Behavioral Neuroscience A
11: Multisensory Integration

Richard Veale

Graduate School of Medicine
Kyoto University

https://youtu.be/BCSj2LzrwXk

Lecture Video at above link.

Multisensory Integration

Multisensory integration

- To learn about some instances in which we combine information from multiple sensory modalities (visual, auditory, somatosensory).

- To learn about brain structures that respond to several modalities.

- To understand the circumstances under which these brain areas respond to multiple modalities.

Today's contents

<u>Multisensory integration</u>

Objects around us stimulate not only one but multiple senses…

- Audiovisual integration of speech

- Neurophysiology of multisensory integration

- Brain areas involved in multisensory integration

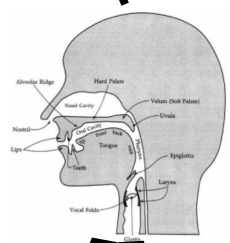
- Audiovisual integration: dominance of vision for spatial location – of audition for time

- Synesthesia

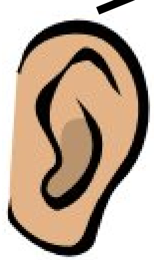
McGurk Effect

Integration of audiovisual speech signals results in the perception of /da/ when hearing /ba/ and seeing /ga/ (see video):

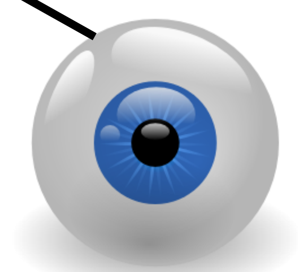
Perception: /da/



Possibly: Representation as intended articulatory gesture.



