Audible smiles and frowns during sentence comprehension

Anne van Leeuwen, Hugo Quené and Jos van Berkum
Utrecht institute of Linguistics OTS, Utrecht University

The issue

When we speak, we use more than words alone.
Speaking = linguistic content + facial expression + posture.
Speaker stance = speaker's attitude towards content of what speaker says, expressed by e.g. facial expression + posture.

For example:
[+ smile ☺]: 'we're having a baby' → positive attitude
[+ frown ☹]: 'we're having a baby' → negative attitude

Using Event Related Potentials (ERPs), we investigated the following questions:

1. Do listeners pick up on subtle *audible* cues to speaker stance, e.g. whether speaker is smiling or frowning?
2. If listeners pick up on these audible cues, what do they do with those cues?
   - Does an audible smile or frown lead them to simply expect a positively or negatively valenced word?
   - Or can those expectations be more sophisticated? Do they go beyond simple sound-word valence associations, taking into account the compositional meaning of the entire unfolding utterance?

The experiment

Participants (n=34) listened to Dutch sentences (spoken by male speaker). No additional task. Factors varied within sentences:
Expressed valence (manipulated using Praat's LPC-8; Quené et al., 2012):
   - ☺ smiling expression
   - ☹ frowning expression
Typical valence of event for the subject:
   - Positive event or situation for subject
   - Negative event or situation for subject
Perspective:
   - First person: the event happened to the speaker (Panel A)
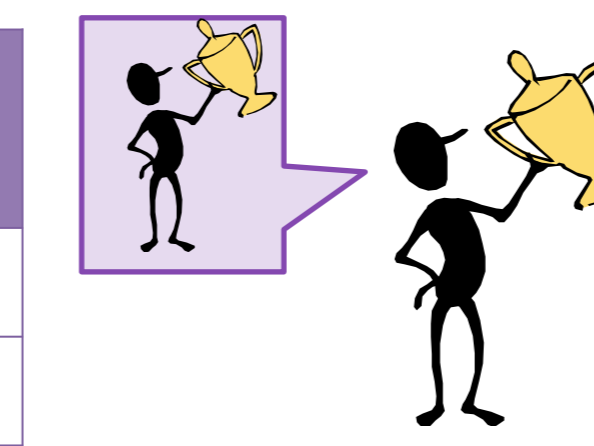   - Third person: the event happened to someone else (Panel C)

Predictions

In general, we predicted that listeners pick up on the audible cues to speaker stance. These cues should lead listeners to expect something of corresponding valence.

1. For utterances describing an event that happened to the speaker (Panel A), we predicted a clear mismatch effect for words whose valence does not match speaker's expression (as, e.g., happy events involving the speaker should make him smile rather than frown).
2. For utterances describing an event that happened to someone else (Panel C), we did not predict such clear processing costs for mismatching words, at least not with the same magnitude (as it is unclear how the speaker, on average, evaluates positive or negative events that happened to someone else).

