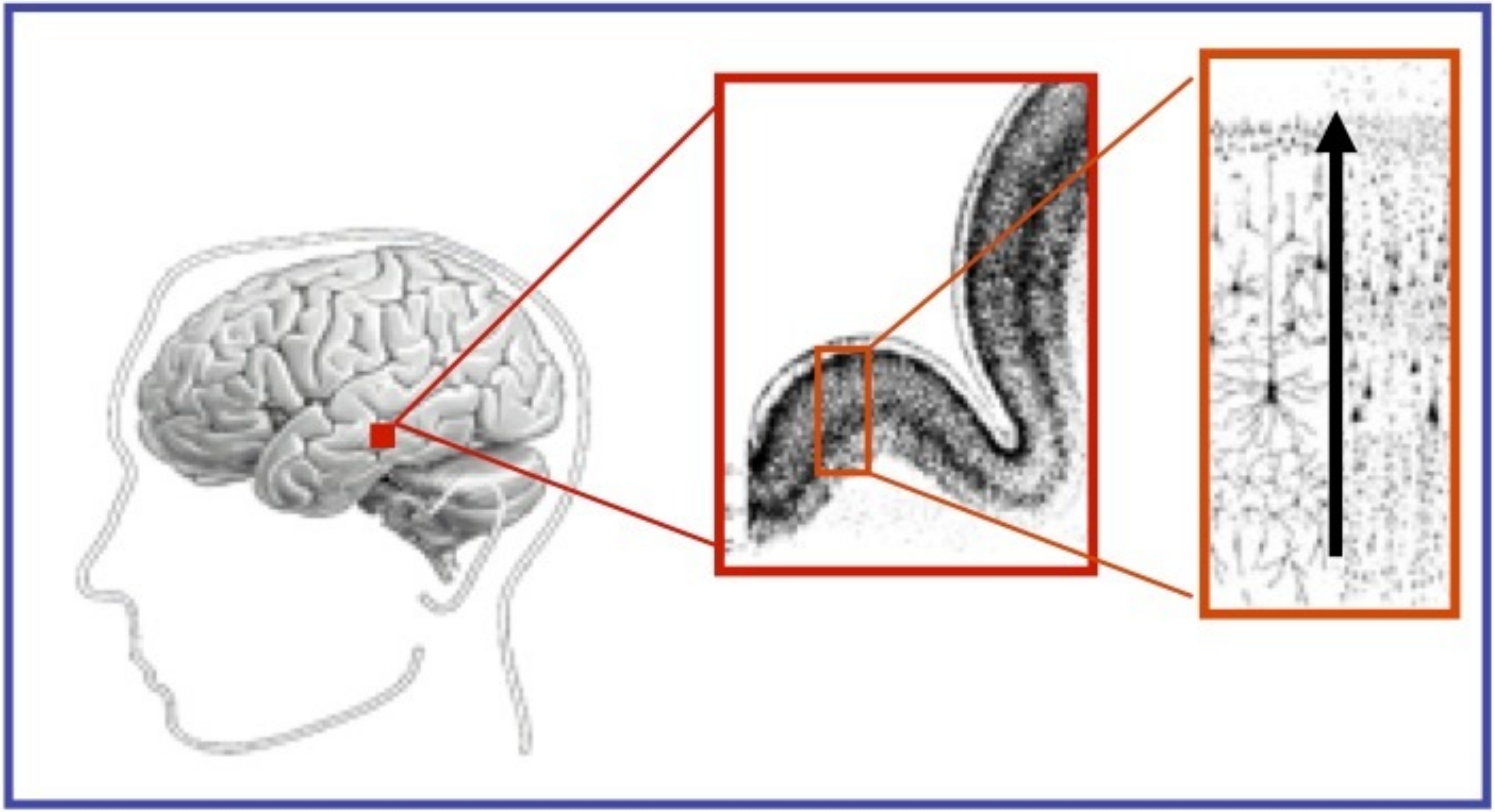
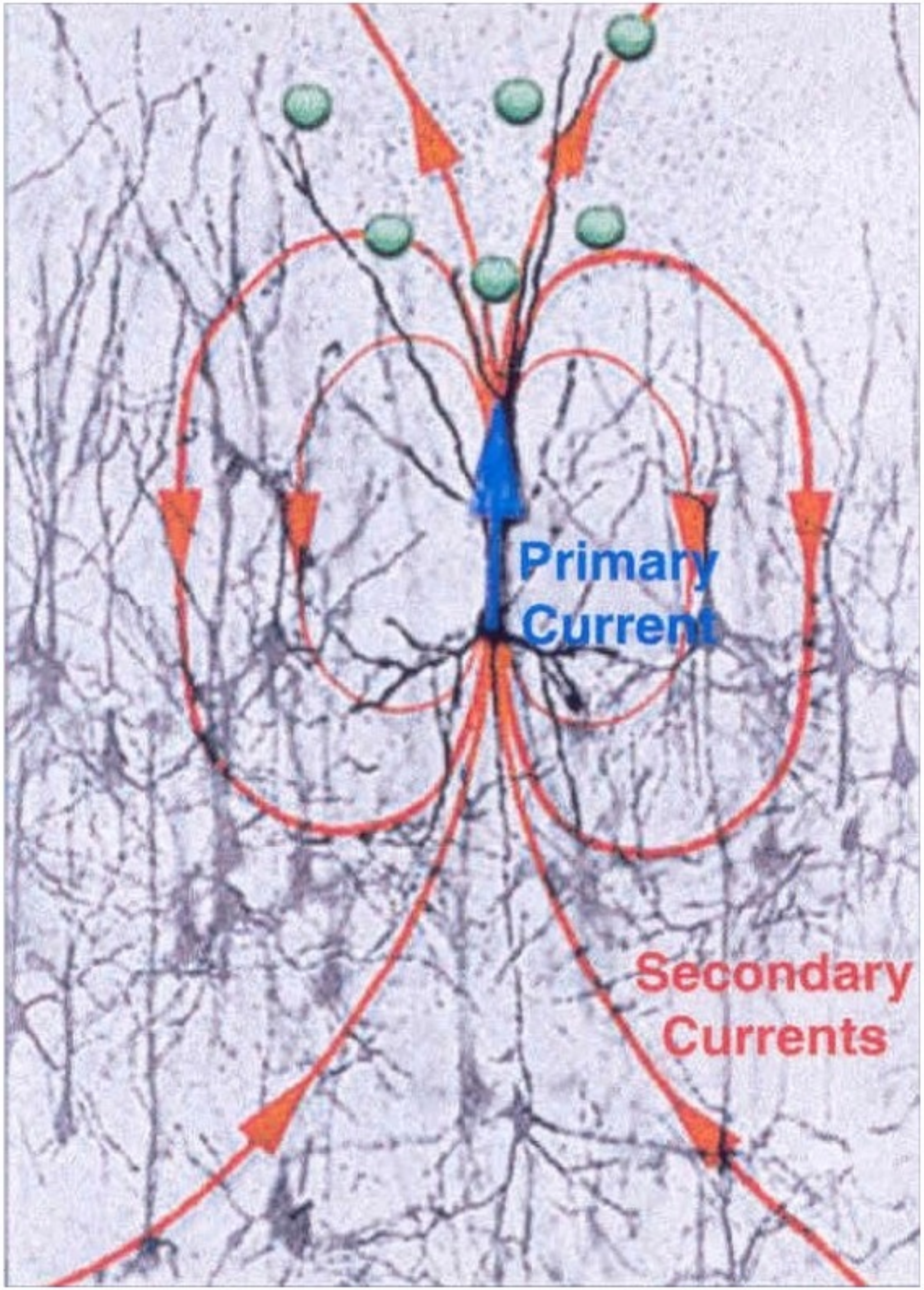
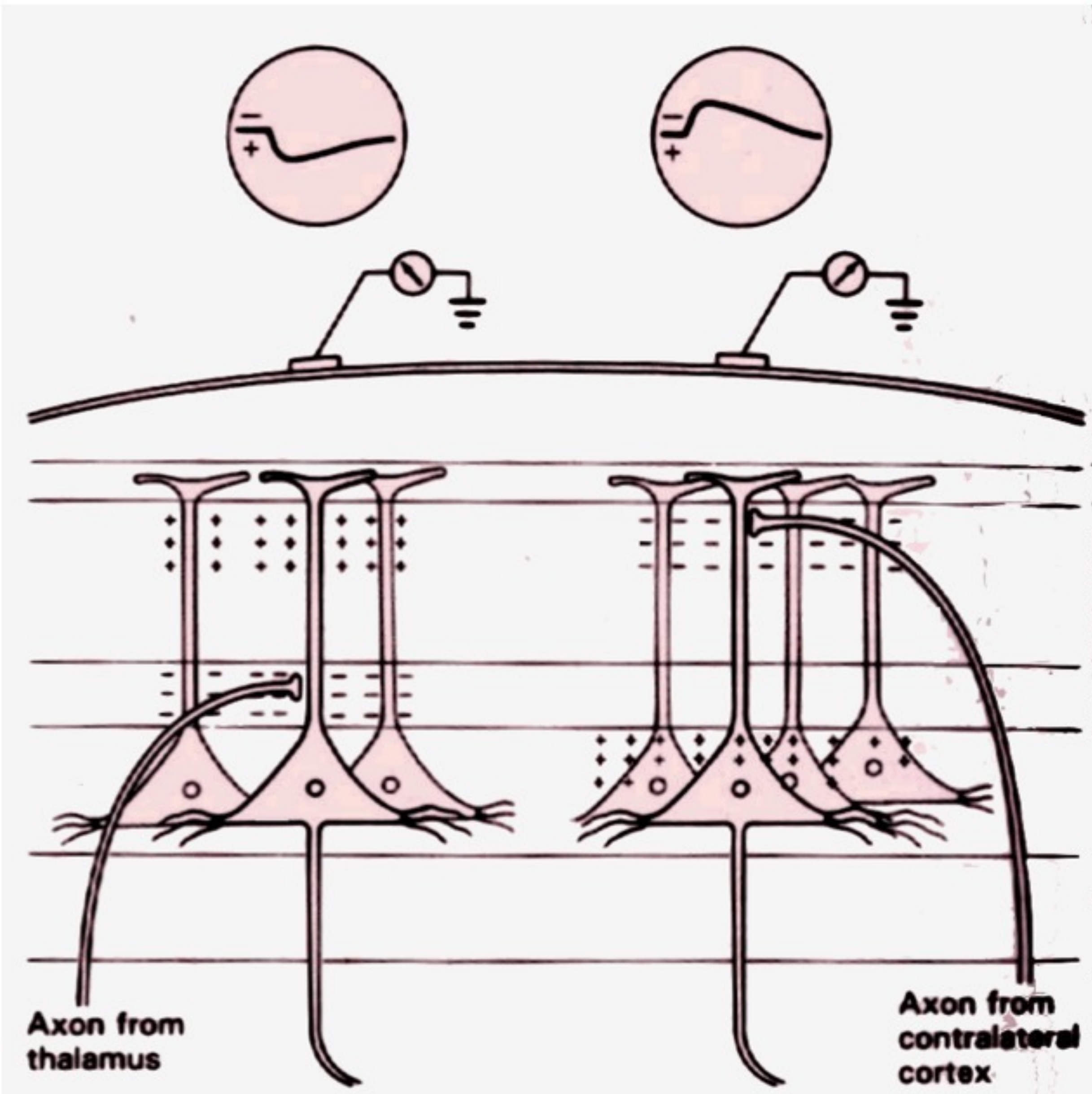


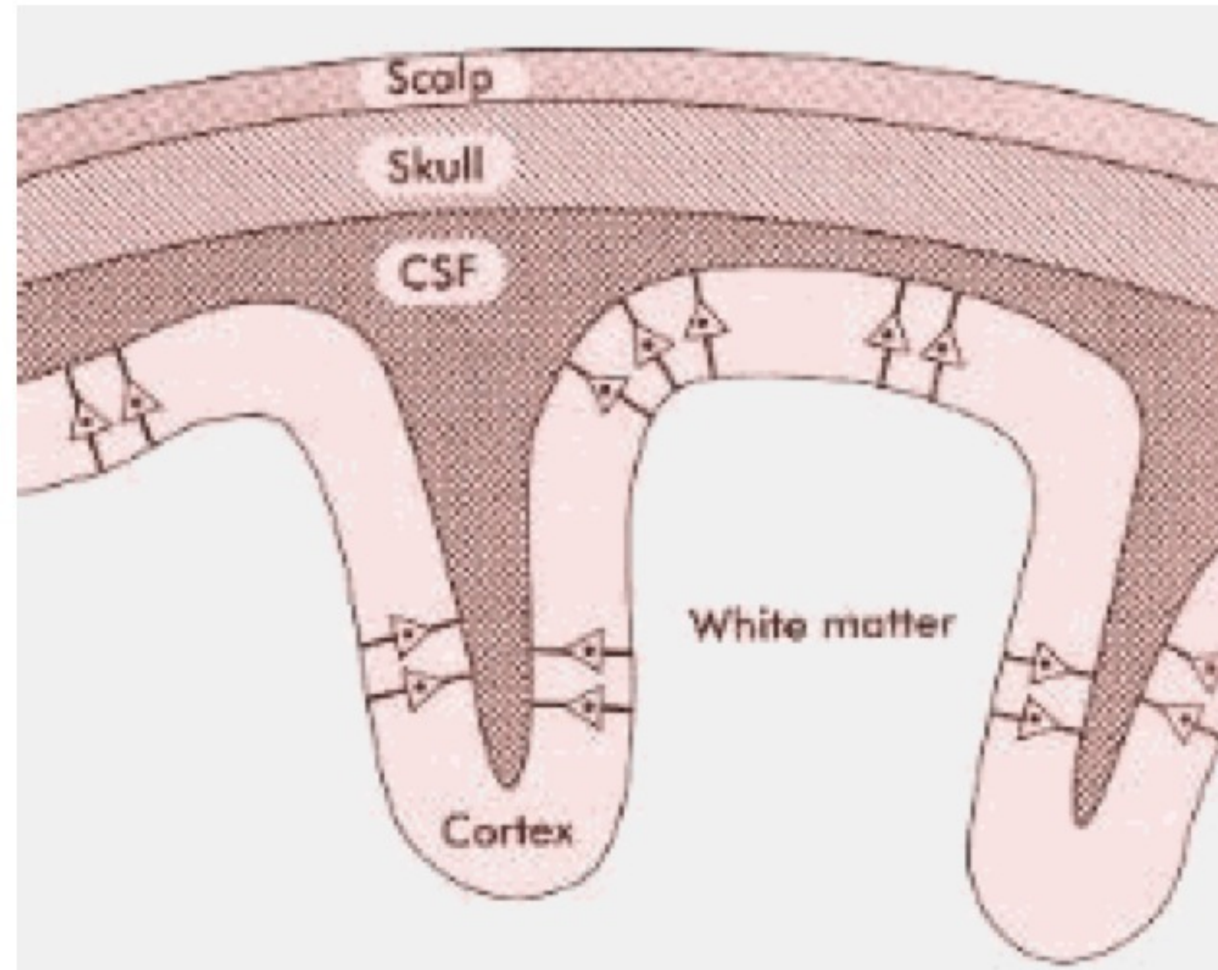
Figure 4.24 EEG profiles obtained during various states of consciousness. From Kolb and Whishaw (1986) after Penfield and Jasper (1954).

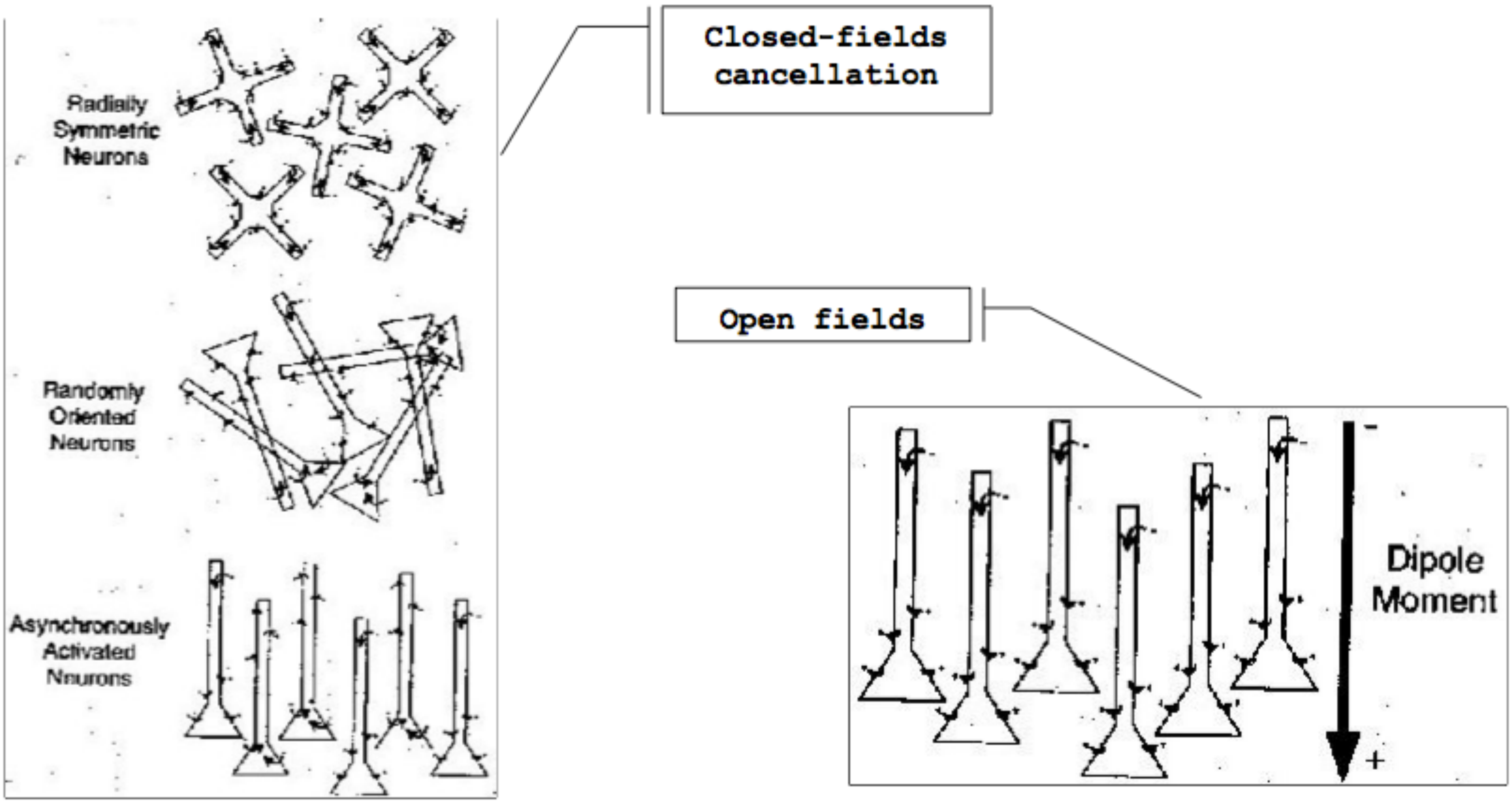
WHERE DO ELECTROMAGNETIC SIGNALS COME FROM?

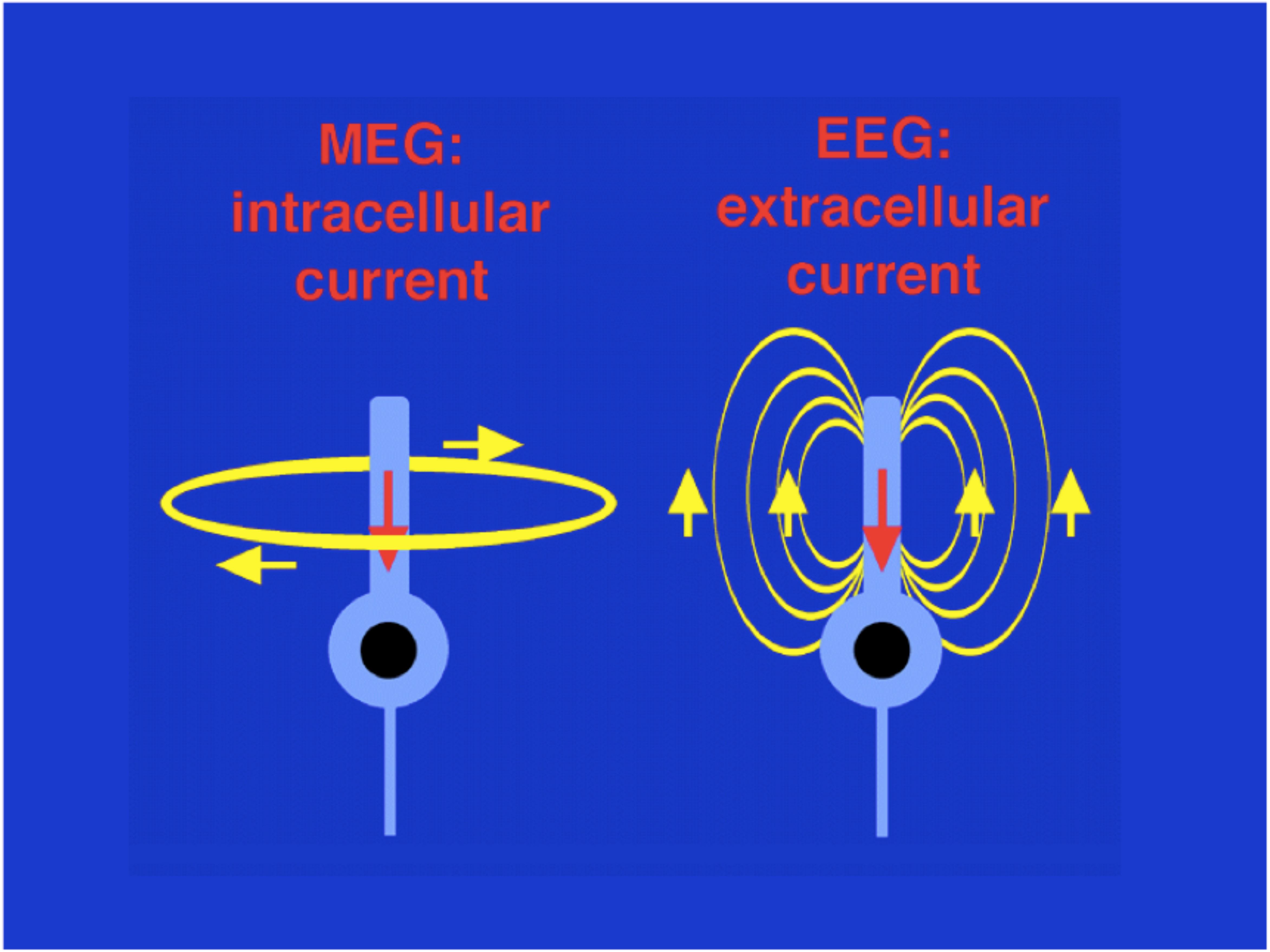



Primary Current
Secondary Currents


Axon from thalamus
Axon from contralateral cortex

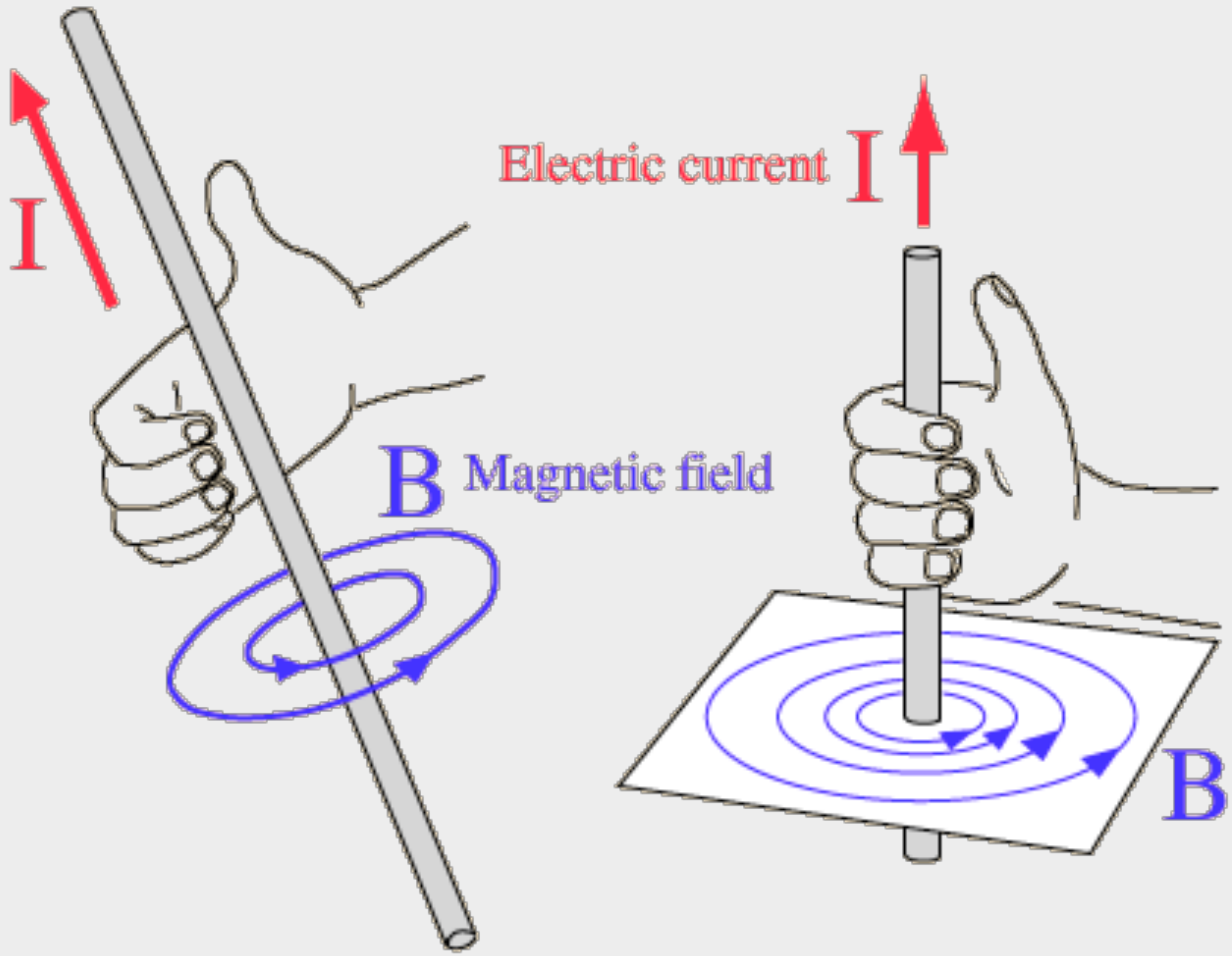

Scalp
Skull
CSF
White matter
Cortex


Radially Symmetric Neurons
Randomly Oriented Neurons
Asynchronously Activated Nnurons
Closed-fields cancellation
Open fields
-
Dipole Moment
+


MEG:
intracellular
current
EEG:
extracellular
current

An electric current creates a magnetic field around it.

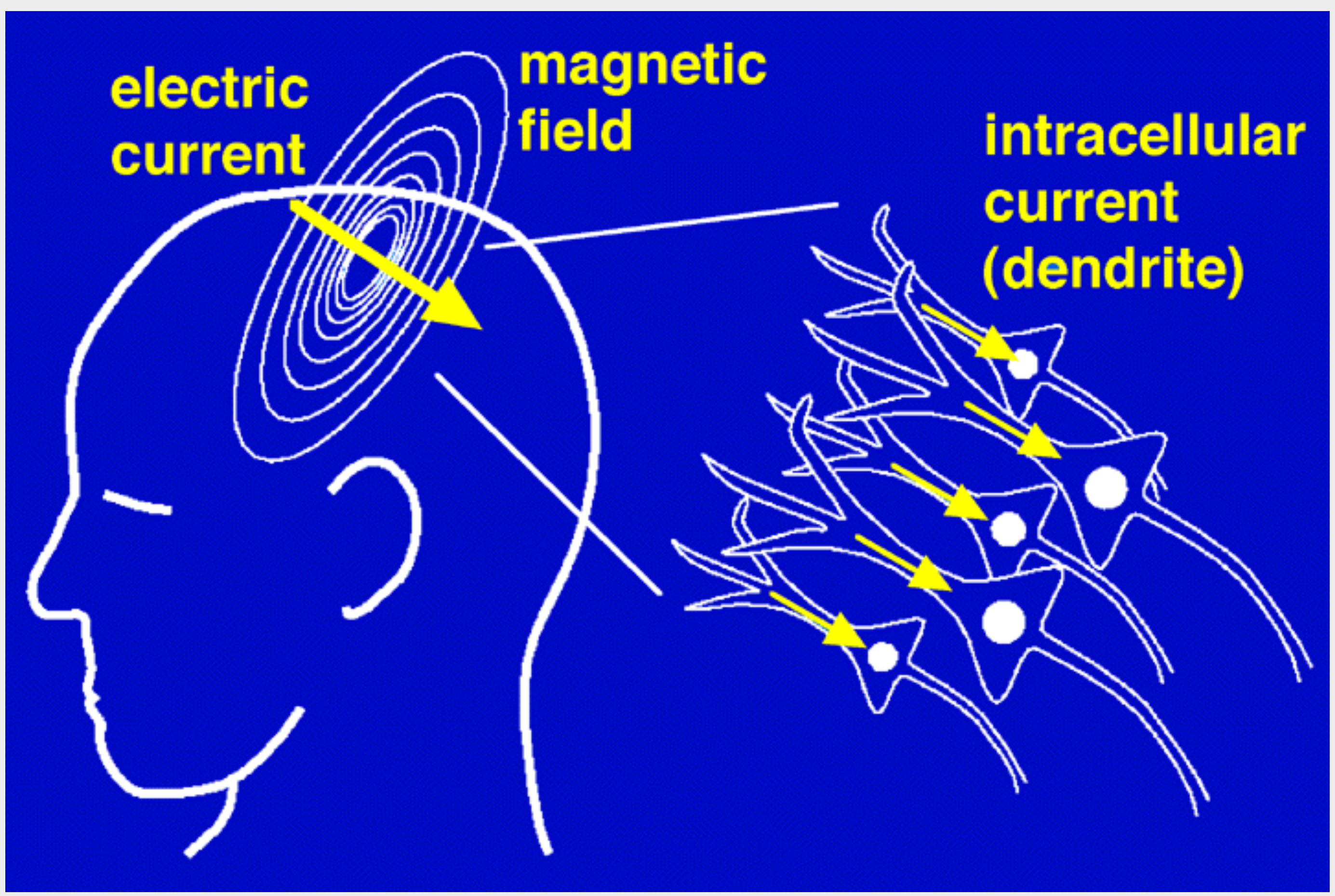
The right-hand rule:
When the thumb of the right hand is pointing in the direction of the current, the fingers of the right hand curl in the direction of the magnetic field



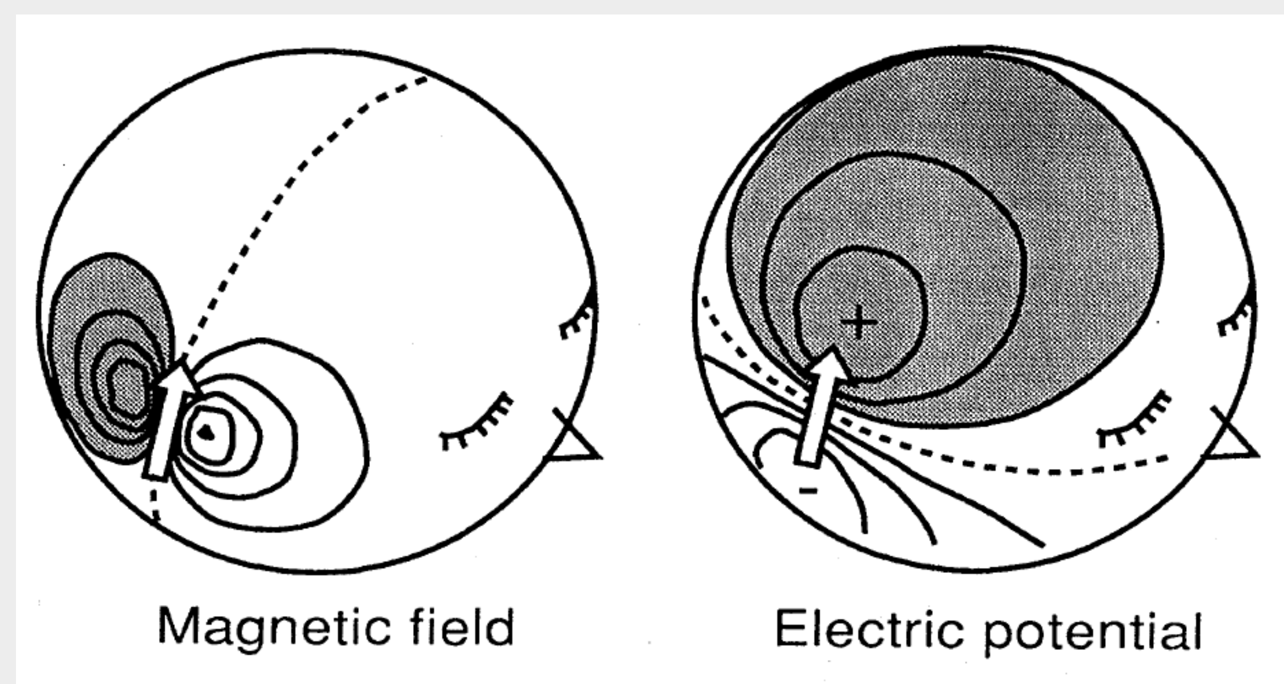
ELECTROMAGNETISM

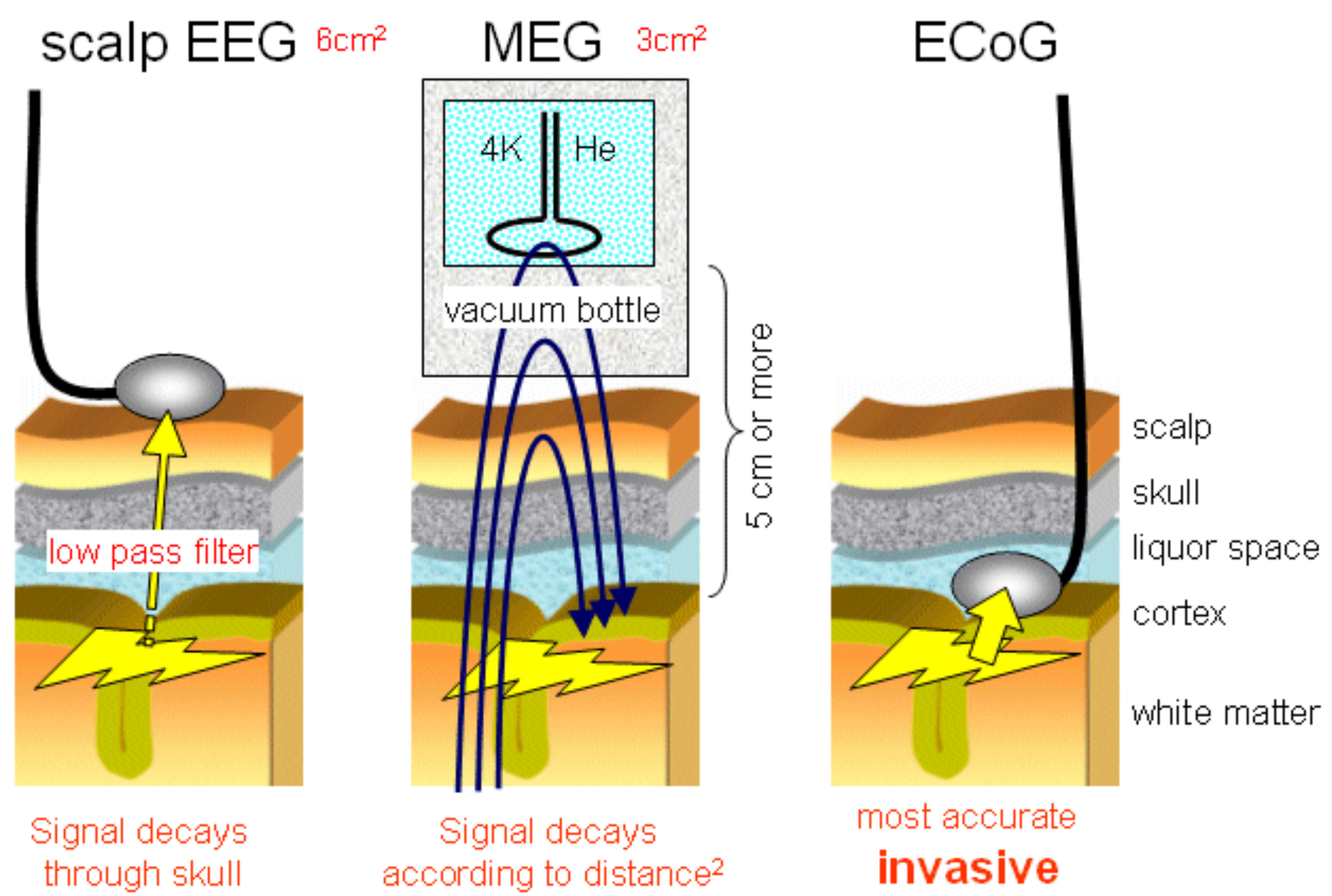
EEG (electroencephalography): electric potentials

