

# Class 2 Intro to MMN and phon

EGG course 1

Arild Hestvik

Department of Linguistics and Cognitive Science

University of Delaware, USA

# Back to language: Mapping sound to meaning

- Phones: {[t], [tʰ], [tʰ], ....}    phonemes: /t/
- How do we measure whether an organism has phones and/or phonemes?
- Externally observable behavior by the “user”:
  - reaction time, decisions, comprehension
  - Note: behavior by itself implies a neurally instantiated symbolic system (Piantadosi and Gallistel 2024)
- Internally observable behavior:
  - in order to say “that was a /d/!” your brain first has to decode an acoustic signal and map it to a mental representation of the linguistic unit
  - Observe: Blood flow, energy use (oxygenation/deoxygenation), electrical transmission of signals between neurons, population-level neural activity (oscillations, **event-related potentials**)

# Brain level measures

• Knowledge state:

Mind/brain representation

/d/

← →

?

-we can ask, can we measure the ? inside the brain that corresponds to the abstract knowledge unit?

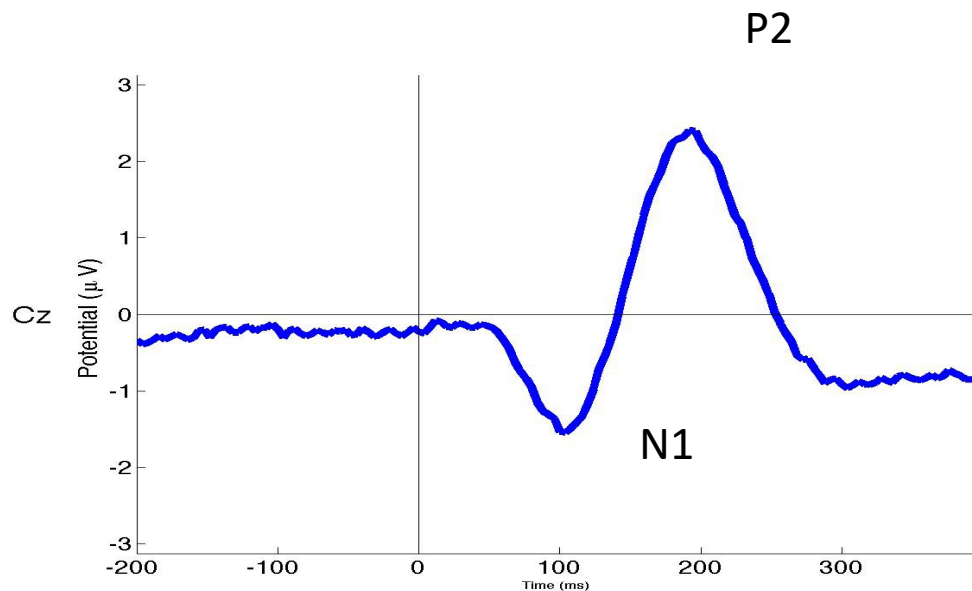
-**what do we expect to find?** Some isomorphic encoding of the same unit (in terms of the vocabulary of the brain level behavior--Poeppel)

# Brain level measures

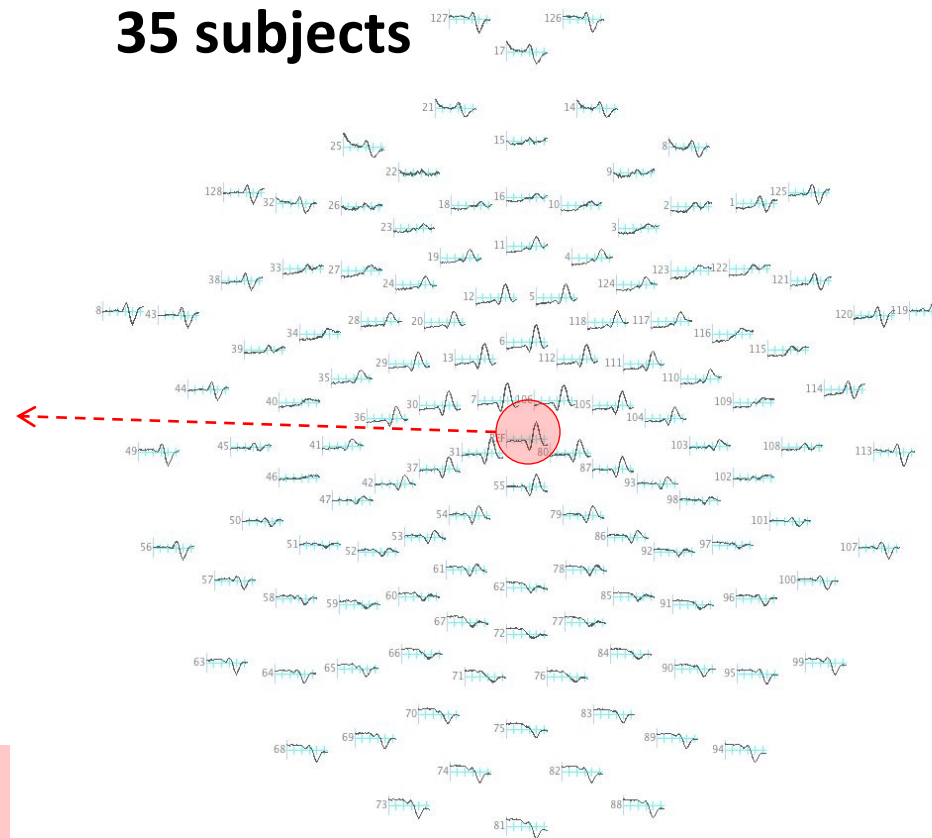
- Always ask: **Why bother?** Looking in brain is so much more work/harder...
- a scientific question in its own right: how is the knowledge system actually implemented (in humans/martians/Apple's "Siri"/chatGPT)
- we might find evidence about the nature of the knowledge system that is not discernable in externally observed behavior

# Recap: The Auditory Evoked Potential!

## Single electrode (Cz)



## 35 subjects



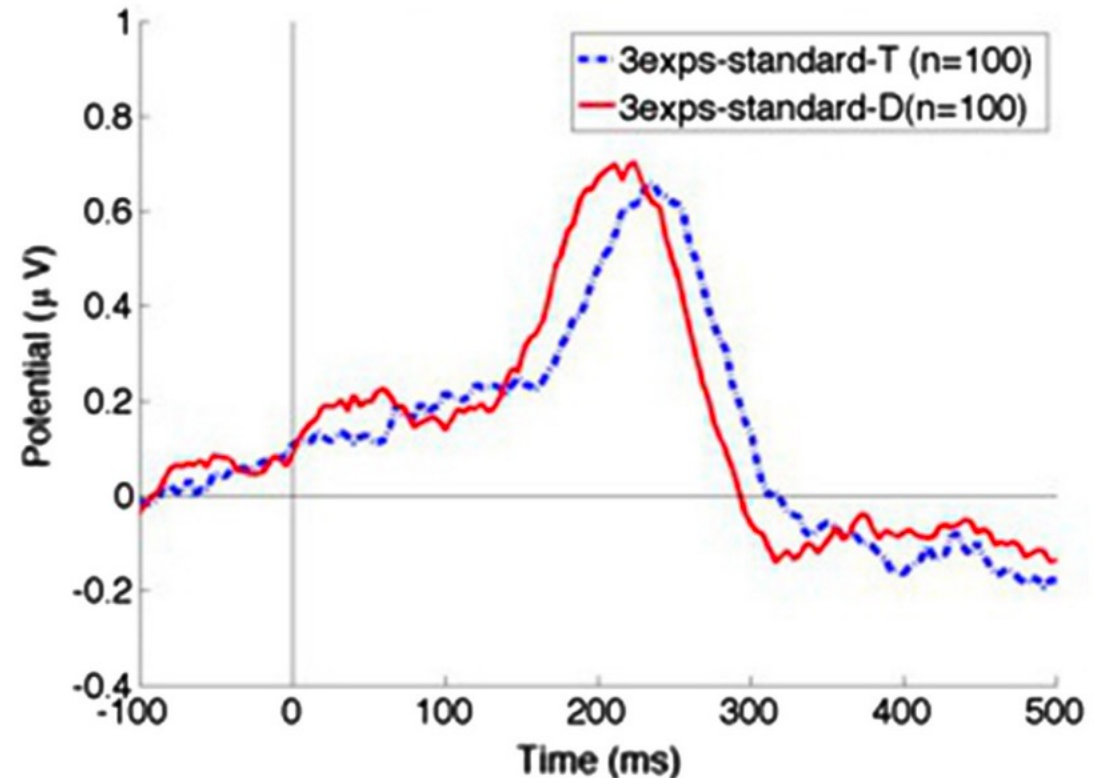
Note: This is an average of thousands of trials; it is not how a single response to the single initial sound in “the...” looks like in your brain!

## 2 important questions and answers

- Is auditory cortex is sensitive to some acoustic/phonetic properties that are NOT important for making LEXICAL distinctions?
  - --> YES
- Can we measure whether having a specific phoneme inventory modulate that response?
  - → YES

# gradient/phonetic observables important for LEXICAL distinctions

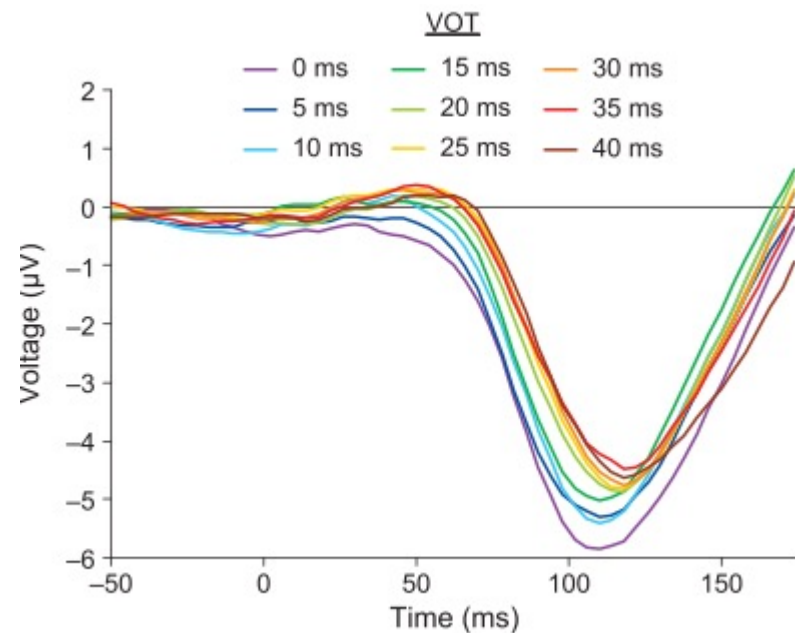
- VOT difference that separates two phonemes results in P2 latency difference



**Fig. 10.** Latency difference (but not significant amplitude difference) between the mean Auditory Evoked Potential to /d/ vs. /t/ (averaged across all tokens and all subjects in all three experiments).