Event happened to speaker

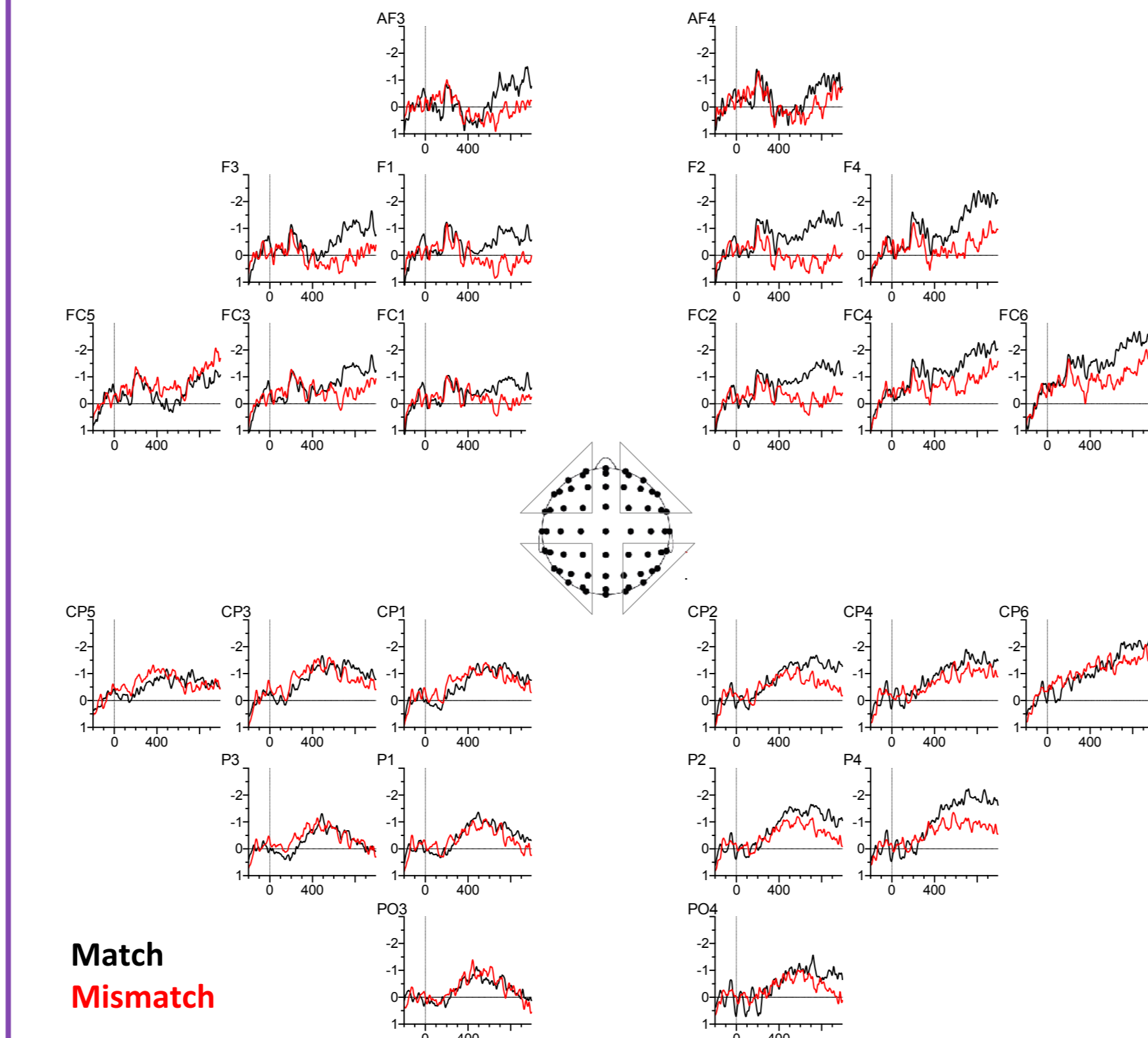### Example items – event happened to speaker

Panel A

Condition	Expressed valence	Typical valence of event for subject	Example sentence
Match	Smile ☺	Positive	'I have received a prize'
Mismatch	Frown ☹	Positive	'I have received a prize'

Results

Panel B

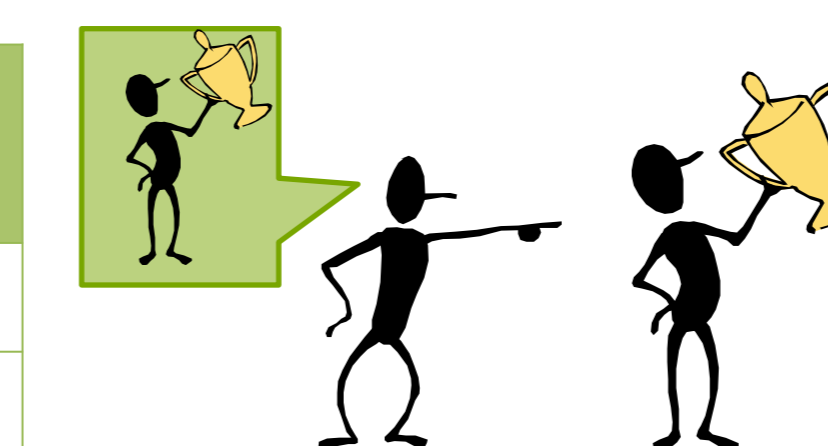ERPs to matching and mismatching items   Scalp distribution mismatch effect (mismatch - match)