MEG (magnetoencephalography): magnetic fields



MEG EEG

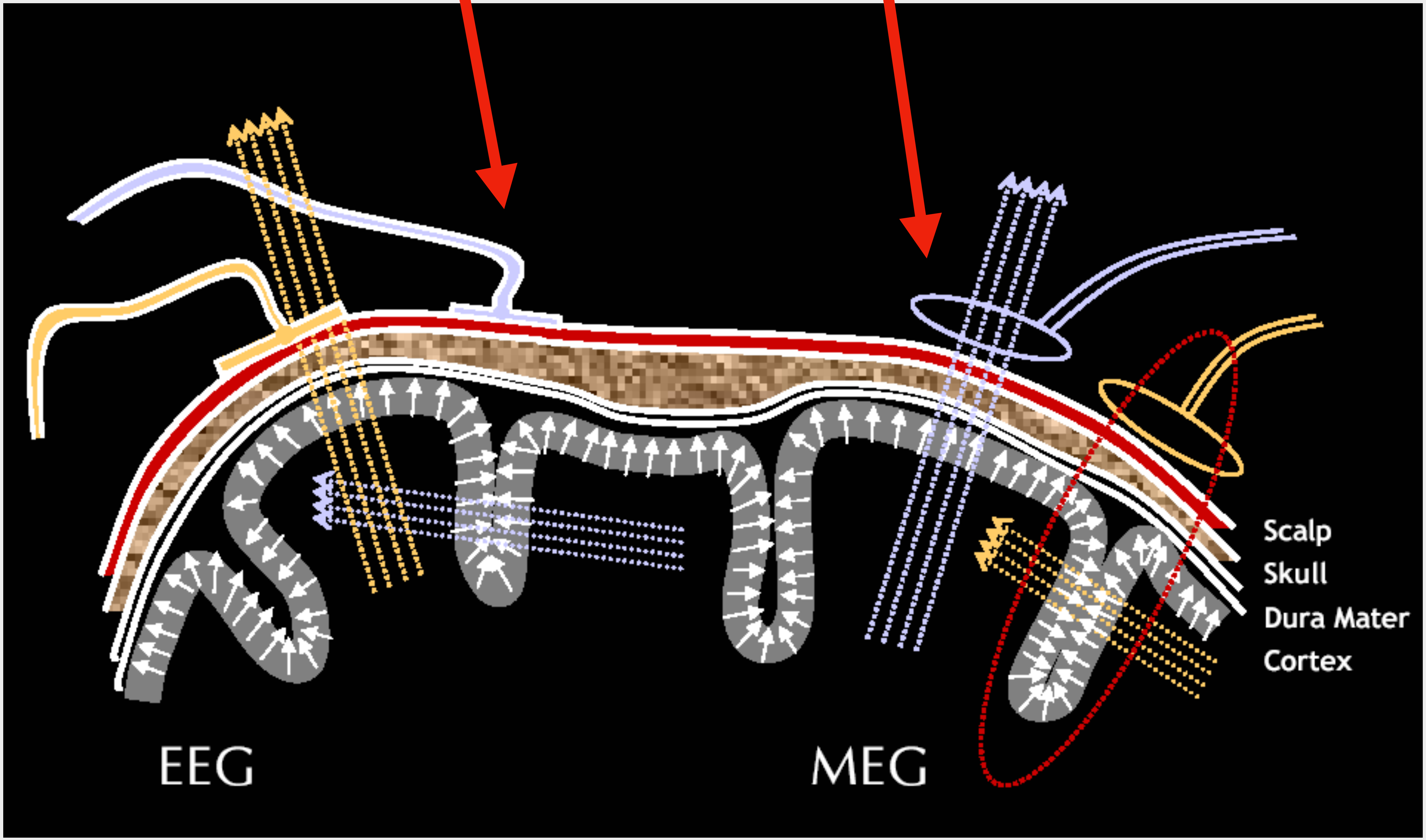




scalp EEG 6cm2
MEG 3cm2
ECoG
4K He
vacuum bottle
5 cm or more
low pass filter
scalp
skull
liquor space
cortex
white matter
Signal decays through skull
Signal decays according to distance2
most accurate
invasive

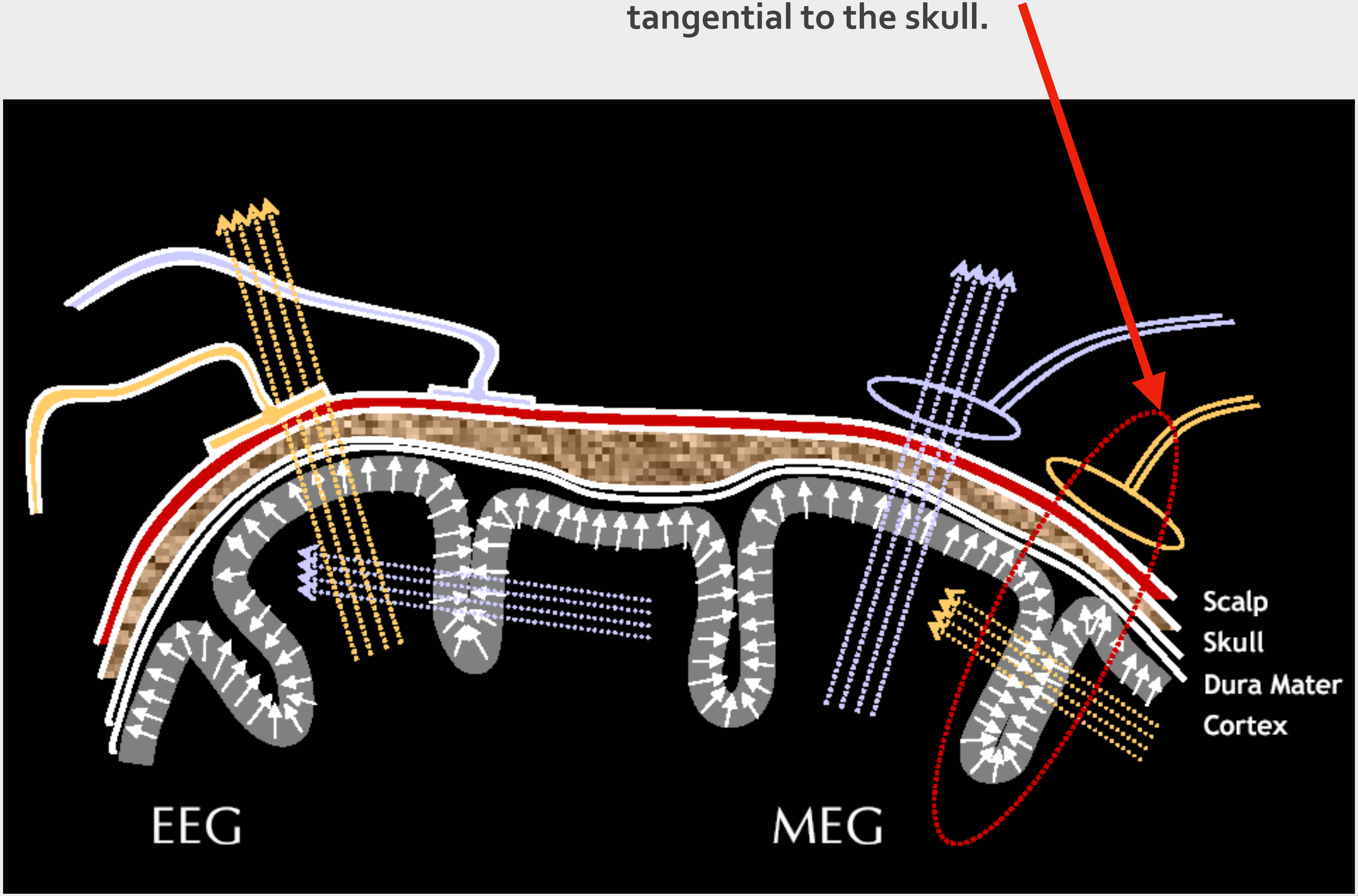
EEG electrodes on the scalp

MEG sensors outside the head, in a tank containing liquid helium to enhance superconductivity



Source: http://www.allgpsy.unizh.ch/graduate/mat/180102/Lecture1.pdf

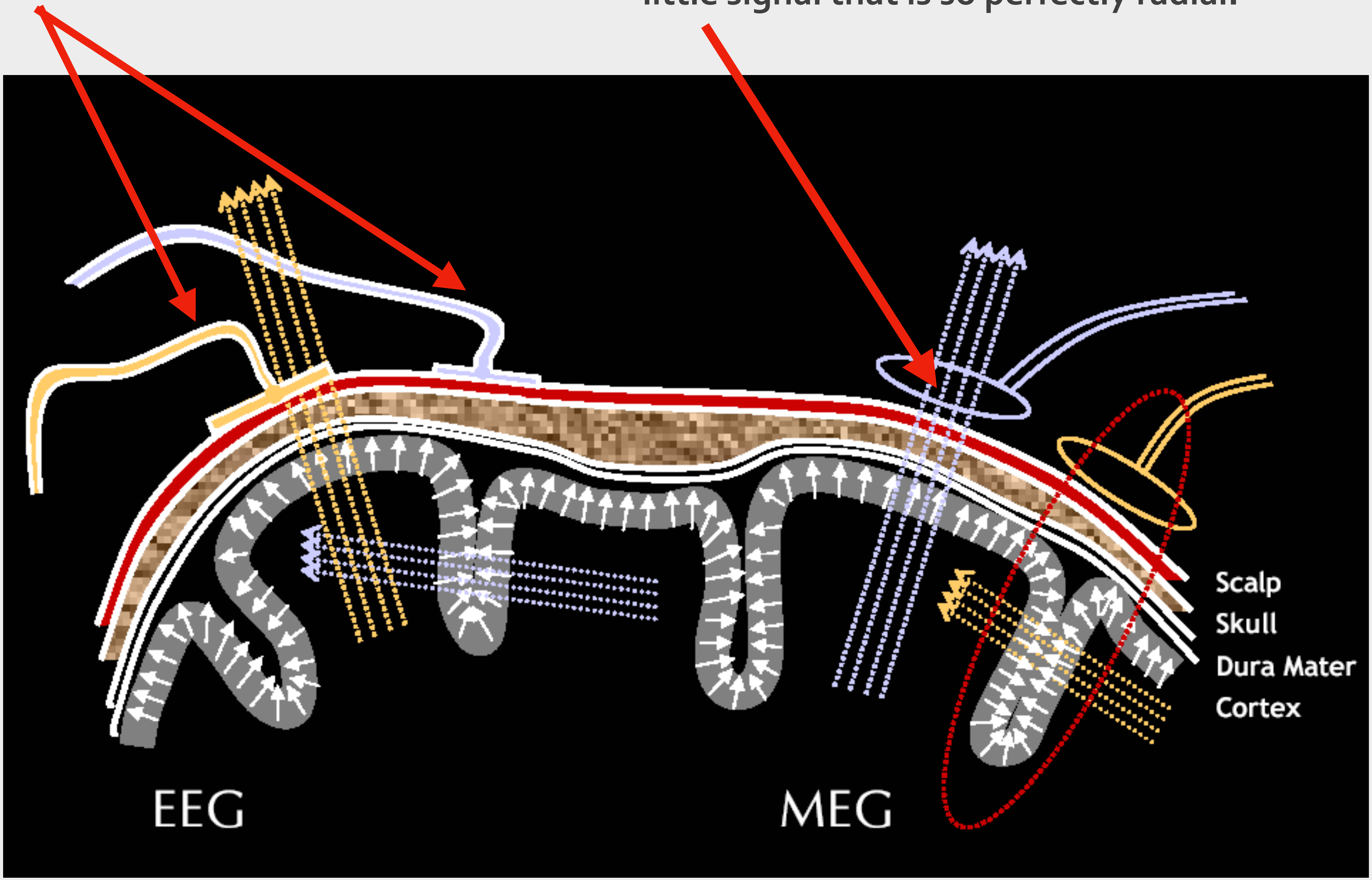
MEG signal is dominated by currents oriented tangential to the skull.



Source: http://www.allgpsy.unizh.ch/graduate/mat/180102/Lecture1.pdf

EEG picks up tangentially and radially oriented currents equally.

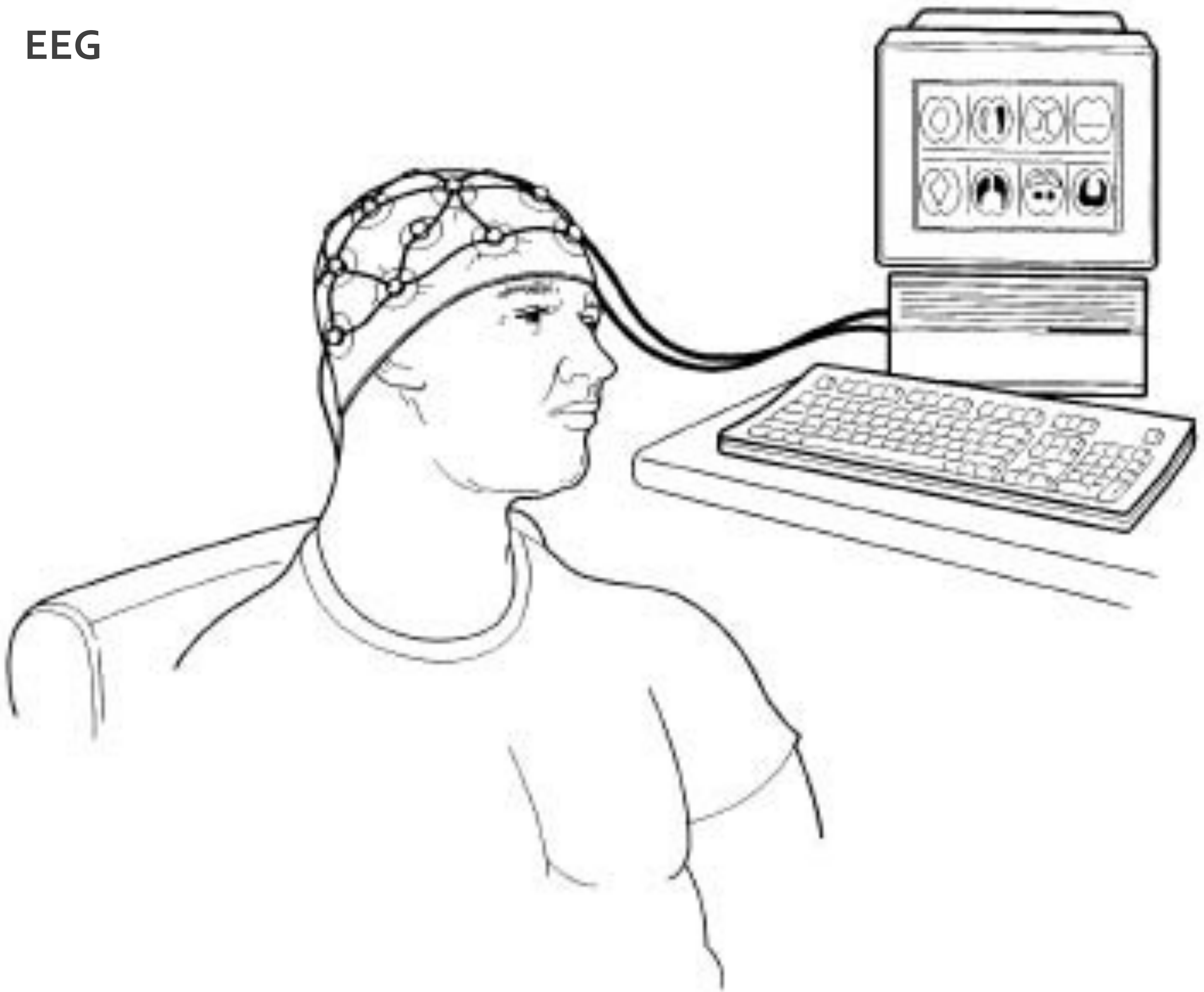
Currents oriented perfectly radial to the skull are missed in MEG. But there is very little signal that is so perfectly radial.

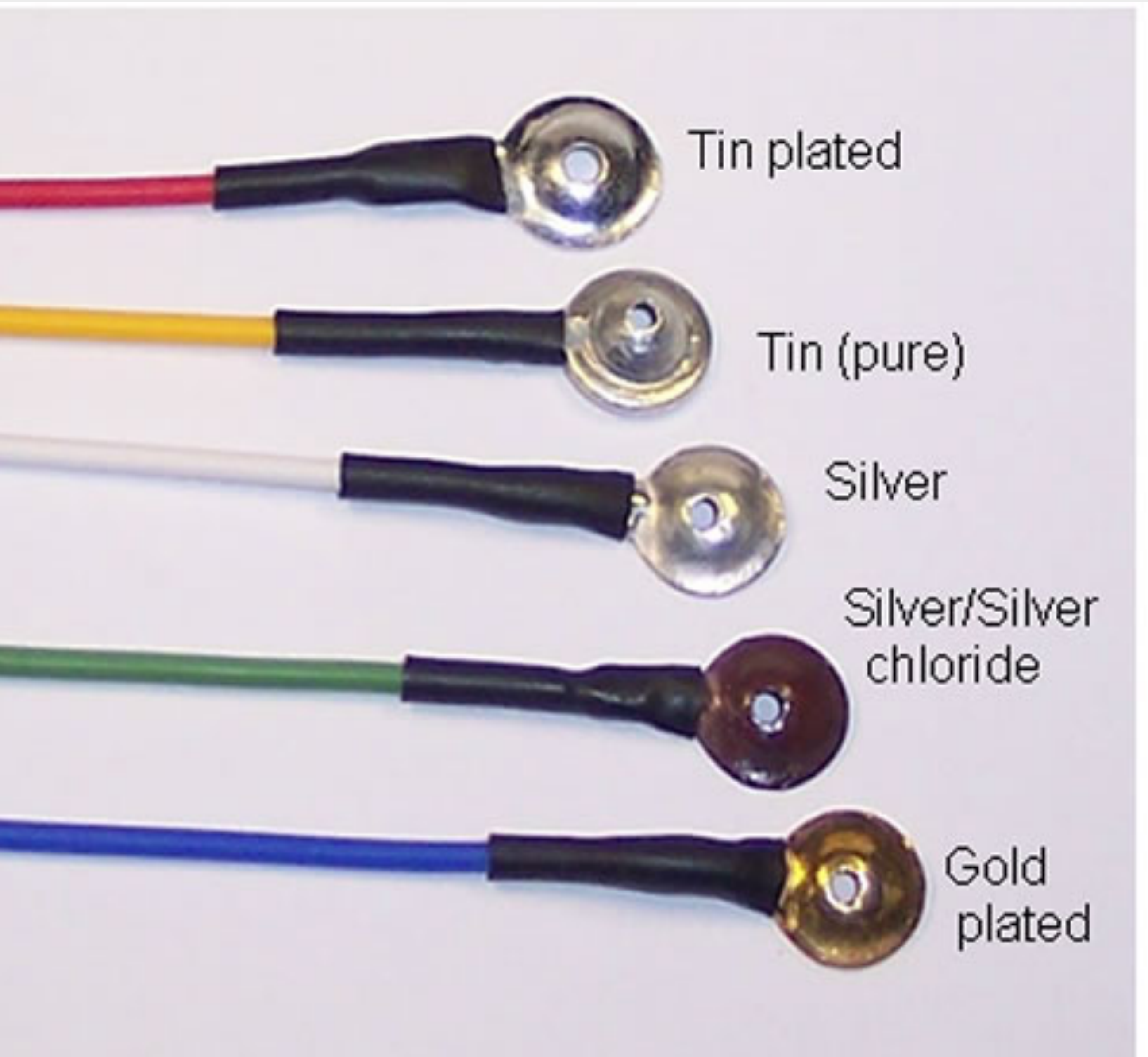


Source: http://www.allgpsy.unizh.ch/graduate/mat/180102/Lecture1.pdf

THE EEG

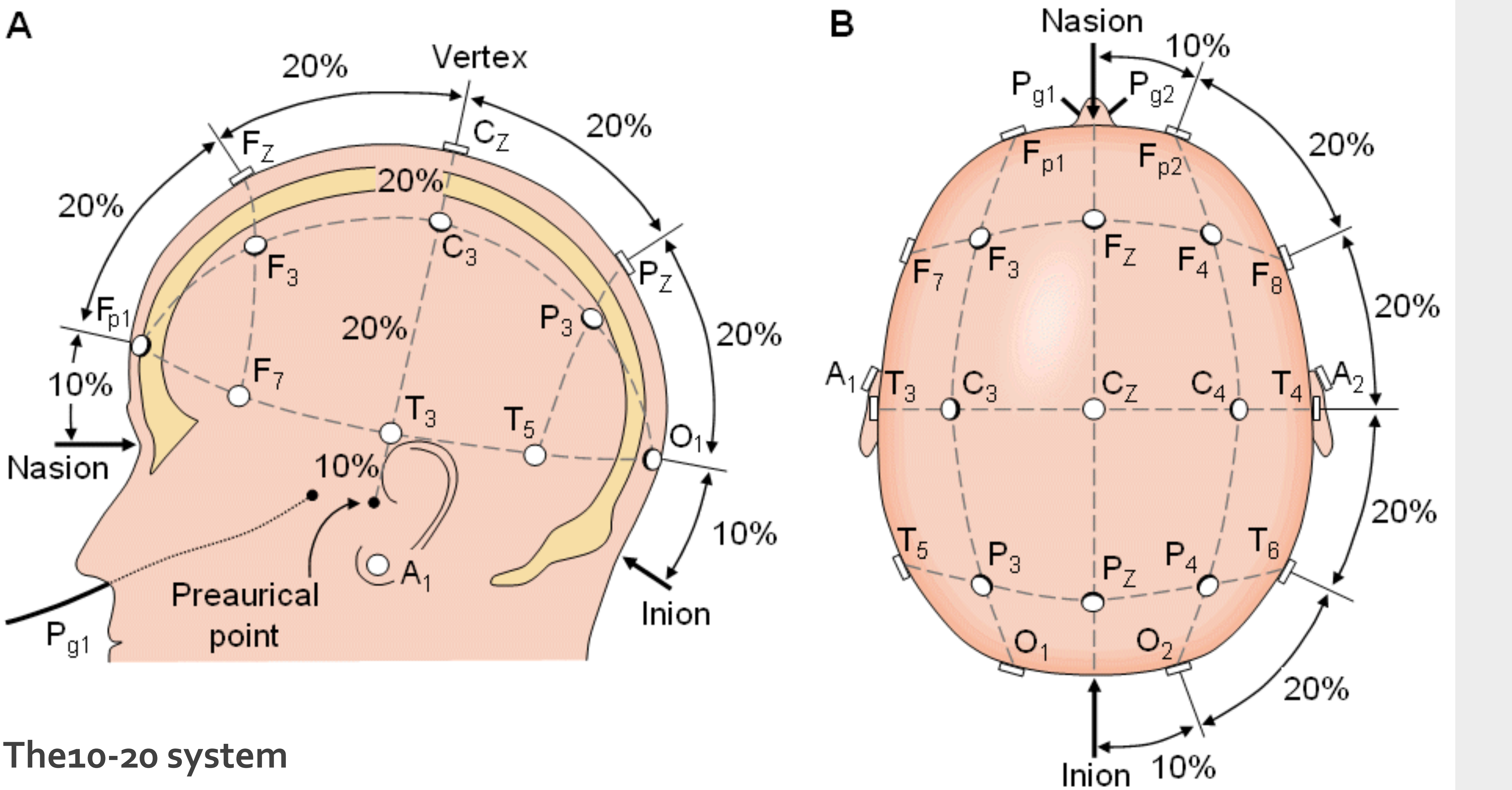
EEG



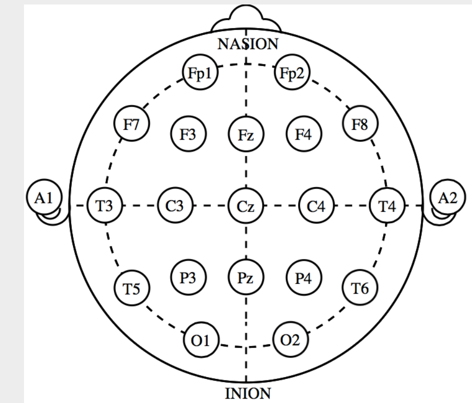


Metal electrodes are attached to an elastic cap

Conducting gel is apply to connect electrodes to the scalp

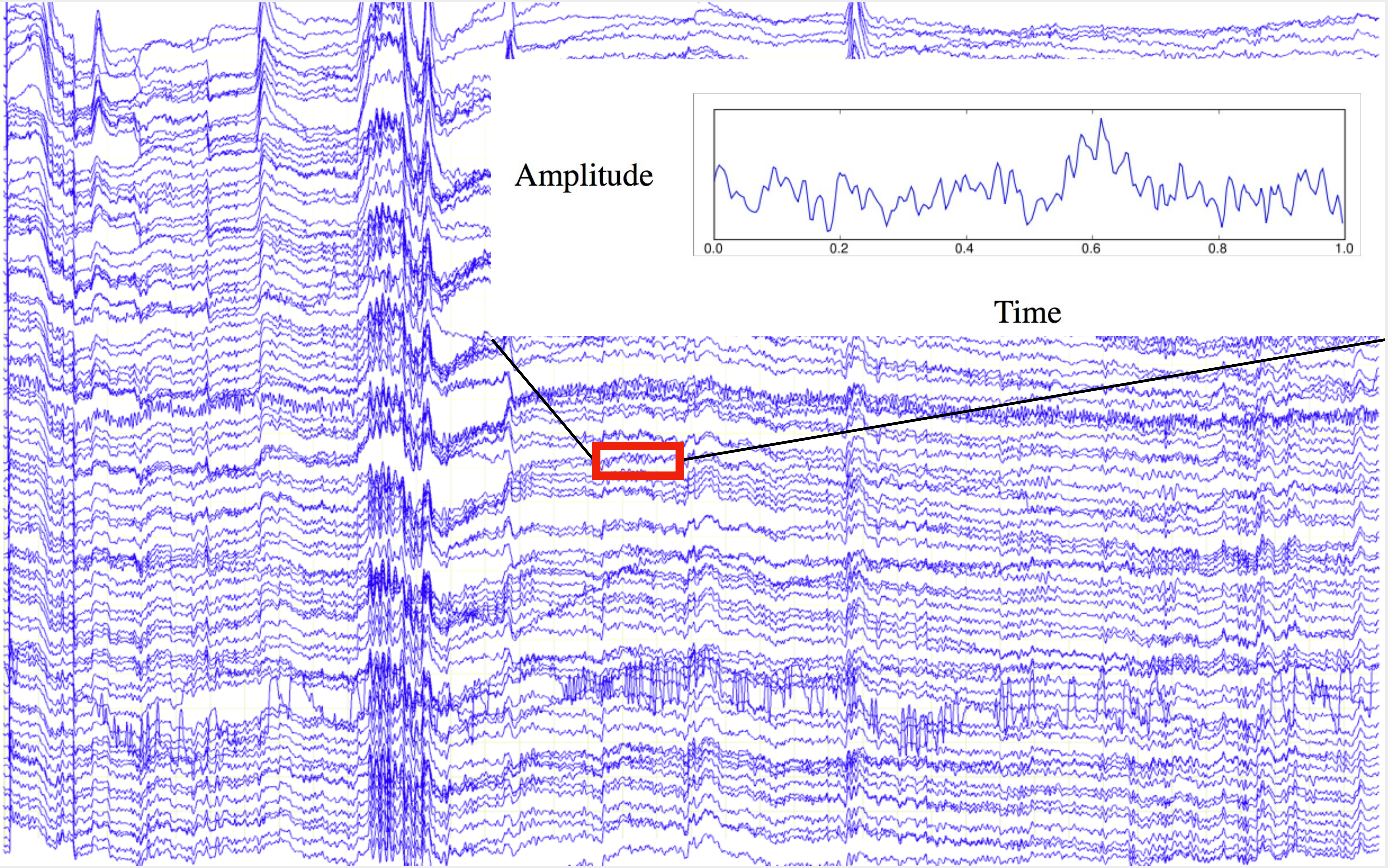


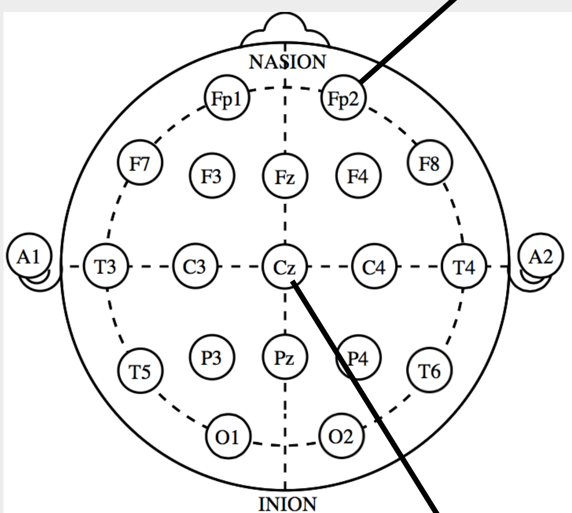
The10-20 system



Amplifier

A/D converter

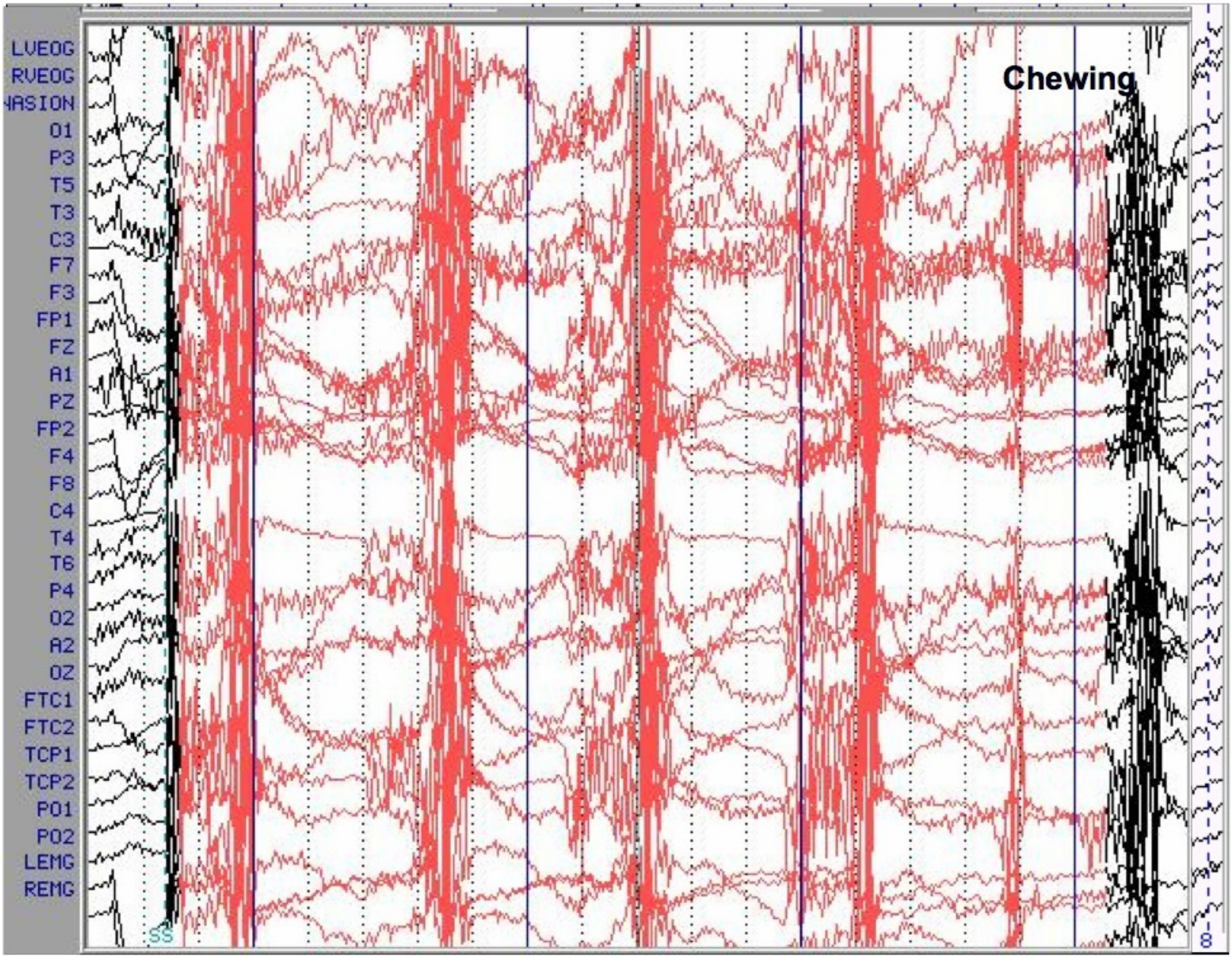

Amplitude
0.0
0.2
0.4
0.6
0.8
1.0
Time



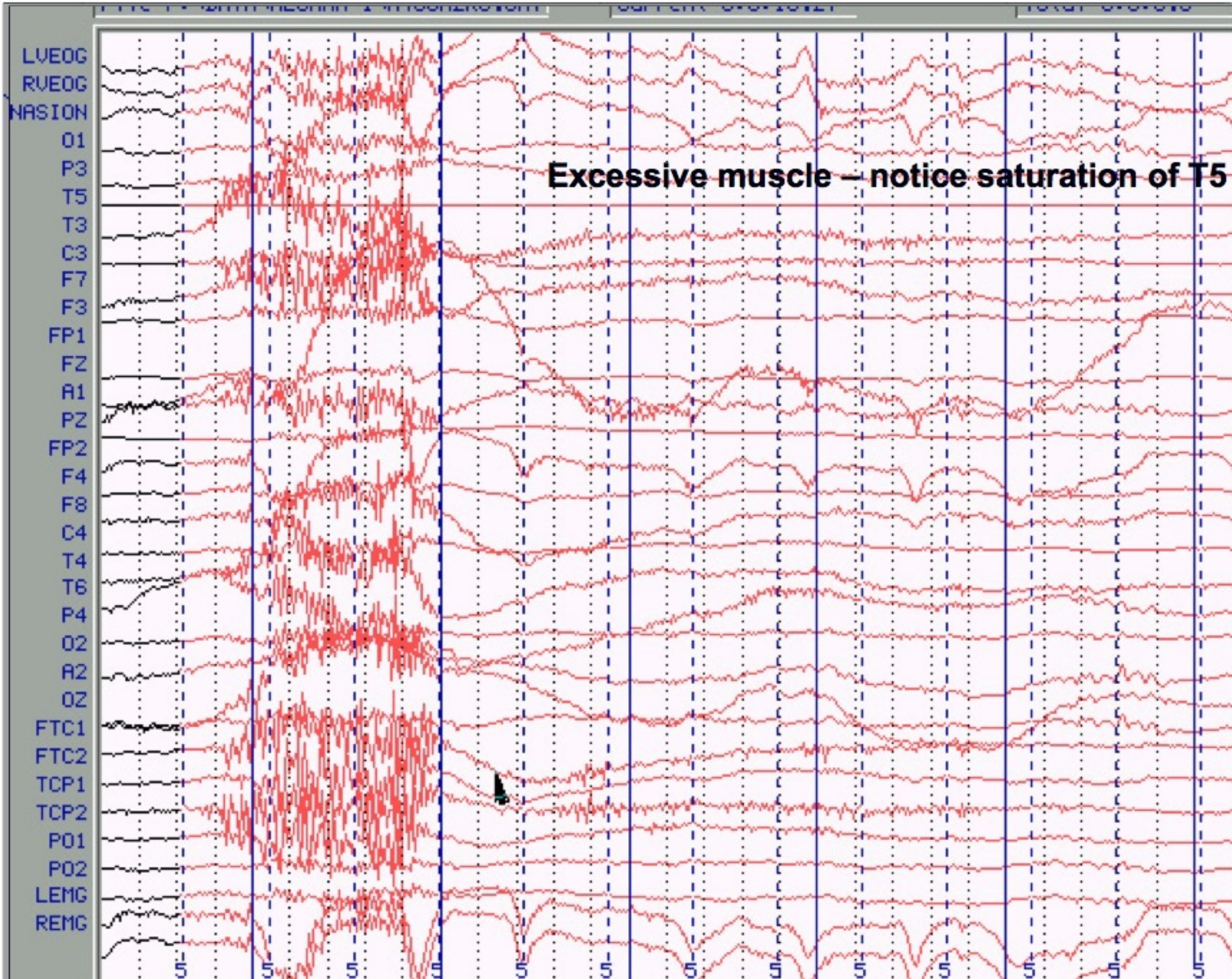
Electroencephalogram (EEG)