# gradient/phonetic observables NOT necessarily important for LEXICAL distinctions

## Toscano et al 2010: N1 amplitude modulation related to VOT



Brain can tell how much VOT you used!

But:

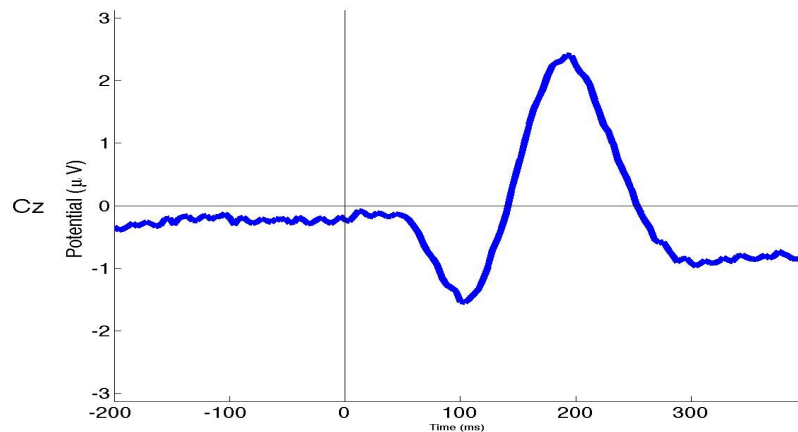
A: Whether you pronounce “dog” with 0ms VOT or 15m VOT makes no difference for activating the lexical item DOG, brain don’t care!

B: but it could make difference for whether you activate the mental concept “Speaker is a foreigner!” or not

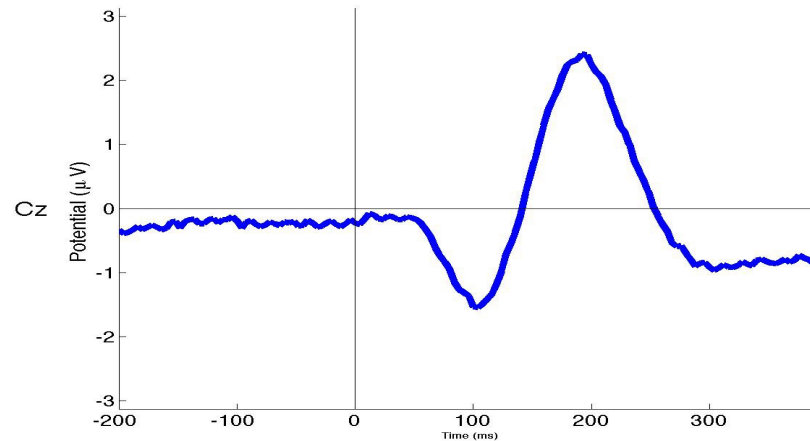
# Grammatical vs. physical distinctions

- Suppose we instead presented a voiceless “th-” (as in *think*), can we tell from the AEP whether the brain (a) respond differently or (b) responded in a way that would distinguish between voiced/voiceless?
- → **No**. AEP just says “I got a sound”. There is no unique absolute ERP signature for each sound (but Chang’s group has found populations of neurons responding to specific phon features)

[the]



[think]



Given this, how tell whether a stimulus distinction is recognized by brain as reflecting different stored PHONEMES?

→ i.e., how do an experiment that shows that /the/ and /think/ induce different neural responses?

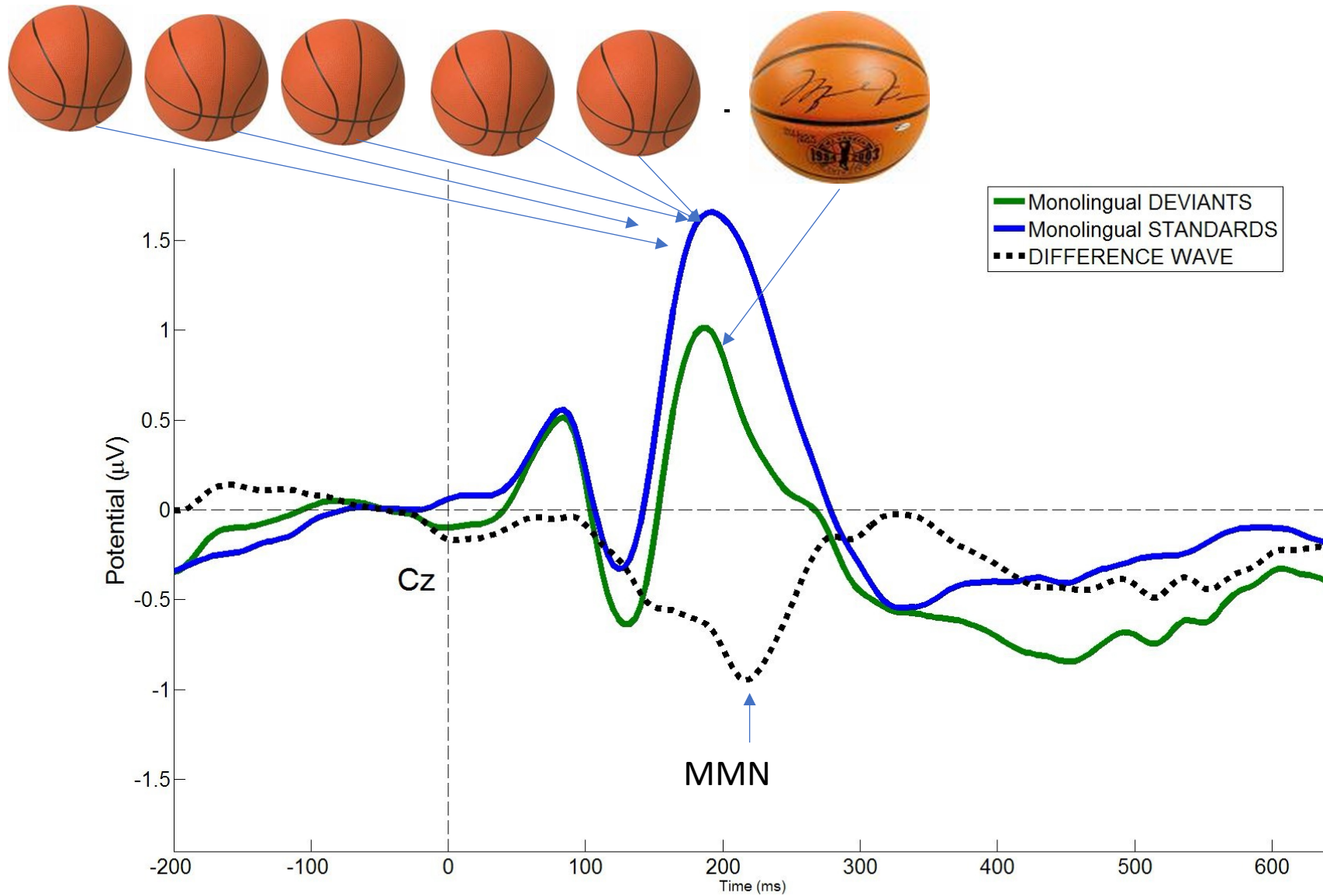
# Answer: a change detection mechanism!

- There is a way to measure whether auditory cortex detected a difference between two sounds, such that sound A  $\neq$  sound B
- (which begs the question, what counts as "not identical")