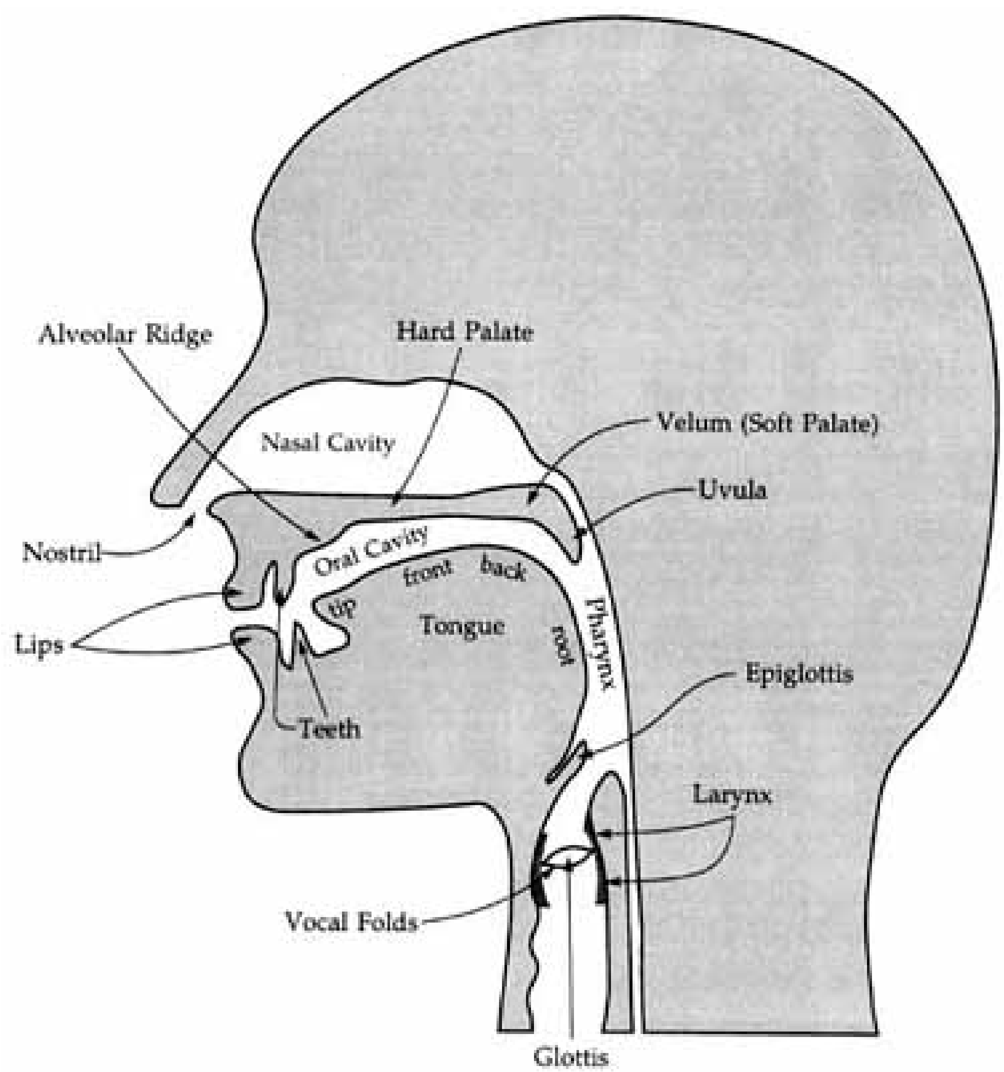
/ba/



/ga/

Reference: McGurk and McDonald, Nature, 1976

McGurk Effect: Explanation



/ba/ /da/ /ga/
are plosive consonants:

The air-stream from the lung is stopped and then released to produce these sounds.