ARTIFACTS IN THE EEG SIGNALS

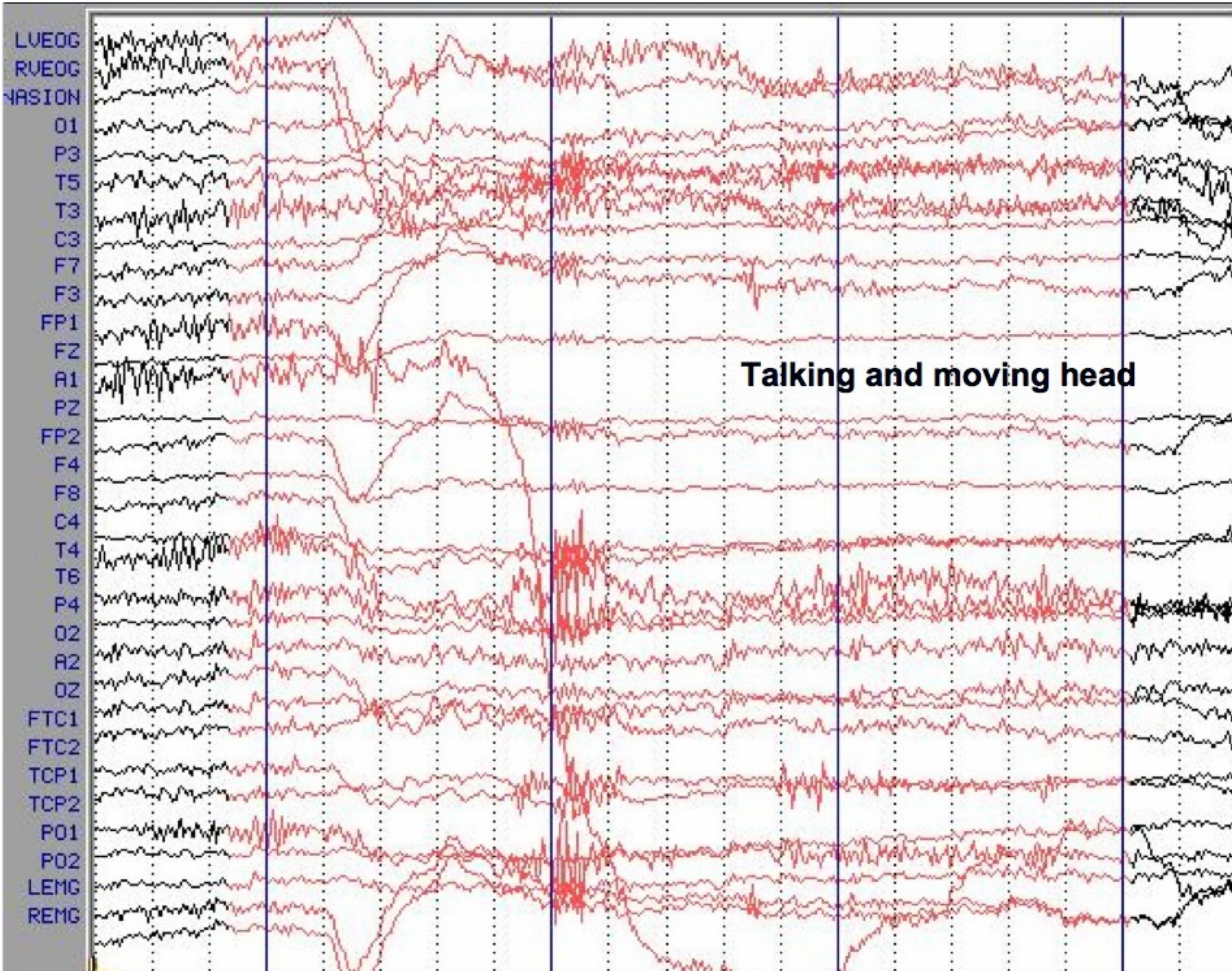

Movement in reference lead
LVEOG
RVEOG
NASION
O1
P3
T5
T3
C3
F7
F3
FP1
FZ
A1
PZ
FP2
F4
F8
C4
T4
T6
P4
O2
A2
OZ
FTC1
FTC2
TCP1
TCP2
PO1
PO2
LEMG
REMG



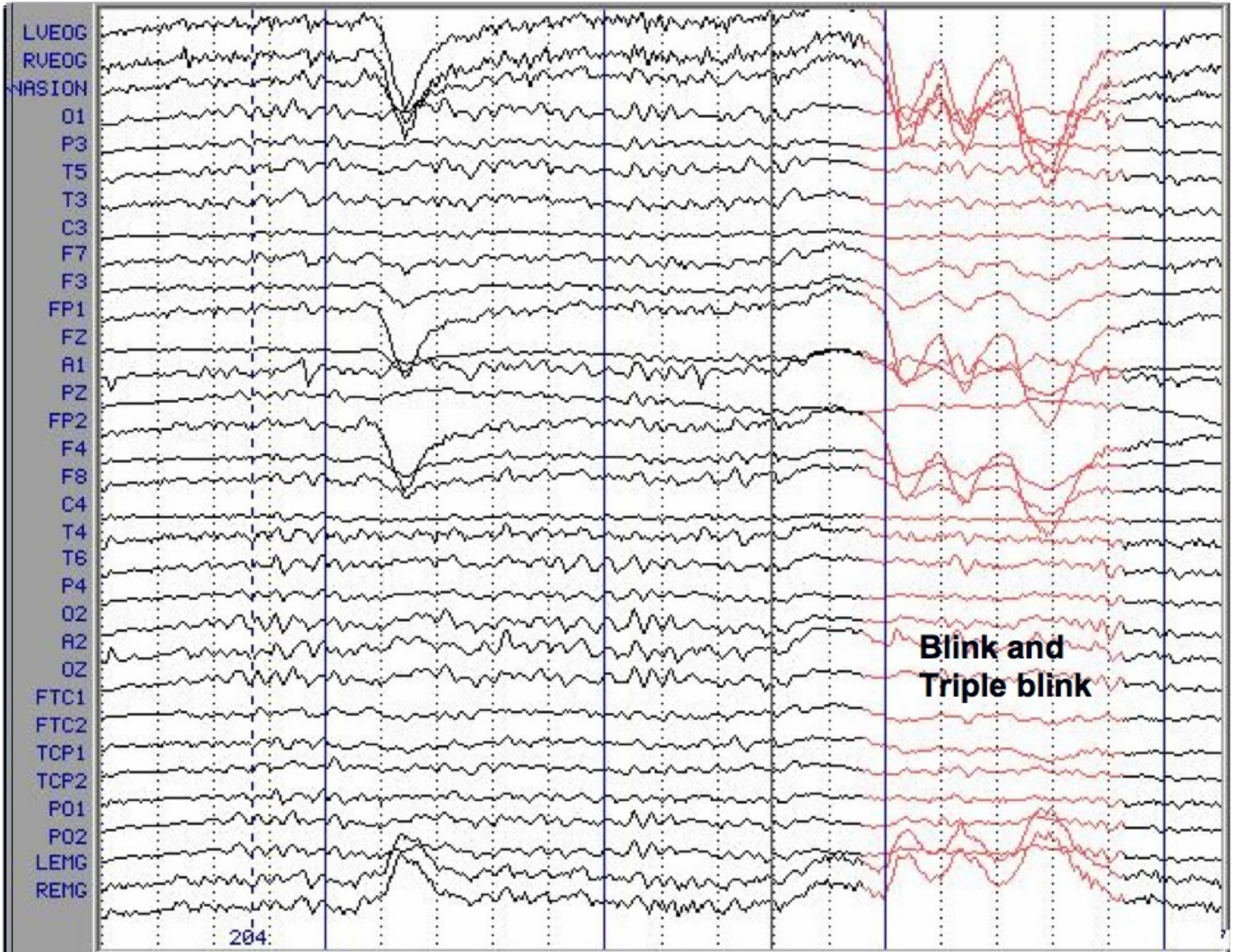
Chewing
LVEOG
RVEOG
NASION
O1
P3
T5
T3
C3
F7
F3
FP1
FZ
A1
PZ
FP2
F4
F8
C4
T4
T6
P4
O2
A2
OZ
FTC1
FTC2
TCP1
TCP2
PO1
PO2
LEMG
REMG
SS
8

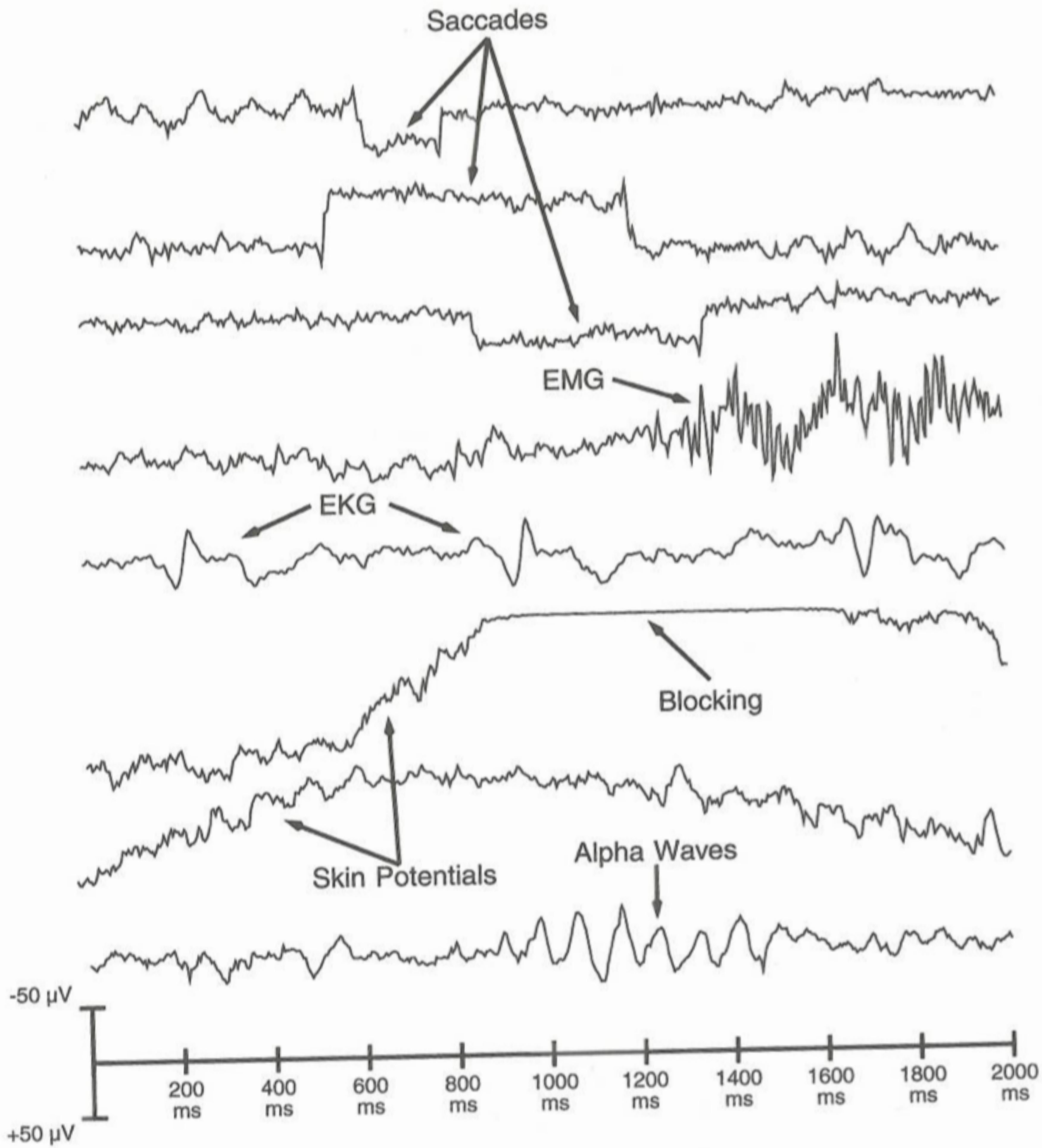


LVEOG
RVEOG
NASION
O1
P3
T5
T3
C3
F7
F3
FP1
FZ
A1
PZ
FP2
F4
F8
C4
T4
T6
P4
O2
A2
OZ
FTC1
FTC2
TCP1
TCP2
PO1
PO2
LEMG
REMG
Excessive muscle – notice saturation of T5



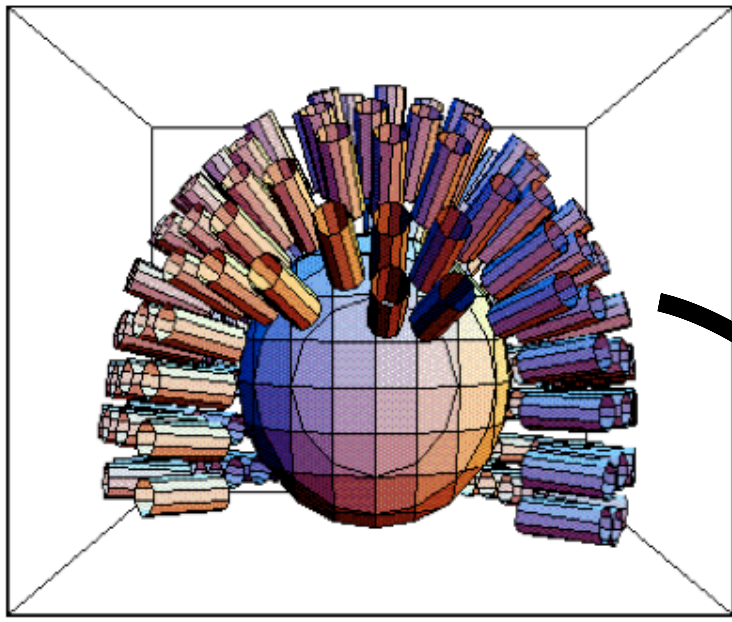
LVEOG
RVEOG
NASION
O1
P3
T5
T3
C3
F7
F3
FP1
FZ
A1
PZ
FP2
F4
F8
C4
T4
T6
P4
O2
A2
OZ
FTC1
FTC2
TCP1
TCP2
PO1
PO2
LEMG
REMG
Talking and moving head


LVEOG
RVEOG
NASION
O1
P3
T5
T3
C3
F7
F3
FP1
FZ
A1
PZ
FP2
F4
F8
C4
T4
T6
P4
O2
A2
OZ
FTC1
FTC2
TCP1
TCP2
PO1
PO2
LEMG
REMG
204
Blink and
Triple blink

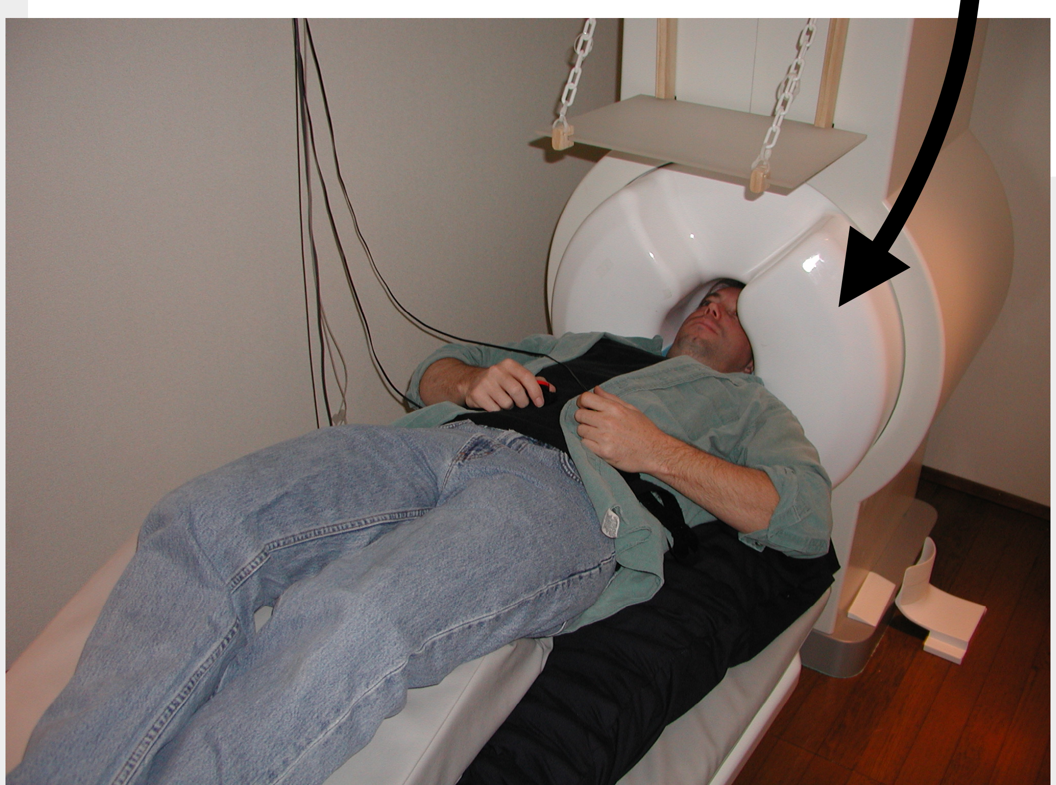

Saccades
EMG
EKG
Blocking
Skin Potentials
Alpha Waves
-50 µV
+50 µV
200 ms
400 ms
600 ms
800 ms
1000 ms
1200 ms
1400 ms
1600 ms
1800 ms
2000 ms

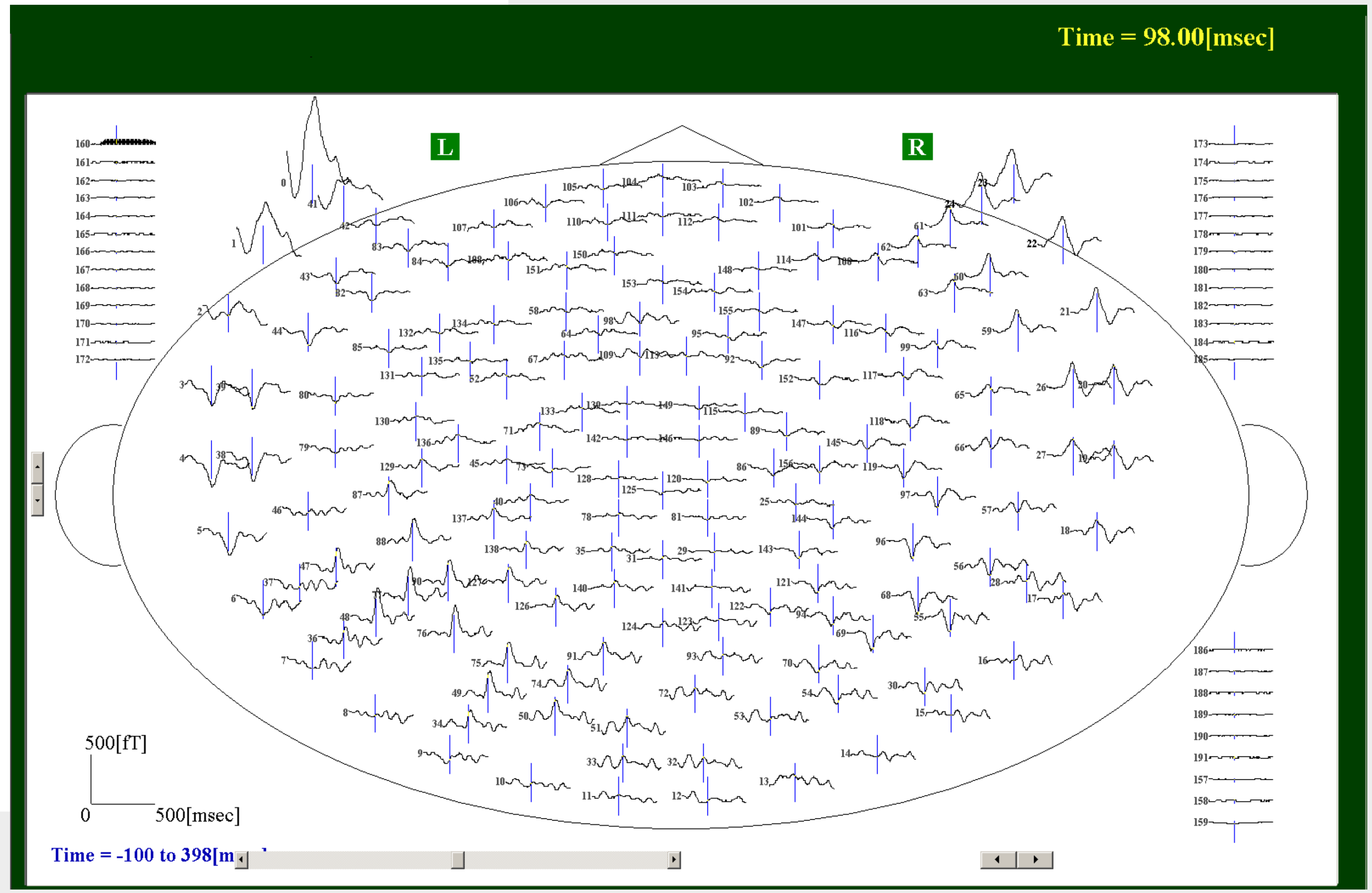
THE MEG

MEG

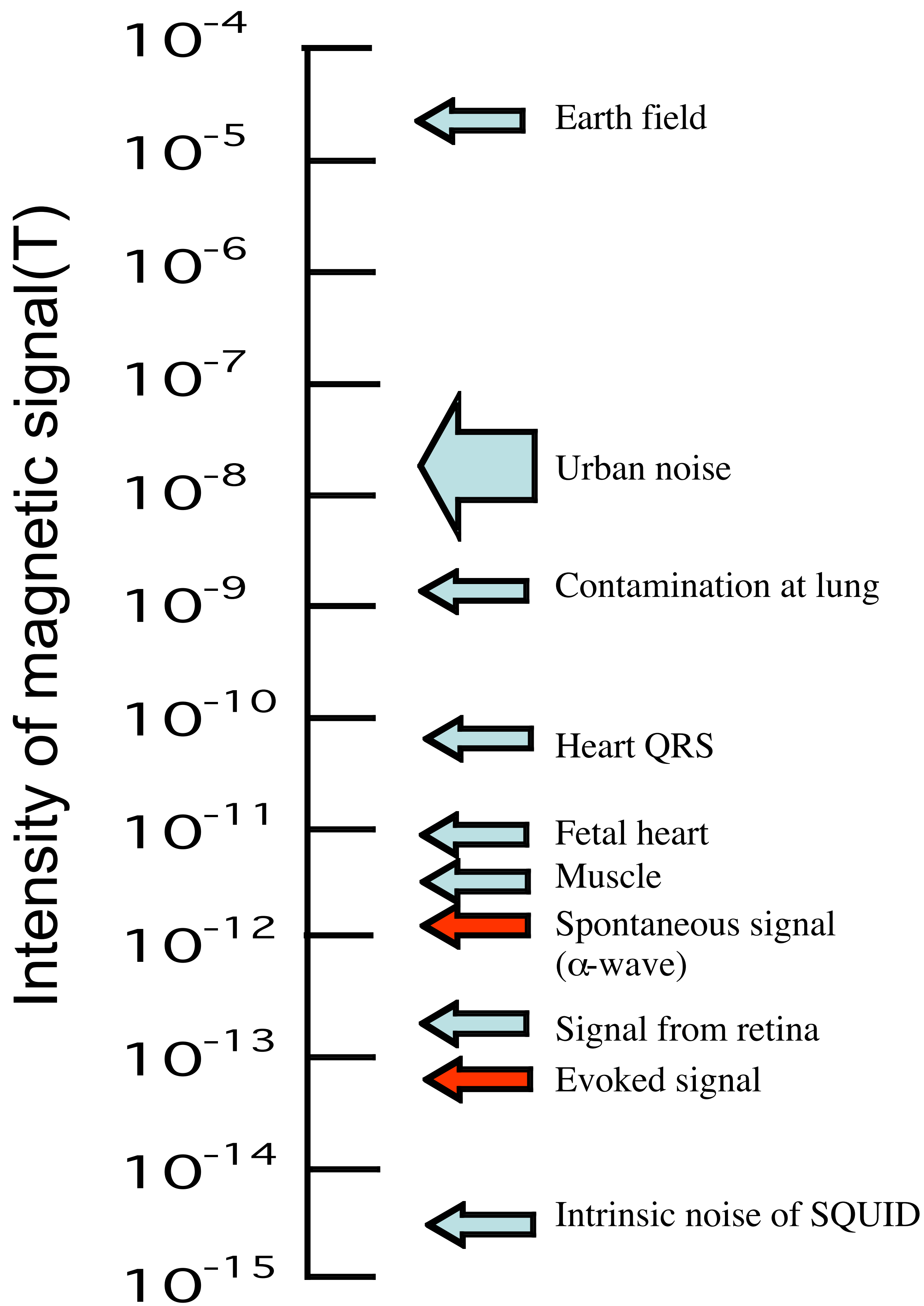


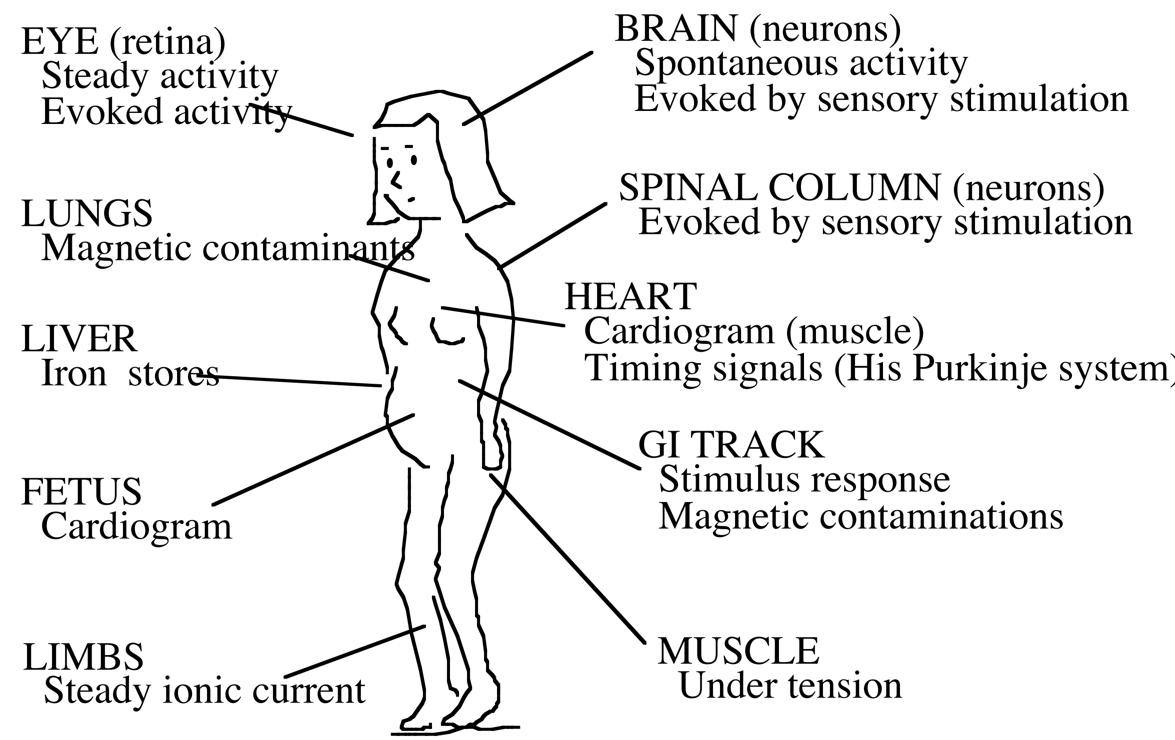
front view





How small are the electromagnetic signals from the brain?



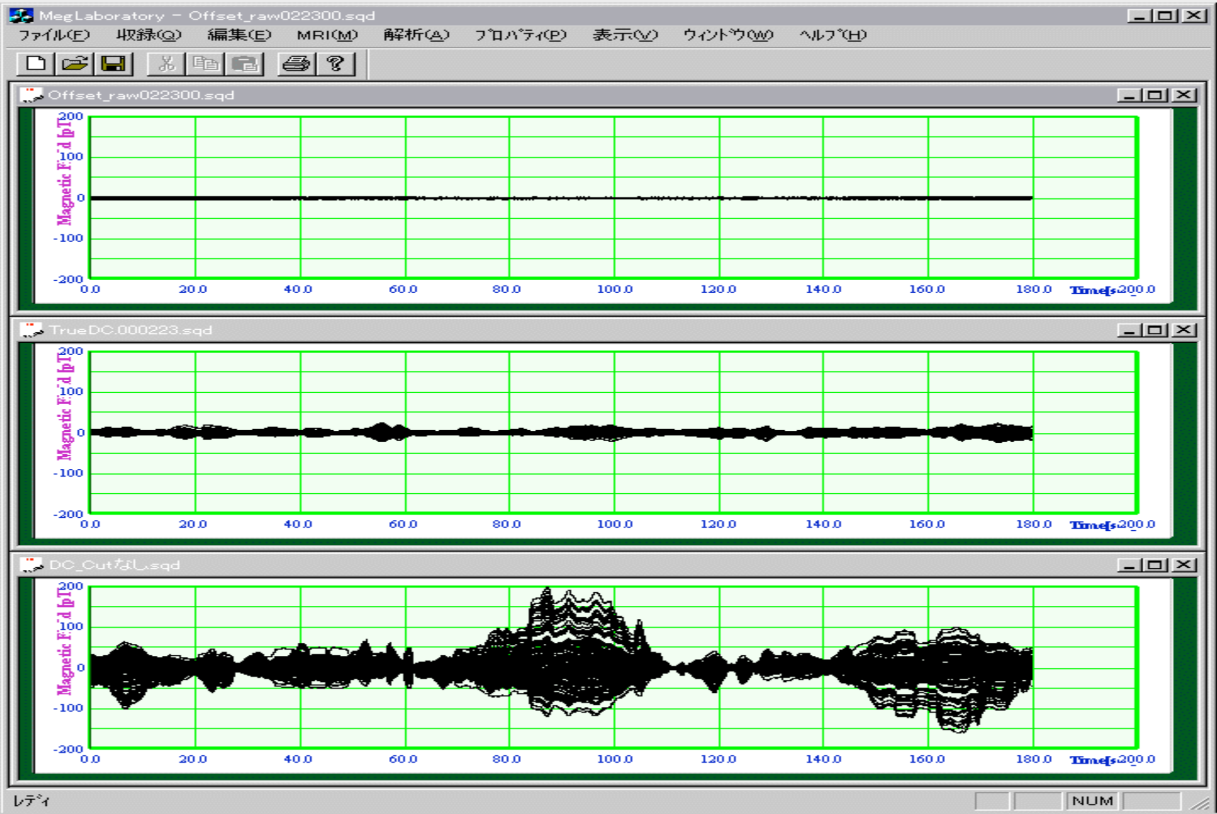


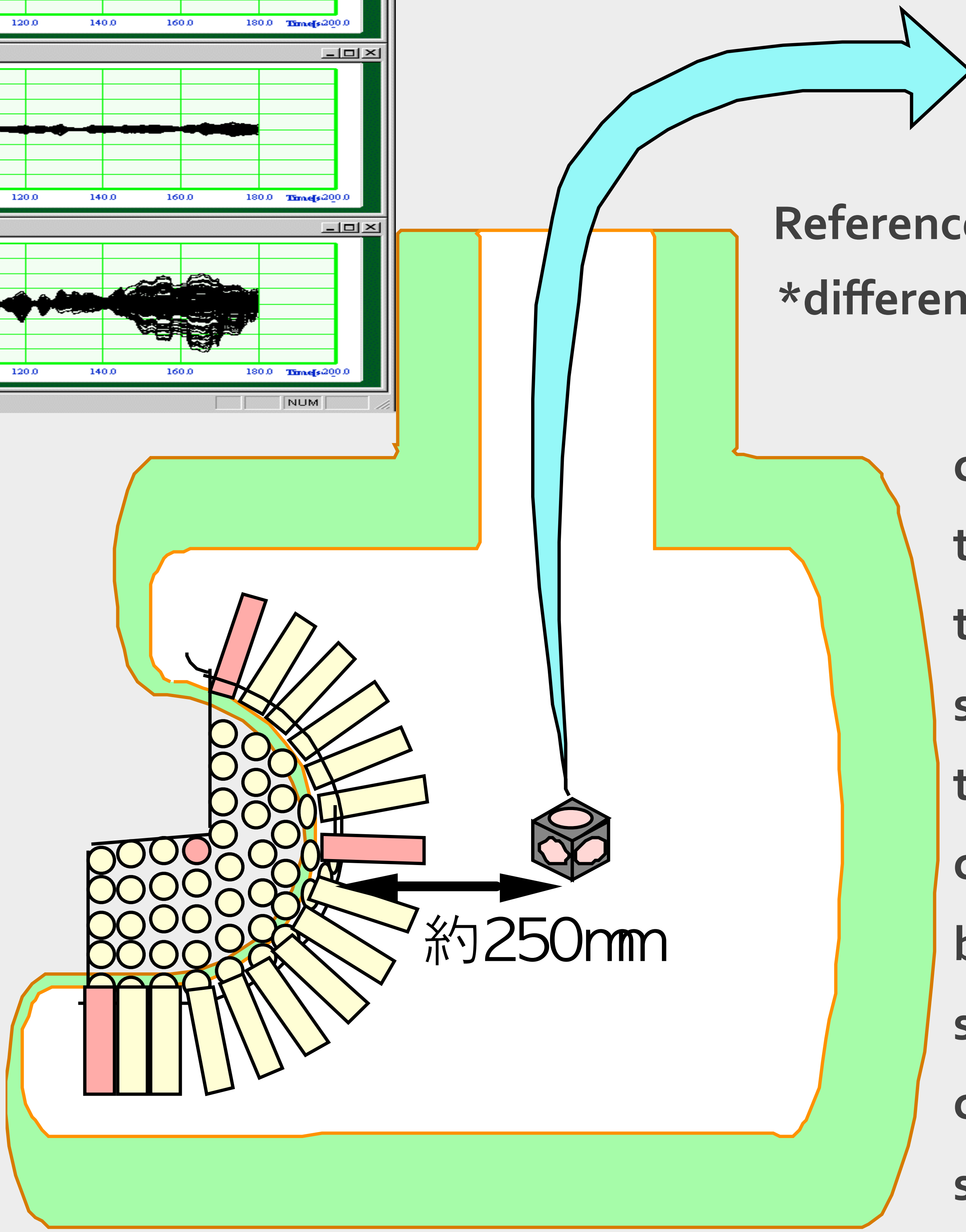
Biomagnetism

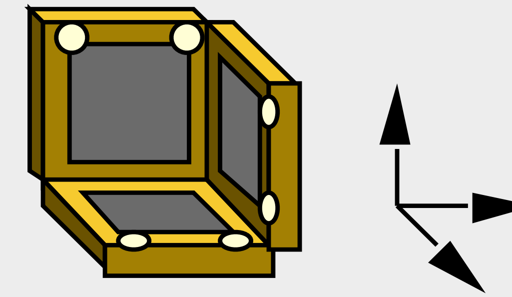
requires high sensitivity sensors
(low noise amplification and high gain)

Noise reduction: Magnetically Shielded Room - MSR

Besides the gradiometers and the reference sensors.