# An oddball paradigm



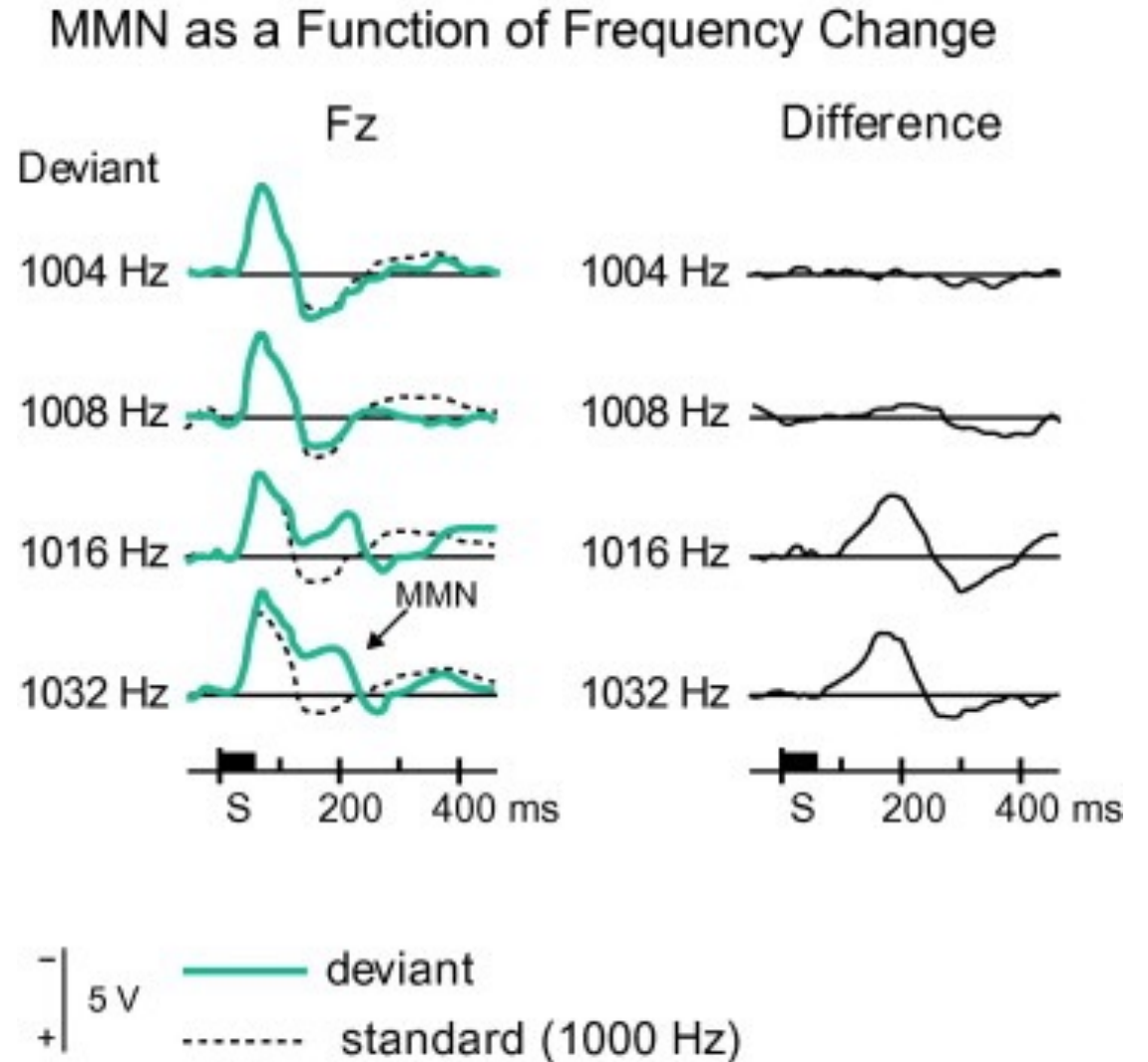
# Adult MMN



Even if the oddball is the same physical sound, cf. identity MMN

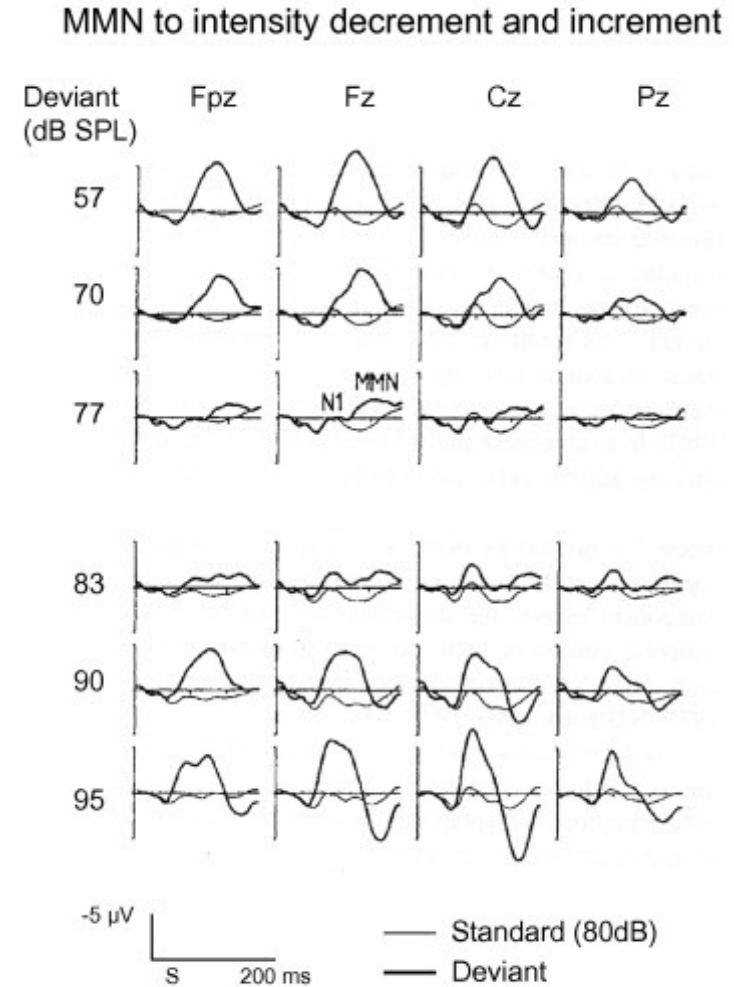
# MMN can be derived for any physical difference detectable by auditory cortex

Pure tone frequency differences are detectable by MMN



MMN can be derived for any physical difference detectable by auditory cortex

Intensity differences are detectable by MMN

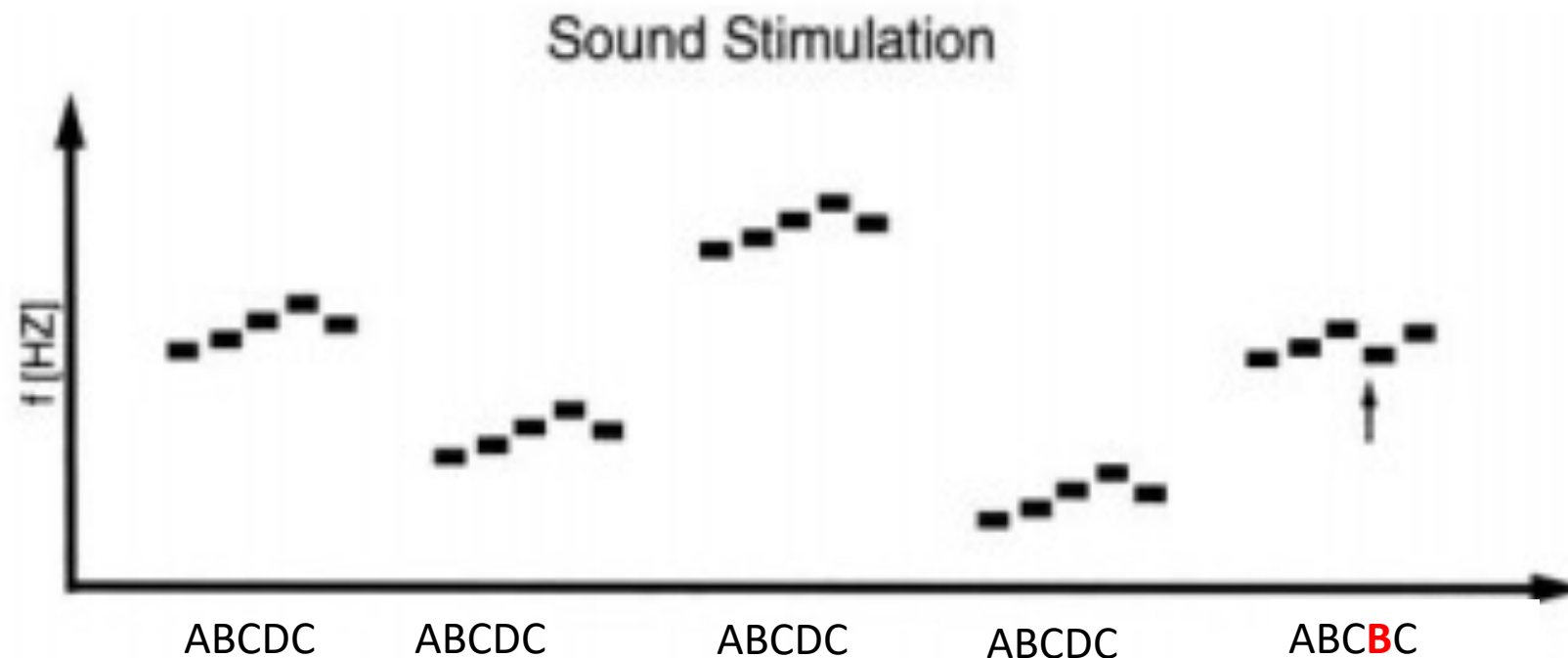


# General auditory processing

- MMN generated by
  - Pure tones, pitch, intensity, duration, ISI,...
- MMN is typically a *discrimination* measure
- **Can also be thought of as an error signal related to prediction error**

# MMN is not necessarily a response to a single physically different stimulus!

- Stimuli: 12 different melody-like tone patterns, differing in pitch frequencies (different keys): PATTERN VIOLATION



# MMN is not necessarily a response to a single physically different stimulus!

- Standard train



- Omission train



Omission MMN!

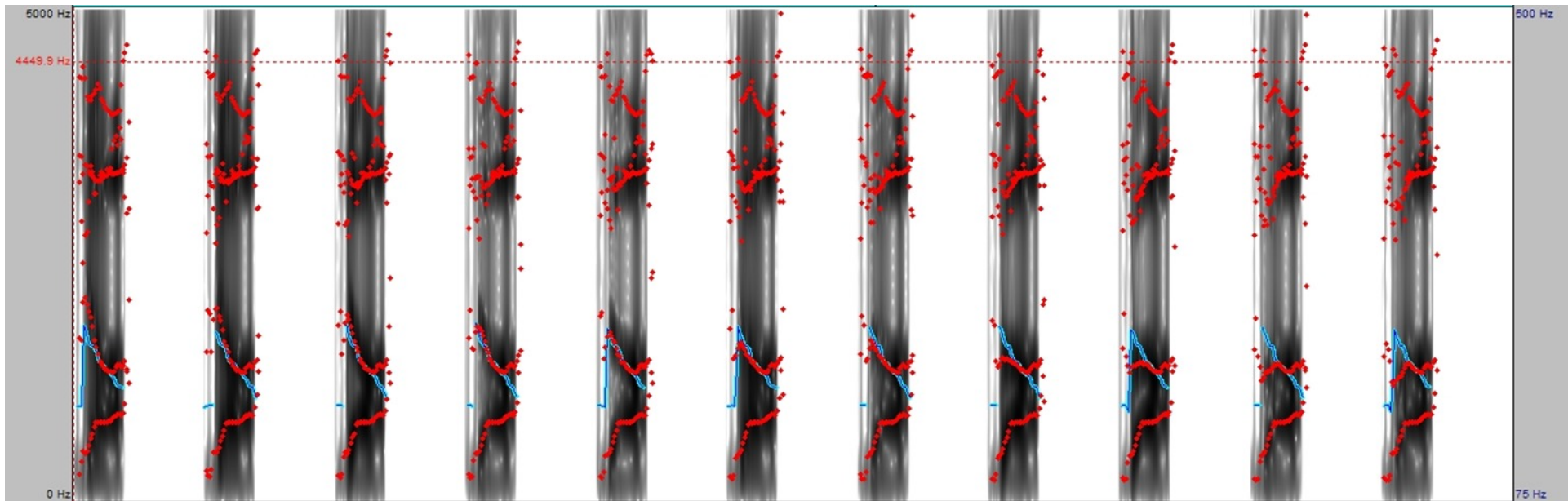
How use MMN for linguistics?

# Add grammatical content to acoustic difference

- We can use MMN to measure whether auditory cortex “cares” more about linguistically meaningful physical differences than any physical difference
- E.g., a VOT difference between 0 and 20ms are both voiced stops
  - 0ms = /d/
  - 20ms = /d/
  - 40ms = /t/
- Same VOT distance but crossing category boundaries: [ADD]

# Continuous formant change in 2AFC

- da - ba sequence



[da]