Match
Mismatch
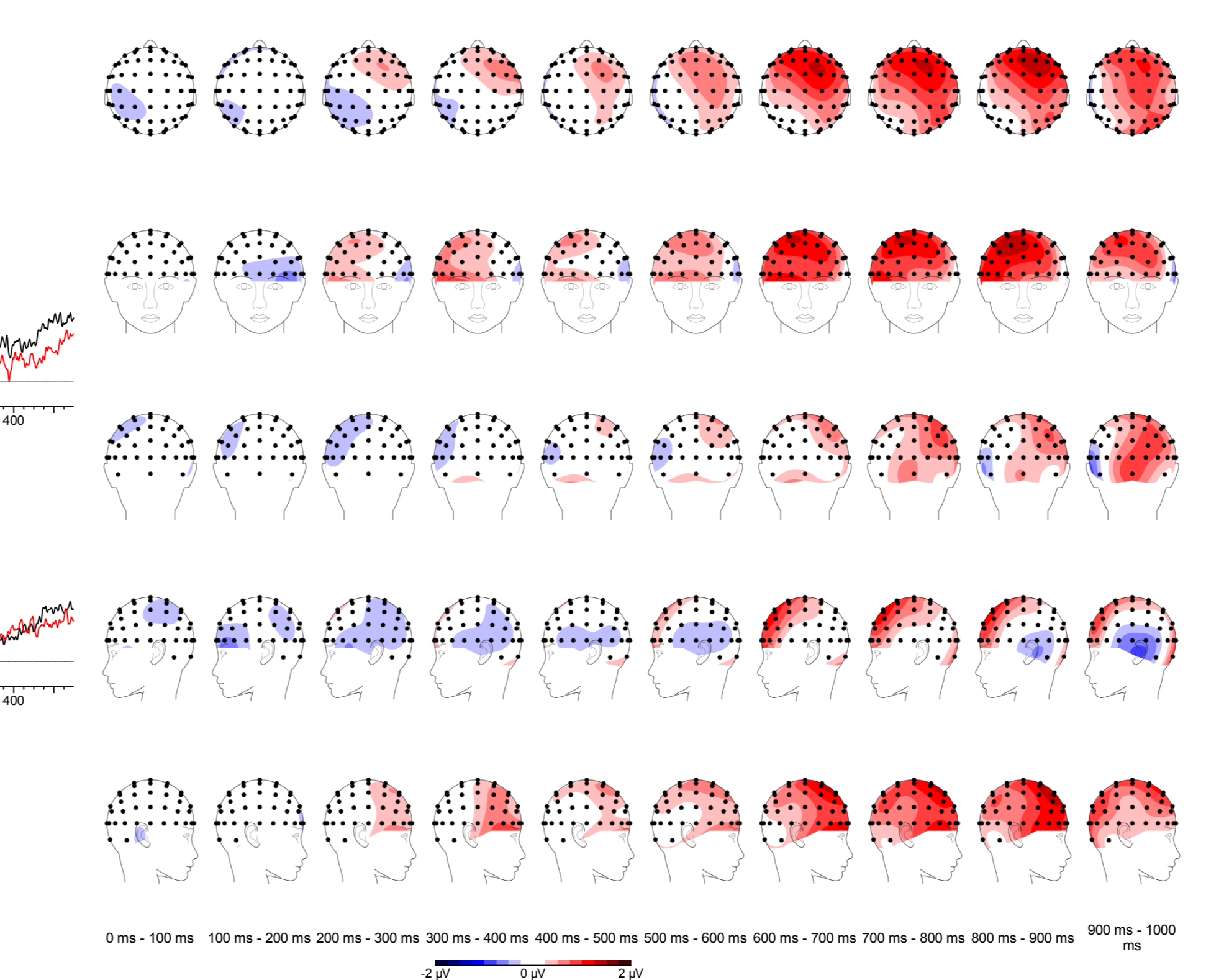