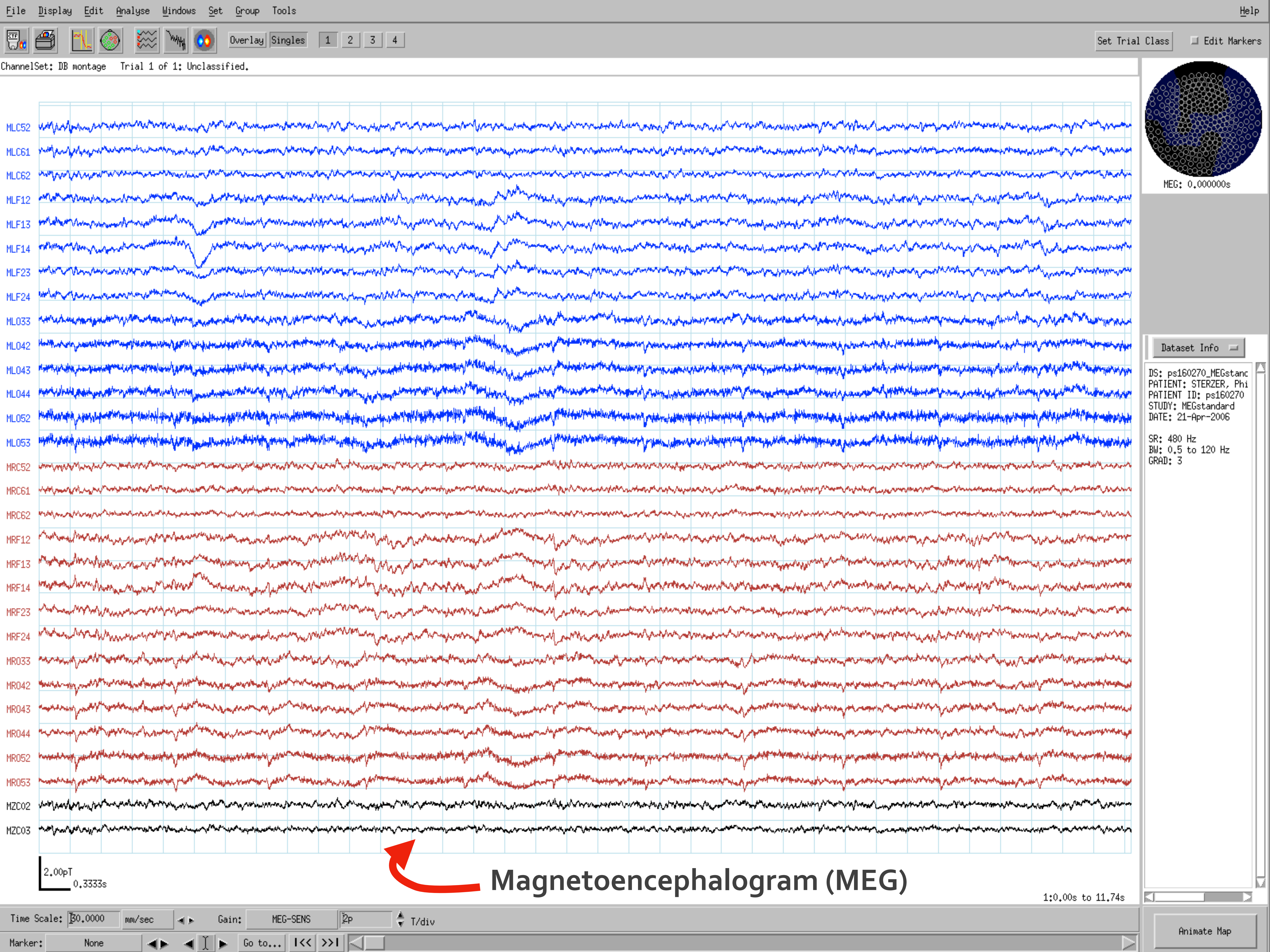




Reference sensors*

*different from EEG reference

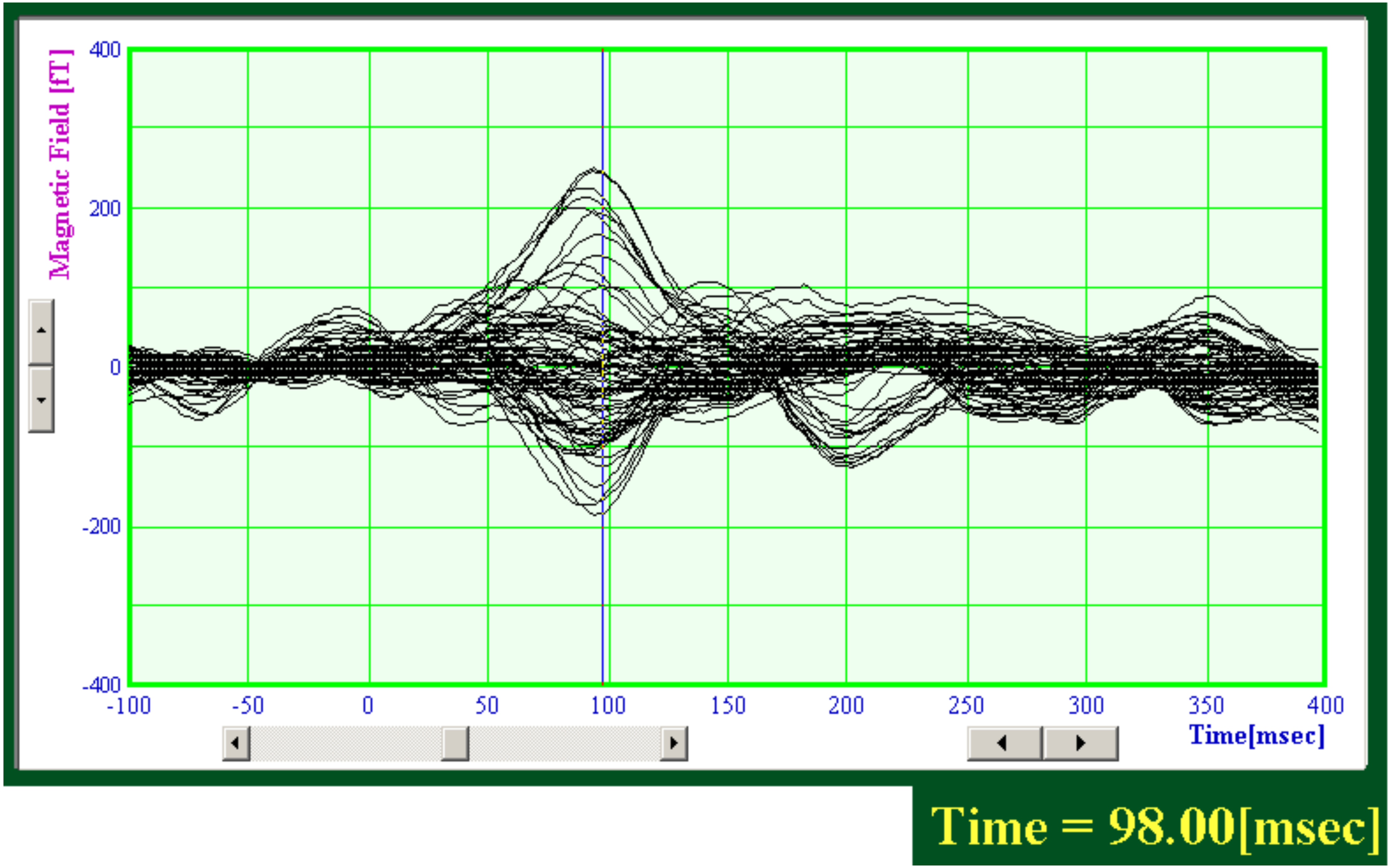
close enough to the other sensors to pick up the same noise that they are contaminated by, but far enough such that they capture no brain signals



File Display Edit Analyse Windows Set Group Tools
Help
Overlay Singles 1 2 3 4
Set Trial Class
Edit Markers
ChannelSet: DB montage Trial 1 of 1: Unclassified.
MEG: 0.000000s
MLC52
MLC61
MLC62
MLF12
MLF13
MLF14
MLF23
MLF24
MLO33
MLO42
MLO43
MLO44
MLO52
MLO53
MRC52
MRC61
MRC62
MRF12
MRF13
MRF14
MRF23
MRF24
MRO33
MRO42
MRO43
MRO44
MRO52
MRO53
MZC02
MZC03
Dataset Info
DS: ps160270_MEGstanc
PATIENT: STERZER, Phi
PATIENT ID: ps160270
STUDY: MEGstandard
DATE: 21-Apr-2006
SR: 480 Hz
BW: 0.5 to 120 Hz
GRAD: 3
2.00pT
0.3333s
Magnetoencephalogram (MEG)
1:0.00s to 11.74s
Time Scale: 30.0000 mm/sec
Gain: MEG-SENS 2p T/div
Marker: None
Go to...
Animate Map

Exemple of a butterfly plot: All sensors over right termporal lobe superposed

Peak at 98 ms post-onset of acoustic stimulus presentation



EVENT RELATED POTENTIALS (ERP): EEG

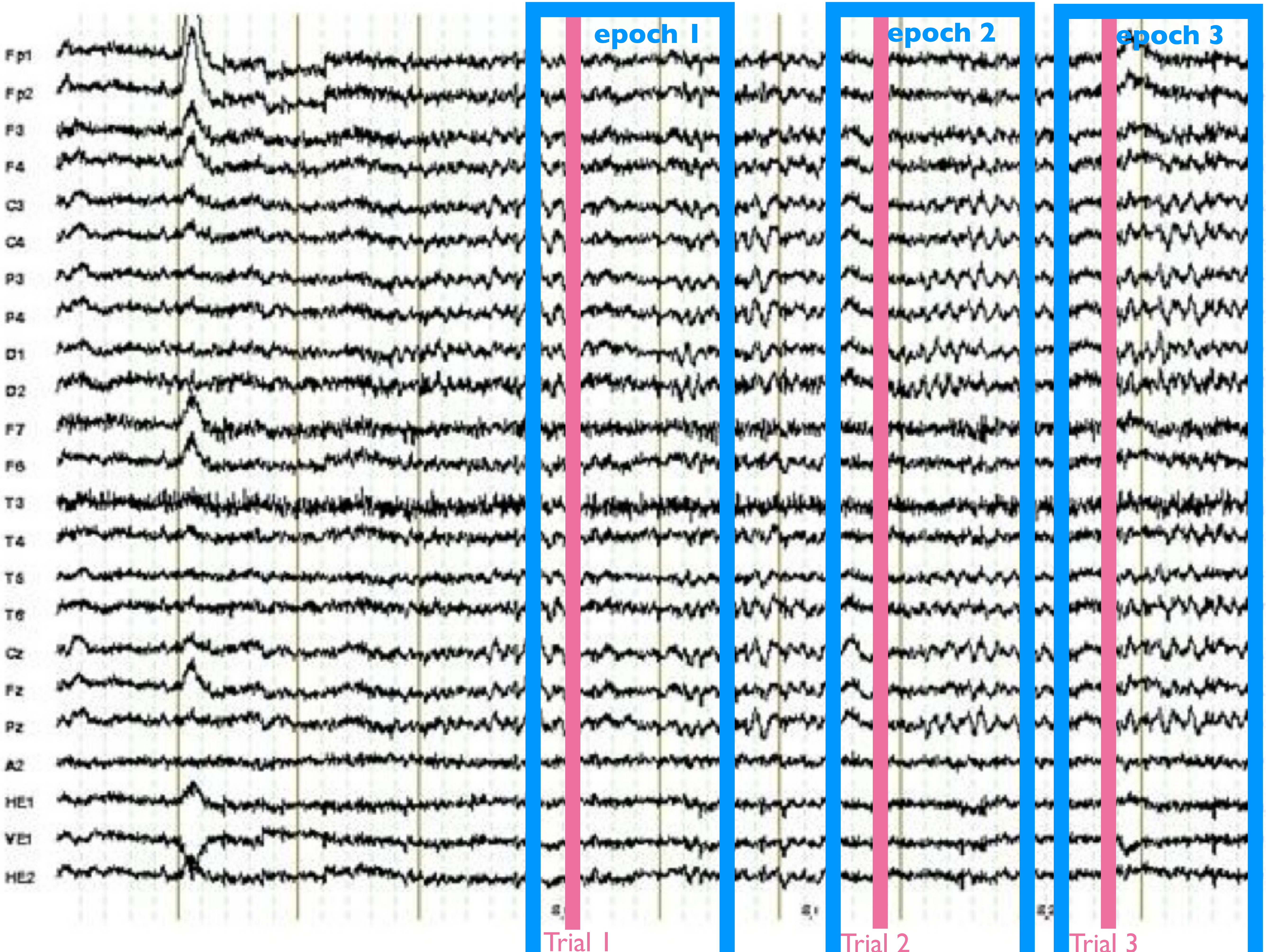
EVENT RELATED FIELDS (ERF): MEG


house


blicket


tiger

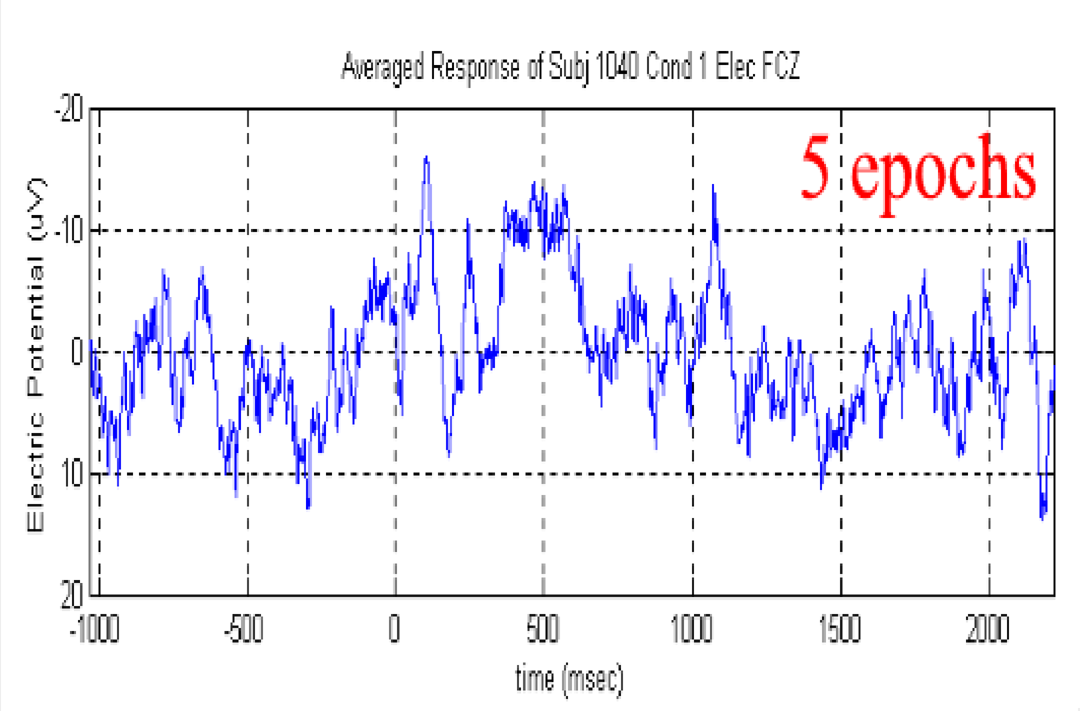

finger

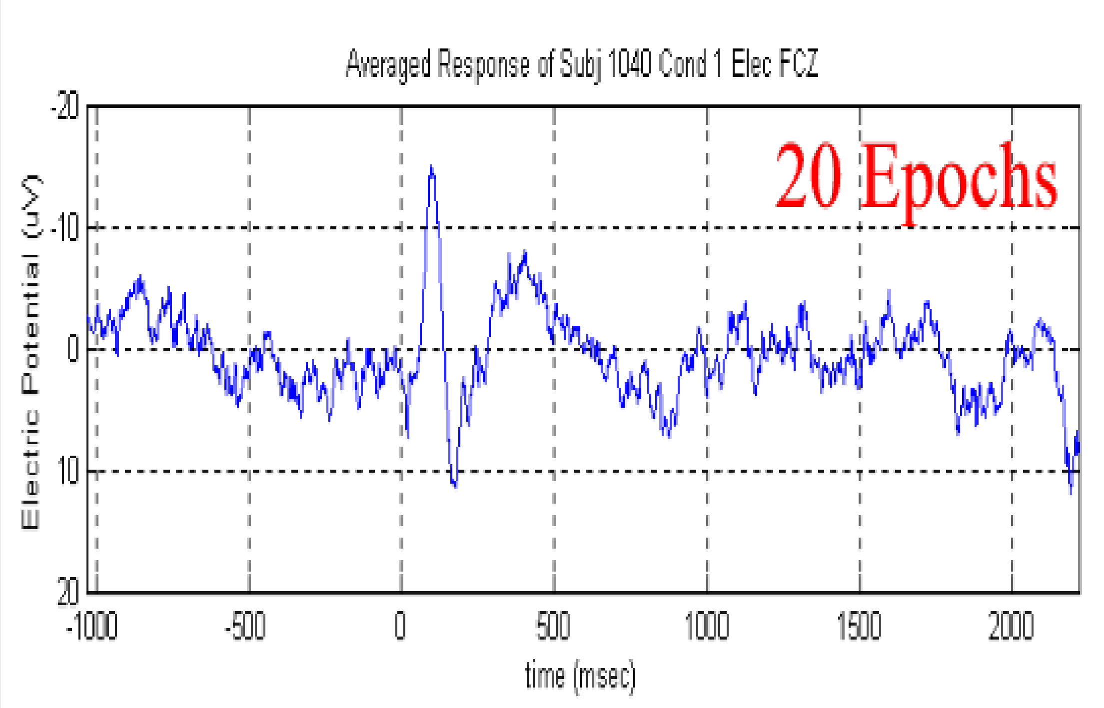



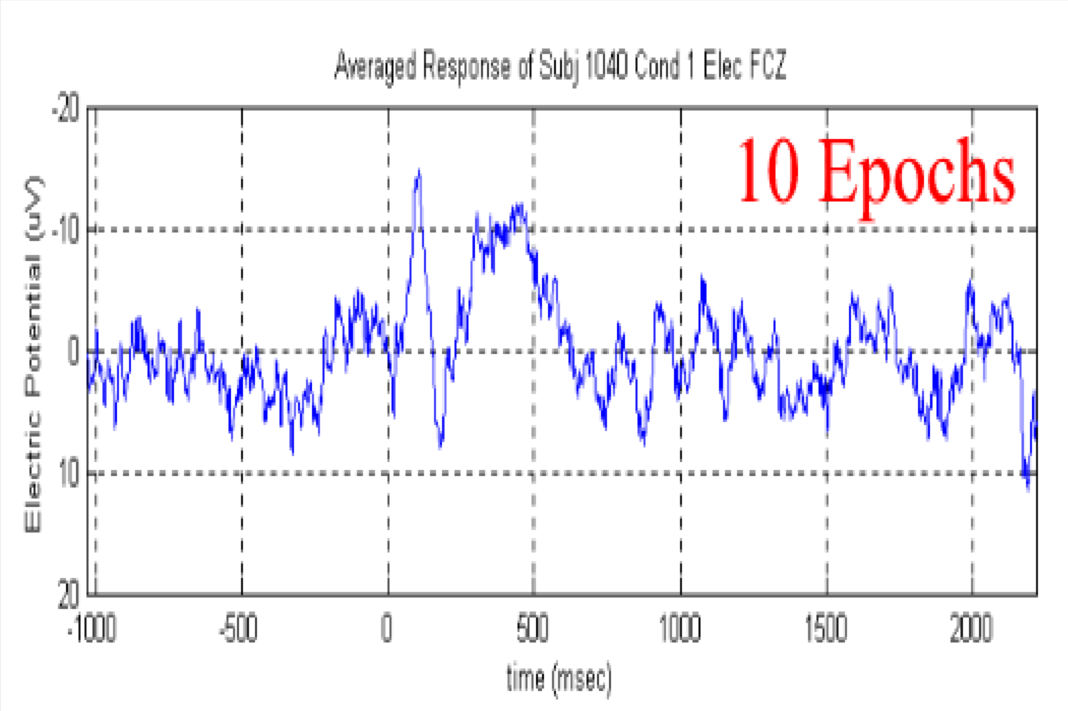
Fp1
Fp2
F3
F4
C3
C4
P3
P4
O1
O2
F7
F8
T3
T4
T5
T6
Cz
Fz
Pz
A2
HE1
VE1
HE2
epoch 1
epoch 2
epoch 3
Trial 1
Trial 2
Trial 3

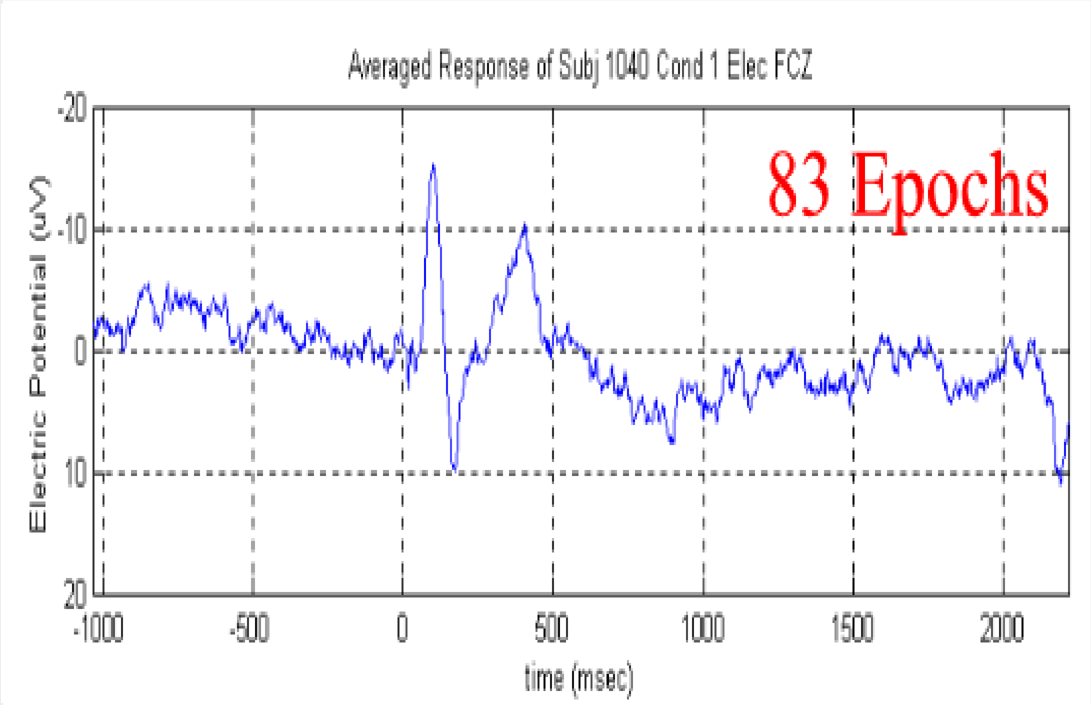
EVENT-RELATED BRAIN POTENTIALS

Epoch means

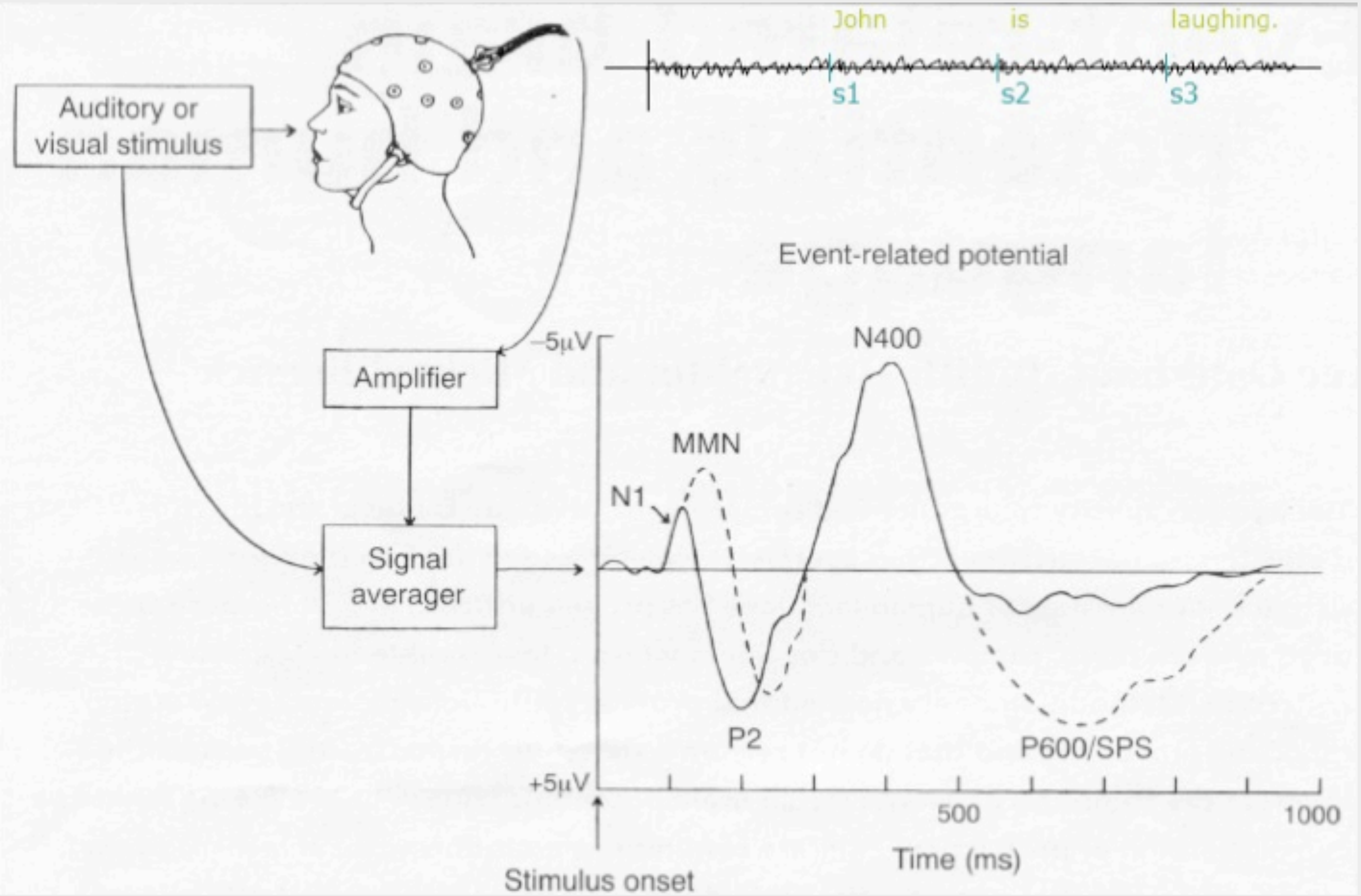




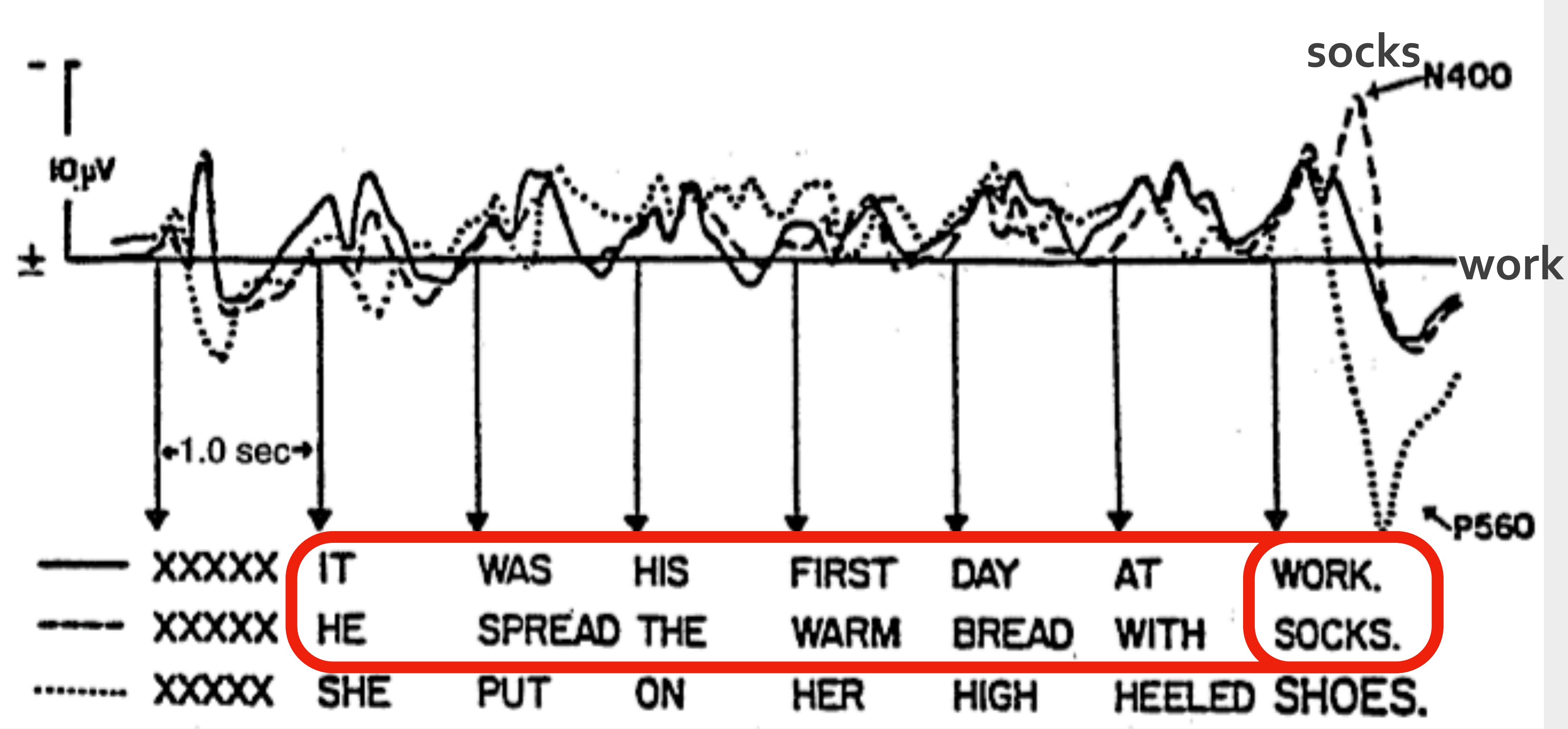




EVENT-RELATED BRAIN POTENTIALS



EVENT-RELATED BRAIN POTENTIALS



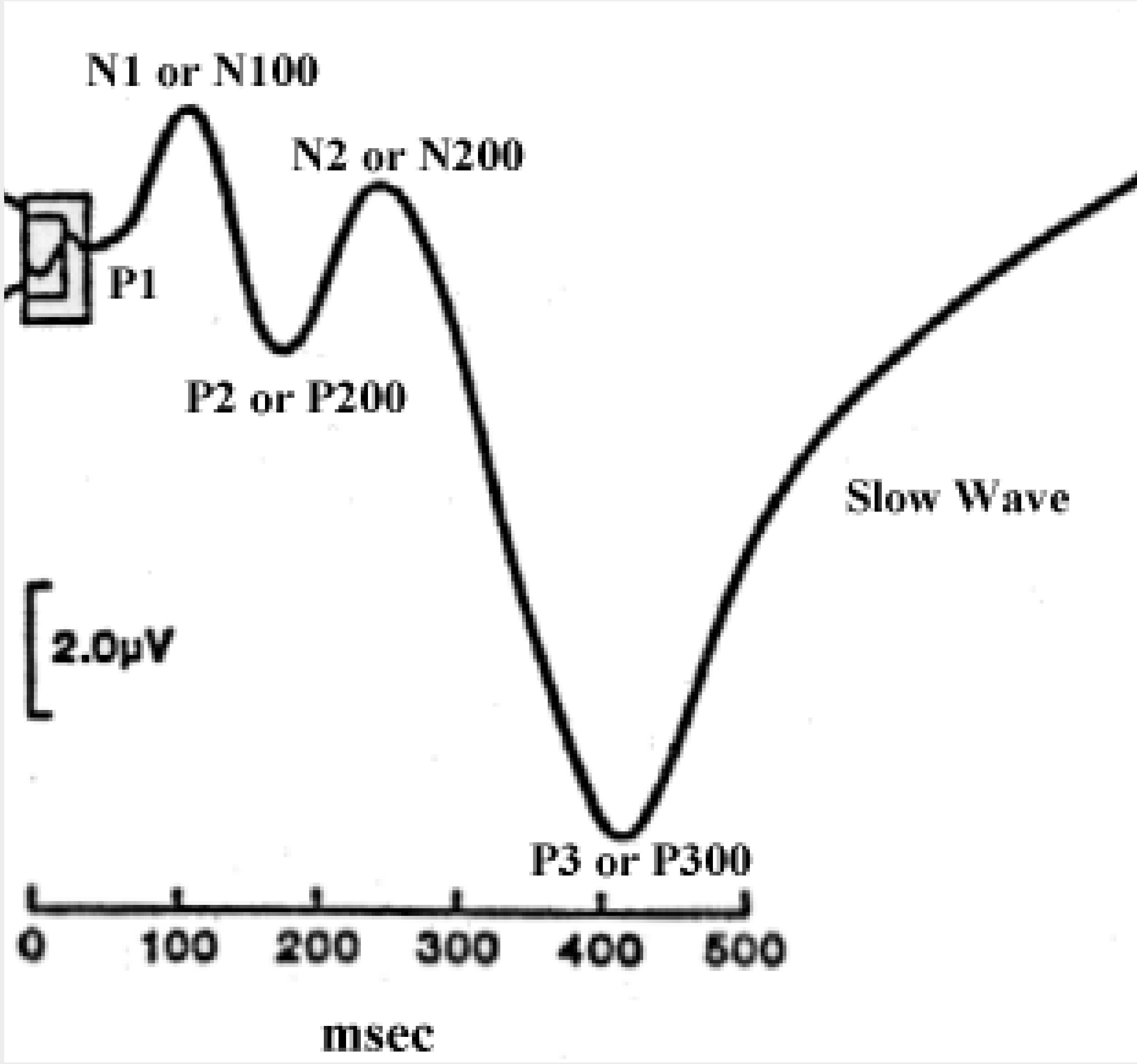
Kutas & Hillyard, 1980

LABELLING OF ERP COMPONENTS

P or **N**: whether the component is negative or positive going.

Number after the letter: Approximate peak latency of the components. 1, 2, 3, etc. are short for 100ms, 200ms, 300ms and so forth.

Traditionally, negative is plotted up and positive down.