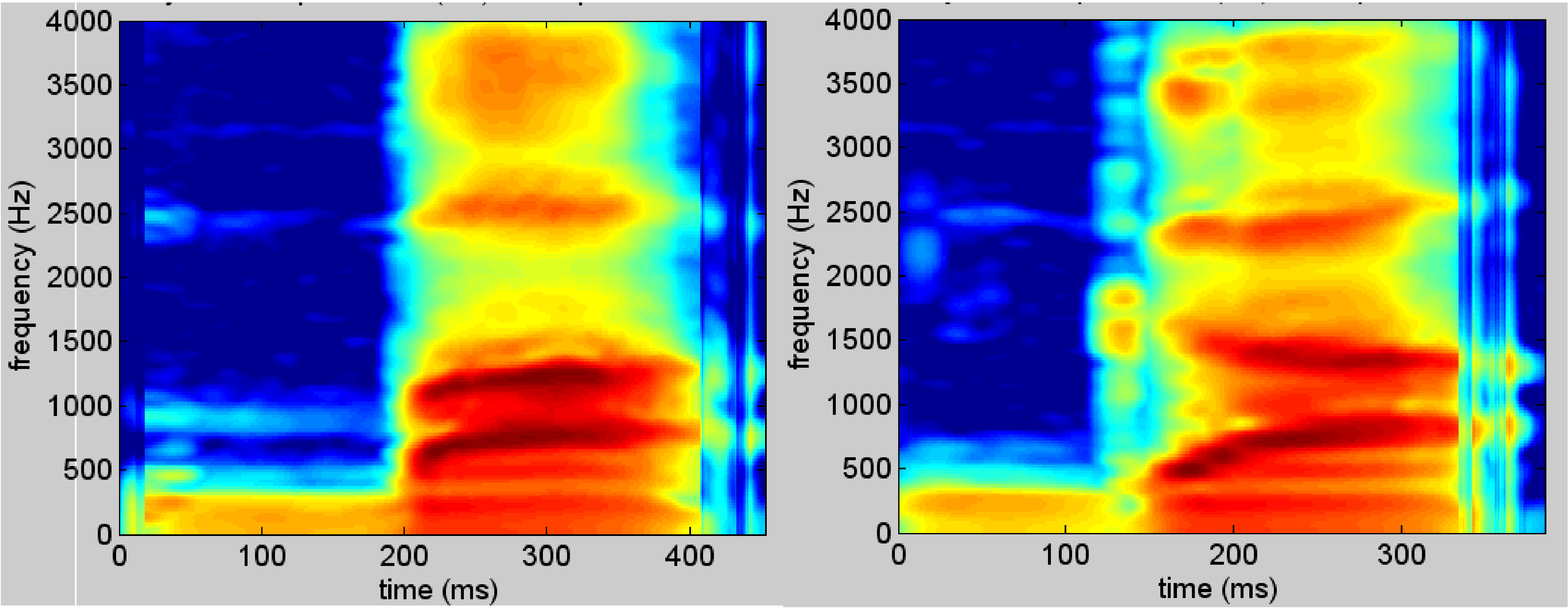
/ba/: lip

/da/: alveolar ridge

/ga/: velum

McGurk Effect: Explanation

Time-frequency plots of speech sounds

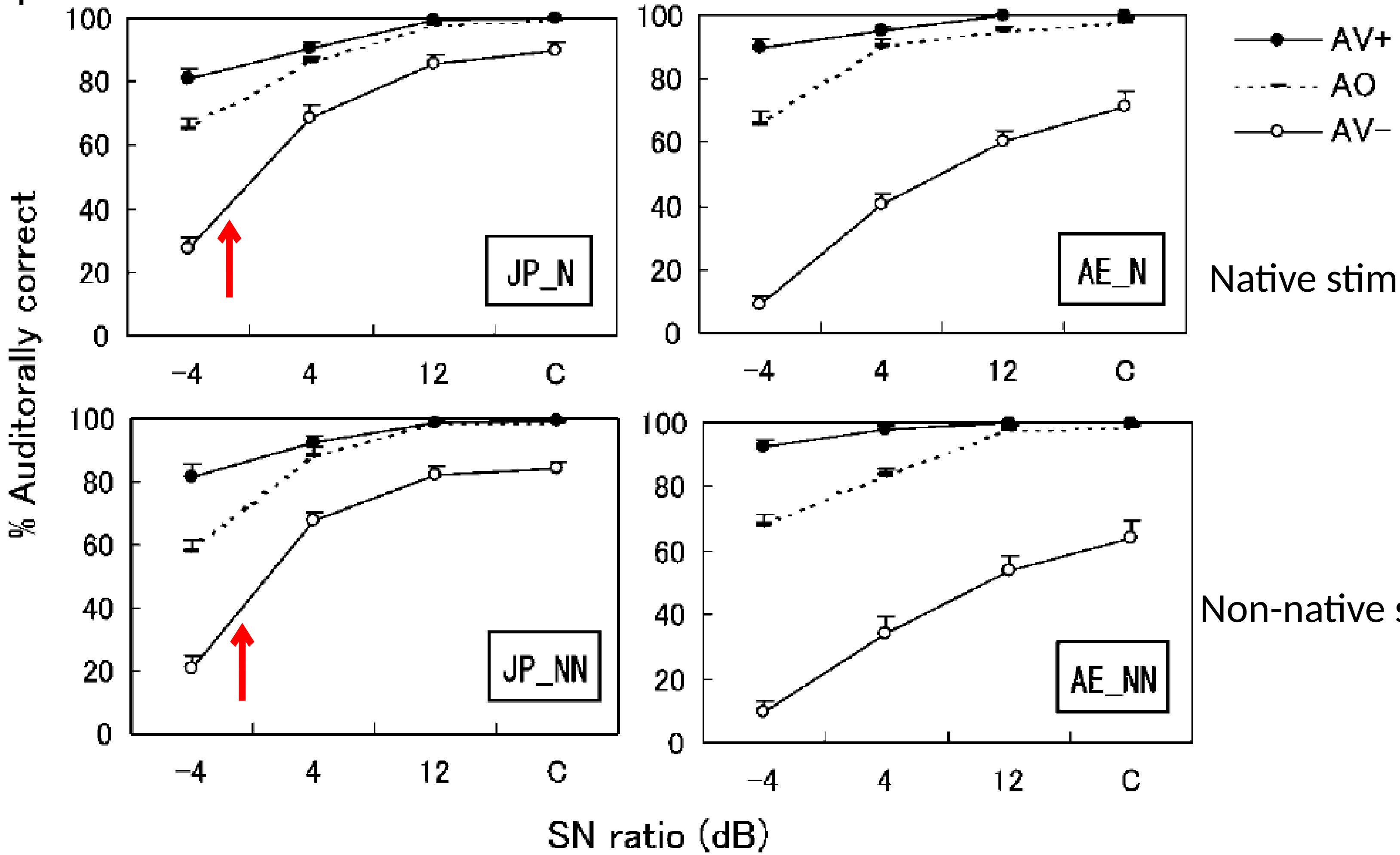


/ba/: lip

/ga/: velum

Japanese ↔ English use mouth more?

Japanese speakers show smaller visual influence (curves indicated by red arrow are shifted up) during audiovisual speech perception compared to (American) English speakers.