# Sharma et al 1992: Behavioral discrimination functions

- Formants continuously varied
- task: Identity the sound

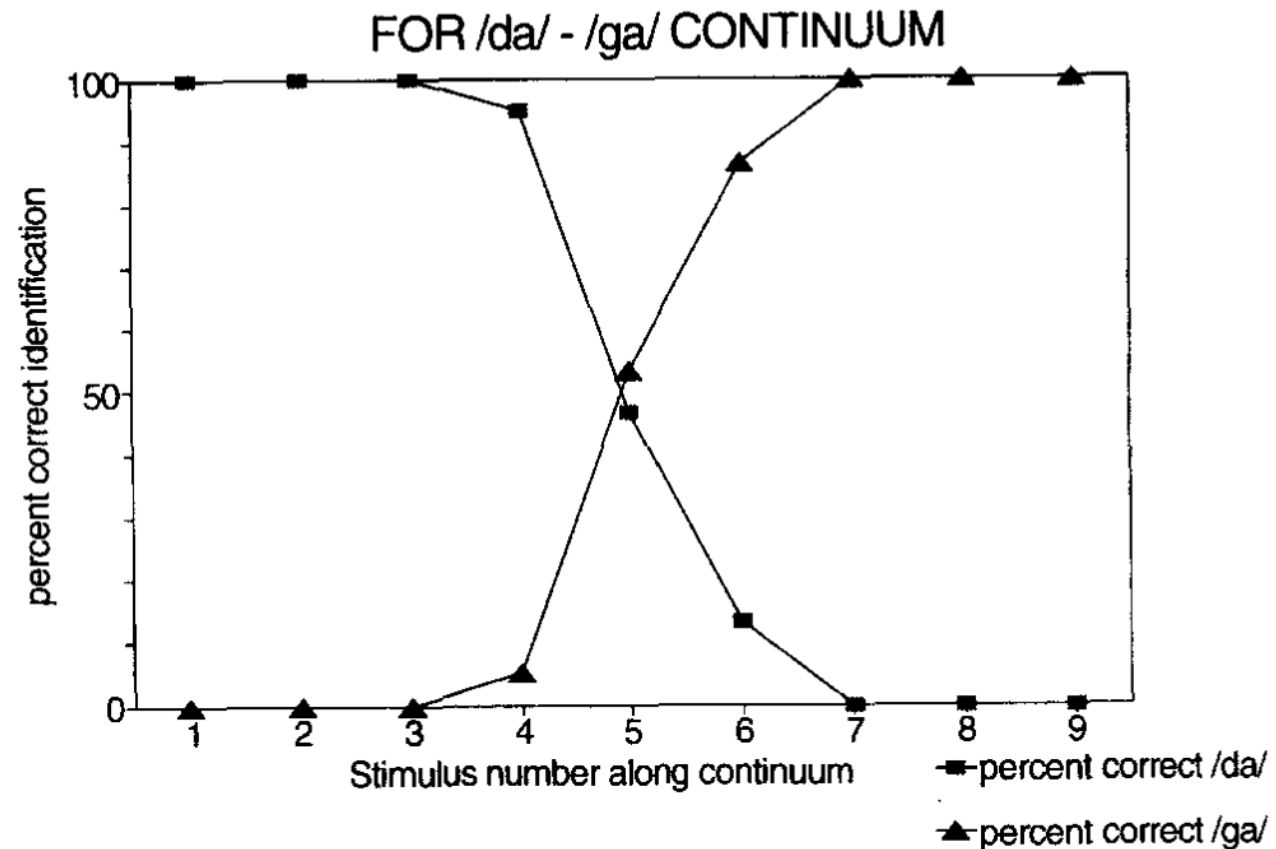


Fig. 2. Mean identification functions for the /da/-/ga/ continuum. Stimuli 1-4 were identified as /da/ and 6-9 as /ga/. Stimulus 5 is at the category boundary.

# Sharma et al 1992

→ Physical distance  
From 2-4 = 4-6

Within-category: 2-4  
Between-category 4-6

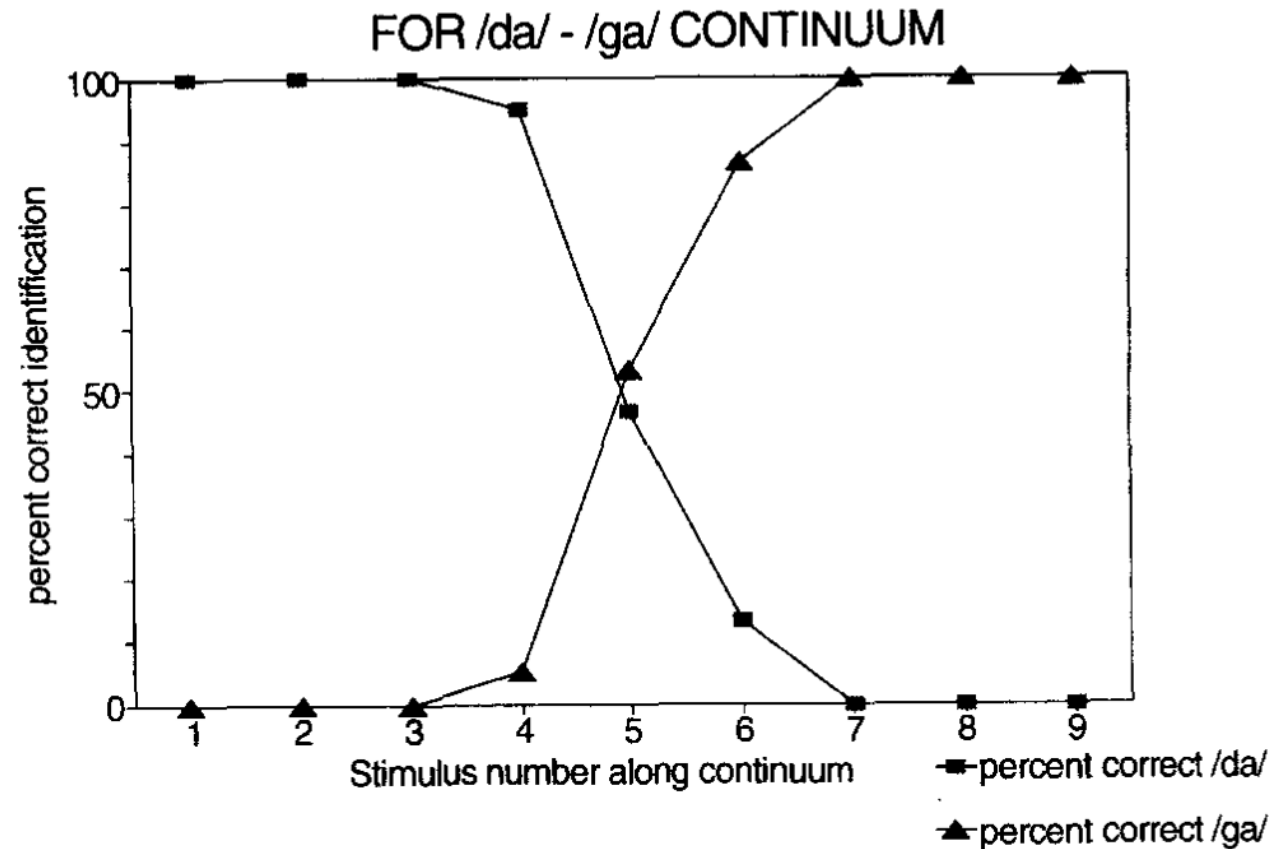


Fig. 2. Mean identification functions for the /da/-/ga/ continuum. Stimuli 1-4 were identified as /da/ and 6-9 as /ga/. Stimulus 5 is at the category boundary.

# Sharma et al 1992

→ Physical distance

From 2-4 = 4-6

Within-category: 2-4

Between-category 4-6

**Same MMN!**

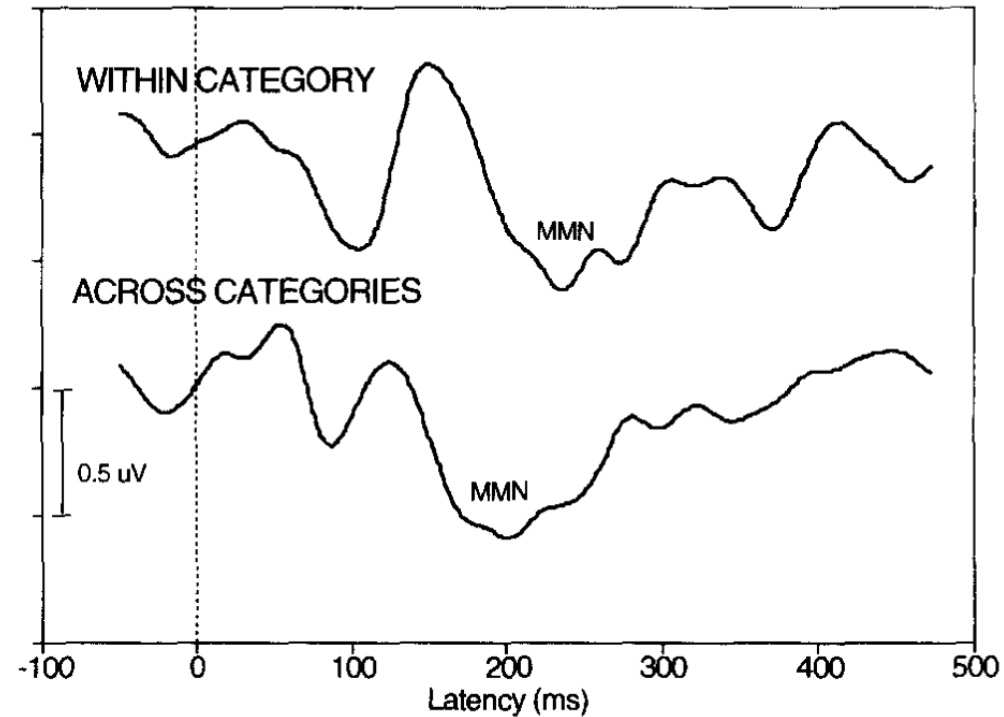


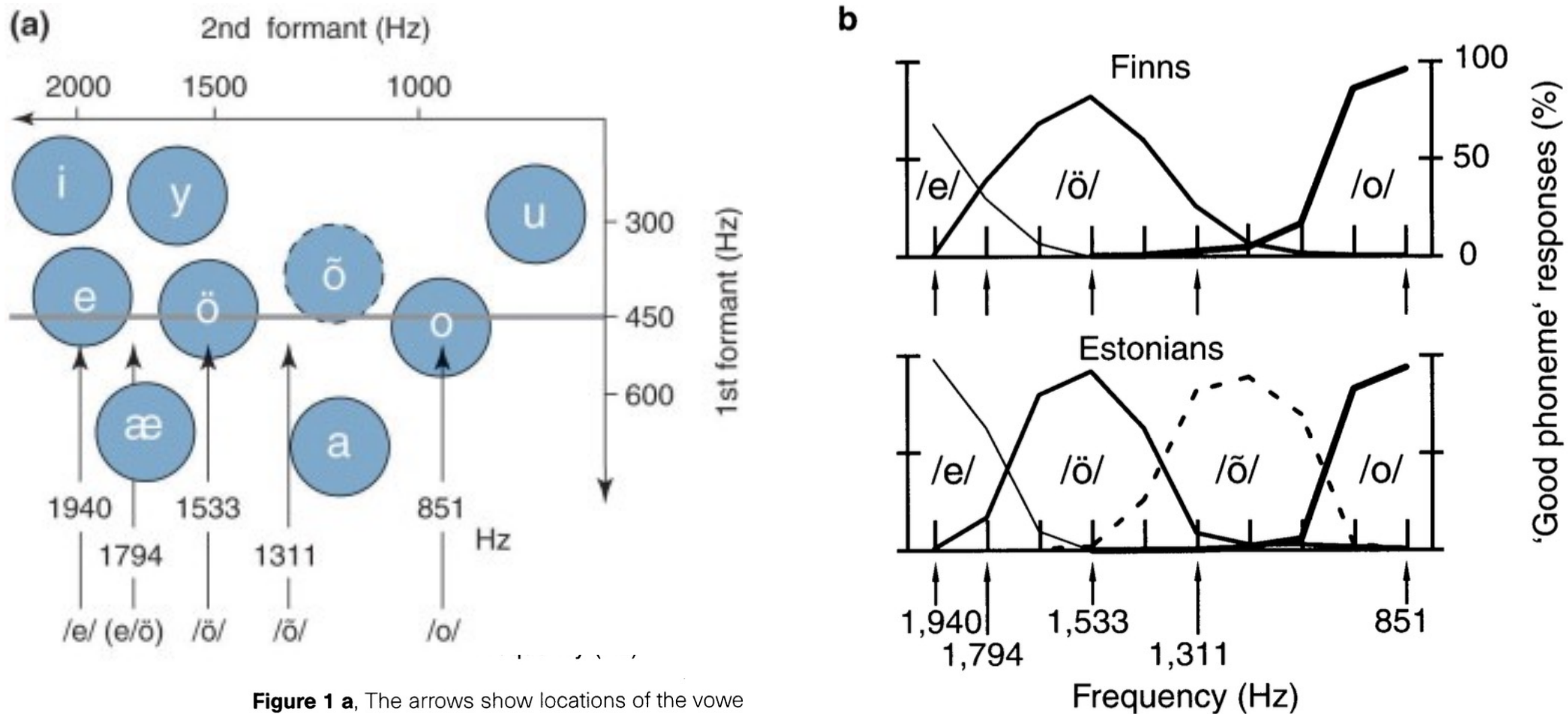
Fig. 4. Grand average difference waves (all subjects) in the within and across category conditions. The MMN is evident between 200–300 msec in both conditions.