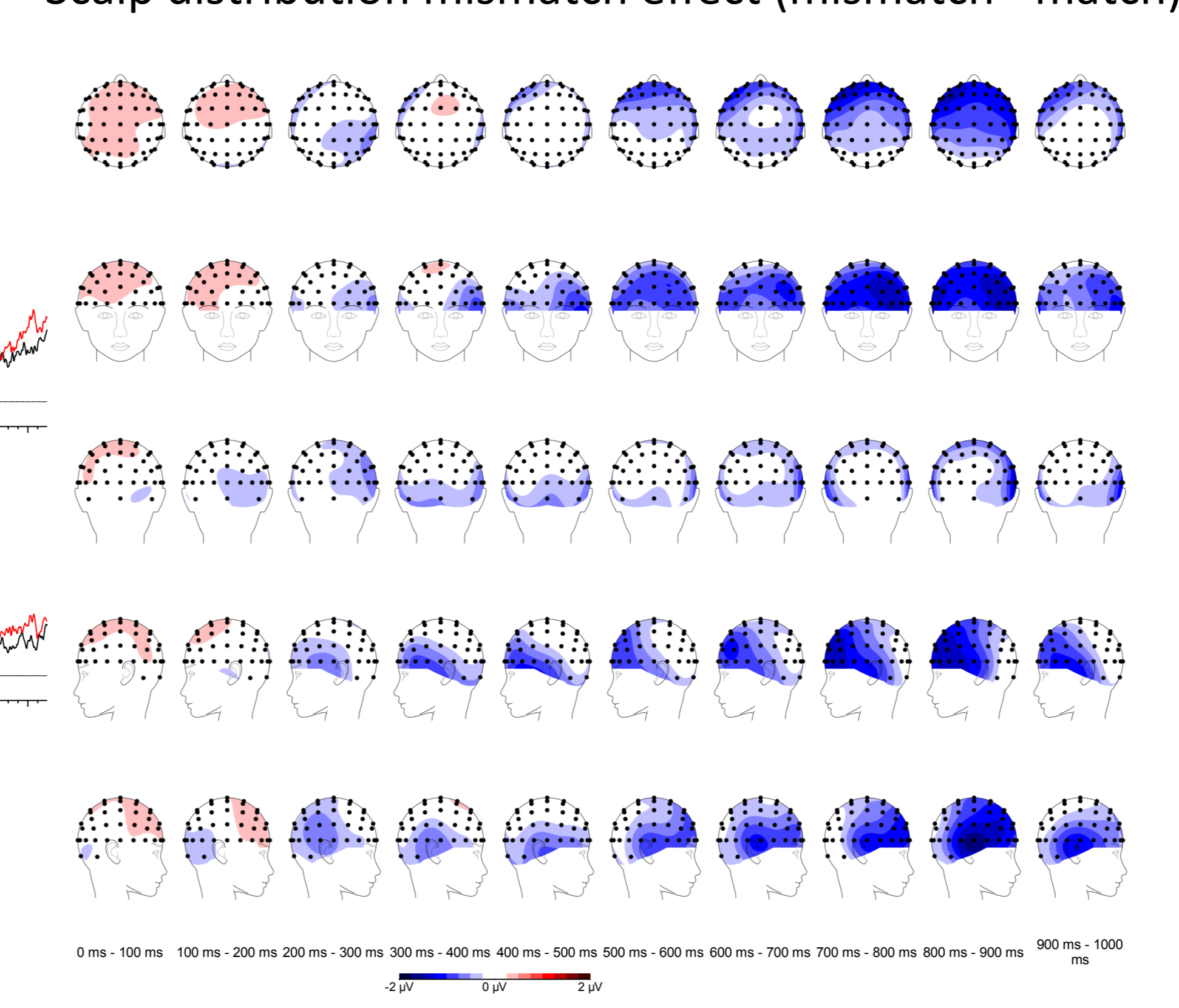**Event happened to someone else**