Native stimuli

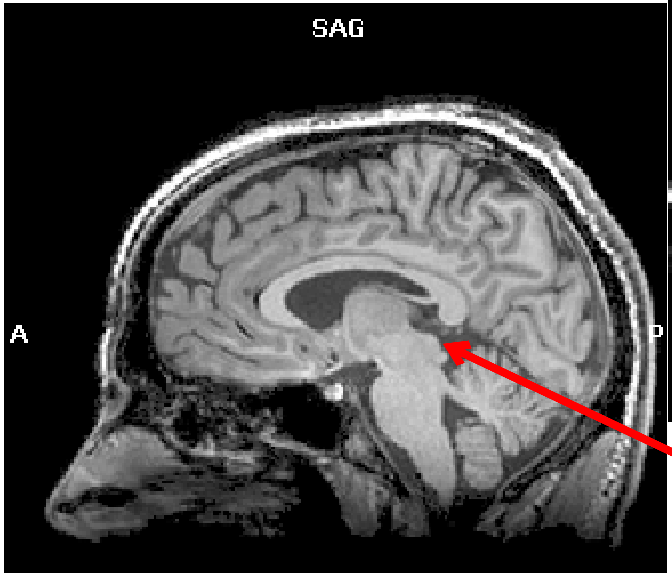
Non-native stimuli

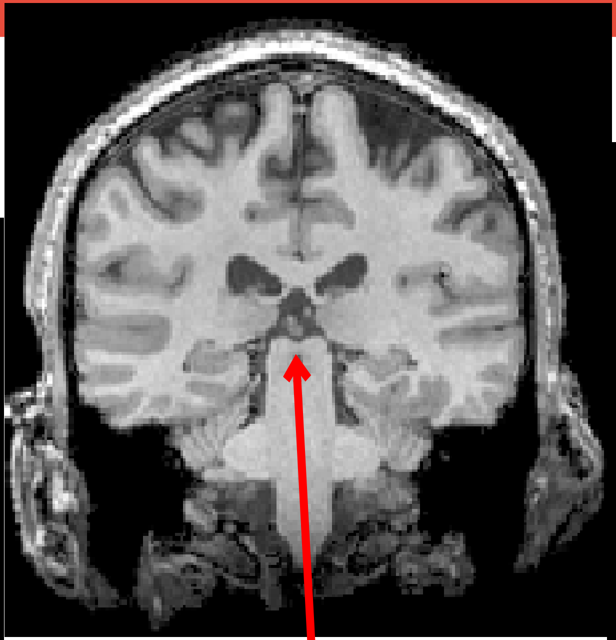
Japanese speakers

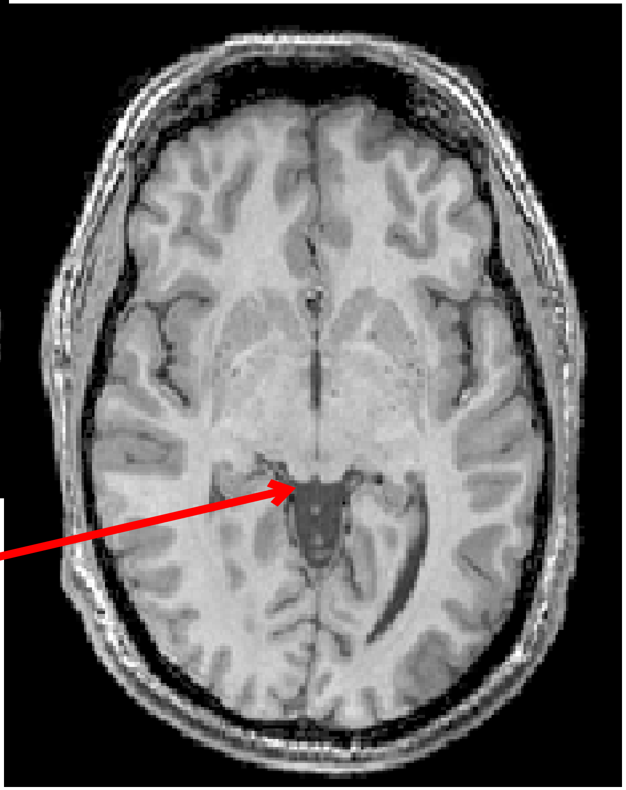
American English speakers

Sekiyama and Burnham, Developmental Science, 2008

Superior Colliculus integrates sight and sound direction?