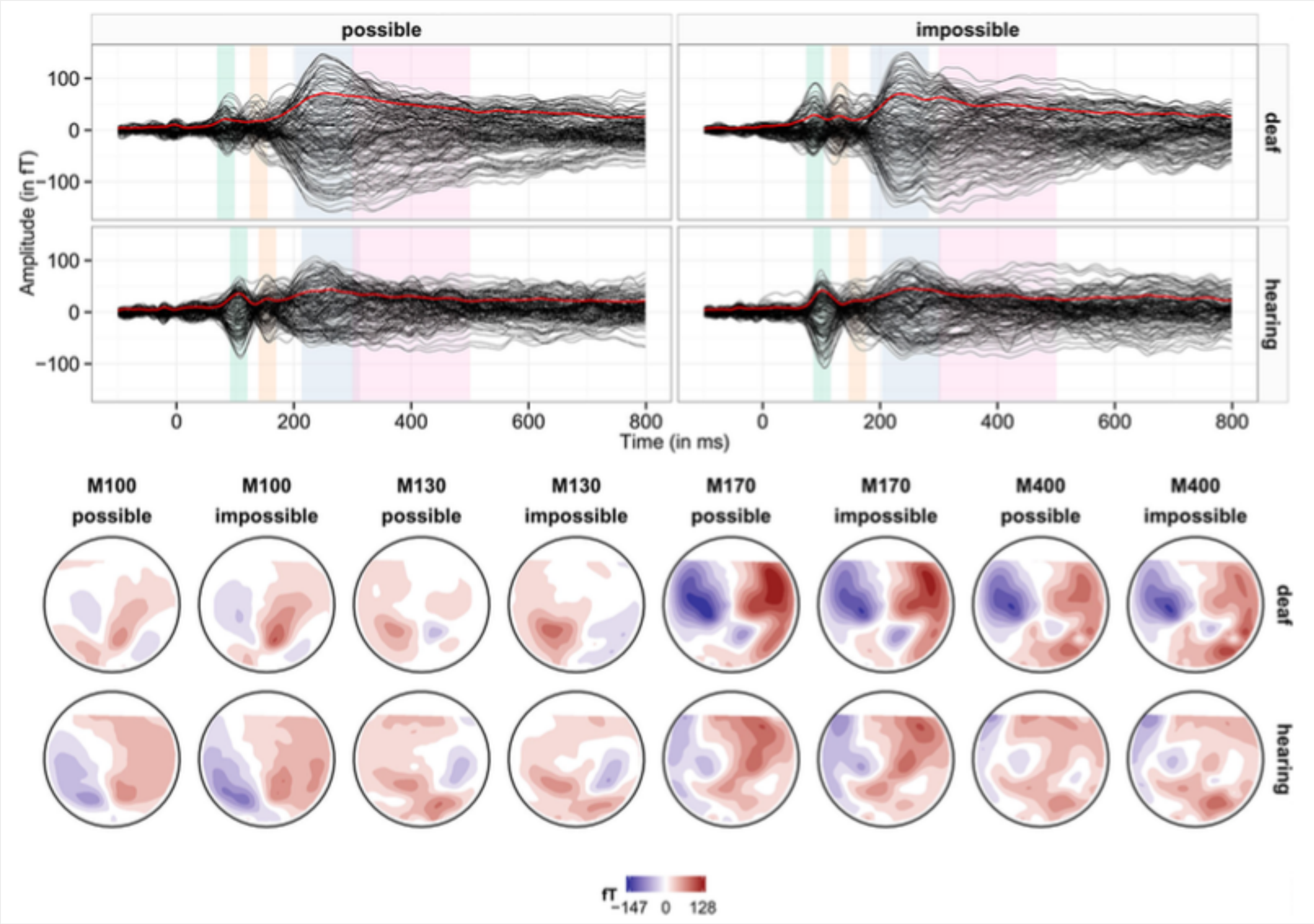


http://neurocog.psy.tufts.edu/images/ERP_components.gif

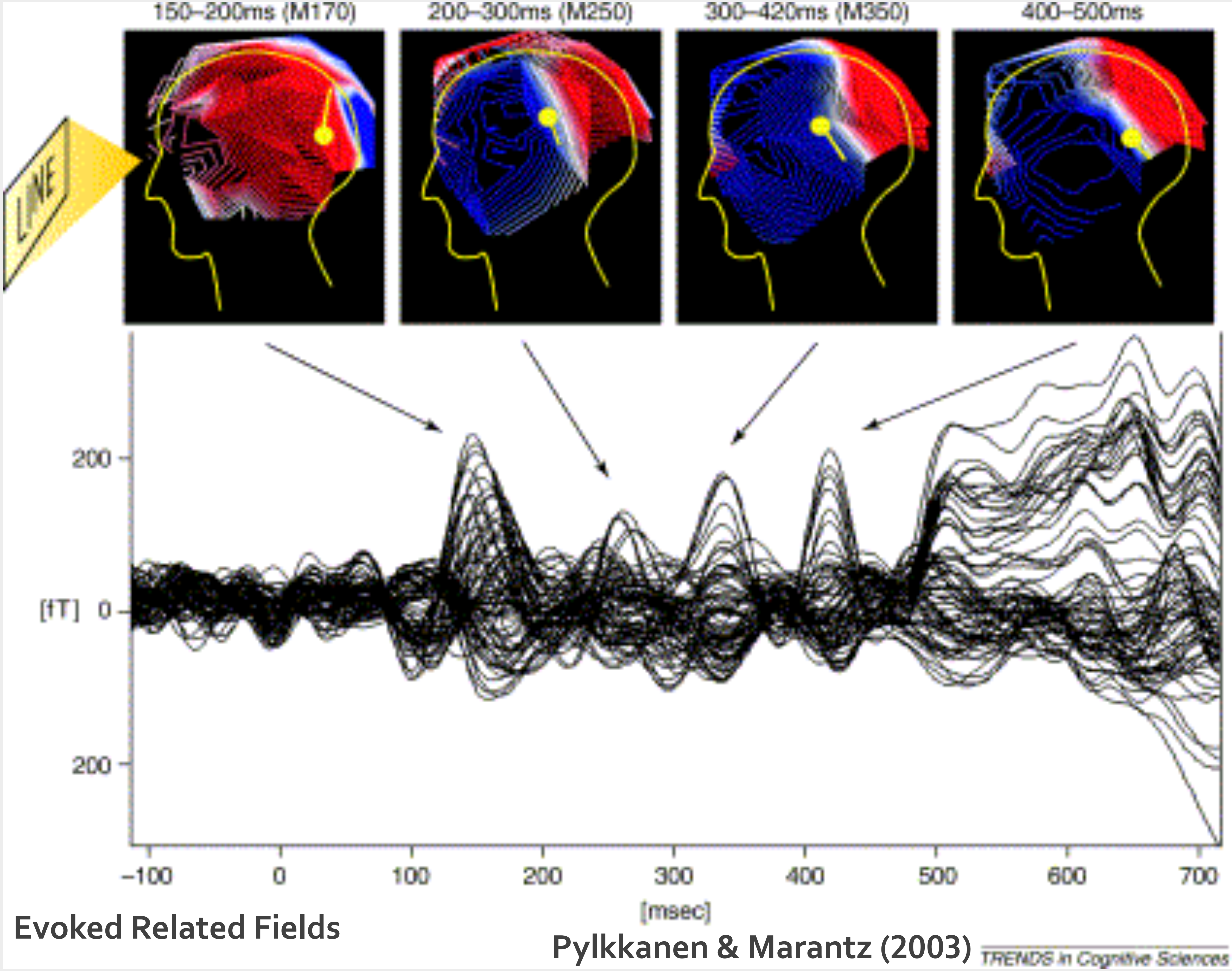
Labelling of MEG components

M: signal is a mix of outgoing and ingoing magnetic flux, no single polarity.

Number after the letter: Approximate peak latency of the components. 1, 2, 3, etc. are short for 100ms, 200ms, 300ms and so forth.



Almeida, Poeppel & Corina, 2015



Evoked Related Fields

Pylkkanen & Marantz (2003)

TEMPORAL AND SPATIAL RESOLUTION OF EEG

Millisecond temporal resolution.

Localization of neural generators complicated.

- Different tissues and the skull differ in their conductivity: Electric potentials do not pass through these structures undistorted.

  - Localization requires realistic head models.

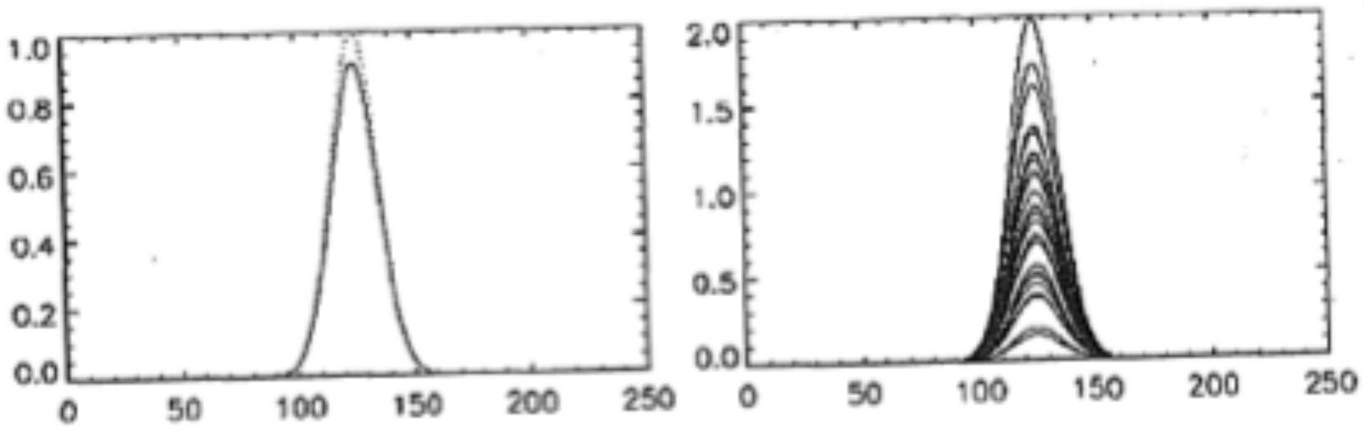


Figure 10.1
Effect of amplitude variability. (*Left*) The simulated ERP component is drawn with a dotted line and the average with a solid line. (*Right*) Simulated component on each single trial (the *y*-axis differs from the left plot). The single-trial amplitudes were drawn from a Gaussian distribution with a mean of 1.0 and a standard deviation of 1.0.

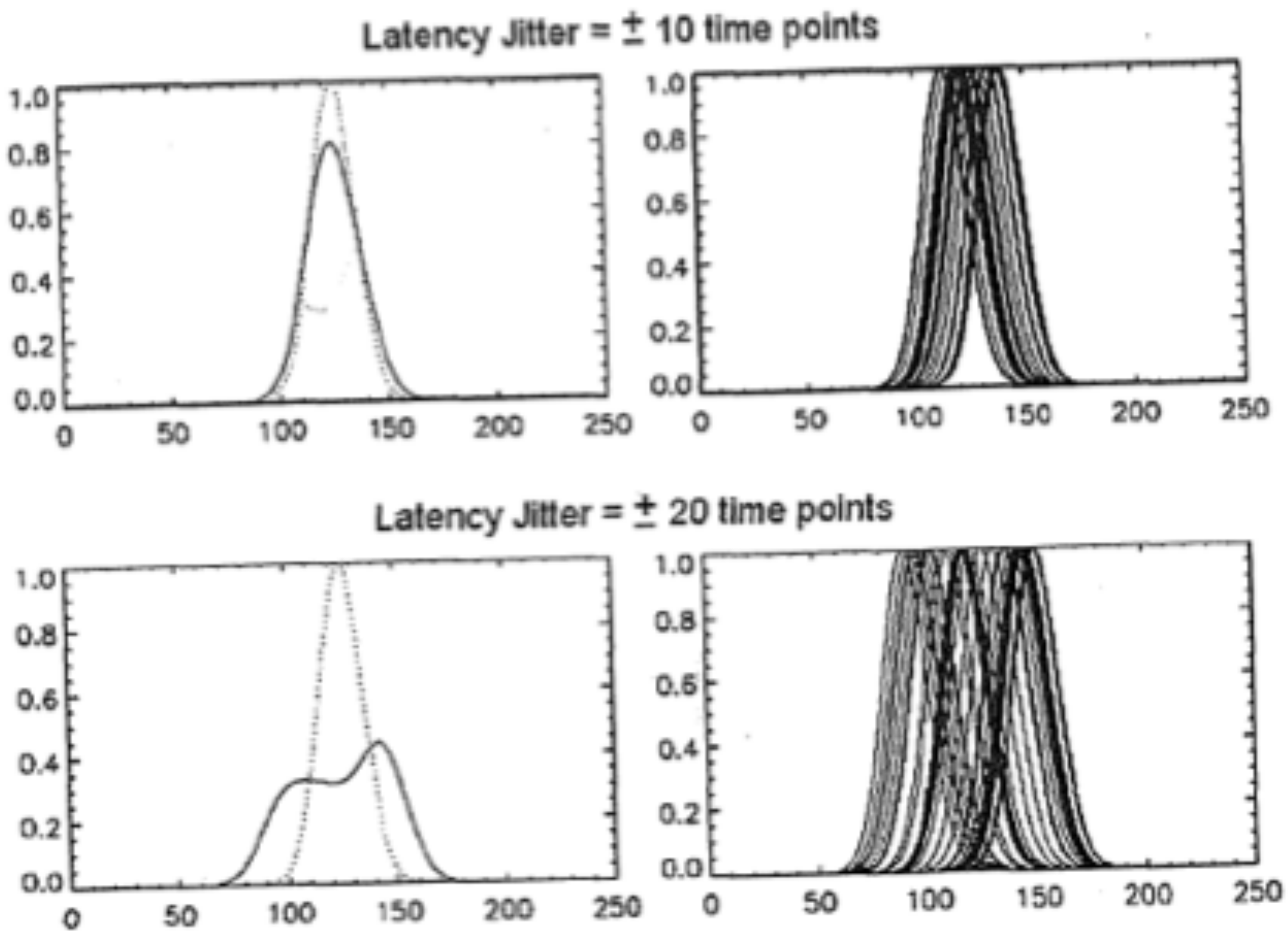
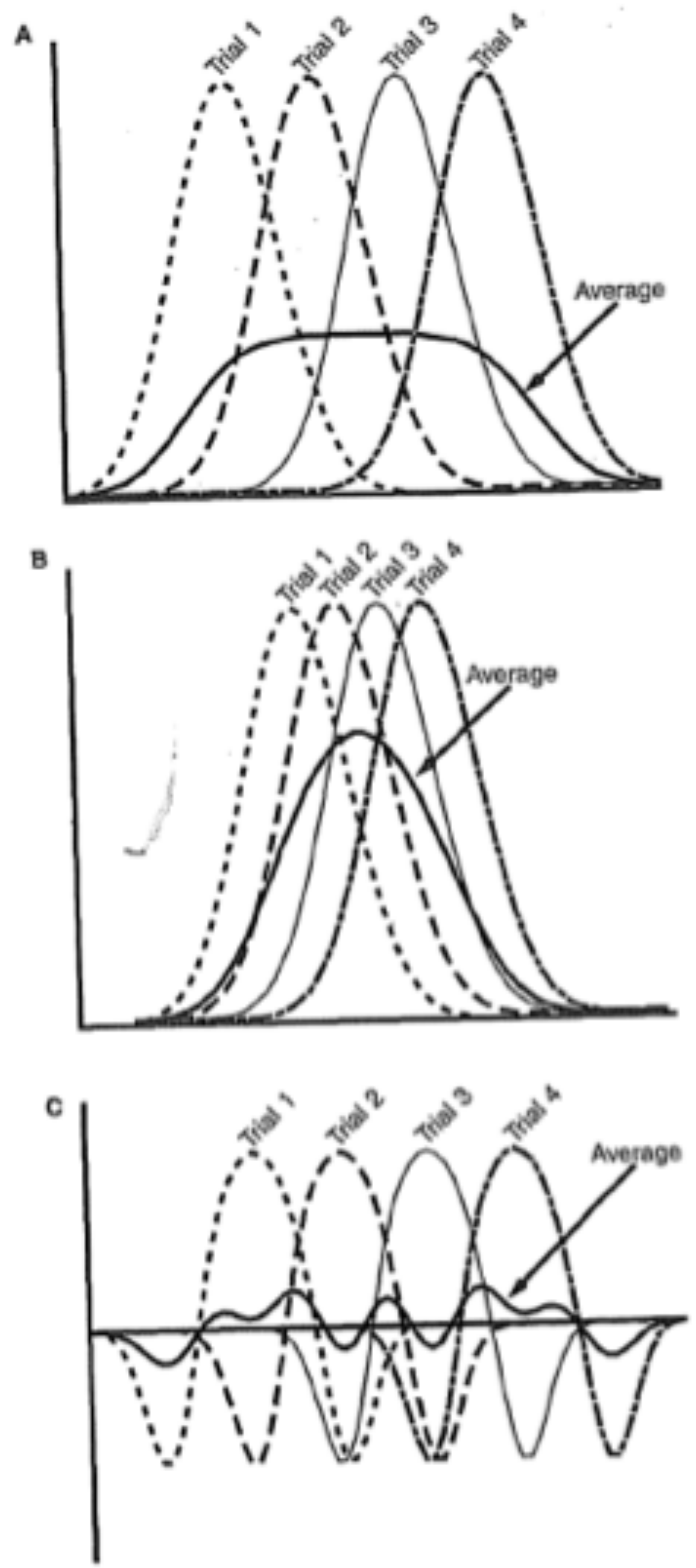


Figure 10.2
Effects of latency variability. (*Left*) The simulated ERP component is drawn with a dotted line and the average with a solid line. (*Right*) Simulated component on each single trial.

Assumption of ERPs:

time-locking!

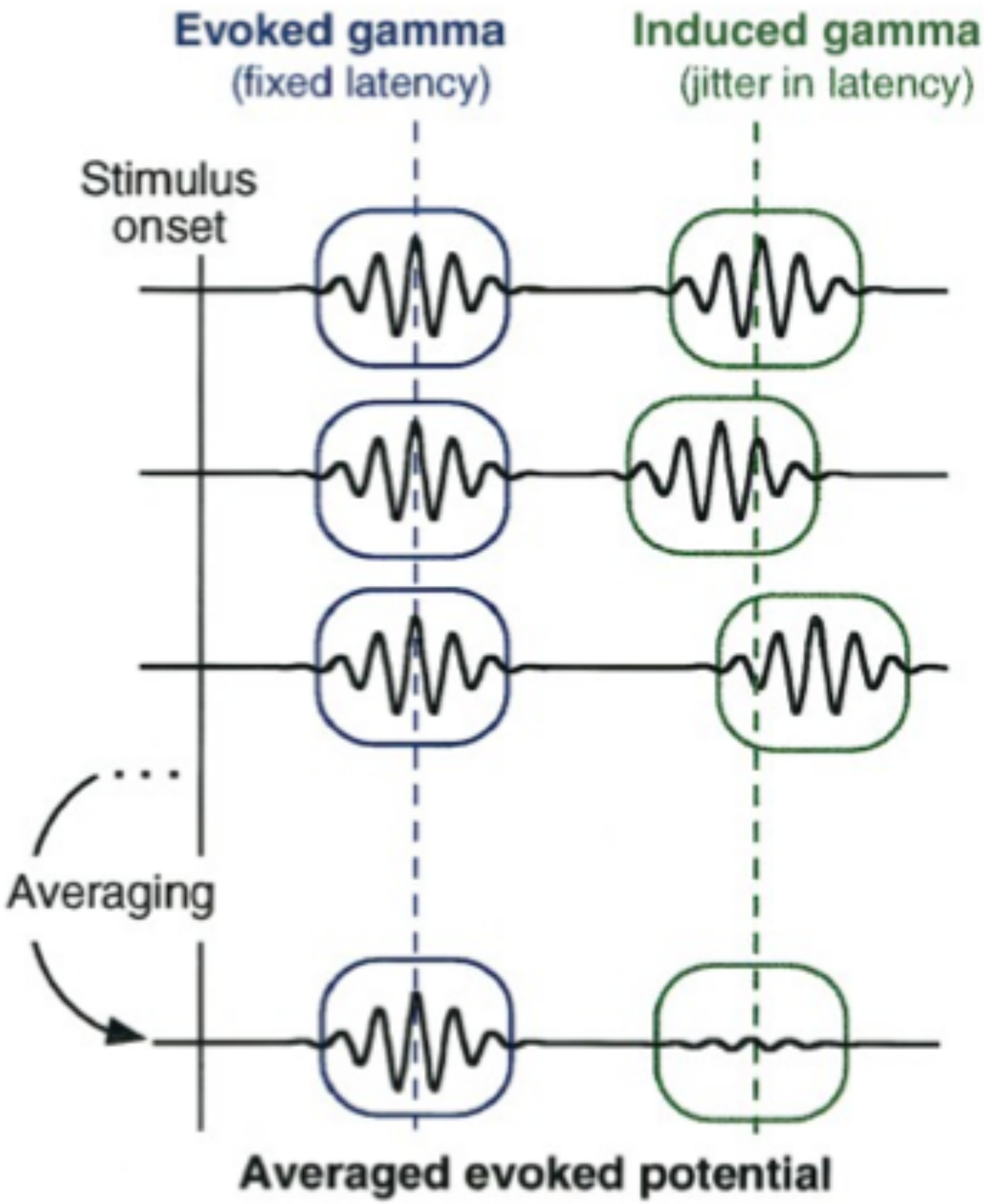
from Handy (ed.) (2005)



Assumption of ERPs:

time-locking!

from Luck (2005)



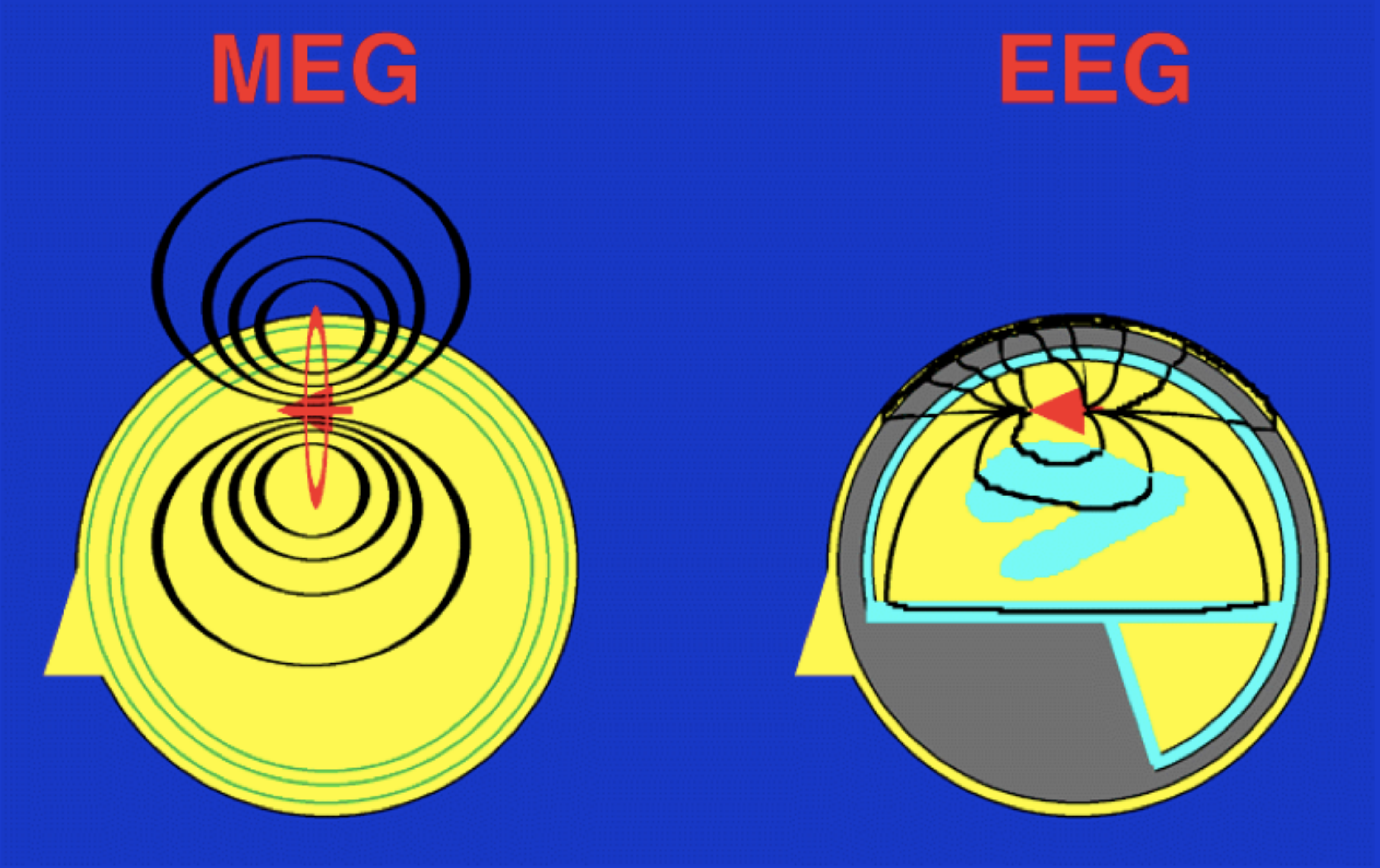
Assumption of ERPs:

time-locking!

From Tallon-Baudry et al (1999)

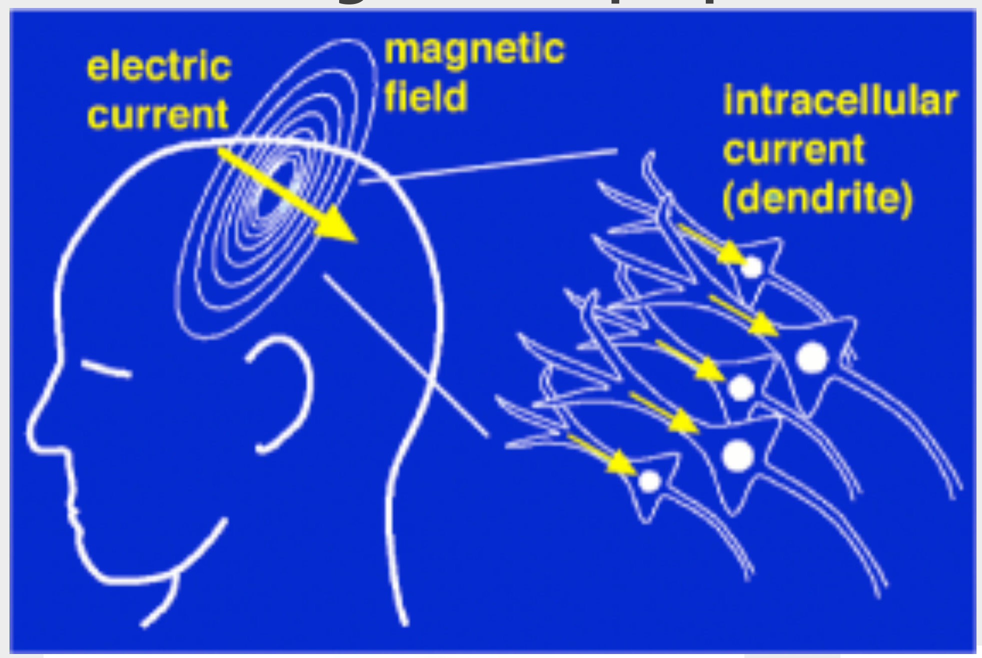
SIMILARITIES AND DIFFERENCES BETWEEN EEG AND MEG AND BETWEEN ERPS AND ERFS

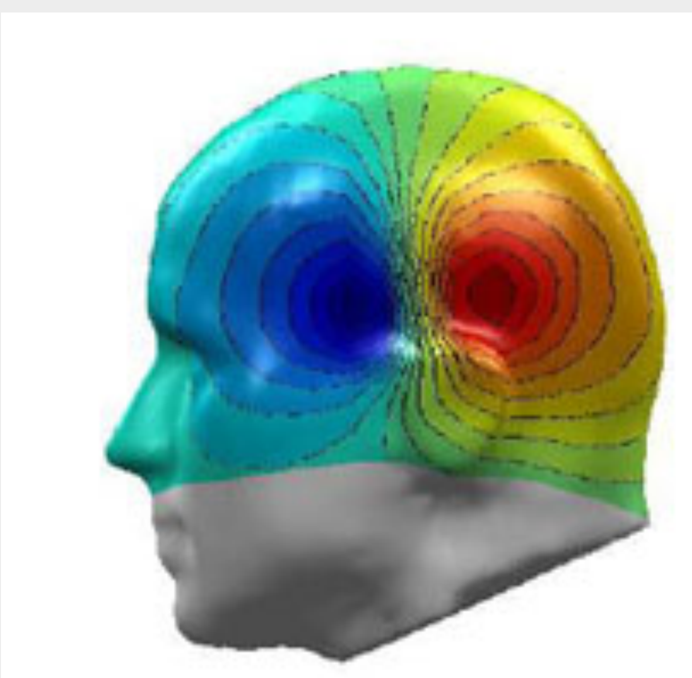
EEG more 'smeared' than MEG



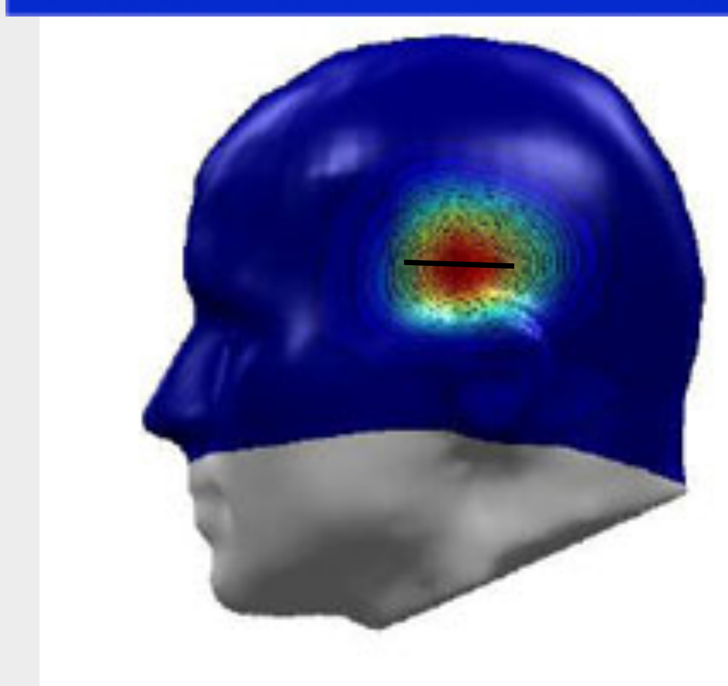
ELECTRO X MAGNETOENCEPHALOGRAM

EEG and MEG signals are perpendicular

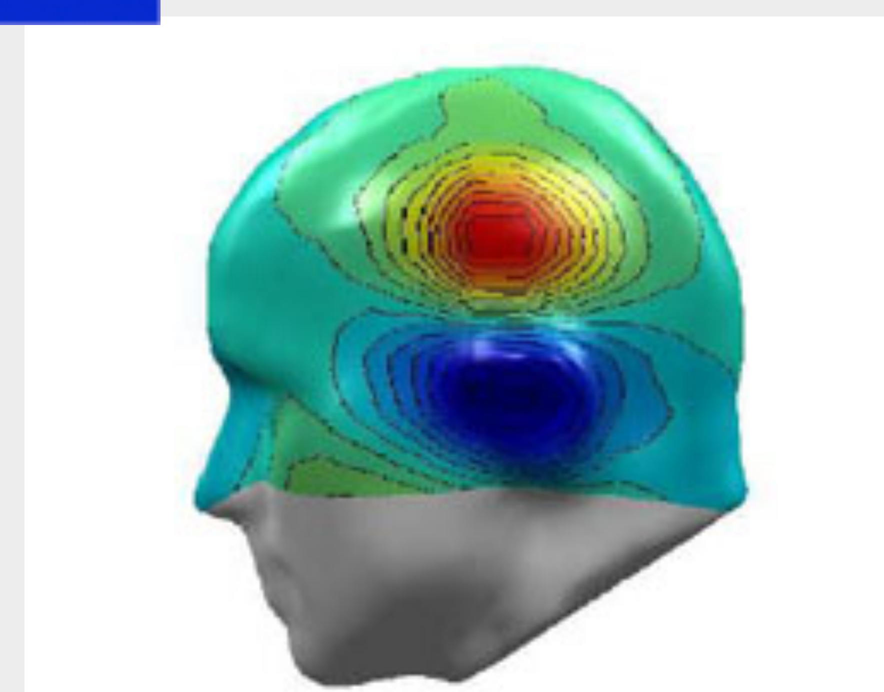




EEG

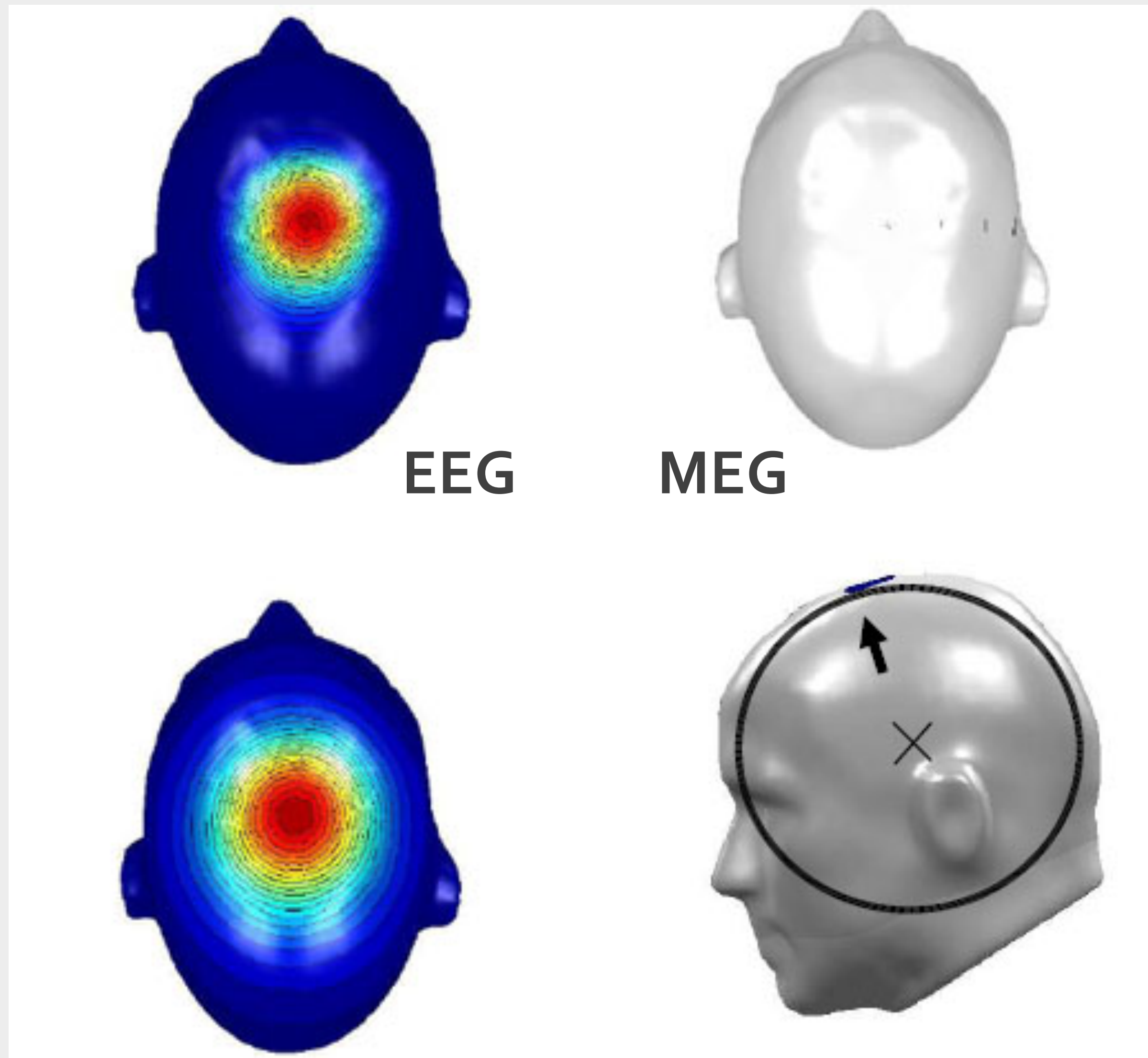


dipolar source

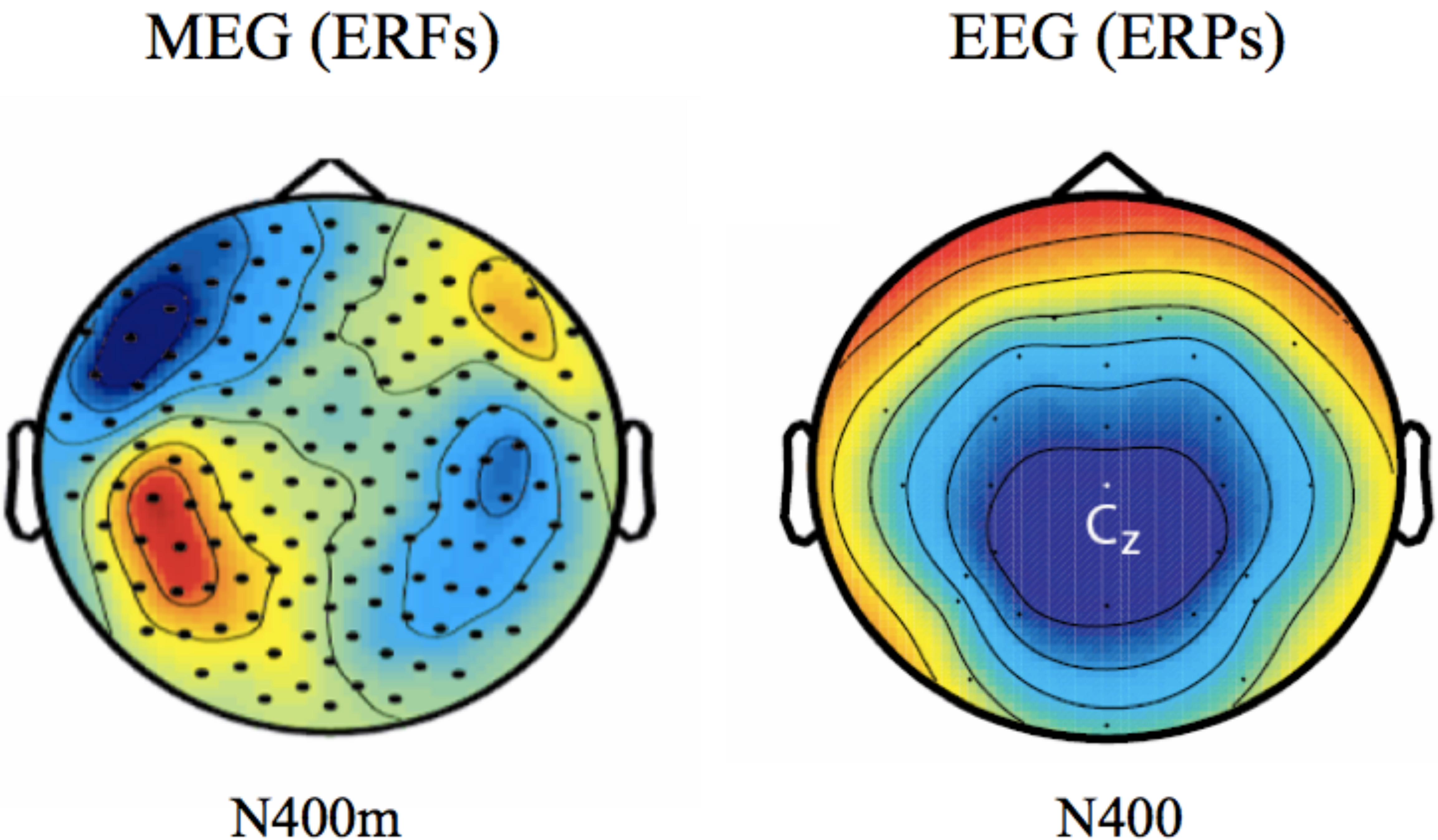


MEG

MEG not sensitive to radial source orientations



dipolar source


MEG (ERFs)
EEG (ERPs)
Cz
N400m
N400

02

THE PROBLEM TODAY

SPEECH PERCEPTION *VS* LEXICAL STORAGE

Long history of research on speech perception

- **Goal:** mapping an acoustic event to a phonetic interpretation
- **General assumption:** that's all there is!