Example items – event happened to someone else

Panel C

Condition	Expressed valence	Typical valence of event for subject	Example sentence
Match	Smile ☺	Positive	'He has received a prize'
Mismatch	Frown ☹	Positive	'He has received a prize'
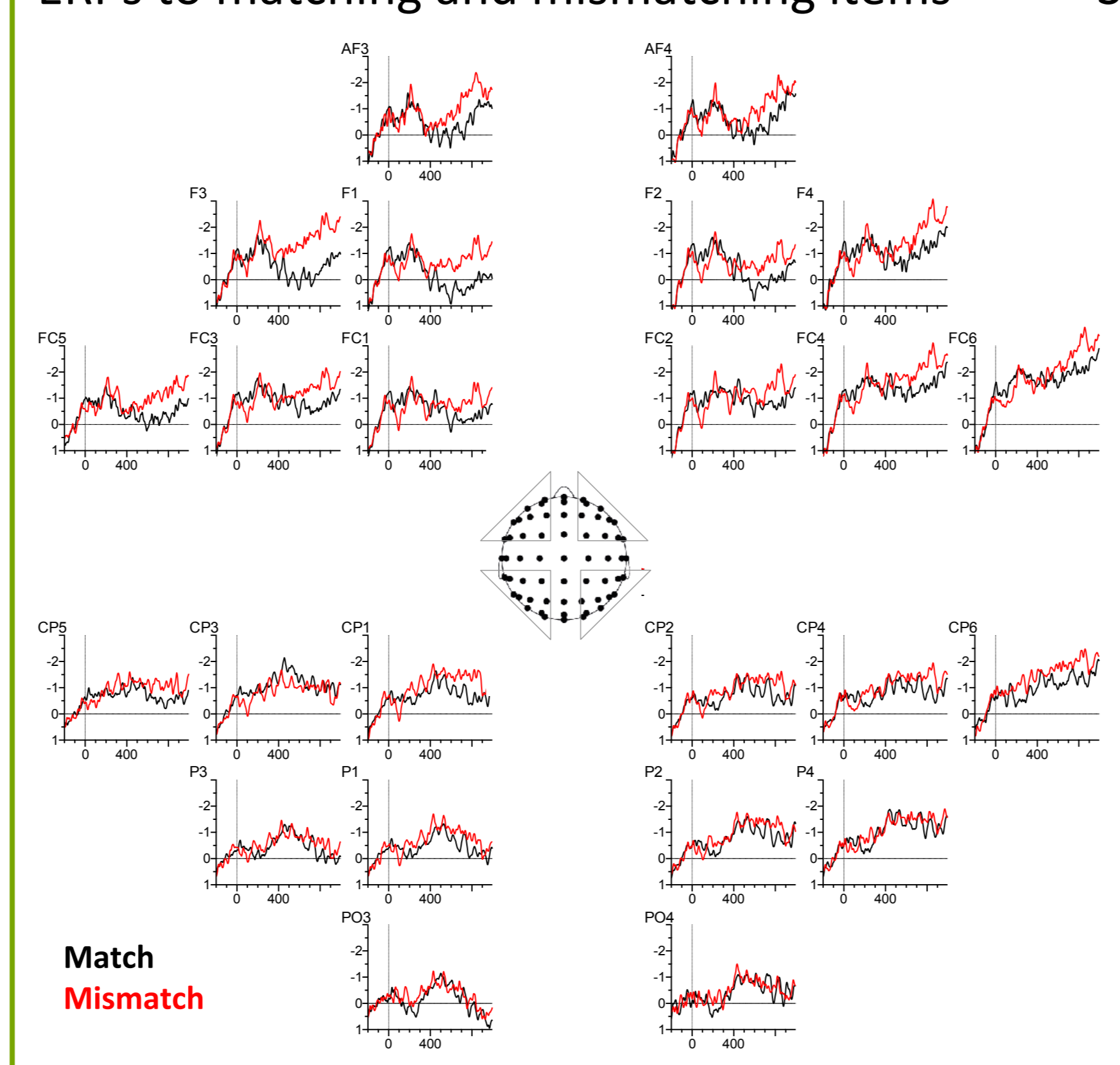
Results

Panel D

ERPs to matching and mismatching items   Scalp distribution mismatch effect (mismatch - match)

Match
Mismatch

Findings:

Events that happened to the speaker (Panel A & B):
- Mismatch effect: ERP positivity around 600-900ms
- This indicates that listeners **immediately pick up on audible cues to speaker stance**. Listeners rapidly detect if speaker's audible smile or frown mismatches the valence of an event that happened *to the speaker*.

Events that happened to someone else (Panel C & D):
- Mismatch effect: ERP negativity around 500-900ms
- Again shows that **listeners pick up on audible cues to speaker stance.**

Comparison of utterance types (Panel B vs. Panel D):
- The magnitude and direction of the mismatch effect depends on whether the positive or negative event happened to the speaker or to someone else.
- Audible smiling or frowning does not simply lead to expectancy of words of matching valence: **listeners take into account *about whom* something is said.**
- This shows that audible smiles or frowns are rapidly related to a sophisticated model of what is said.

Open questions

- Why is it that we do not find a N400 effect, while the N400 is generally involved in conceptual-semantic processing? And why are these mismatch effects mainly frontally distributed?
- Why do the mismatching sentences elicit opposite effects for sentences describing the speaker ("I") or someone else ("he")? What are the underlying processes?

References

Quené, H., Semin, G. R., & Foroni, F. (2012). Audible smiles and frowns affect speech comprehension. Speech Communication.

Contact information

Anne van Leeuwen, Utrecht University, Utrecht institute of Linguistics OTS. Email: a.r.vanleeuwen@uu.nl