**But Sams et al 1990 did find a difference in similar experiment, so results are mixed**

# Finnish vs Estonian vowels + native vs non-native speakers

		Standard	deviant	Prototypical in L1?
Finnish L1		/e/	/o/, /ö/	YES
			/õ/	NO
Estonian L1		/e/	/o/, /ö/	YES
			/õ/	YES

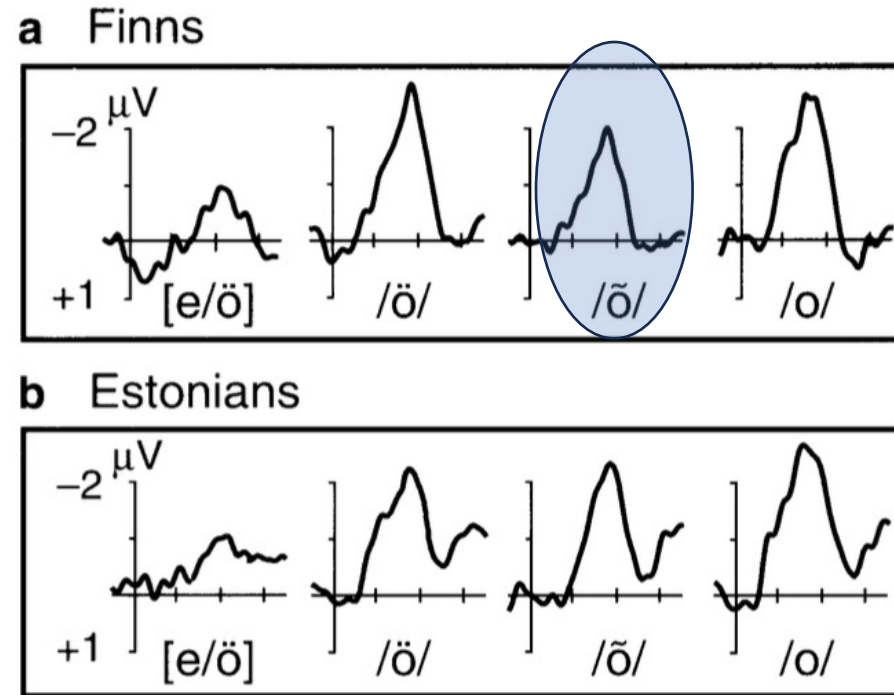
Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, 385(6615), 432–434. <https://doi.org/10.1038/385432a0>



**Figure 1 a**, The arrows show locations of the vowels. The standard stimulus was the Finnish and Estonian prototype /e/; the deviant stimulus [e/ö] was an intermediate between /e/ and /ö/; the deviant stimulus /ö/ was a prototype shared by the two languages; the deviant stimulus /õ/ was an intermediate between the Finnish and Estonian prototypes /ö/ and /o/ (and closely corresponded to the Estonian prototype /õ/); the deviant stimulus /o/ was a prototype in both languages. **b**, The percentages of 'good phoneme' /e/, /ö/ and /o/ responses in Finns (top) and of 'good phoneme' /e/, /ö/, /õ/ and /o/ responses in Estonians (bottom) in the behavioural task. The ten stimuli used in the judgement task are indicated with tick marks. The frequency scales are logarithmic.

# Finnish vs Estonian vowels X speaker groups

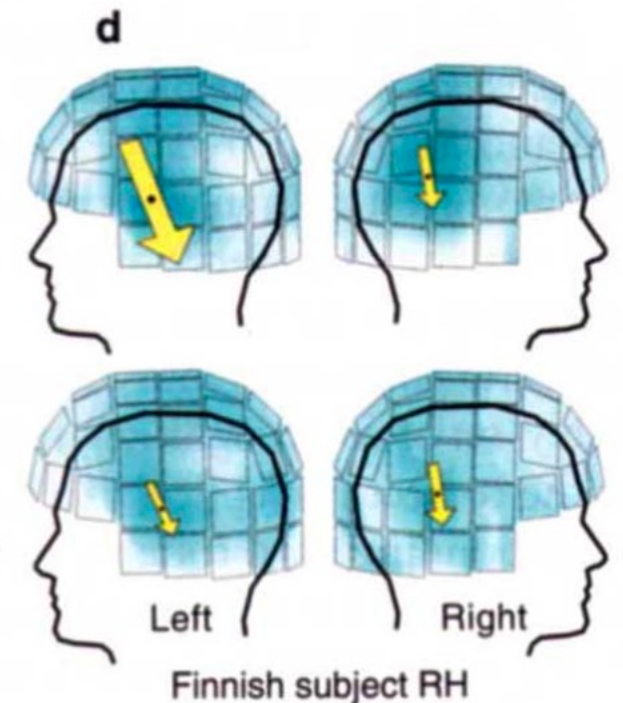
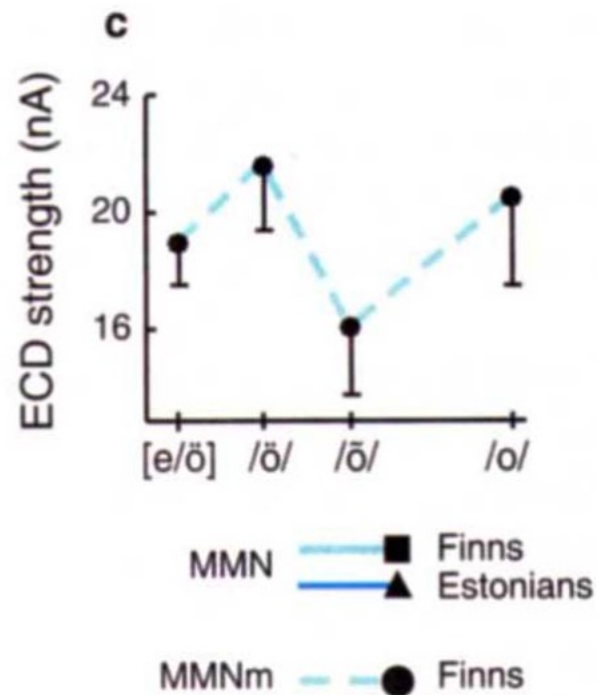
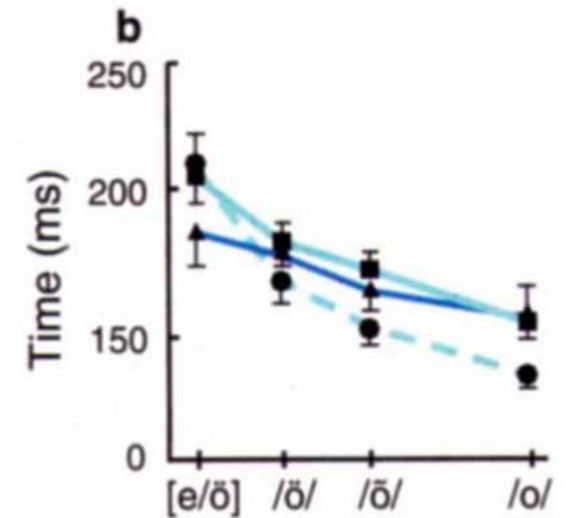
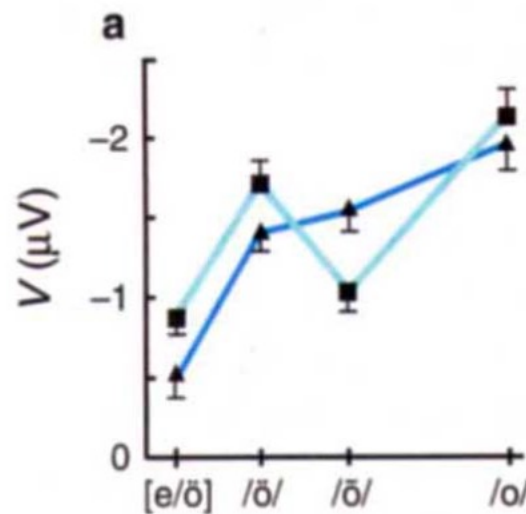
NB! Plotting negative up!



**Figure 3** The amplitude of the mismatch negativity (MMN; frontal Fz electrode; grand-average deviant-standard difference waveforms) reflects the language-specific phoneme categories of the Finnish and Estonian languages. Note that the MMN amplitude for the deviant /õ/, a non-prototype in Finnish, is clearly smaller than that for the adjacent prototype deviants with Finnish subjects (**a**) but not with Estonian subjects (**b**) for whom /õ/ was a prototype.

# Hemisphere differences

- Amplitude of MMN interaction with L1
- Finnish subjects also had larger current dipole in left hemisphere for within-language deviant than out-of-language deviant



# English vs Spanish vowels

- English phonemes:

Spanish

- $/ɪ/$        $/ɛ/$   
“bit”      “bet”

$/i/ = \{ [i], [ɪ]^* \}$

$/e/ = \{ [e], [ɛ], [ɪ]^* \}$

- QUESTION: We have two groups of speakers, native L1 American English and Spanish speakers who acquired English as L2 before 5 years of age. Is Spanish speakers' MMN to deviant  $[ɪ]$  distinguishable from native AmE MMN?
- If yes, they are native-like Eng speakers. If no, they are non-native like.