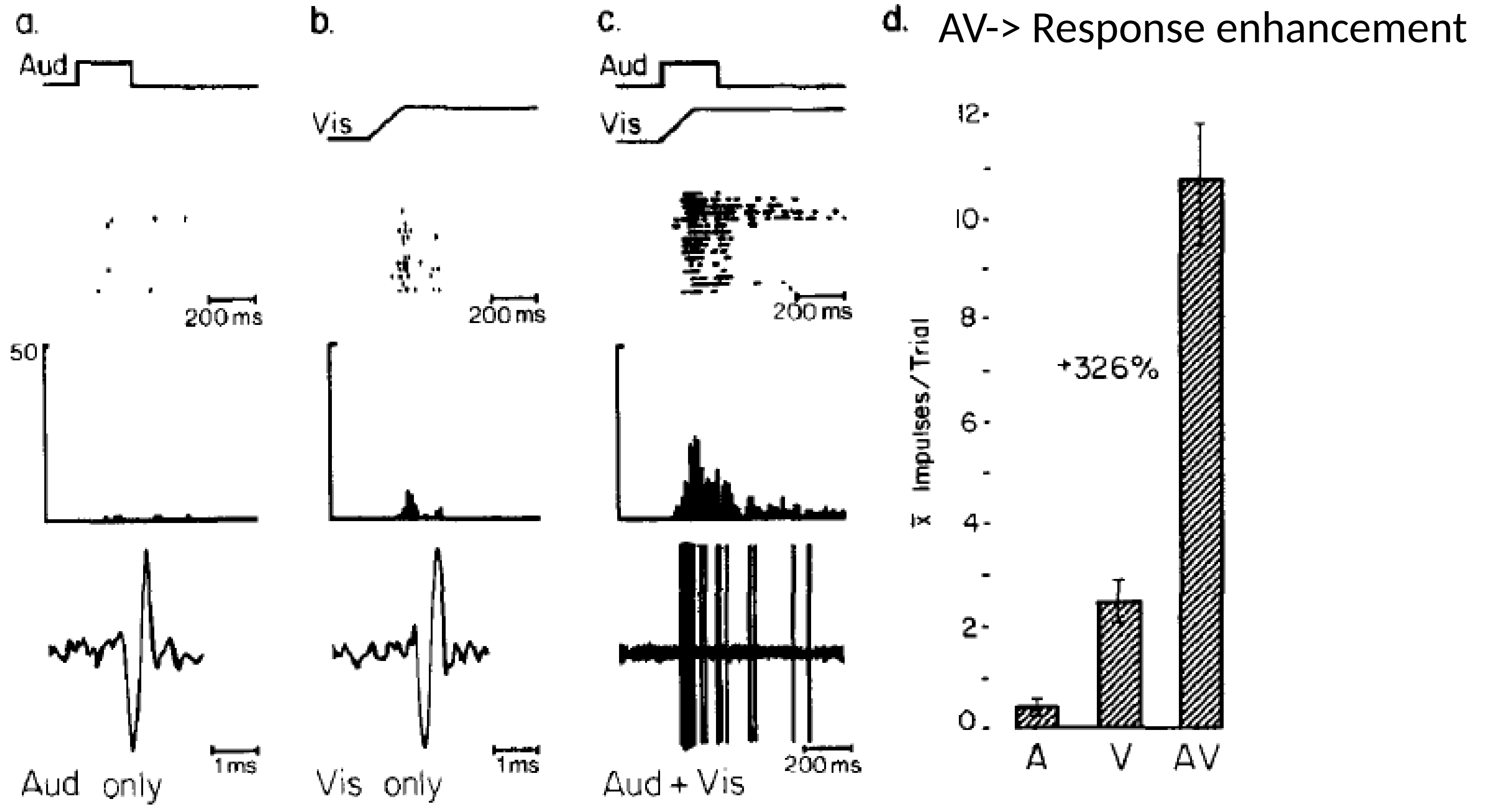


Superior colliculus
(Human)

The superior colliculus (SC) is a midbrain structure with neurons in its deep layers that respond to audio-visual and visual-somatosensory stimuli. These neurons are <u>multimodal</u>. SC plays a key role in orienting behaviors such as moving the head or eyes or allocating spatial attention.