SPEECH PERCEPTION *VS* LEXICAL STORAGE

Long history of research in the sound structure of languages

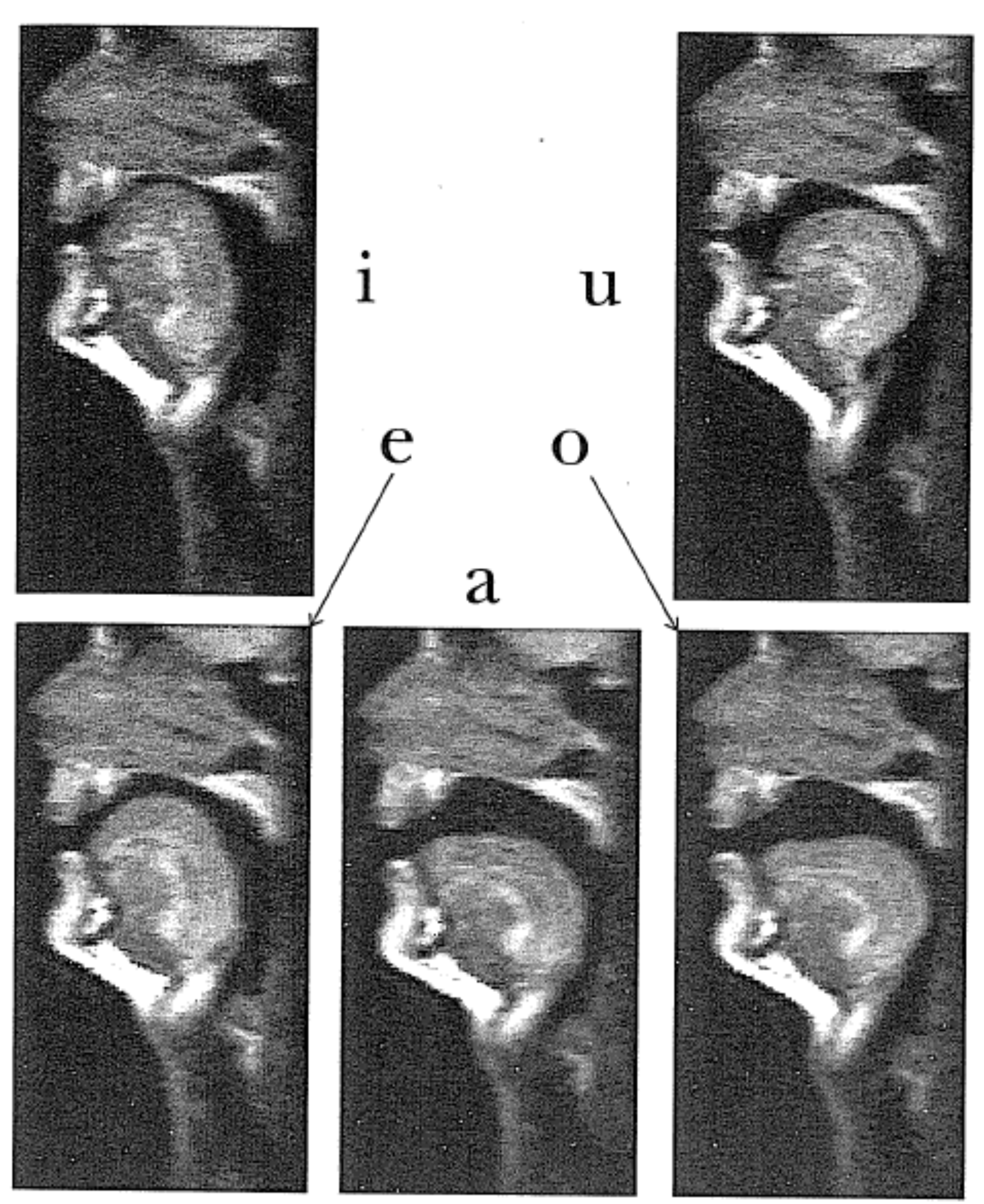
- **Phonology**
  - **Goal:** determine the representational format of a speech sound for lexical storage

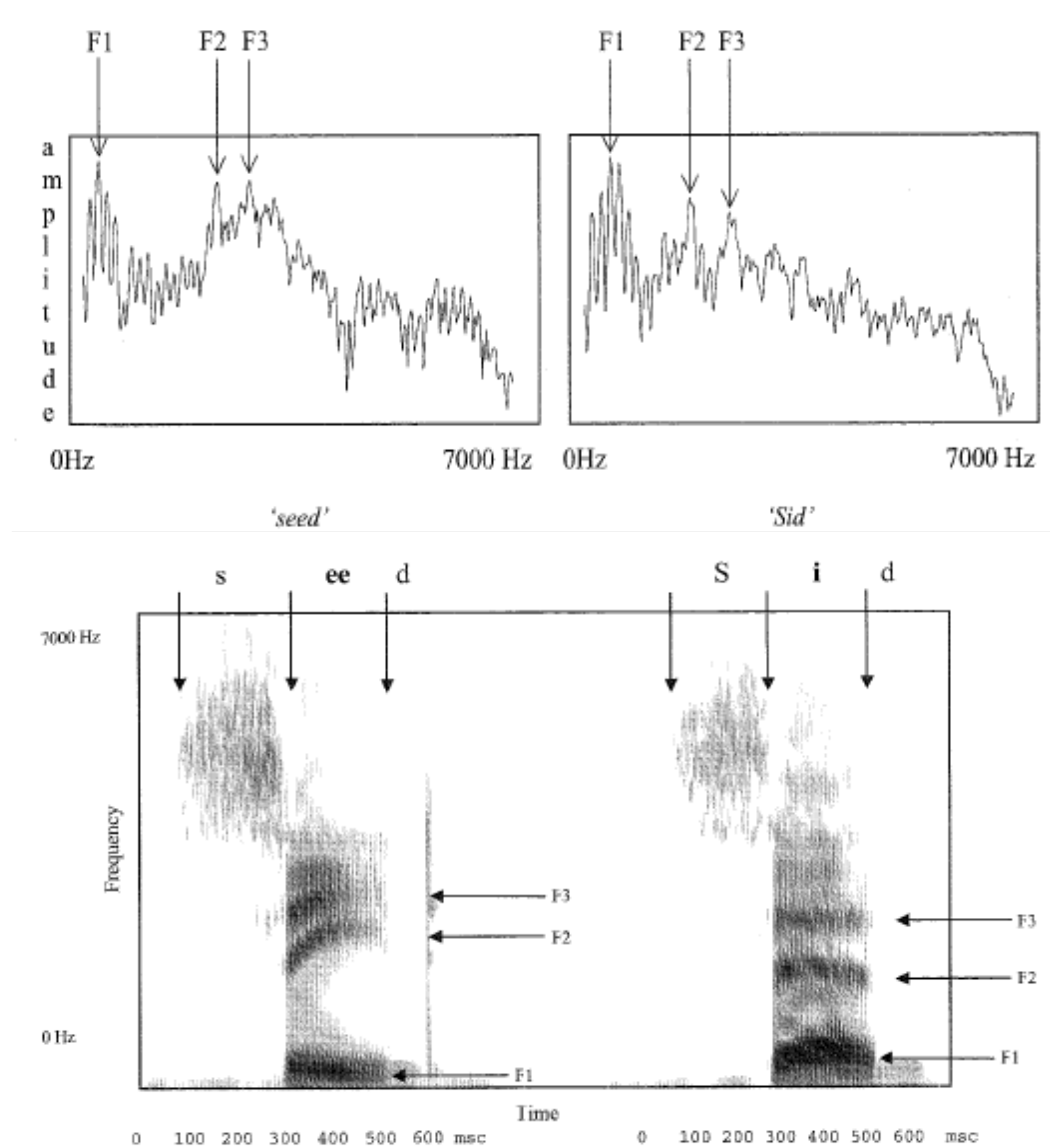
SPEECH PERCEPTION: MAPPING PROBLEM

Nature of mapping

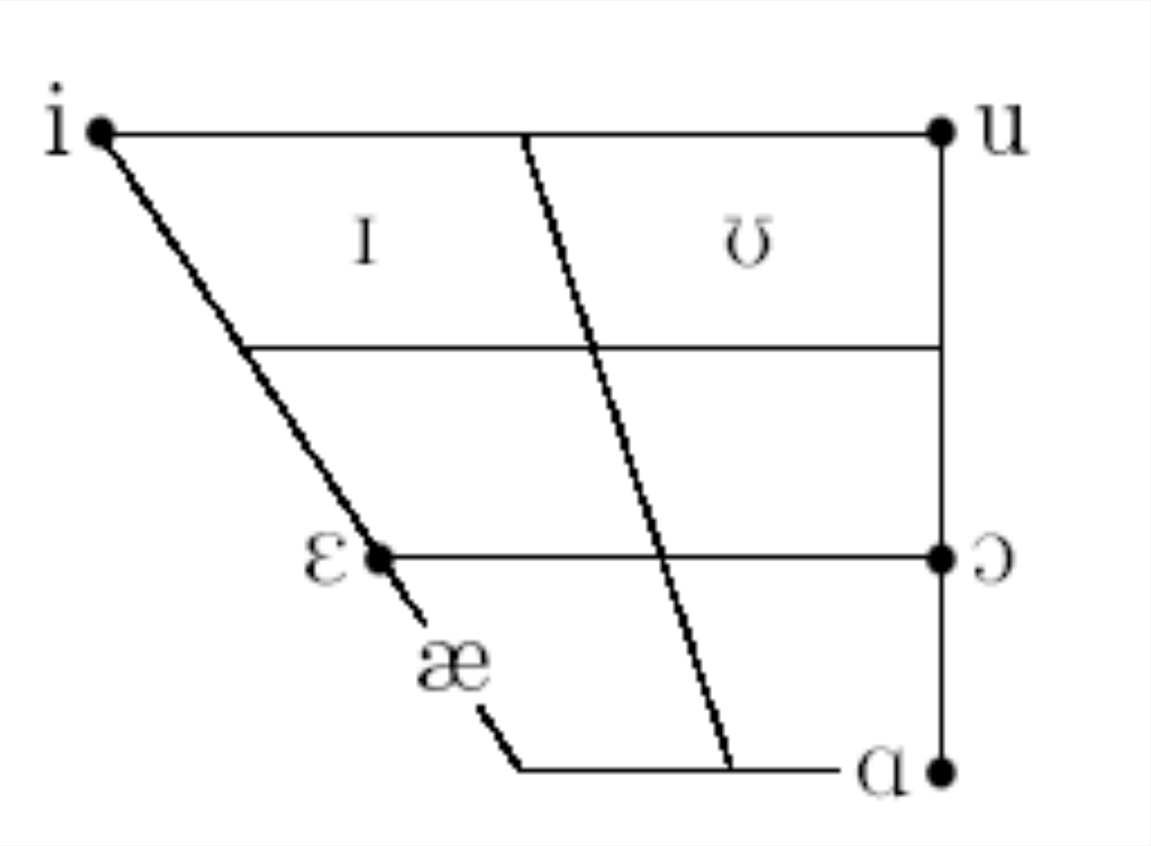
- e.g.: vowels

tongue position creates different filters → Certain frequencies have more energy than others (formants)

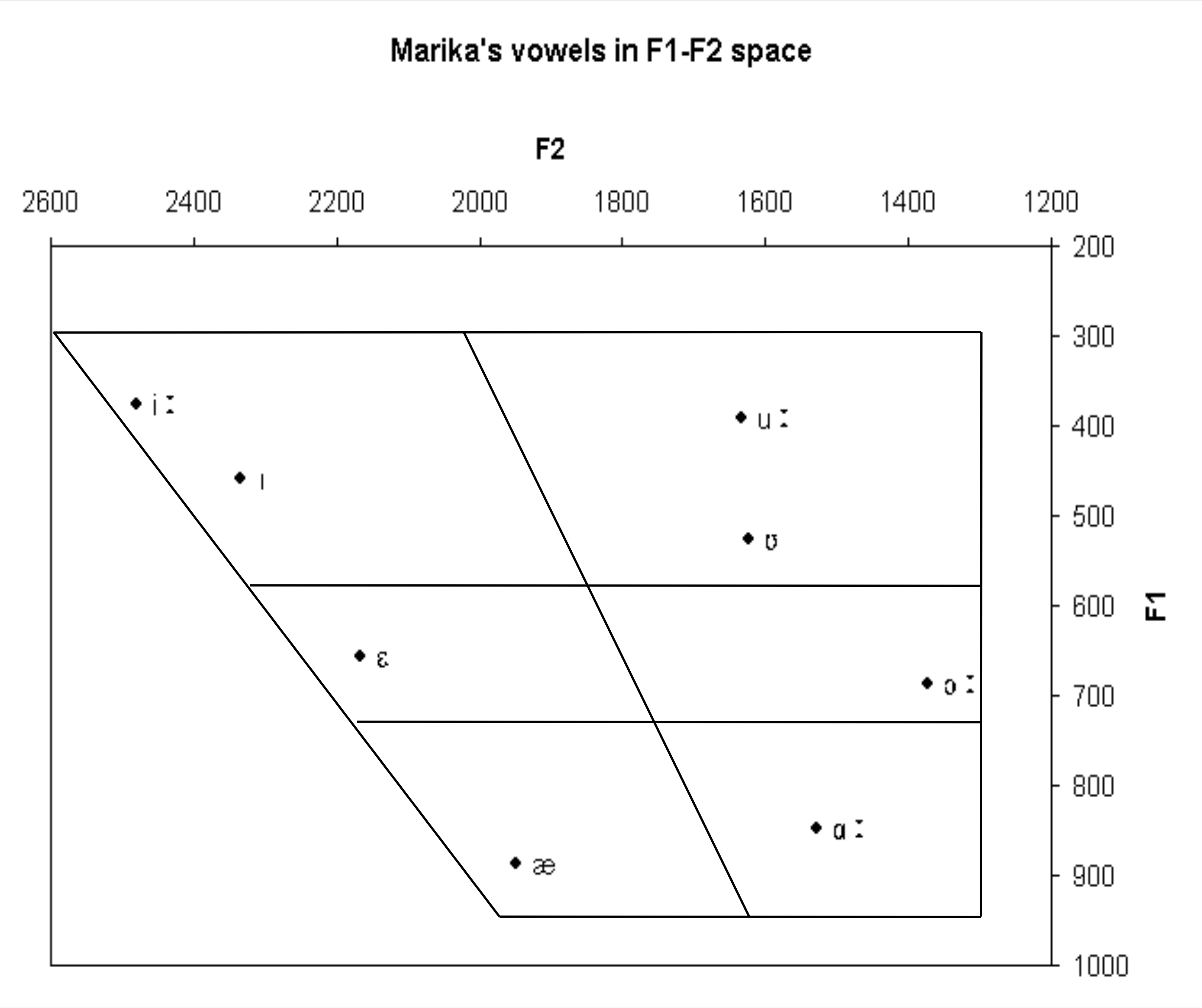




articulatory map (from IPA)



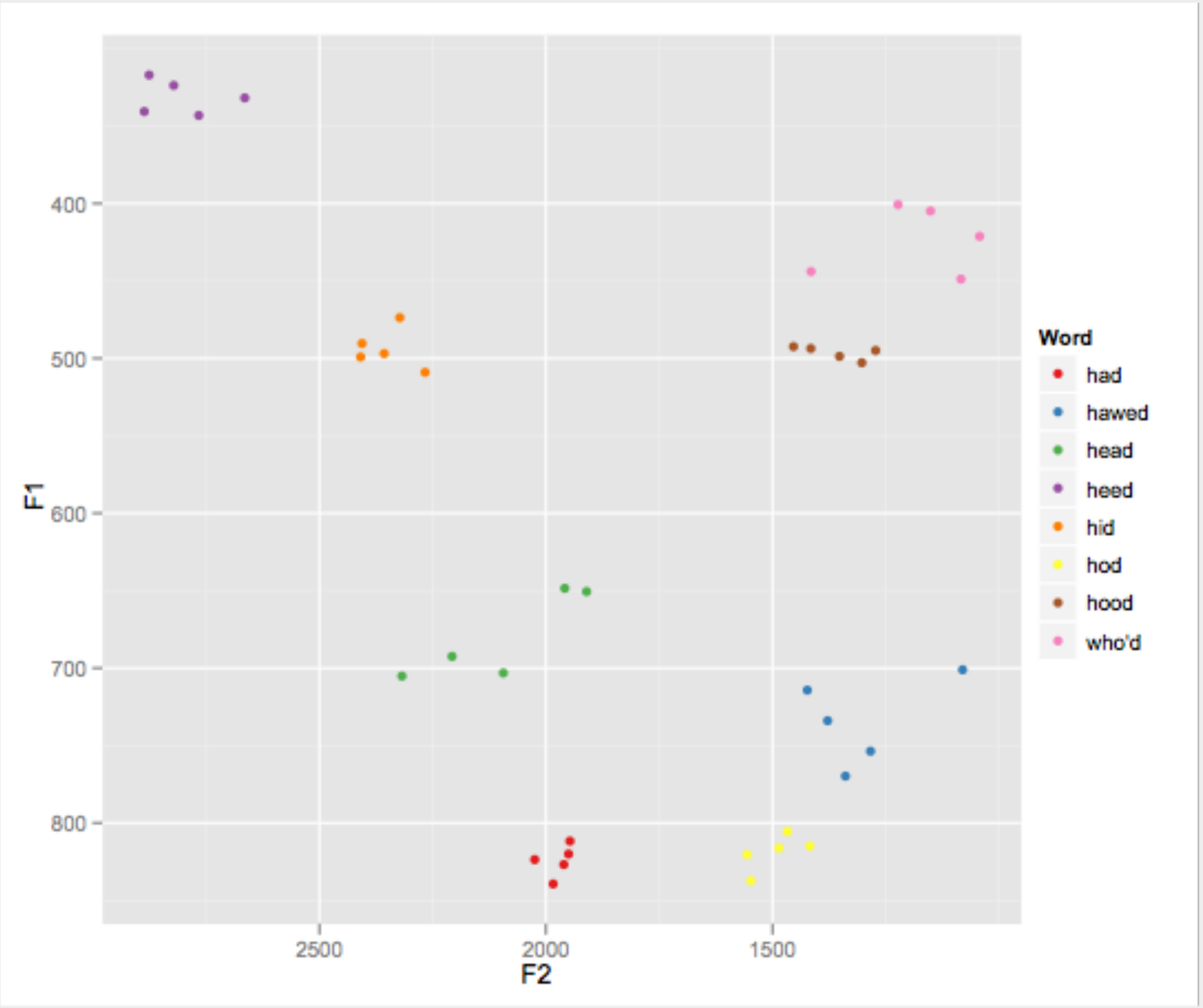
acoustic map (from recordings)



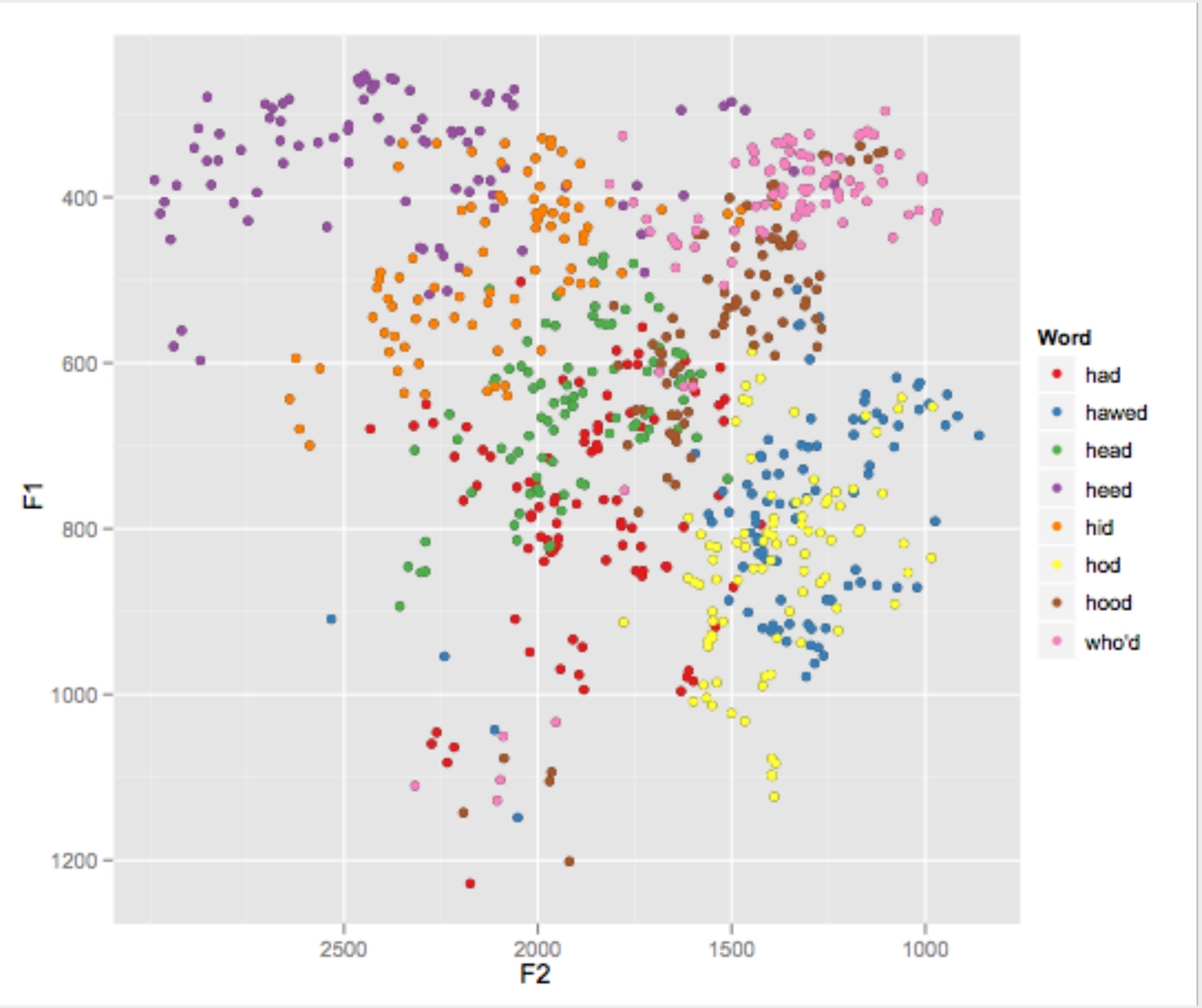
SPEECH PERCEPTION: MAPPING PROBLEM

Problem of variation

Data from one person



Data from several people



HOW MANY “T”S?

Till

Still

Seat

Tree

Butter

Button

PHONETIC TRANSCRIPTION

t - tʰ - t̚ - tʷ - ɾ - ʔ

Till	[tʰ ɪ ɫ]
Still	[s t ɪ ɫ]
Seat	[s iː tʰ] or [s iː t̚], but not [s iː t]
Tree	[tʷɹ̥ iː]
Butter	[b ʌ ɾ ɚ]
Button	[b ʌ ʔ n̩]

DOES ALL THIS VARIATION MATTER?

YES: for pronunciation

- Native vs non-native sounding speech!

Bulk of speech perception research stops here!

- **Map:** acoustic events → phonetic units

Assumption: words are stored as strings of phonetic units

DOES ALL THIS VARIATION MATTER?

NO: for storing different words

- [t] and [p] are also different sounds, but they also help distinguish word forms in english
  - Pill vs Till, Pan vs Tan, Map vs Mat
- You cannot find pairs of words that would differ only in terms of
  - t vs t^h ||| t vs t^w ||| t vs ʔ ||| t vs ɾ

PHONES VS PHONEMES

Large number of “sounds” used in actual pronunciation

- We call these **phones**

Only a subset of phones are lexically contrastive

- We call these **phonemes**

PHONEMES ARE AN ABSTRACTION

Layer of representation that groups distinct phones together

[t - th - tw - ʔ - ɾ] can be represented with a single symbol **/t/**

Crucially, phonemes are language-dependent!

[t - th - tw - ʔ - ɾ] can be represented with a single symbol **/t/**…

- …IN (AMERICAN) ENGLISH, not so in other languages!

PHONEMES ARE AN ABSTRACTION & LANGUAGE DEPENDENT!

English:

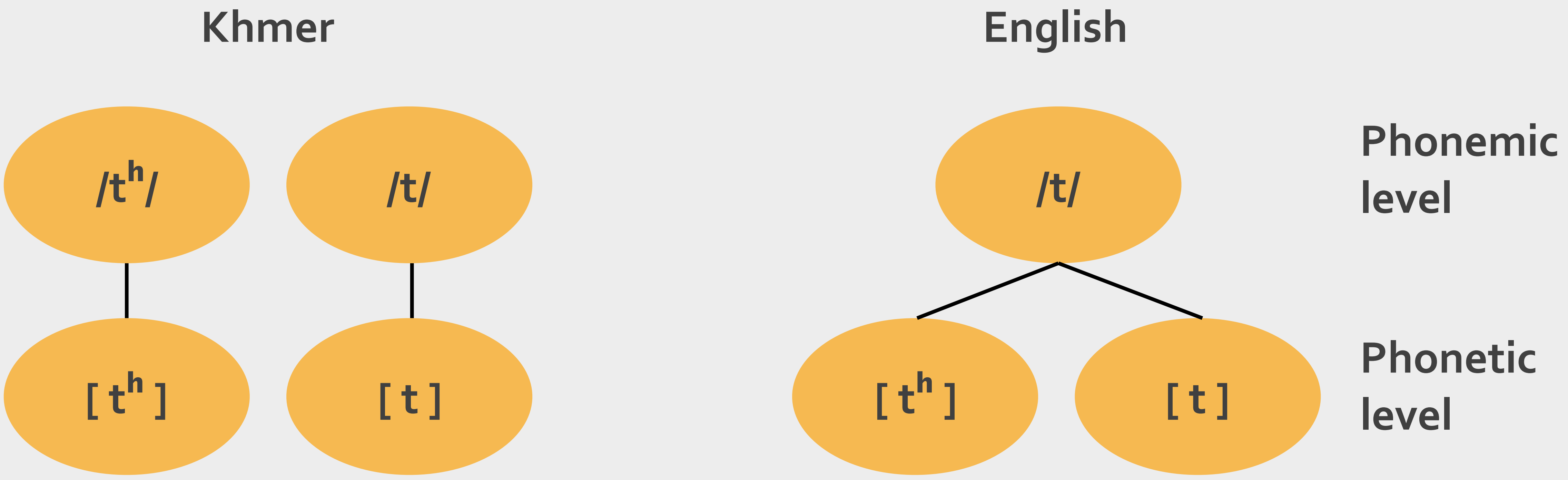
[t^h ɔ p] vs [t ɔ p] = same word (just weird without aspiration)

Khmer:

[t^h ɔ p] = to support [t ɔ p] = be suffocated

PHONEMES ARE AN ABSTRACTION & LANGUAGE DEPENDENT!

[t^h] vs [t]: phonemic in Khmer, but only **phonetic** in English



PHONEMES AND PHONES

General Agreement: You always hear phones

Less Agreement: you store phonemes, not phones

PHONEMES AND PHONES

Phones linked to the same **phoneme** are usually predictable from context

- You get aspiration for **/t/** in syllable initial positions
- You get unaspirated **[t]** after **/s/**
- You get the flap in between vowels
- Etc.
- So maybe it's all about coarticulation?

PHONEMES AND PHONES

Phones linked to the same **phoneme** are usually predictable from context

- So maybe it’s all about coarticulation?

…But sometimes the variation is random(-ish)

- What is the first consonant in words like ”human” or ”huge”?
- Sometimes [h], sometimes [j] and sometimes even [ç]

PHONEMES AND PHONES

Phones linked to the same **phoneme** are usually predictable from context

- So maybe it’s all about coarticulation?

…But sometimes there’s conditioning at a distance

- E.g. Vowel Harmony (here, Hungarian)
- Bab – Babnak; Rum – Rumnak
- Hit – Hitnek; Tök – Töknek

PROBLEM OF TWO LEVELS OF REPRESENTATION

Locus of phenomena?

Example: Place assimilation:

- ”Sweet boy” often sounds as ”sweep boy”
- But ”sweep toy” never sounds as ”sweet toy”
- Phonetically or phonologically conditioned?

BUILDING BLOCKS OF PHONES: DISTINCTIVE FEATURES

[t]

- [+plosive]
- [+coronal]
- [+anterior]
- [-voice]

[d]

- [+plosive]
- [+coronal]
- [+anterior]
- [+voice]

BUILDING BLOCKS OF PHONES: DISTINCTIVE FEATURES

Strong *representational* claim:

- They are the smallest building bocks of language (atoms of speech)

- Used to store sounds in the mind

ERP OF INTEREST:

MISMATCH NEGATIVITY (MMN)

THE MISMATCH NEGATIVITY (MMN)

Stream of acoustic stimuli

- One type is repeated a lot (*standard*)
- The other appears only occasionally (*deviant*)
- MMN is (broadly construed) the difference between the two

Elicited pre-attentively!

- Participants are instructed not to attend on the auditory stream, but on reading a book or watching a silent movie

THE MISMATCH NEGATIVITY (MMN)

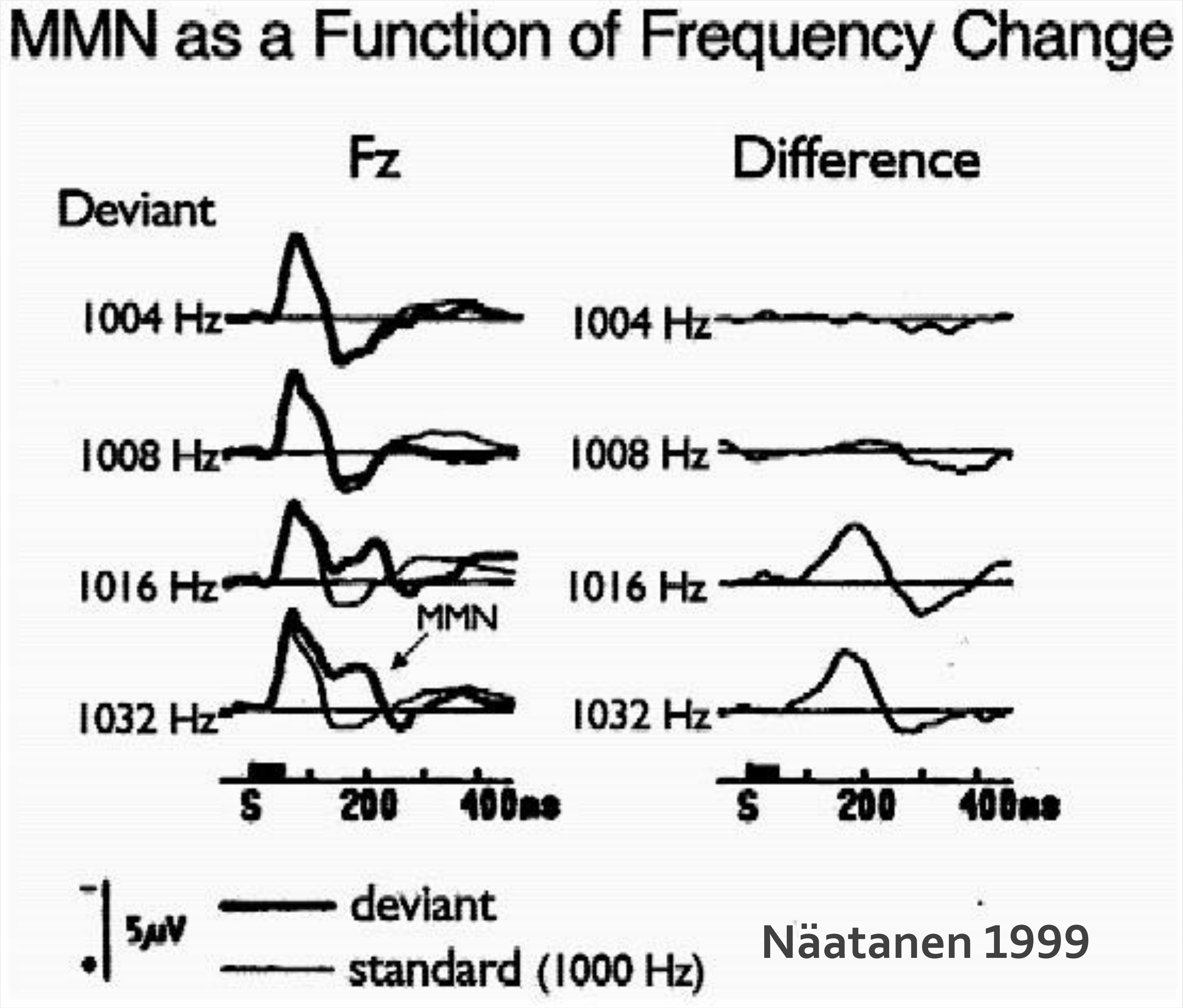
Standard description:

- Automatic auditory difference detection in memory
- The repeated presentation of the standards creates/induces/ retrieves a memory trace of the ongoing stream
- The deviant clashes with that representation
- MMN is the error signal

MMN as error signal

Sensitive to the *magnitude* of the difference between standards and deviants!