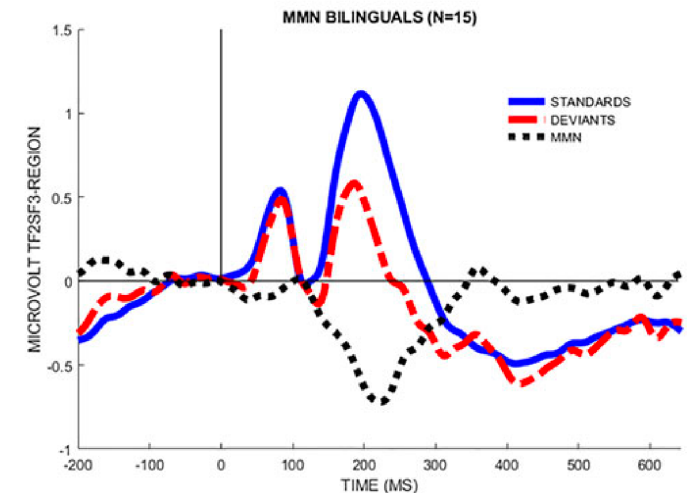
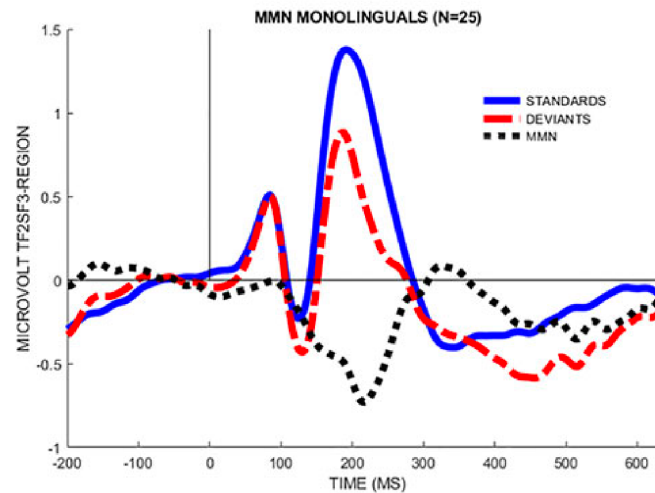
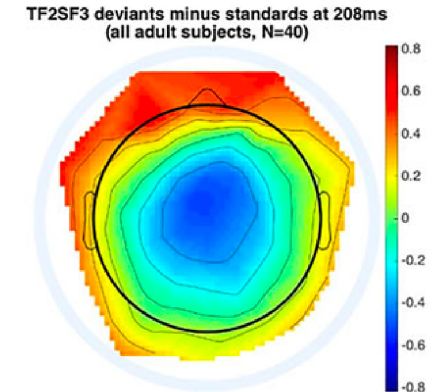
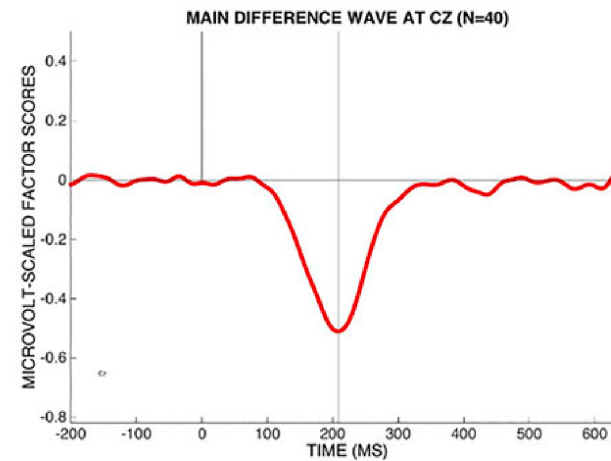
# Design

	Standard	Deviant
Monoling AmE L1	/ɛ/	/ɪ/
Biling So/AmE	/ɛ/	/ɪ/

Plus ATTEND vs IGNORE stimuli, to test whether attention modulates the response

# Results for adult participants

No difference! The Spanish bilinguals effectively have a native-like MMN to the AmE contrast

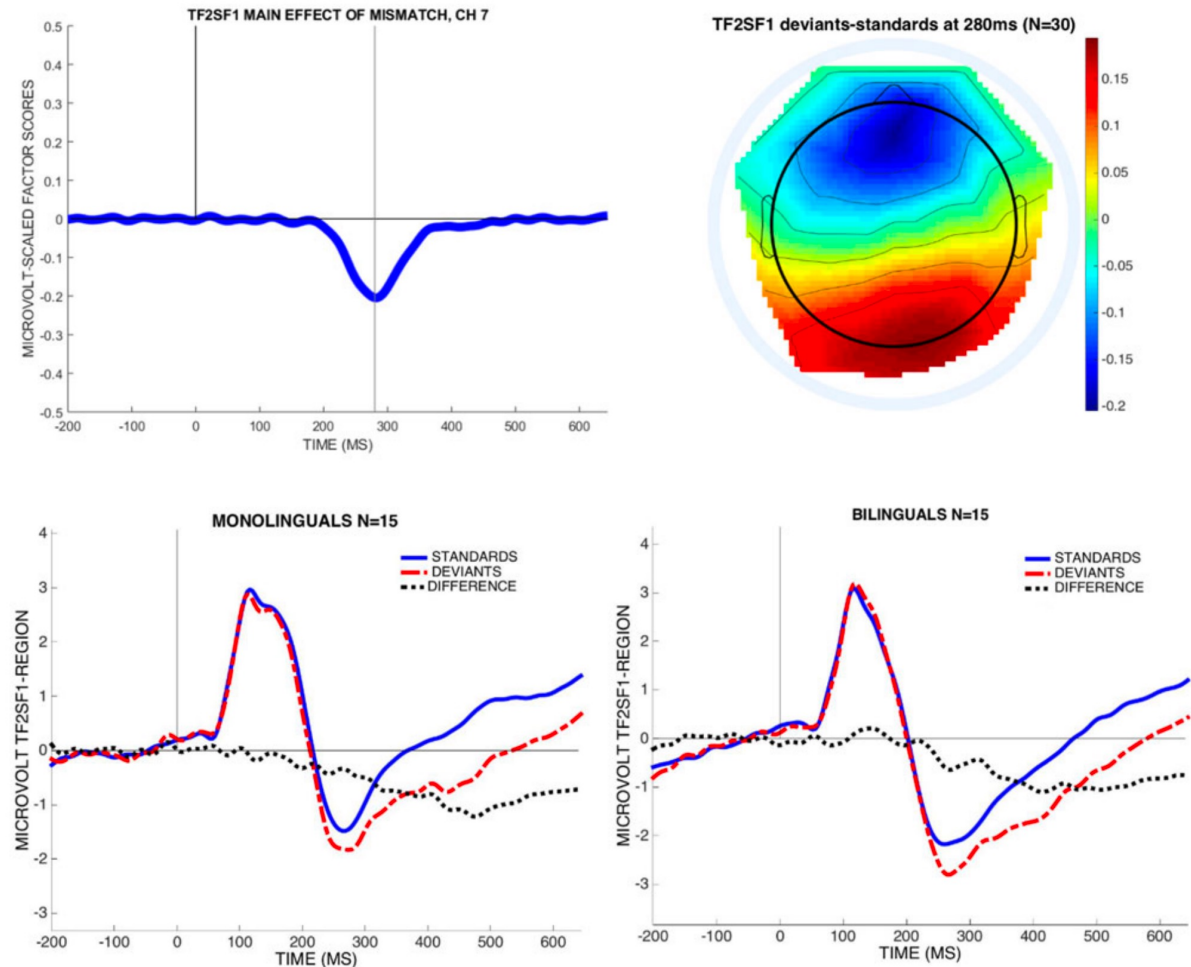


**Fig. 1.** Top row, left panel: Factor score main effect difference wave. Right panel: Spatial distribution of TF2SF3 factor score difference wave. Lower row, left panel: main effect MMN in monolinguals, average of TF2SF3 electrode region; right panel; MMN main effect for bilinguals.

# Results for kids (9-11 year olds)

Same for kids (also acquired English before 5).

However, latency and shape of MMN different in kids

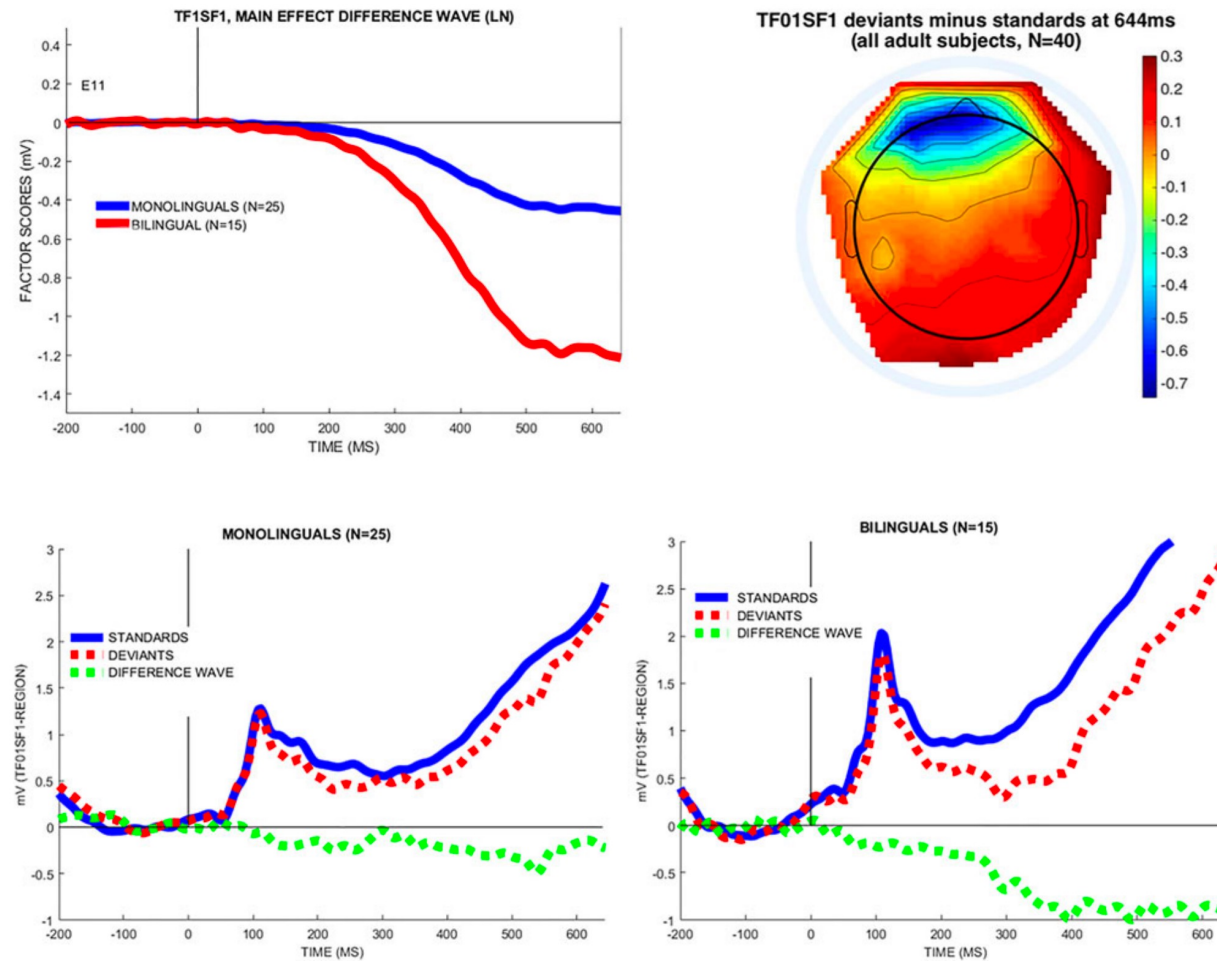


**Fig. 5.** Upper panel: Main effect factor waveform for the first negative peak and spatial distribution of this effect in the factor. Lower panel: mean waveforms for each child group for standard and deviants computed for the region in blue in the topographical plot.

# Inferences

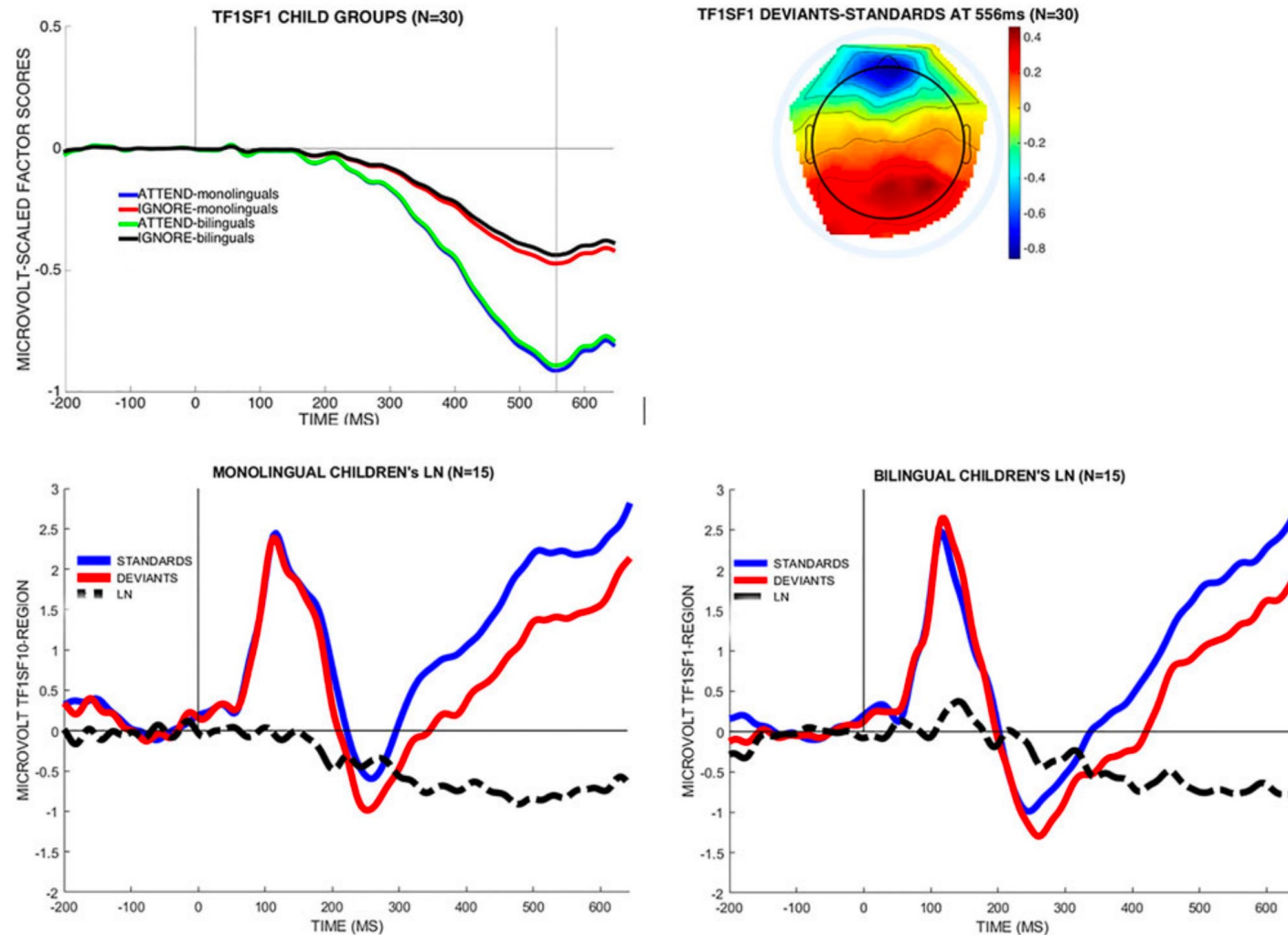
- If you learn an L2 before 5 years of age, you become essentially native like with respect to the automatic detection of a language-relevant speech sound difference in the L2
- HOWEVER: In the adult group, we do see an effect of ATTENTION to the contrast in bilingual adults vs monolingual adults, but not in kids
- → shows up in the “late discriminative negativity” ERP component

# LDN in adults interact with mono vs biling



**Fig. 3.** The top, left panel shows the difference waves for the factor scores for monolinguals vs. bilinguals in the Late Negativity ERP; the right top panel shows the spatial distribution of this main effect difference wave. The bottom row shows the raw voltage average of TF01SF1 electrodes exceeding 0.6 factor loadings (channels 6, 7, 10, 11, 14) for monolinguals (left panel) vs. bilinguals (right panel).

# LDN in kids is same for mono vs biling



**Fig. 6.** Late Negativity in child groups. The left top panel shows the factor waveform by ATTENTION and LANGUAGE; right top panel shows the scalp distribution of the factor. Lower panels show the mean waveforms for the spatial region over time by the standard and deviant stimulus conditions along with the difference waveform.

# Summary Spanish biling study

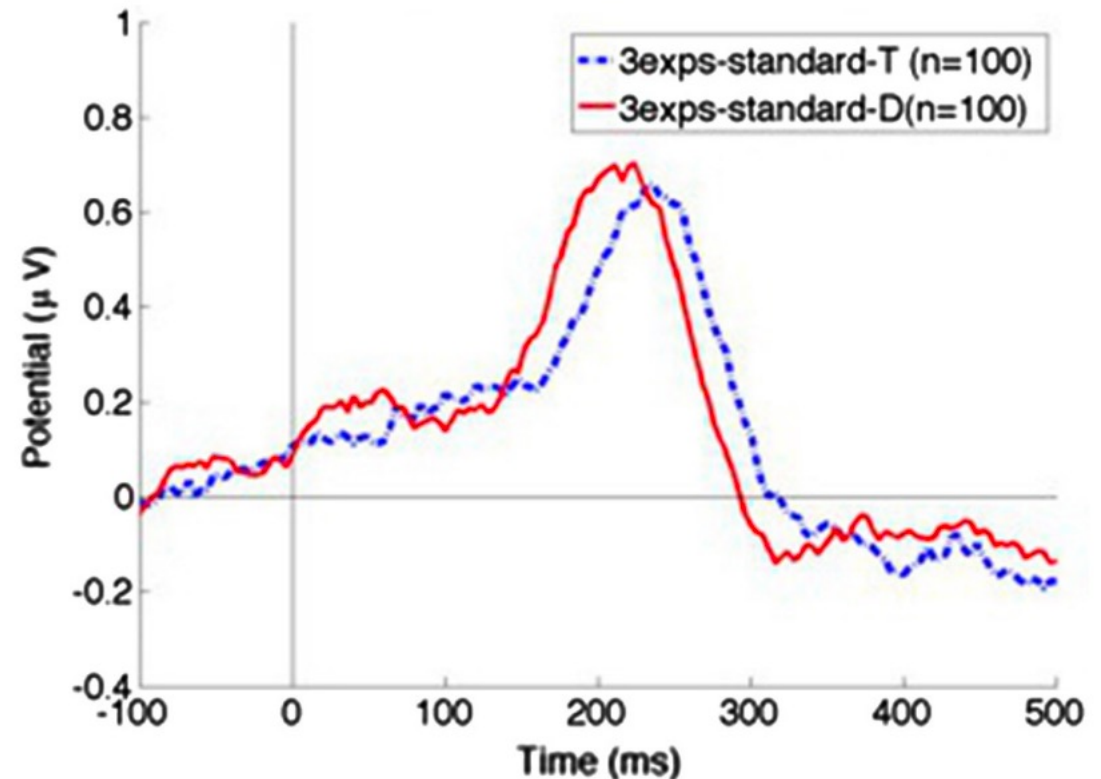
- The early automatic low-level MMN is same for mono and bilinguals
  - Bilinguals have become native-like phonologically
- The later awareness-related LDN is significantly larger for bilingual adults than monolingual adults
  - But not in kids, so this awareness to “which phonology you are listening to” develops with use over time

# NEXT:

- Identity-MMN
  - better method to control for additional inherent stimulus differences

# Controlling for inherent stimulus property confounds

- Remember this?
- If I subtract standard /d/ from standard /t/ here, I'll get something that looks like an MMN 250-350ms!



**Fig. 10.** Latency difference (but not significant amplitude difference) between the mean Auditory Evoked Potential to /d/ vs. /t/ (averaged across all tokens and all subjects in all three experiments).

Control for the possibility that the difference doesn't just arise from different physical responses to different sounds