Larger (acoustic) differences lead to *larger* MMN effects.

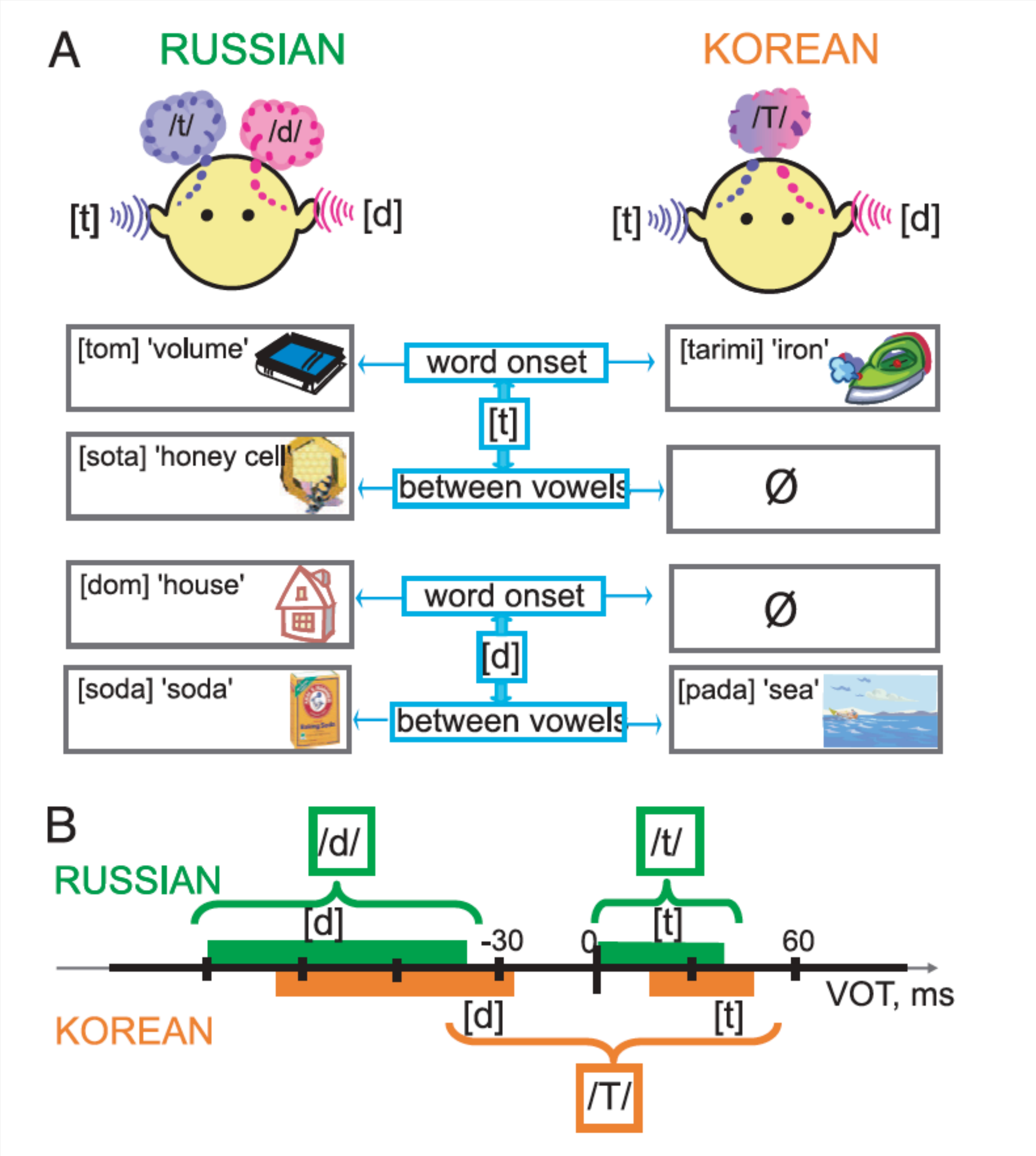


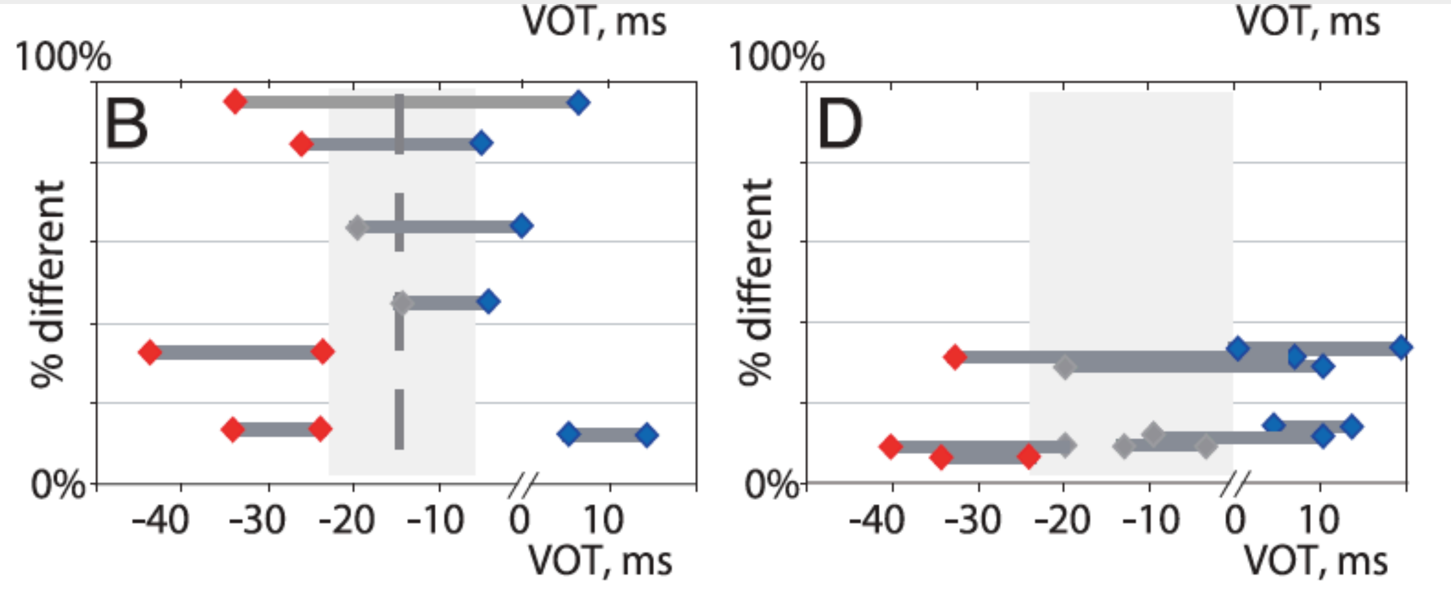
Näatanen 1999

THE MISMATCH NEGATIVITY (MMN)

Crucial question (for language scientists)

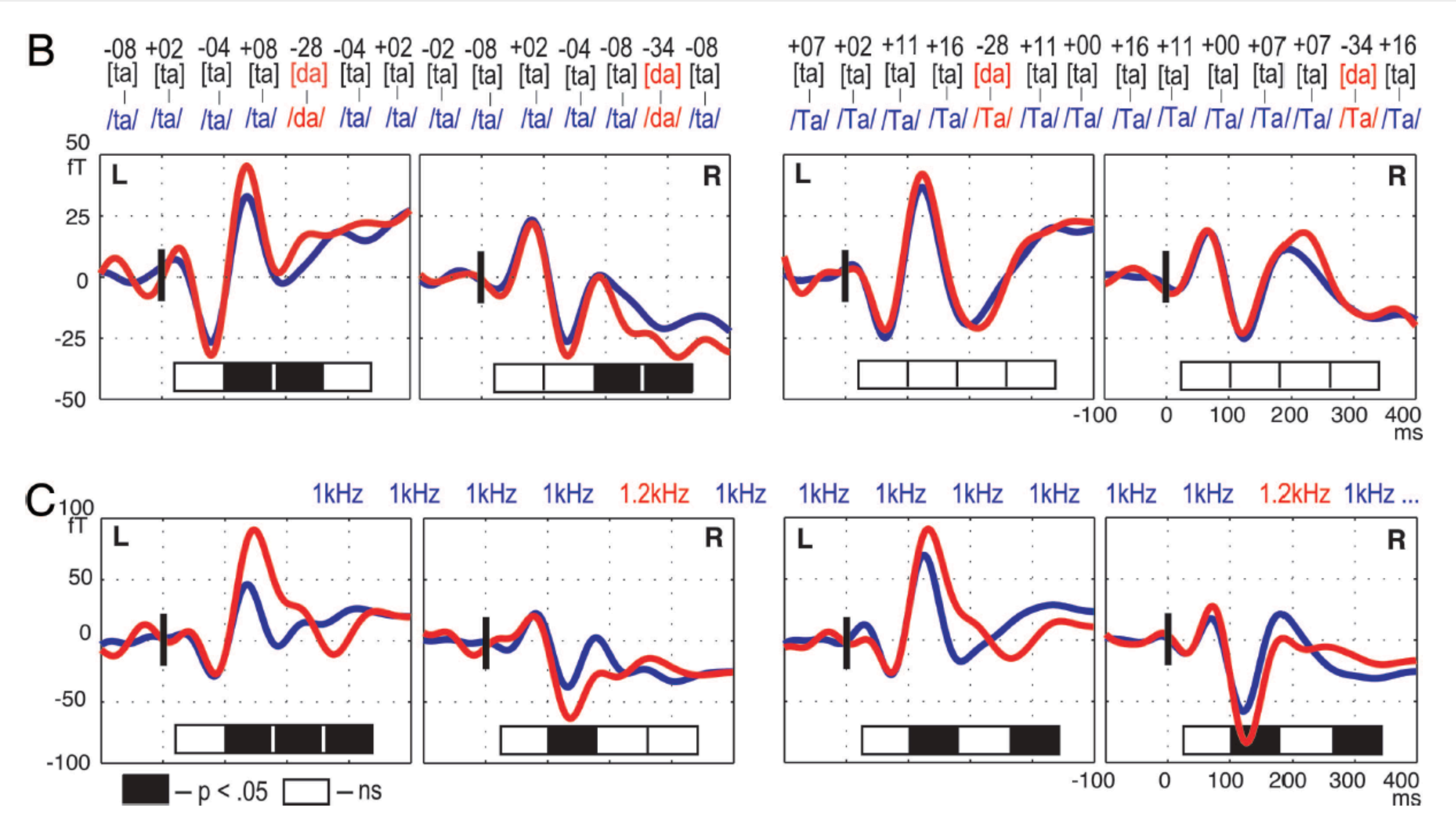
- MMN is sensitive to continuous acoustic differences
- Is it also sensitive to the *categorical* differences at the level of speech perception?
  - Yes





Kazanina et al (2006) PNAS

Russian Korean



QUESTION 1: SPARSE OR FULL CODING?

BUILDING BLOCKS OF PHONES: DISTINCTIVE FEATURES

[t]

- [+plosive]
- [+coronal]
- [+anterior]
- [-voice]

[d]

- [+plosive]
- [+coronal]
- [+anterior]
- [+voice]

BUILDING BLOCKS OF PHONES: DISTINCTIVE FEATURES

Strong *representational* claim:

- They are the smallest building bocks of language (atoms of speech)
- Used to store sounds in the mind

SPARSE OR FULL CODING?

Sounds with **[CORONAL]** place feature **[t, d, s, z, n]** are a bit special

- Very common in the world’s languages
- Often change on historical scale (by assimilation or deletion)
- Assimilate to other sounds, but reverse is not true
  - ”Sweet boy” often sounds as ”sweep boy”
  - But ”sweep toy” never sounds as ”sweet toy”

SPARSE OR FULL CODING?

Sounds with **[CORONAL]** place feature **[t, d, s, z, n]** are a bit special

Proposal 1: some kind of *default* value

Proposal 2: do *not* need to be *stored* in long-term memory

- for e.g.: **Underspecification**

SPARSE OR FULL CODING?

[CORONAL] does not need to be stored in long-term memory (lexical storage)

But you still need it for pronunciation

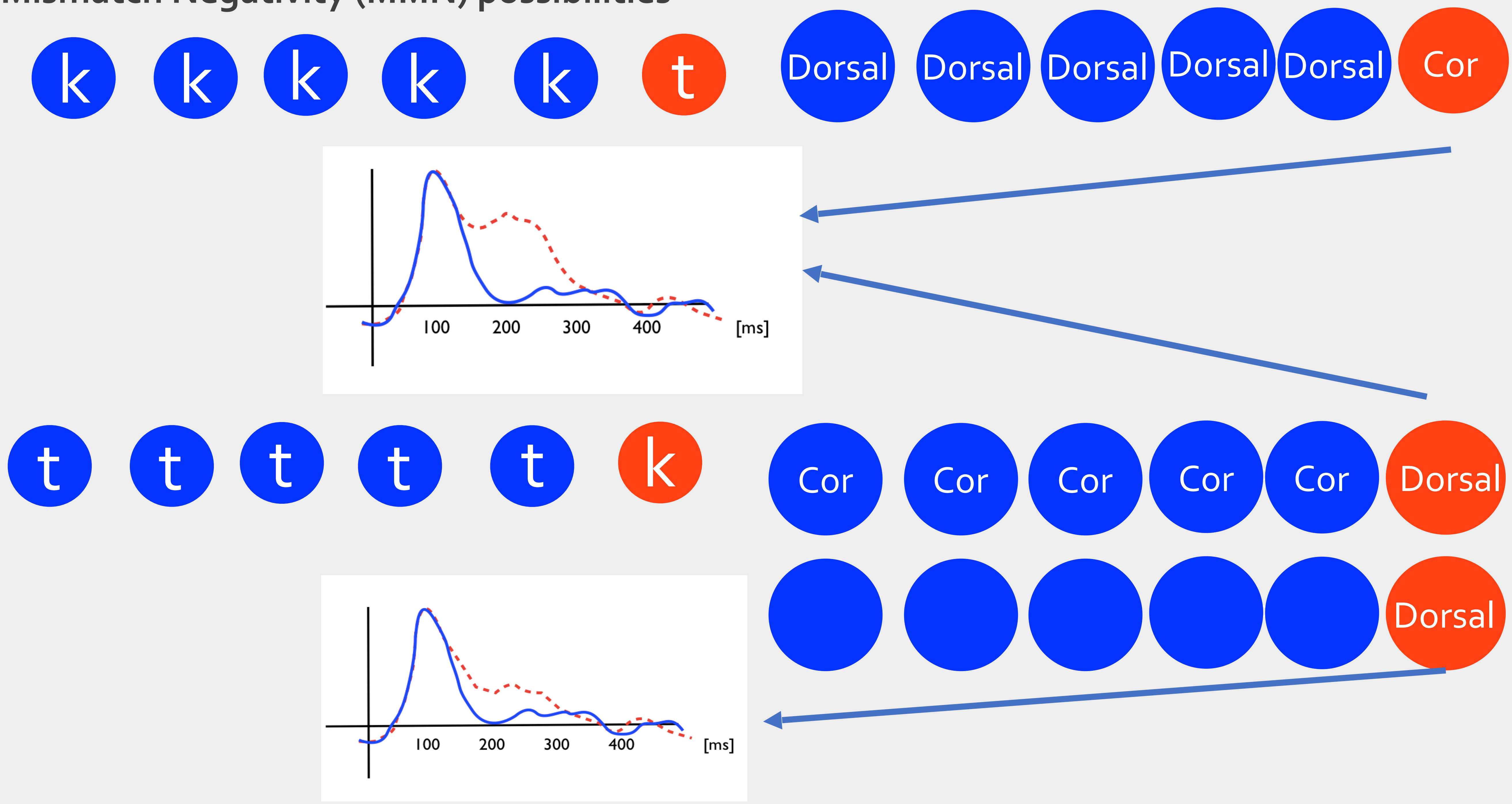
- Inserted by default if place slot is empty

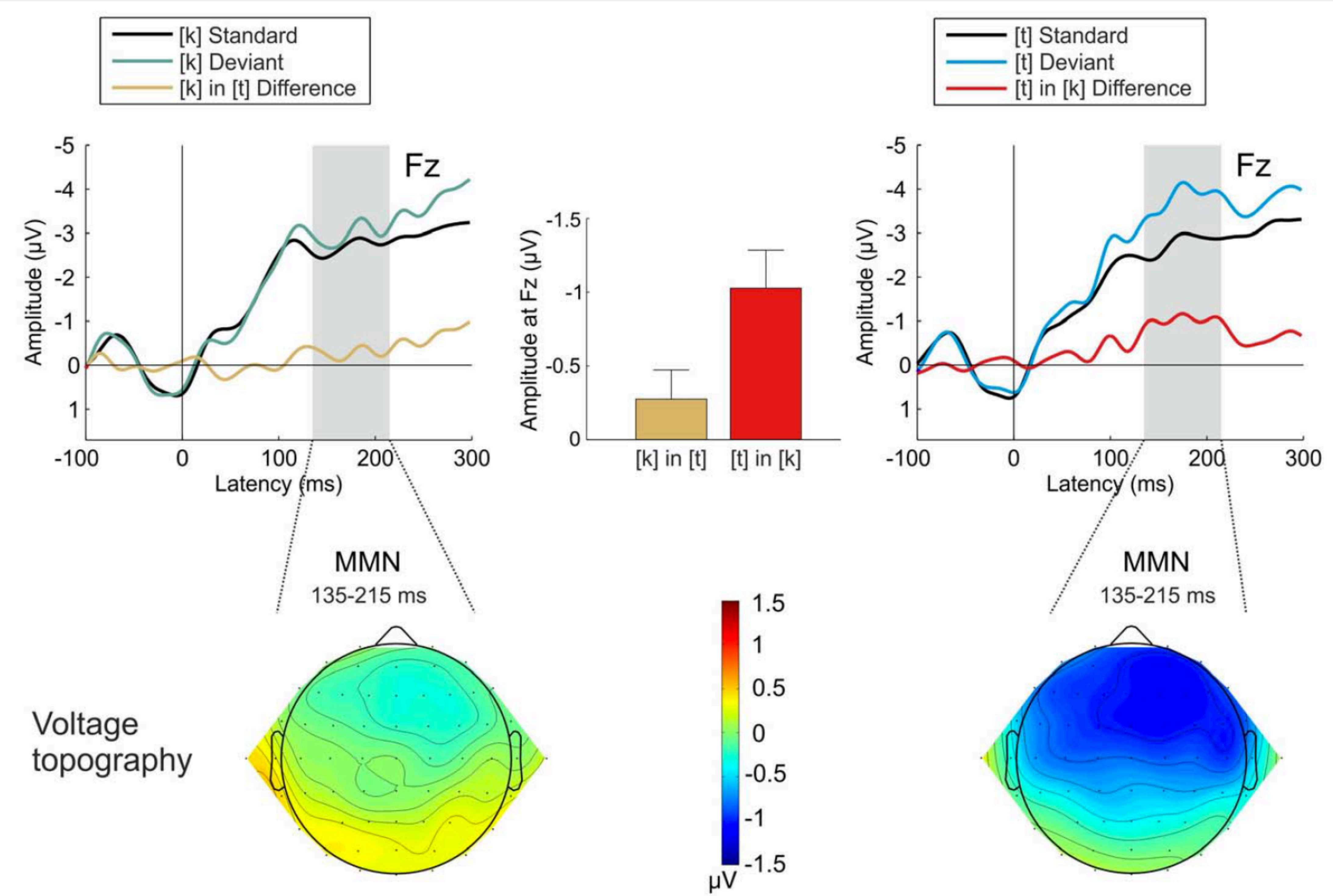
MMN expectations

Phonetic level: [Coronal] [Coronal] [Coronal] [Coronal] [Coronal]

Long Term memory: [Coronal] … -> … [ ]

Mismatch Negativity (MMN) possibilities





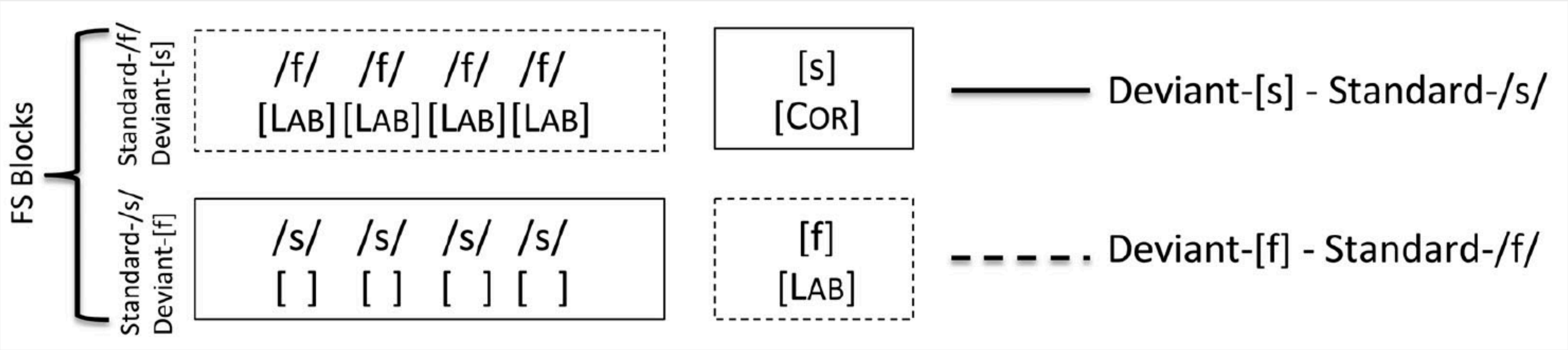
Scharinger et al (2013)

SPARSE OR FULL CODING?

(SCHLUTER, POLITZER-AHLES & ALMEIDA, 2016, *LCN*)

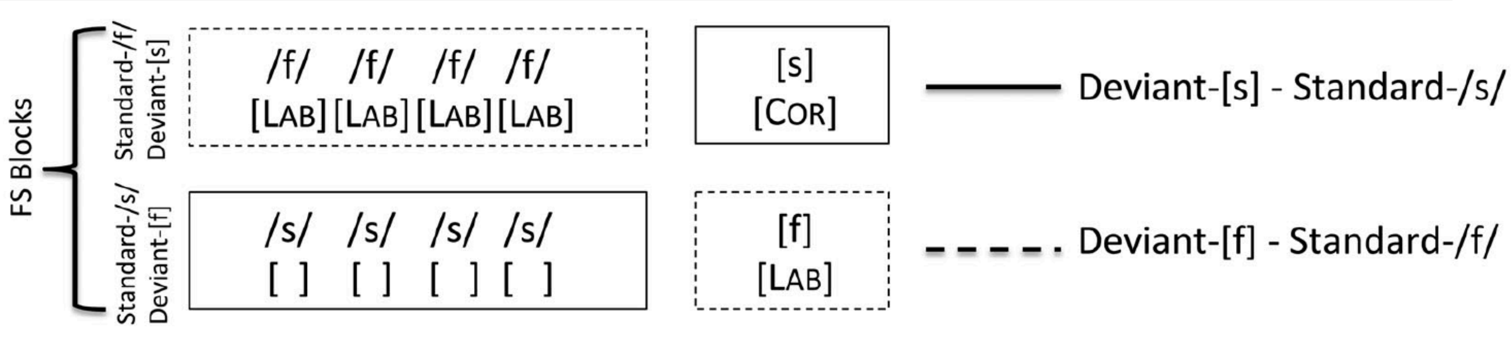
Evidence for [CORONAL] underspecification for stops: t, k

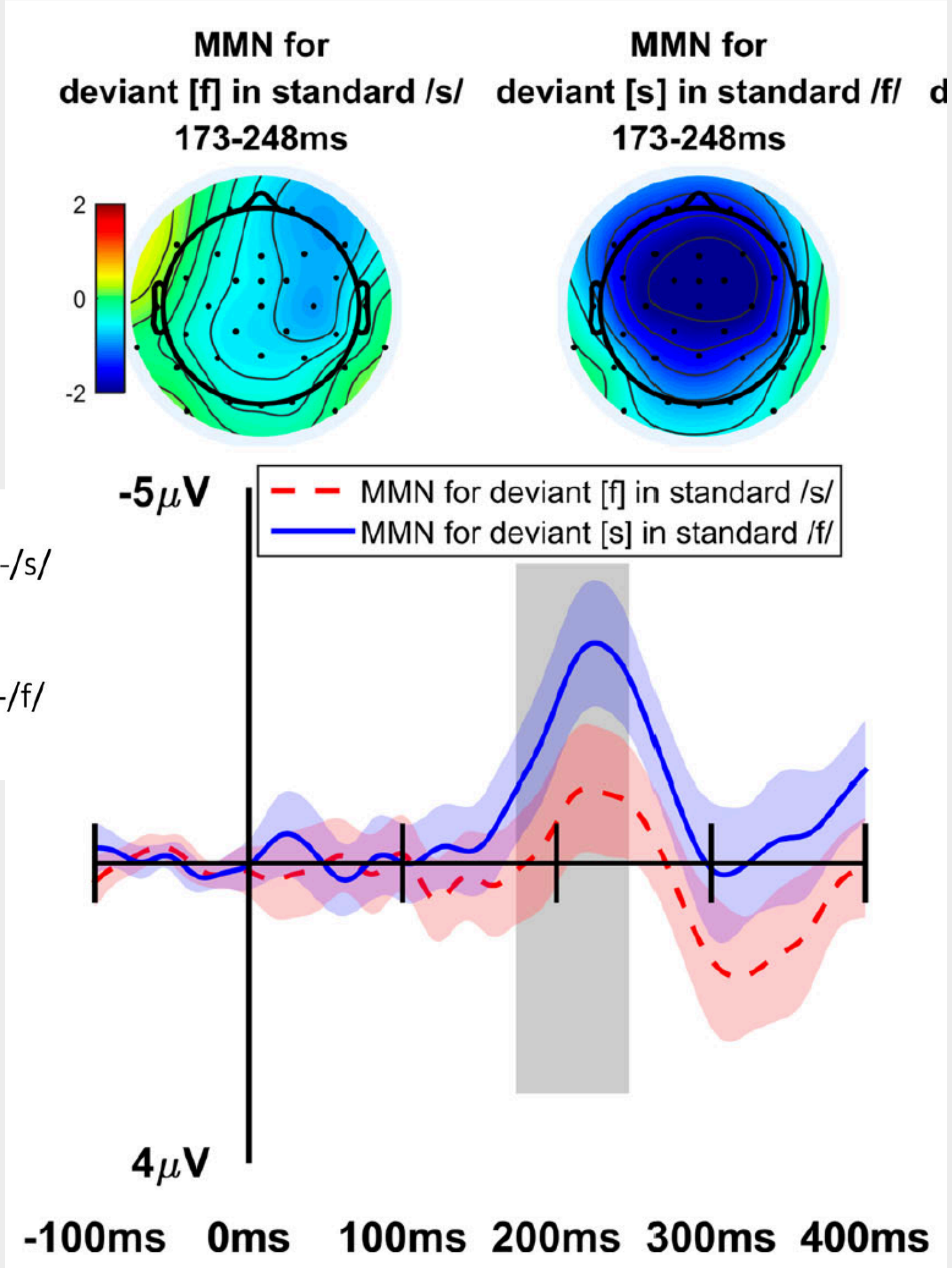
Test 1: Replication with fricatives: s, f



SPARSE OR FULL CODING?

(SCHLUTER, POLITZER-AHLES & ALMEIDA, 2016, *LCN*)





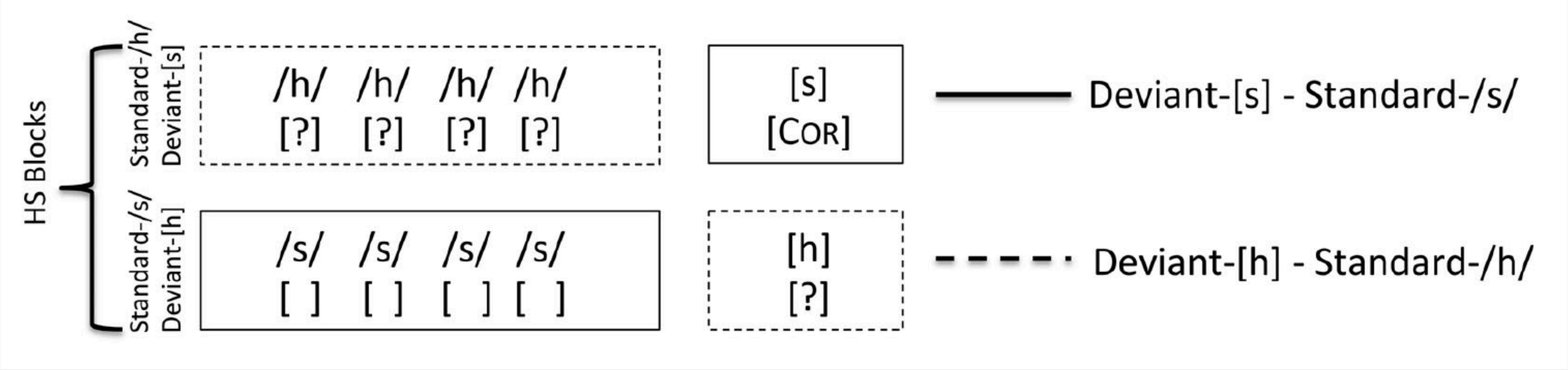
SPARSE OR FULL CODING?

(SCHLUTER, POLITZER-AHLES & ALMEIDA, 2016, *LCN*)

Test 2: Figure out the case of [h]: Also proposed to have a default place

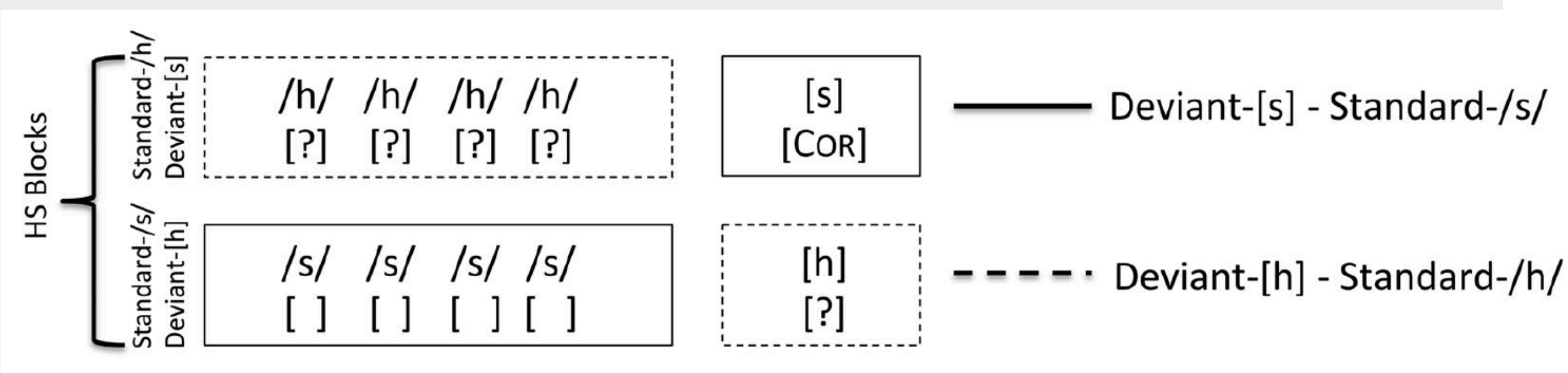
PREDICTION:

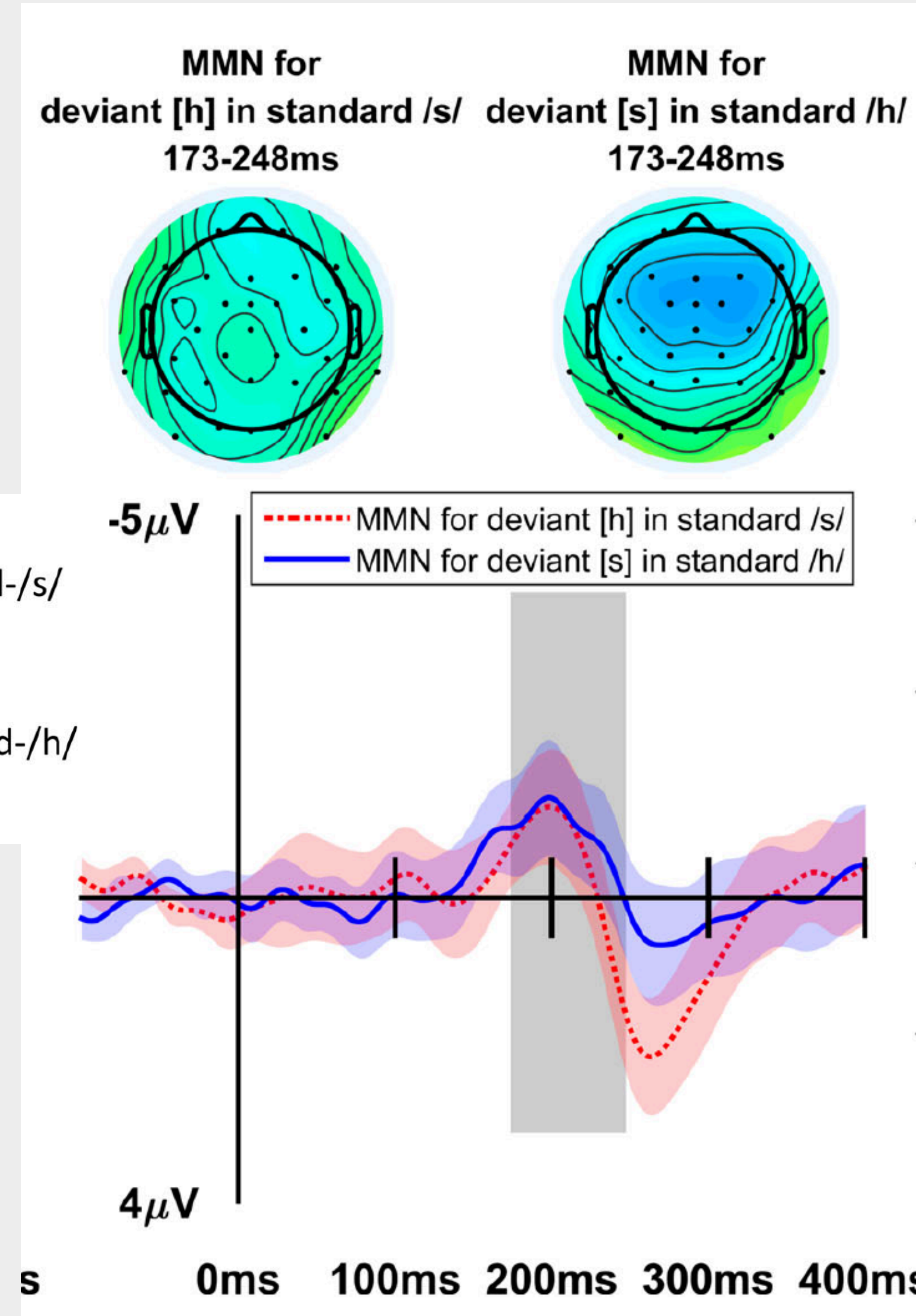
- If [h] has a place feature, then aMMN
- if not, then no aMMN?