- Standard t

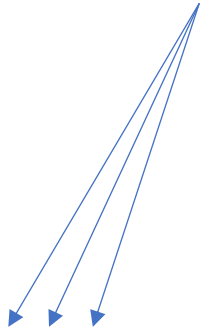
- 

t t t t t t t d

- Standard d

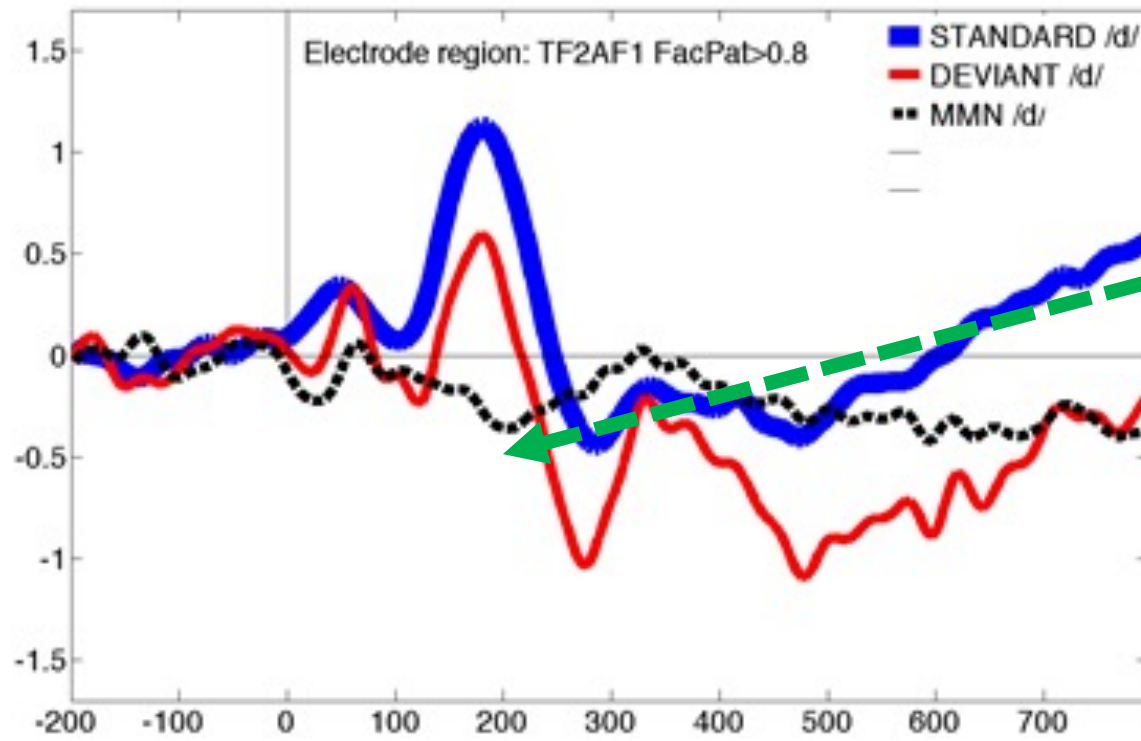
- 

d d d d d d d t



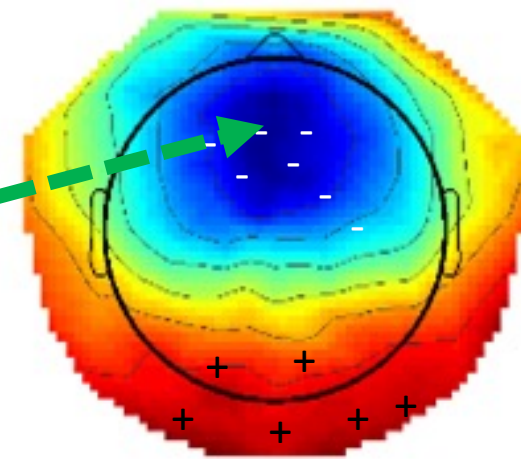
# /d/ as standard compared to /d/ as deviant

TIME: amplitude, latency



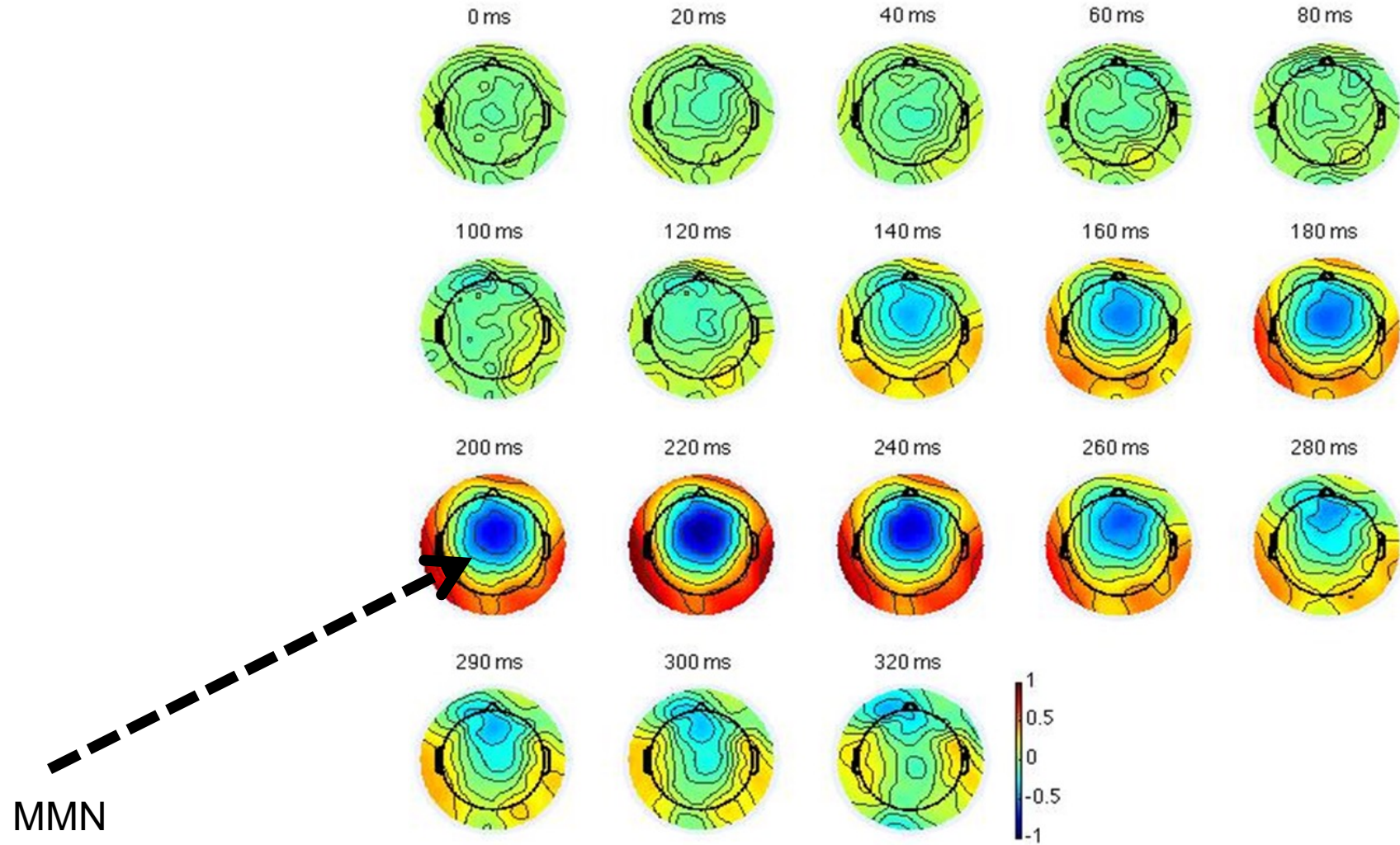
SPACE: location of voltage difference

Exp1-TF3SF1 main effect difference wave

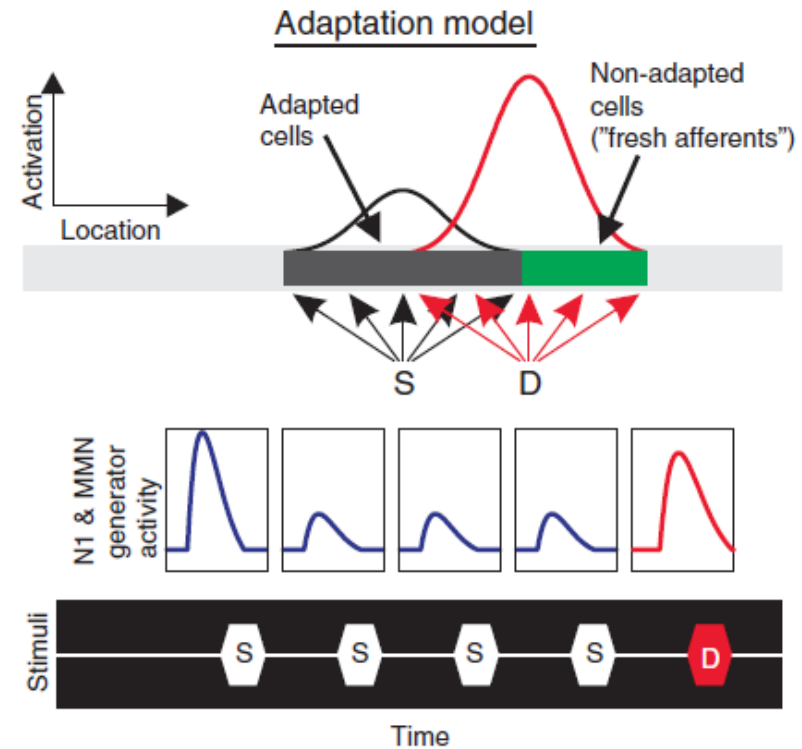
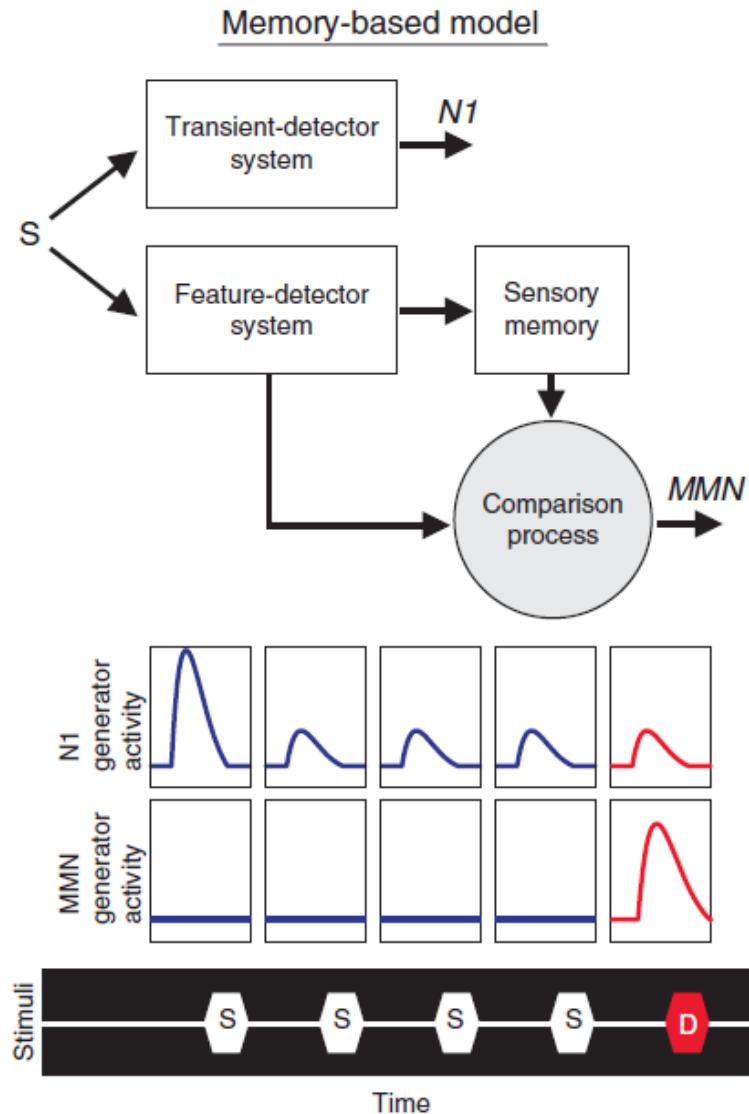


Hestvik and Durvasula 2016

# MMN temporo- spatial properties



# Two views of the mechanism of MMN



Will have difficulty explaining "omission" MMN

May, P. J. C., & Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology*, 47(1), 66–122.

- NEXT CLASS:

- How MMN can be used to measure phonotactic repair mechanisms (illusory epenthesis)
  - Japanese vs French
  - top-down effects on auditory cortex predictions-a Whorfian effect on language itself!
- How the McGurk effect can neutralize or add an MMN
  - Top-down effects on auditory cortex predictions