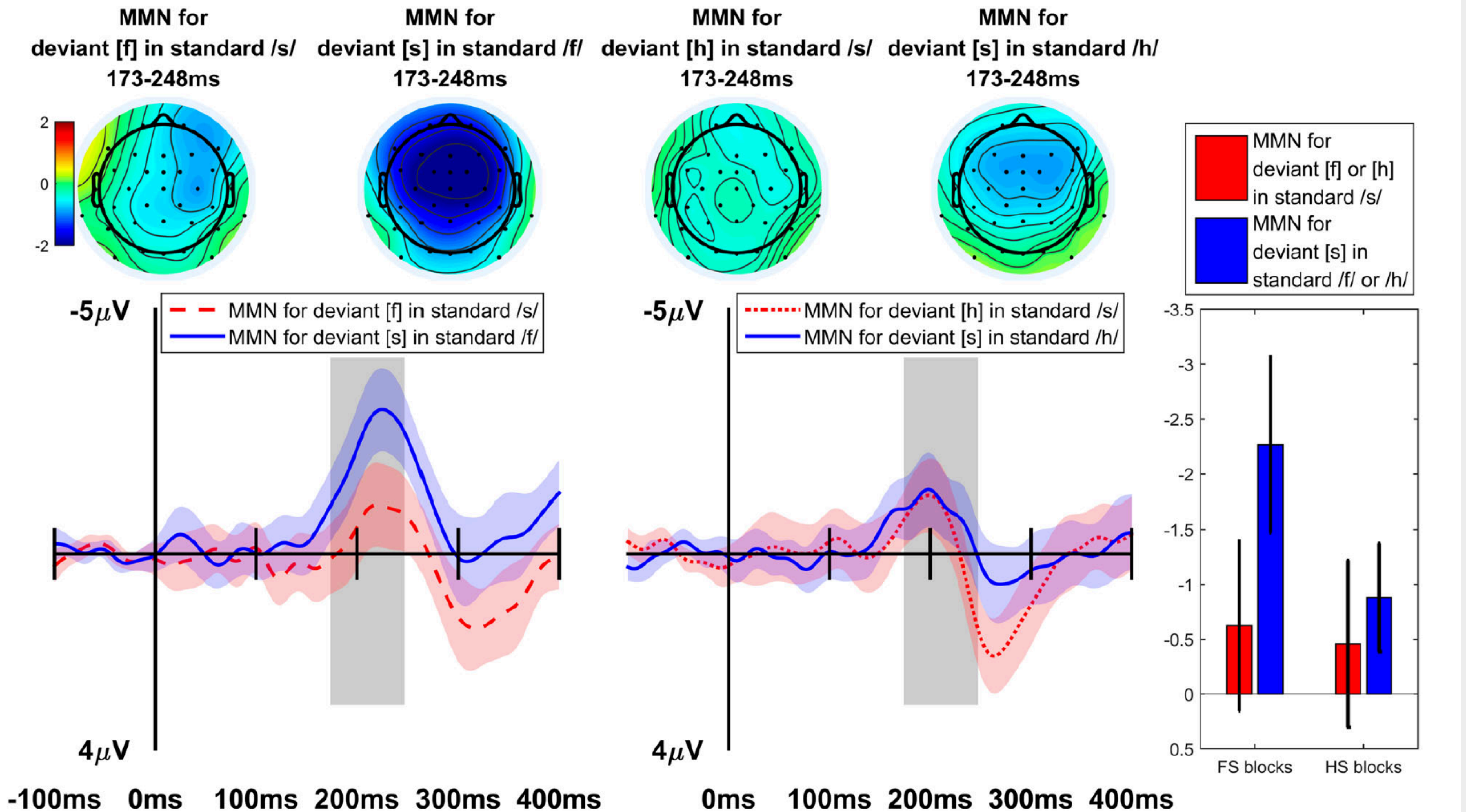


SPARSE OR FULL CODING?

(SCHLUTER, POLITZER-AHLES & ALMEIDA, 2016, *LCN*)



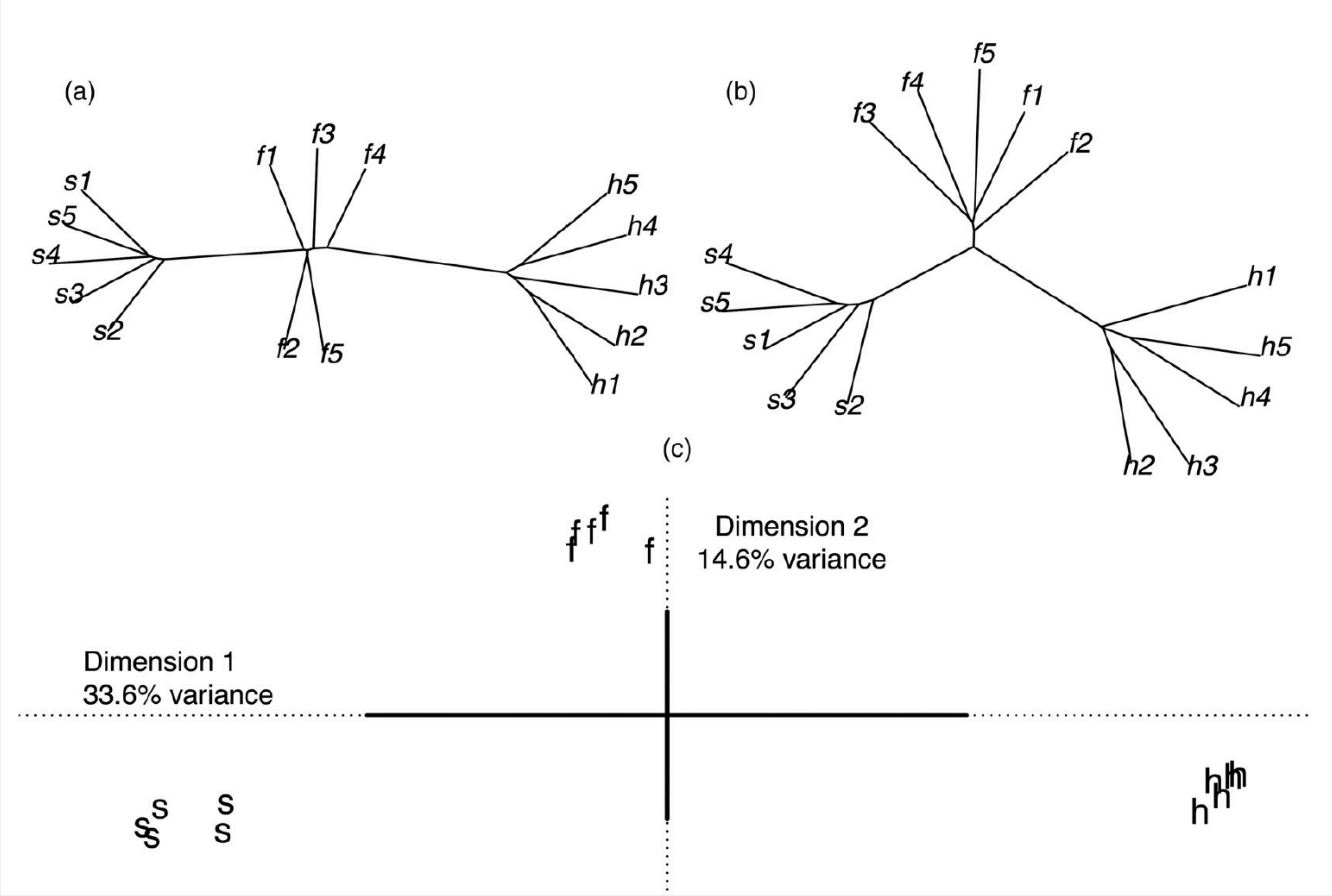




MMN for
deviant [f] in standard /s/
173-248ms
MMN for
deviant [s] in standard /f/
173-248ms
MMN for
deviant [h] in standard /s/
173-248ms
MMN for
deviant [s] in standard /h/
173-248ms
2
0
-2
MMN for deviant [f] or [h] in standard /s/
MMN for deviant [s] in standard /f/ or /h/
-5μV
MMN for deviant [f] in standard /s/
MMN for deviant [s] in standard /f/
4μV
-100ms
0ms
100ms
200ms
300ms
400ms
-5μV
MMN for deviant [h] in standard /s/
MMN for deviant [s] in standard /h/
4μV
0ms
100ms
200ms
300ms
400ms
-3.5
-3
-2.5
-2
-1.5
-1
-0.5
0
0.5
FS blocks
HS blocks

Spectro-temporal dynamic time-warping distance measure

Sliding window dynamic time-warping distance measure



Combined analysis of two distance matrices with DISTATIS

SPARSE OR FULL CODING?

(SCHLUTER, POLITZER-AHLES & ALMEIDA, 2016, *LCN*)

CONCLUSIONS

- Evidence for [CORONAL] underspecification for fricatives as well
- [h] also seems to be underspecified (did not clash with [s])

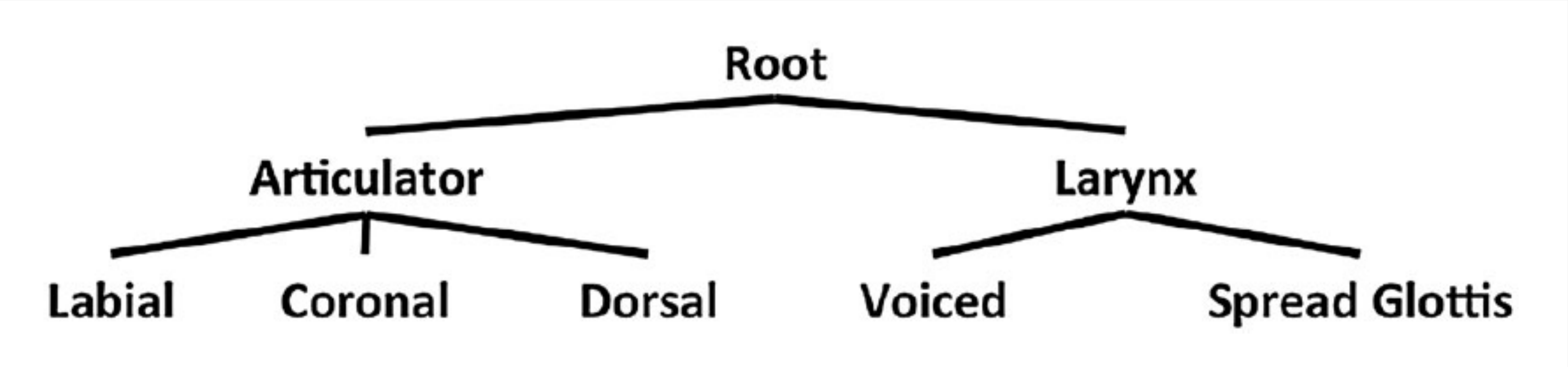
SPARSE OR FULL CODING?

(SCHLUTER, POLITZER-AHLES & ALMEIDA, 2016, *LCN*)

How does a system deal with two default values?

- Old proposal: oral vs laryngeal articulators are in different tiers

- Explains debucalization phenomena (e.g., s→h in some Spanish dialects)



SPARSE OR FULL CODING?

POLITZER-AHLES, SCHLUTER, WU & ALMEIDA (2016). JEP:HPP.

Evidence for suprasegmental (tonal) underspecification as well

Mandarin

- 4 Tones
- T3 assimilates to other tones, other tones never turn into T3

SPARSE OR FULL CODING?

POLITZER-AHLES, SCHLUTER, WU & ALMEIDA (2016). JEP:HPP.

Evidence for suprasegmental (tonal) underspecification as well

Mandarin

Same pattern of results

- T3 shows asymmetric MMN effects
  - in Chinese speakers but not English speakers
- Other Tones do not

QUESTION 2:
HOW *PHONETICALLY ABSTRACT* ARE THESE LONG-TERM REPRESENTATIONS?

HOW MUCH PHONETIC DETAIL IN LONG TERM MEMORY?

(SCHLUTER, POLITZER-AHLES, AL KAABI, & ALMEIDA, 2017, FRONTIERS)

Voicing: Amount of vibration in your vocal folds

Distinguishes pairs of otherwise identical phones:

- p/b
- t/d
- k/g
- f/v
- θ/ð
- s/z
- ʃ/ʒ
- …

HOW MUCH PHONETIC DETAIL IN LONG TERM MEMORY?

(SCHLUTER, POLITZER-AHLES, AL KAABI, & ALMEIDA, 2017, FRONTIERS)

Voicing: Amount of vibration in your vocal folds

Problem: Phonetic implementation varies

- p/b
- t/d
- k/g

  Not voiced in some languages
  e.g.: English (~ 0ms vocal fold vibration)

- f/v
- θ/ð
- s/z
- ʃ/ʒ

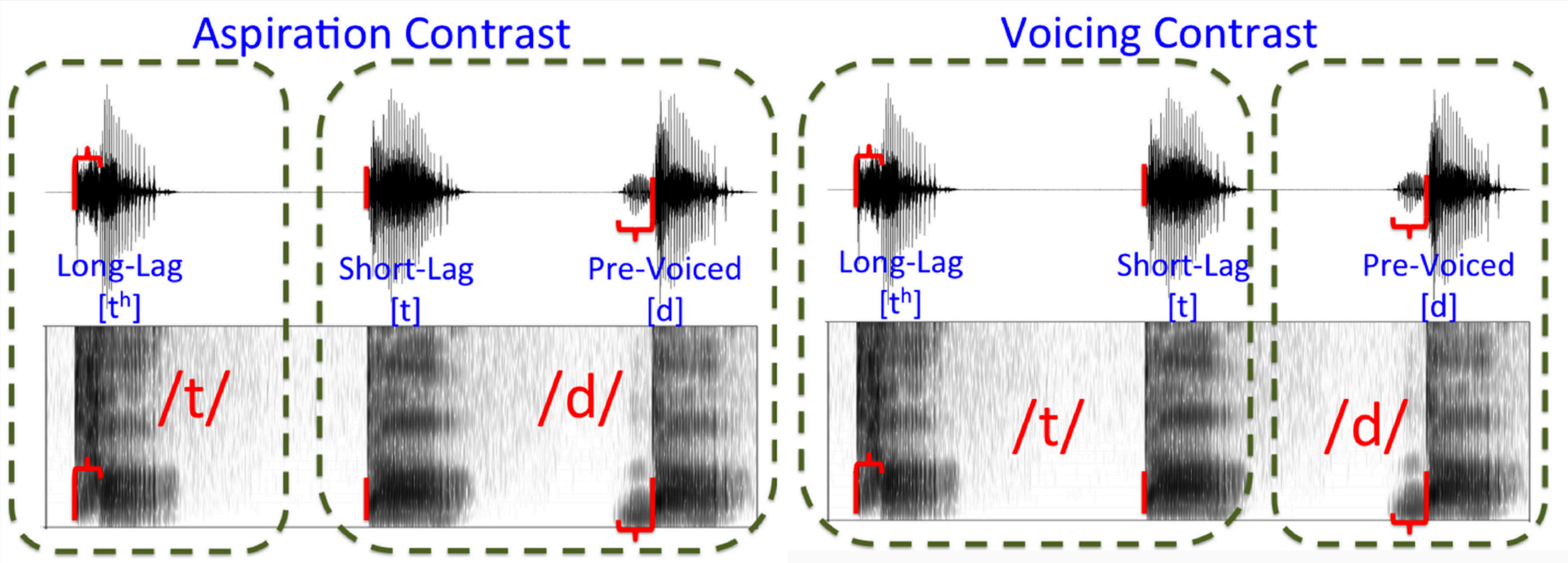
  Always voiced

- …

HOW MUCH PHONETIC DETAIL IN LONG TERM MEMORY?

(SCHLUTER, POLITZER-AHLES, AL KAABI, & ALMEIDA, 2017, FRONTIERS)

Voicing: [SPREAD GLOTTIS] [VOICE] OR [±VOICE]



HOW MUCH PHONETIC DETAIL IN LONG TERM MEMORY?

(SCHLUTER, POLITZER-AHLES, AL KAABI, & ALMEIDA, 2017, FRONTIERS)

Voicing: [SPREAD GLOTTIS] [VOICE] OR [±VOICE]

English

- With phonetic detail: T = [SG] vs D = [-SG] or [   ]
- Hestvik & Durvasula (2016) showed asymmetric MMNs
  - t t t t t d >> d d d d d t
- Compatible with T = [SG] vs D = [   ]

HOW MUCH PHONETIC DETAIL IN LONG TERM MEMORY?

(SCHLUTER, POLITZER-AHLES, AL KAABI, & ALMEIDA, 2017, FRONTIERS)

Voicing: [SPREAD GLOTTIS] [VOICE] OR [±VOICE]

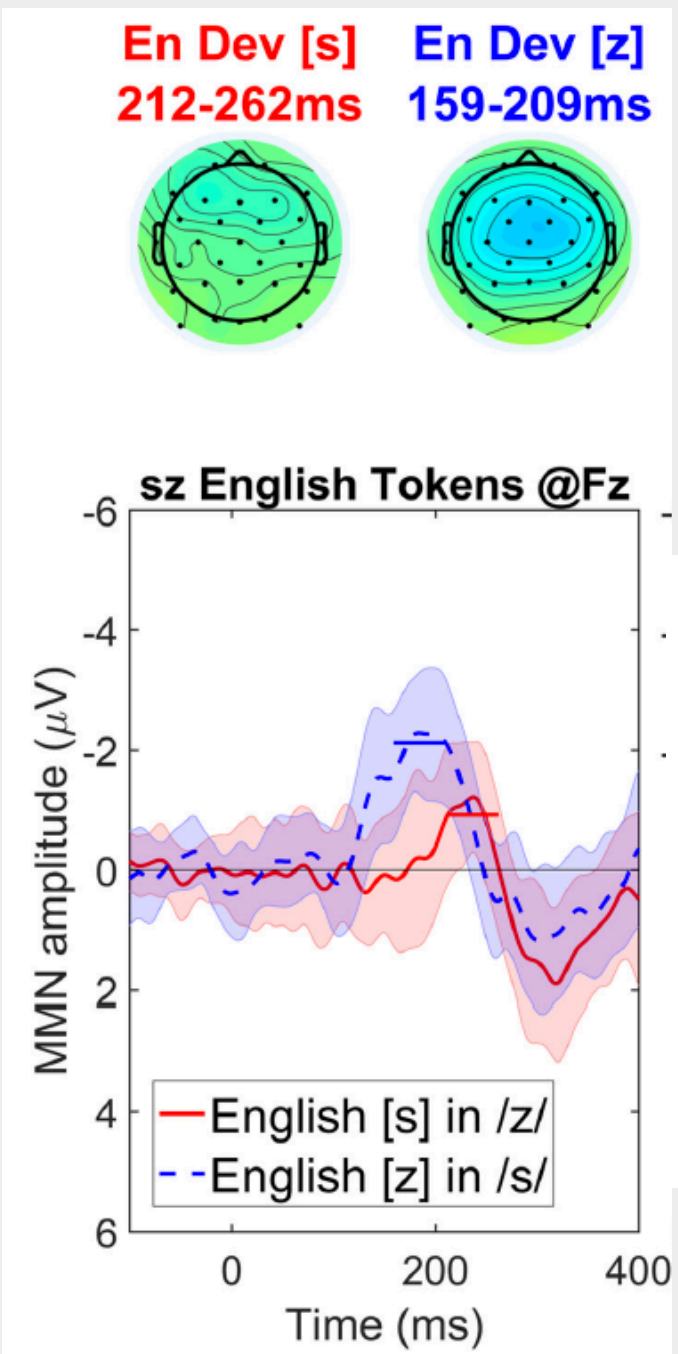
Test 1: English stops vs fricatives

- With phonetic detail: T = [SG] [ ] vs D = [ ] [ ]
- With phonetic detail: S = [ ] [ ] vs Z = [ ] [VOICE]
- PREDICTION: MMN asymmetry should *reverse* compared to Hestvik & Durvasula (2016)

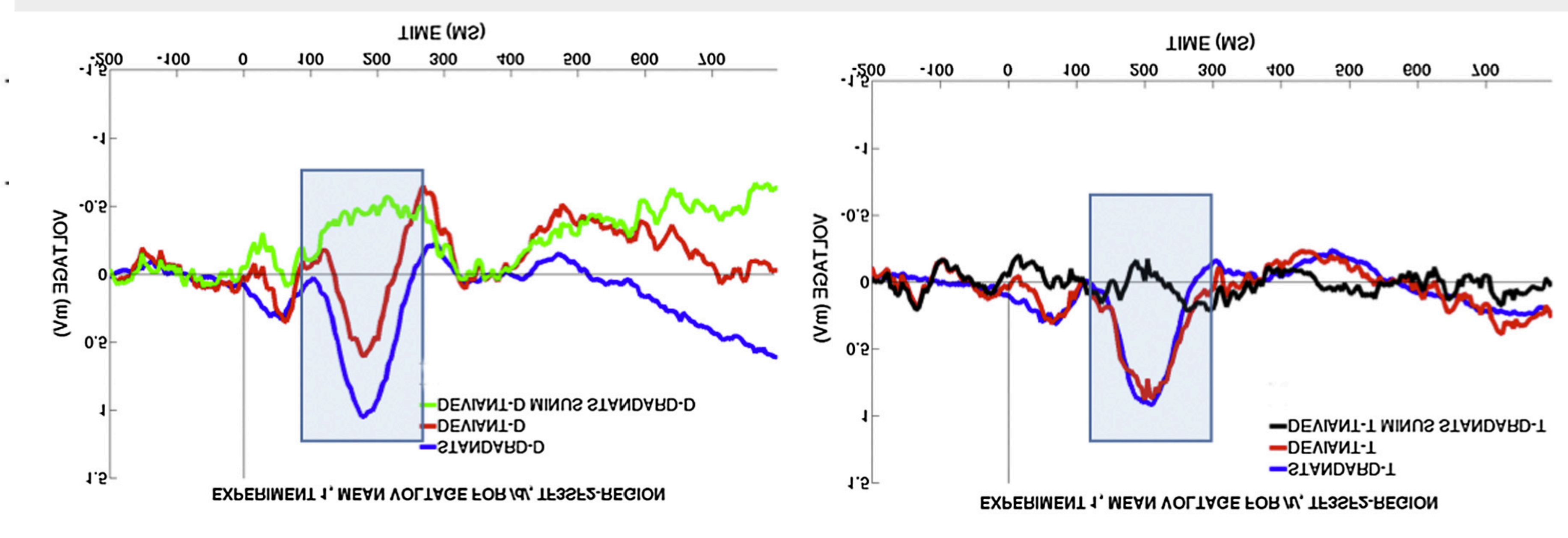
HOW MUCH PHONETIC DETAIL IN LONG TERM MEMORY?

(SCHLUTER, POLITZER-AHLES, AL KAABI, & ALMEIDA, 2017, FRONTIERS)

English speakers



Hestvik & Durvasula (2016)
(comparison is with green and black lines; voiceless lower than voiced difference waves in both cases.)



HOW MUCH PHONETIC DETAIL IN LONG TERM MEMORY?

(SCHLUTER, POLITZER-AHLES, AL KAABI, & ALMEIDA, 2017, FRONTIERS)

Voicing: [SPREAD GLOTTIS] [VOICE] OR [±VOICE]

Test 1: English stops vs fricatives

- With phonetic detail: T = [SG] [ ] vs D = [ ] [ ]
- With phonetic detail: S = [ ] [ ] vs Z = [ ] [VOICE]
- PREDICTION: MMN asymmetry should reverse compared to Hestvik & Durvasula (2016)
  - But it does not

HOW MUCH PHONETIC DETAIL IN LONG TERM MEMORY?

(SCHLUTER, POLITZER-AHLES, AL KAABI, & ALMEIDA, 2017, FRONTIERS)

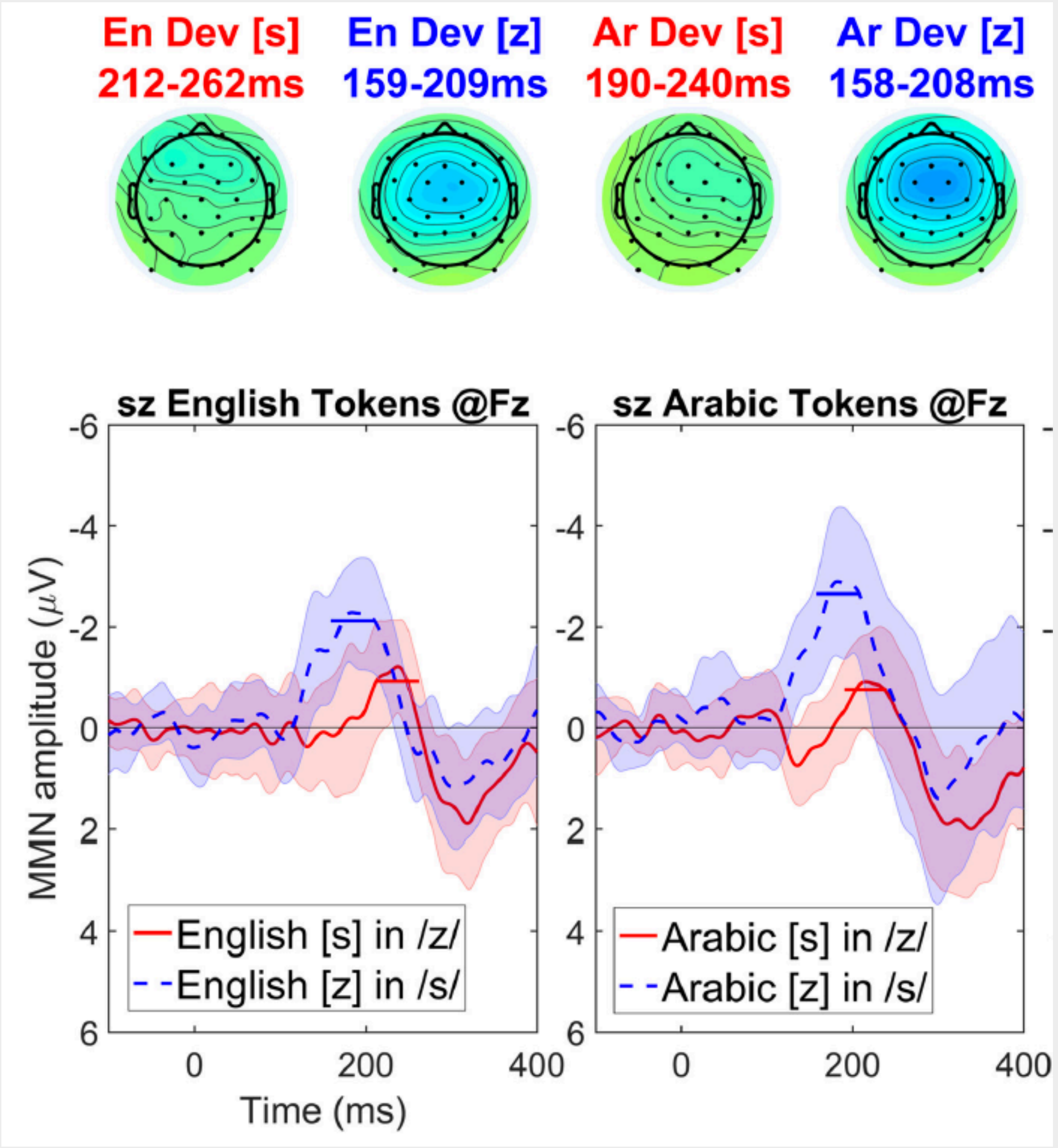
Test 2: English vs {Russian, Arabic}:

- With phonetic detail: [SG] [ ] vs [ ] [VOICE]
  - MMN asymmetry should change according to the language
- Binary, phonetically abstract: [±VOICE] vs [±VOICE]
  - Same MMN pattern across the three languages

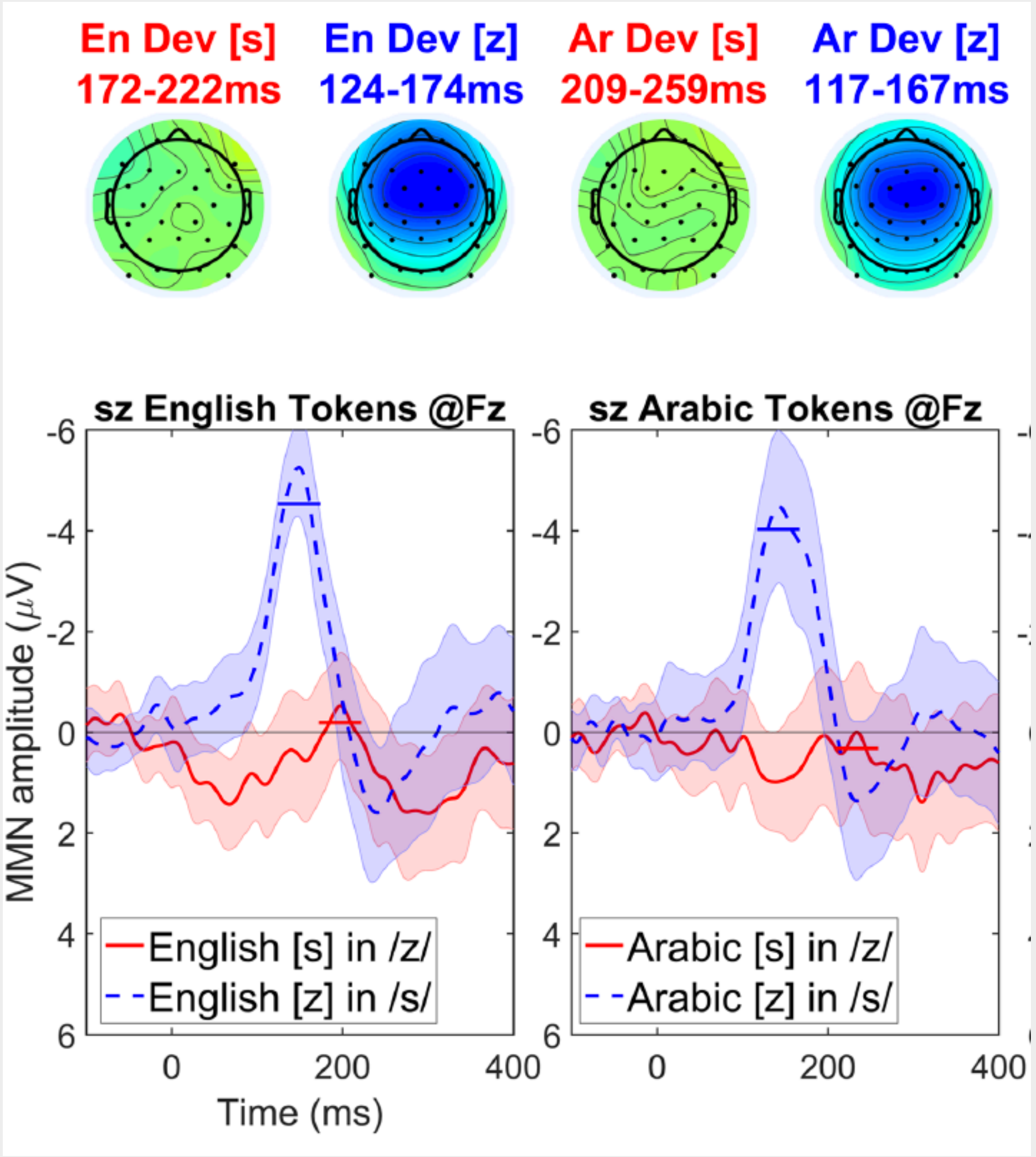
HOW MUCH PHONETIC DETAIL IN LONG TERM MEMORY?

(SCHLUTER, POLITZER-AHLES, AL KAABI, & ALMEIDA, 2017, FRONTIERS)

English speakers



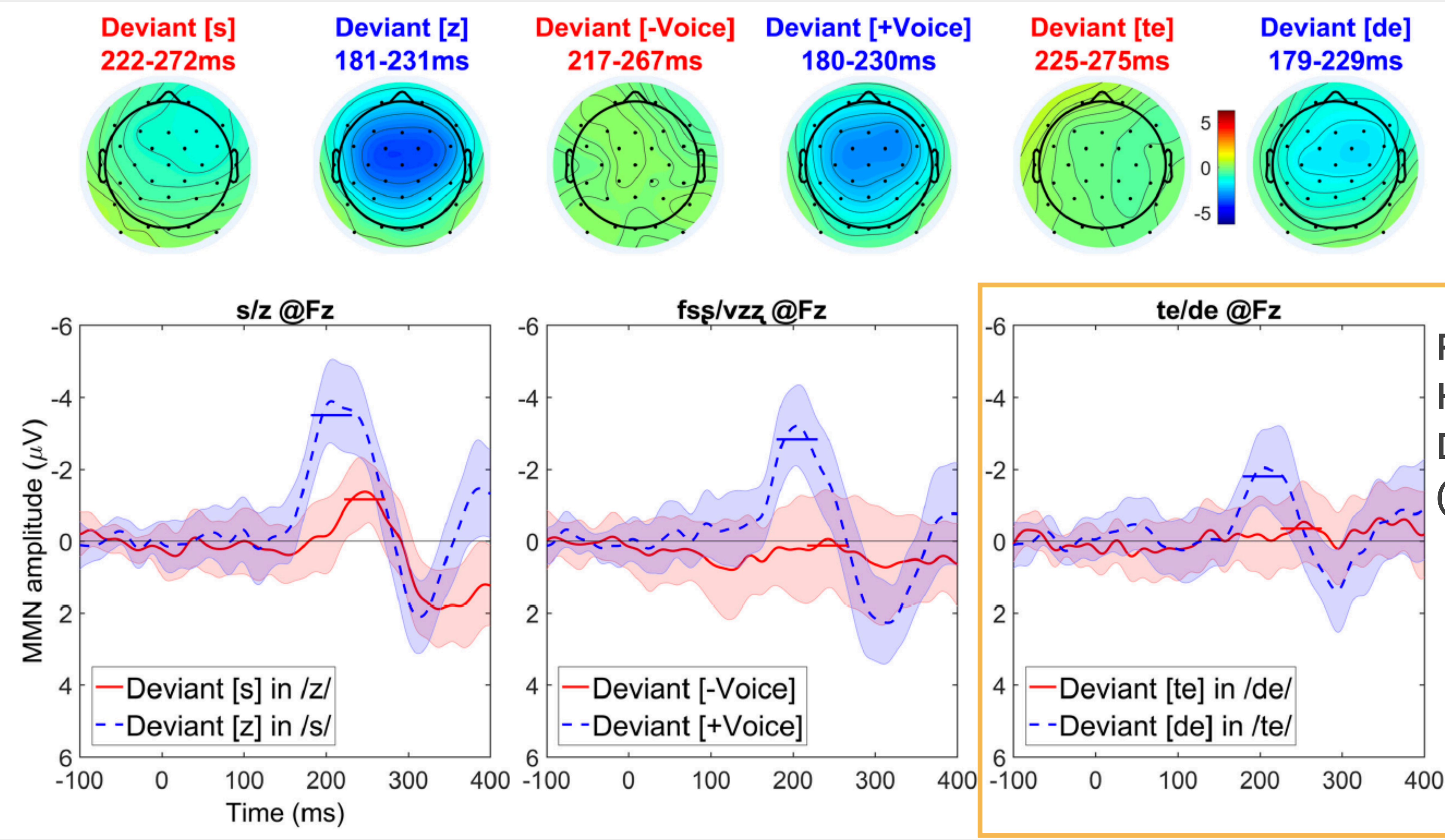
Arabic speakers



HOW MUCH PHONETIC DETAIL IN LONG TERM MEMORY?

(SCHLUTER, POLITZER-AHLES, AL KAABI, & ALMEIDA, 2017, FRONTIERS)

Russian speakers



Replicated Hestvik & Durvasula (2016)

HOW MUCH PHONETIC DETAIL IN LONG TERM MEMORY?

(SCHLUTER, POLITZER-AHLES, AL KAABI, & ALMEIDA, 2017, FRONTIERS)

Test 2: English vs {Russian, Arabic}:

- Same MMN pattern across the three languages
- Marking seems to be [ ] vs [-VOICE] or [ ] vs [ ] [SG]
  - Crucially, it's whatever feature is responsible for the *unvoiced* series that seems to be marked

03
CONCLUSIONS

CONCLUSIONS

Question 1: sparse vs redundant coding?

- Sparse
  - Certain features are abstracted away from the lexicon

Question 2: How much phonetic detail is encoded in long-term memory?

- LTM representations are phonetically quite abstract
  - phonetic details are abstracted away in the lexicon

Kevin Schluter

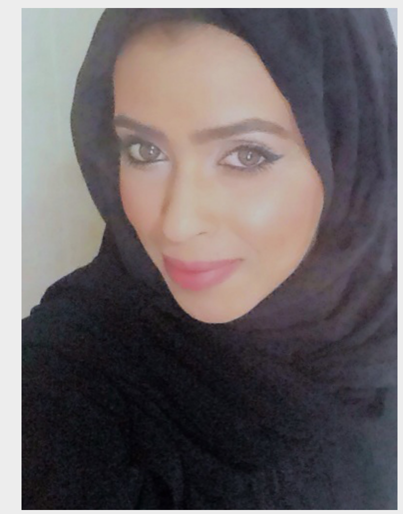
University of Minnesota

Stephen Politzer-Ahles

Hong Kong Polytechnic University

Kefei "Fisher" Wu

NYUAD'16



Meera Al Kaabi

UAE University

THANK YOU!
ANY QUESTIONS?