Superior Colliculus: Stronger if sight and sound come from same direction

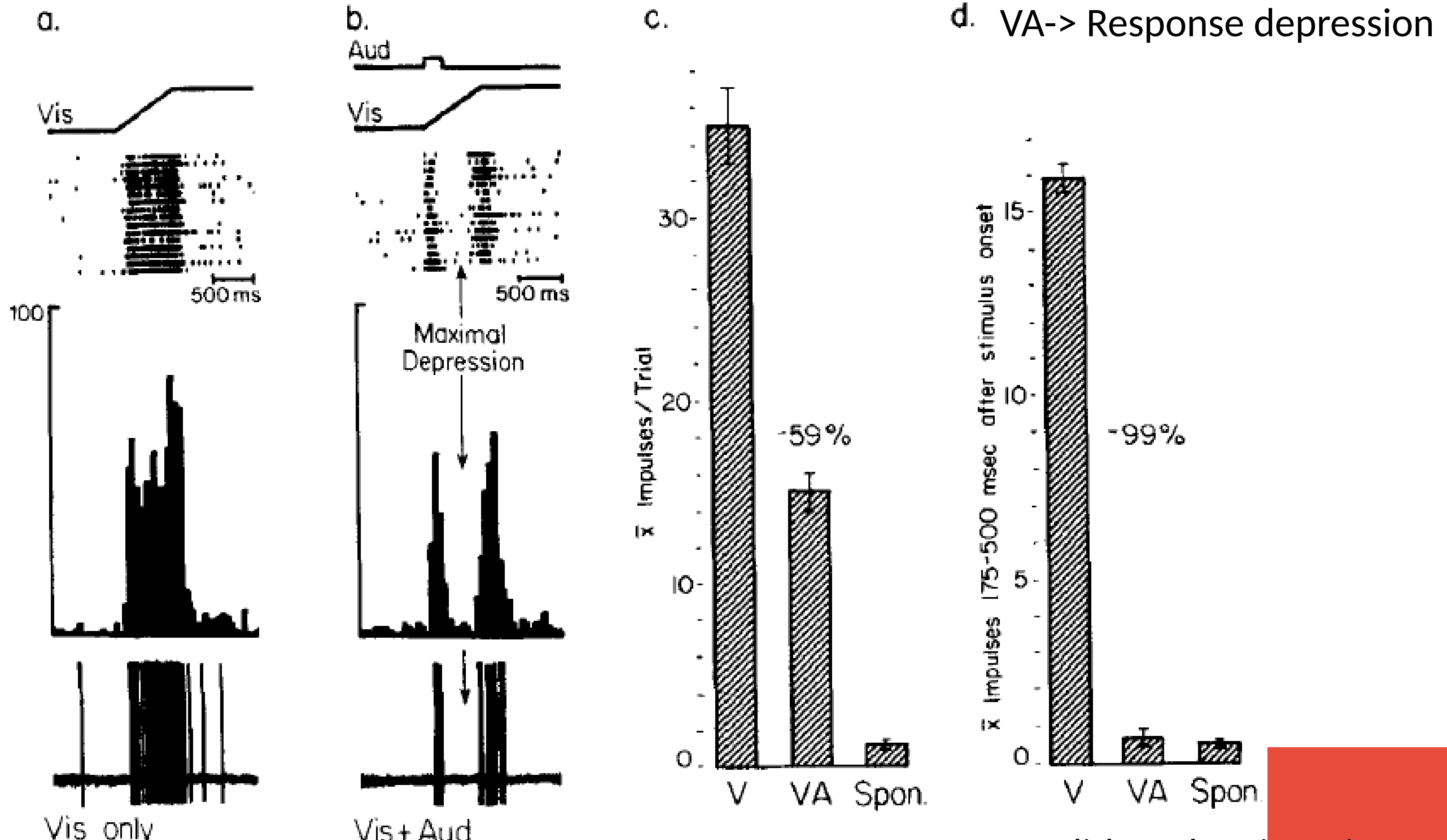
Recordings of some cat superior colliculus (deep layers) neurons showed response enhancement in response to auditory and visual stimuli. This response enhancement was superadditive (AV>A+V).



Meredith and Stein, Science, 1983

Superior Colliculus: Weaker if sight and sound from different directions

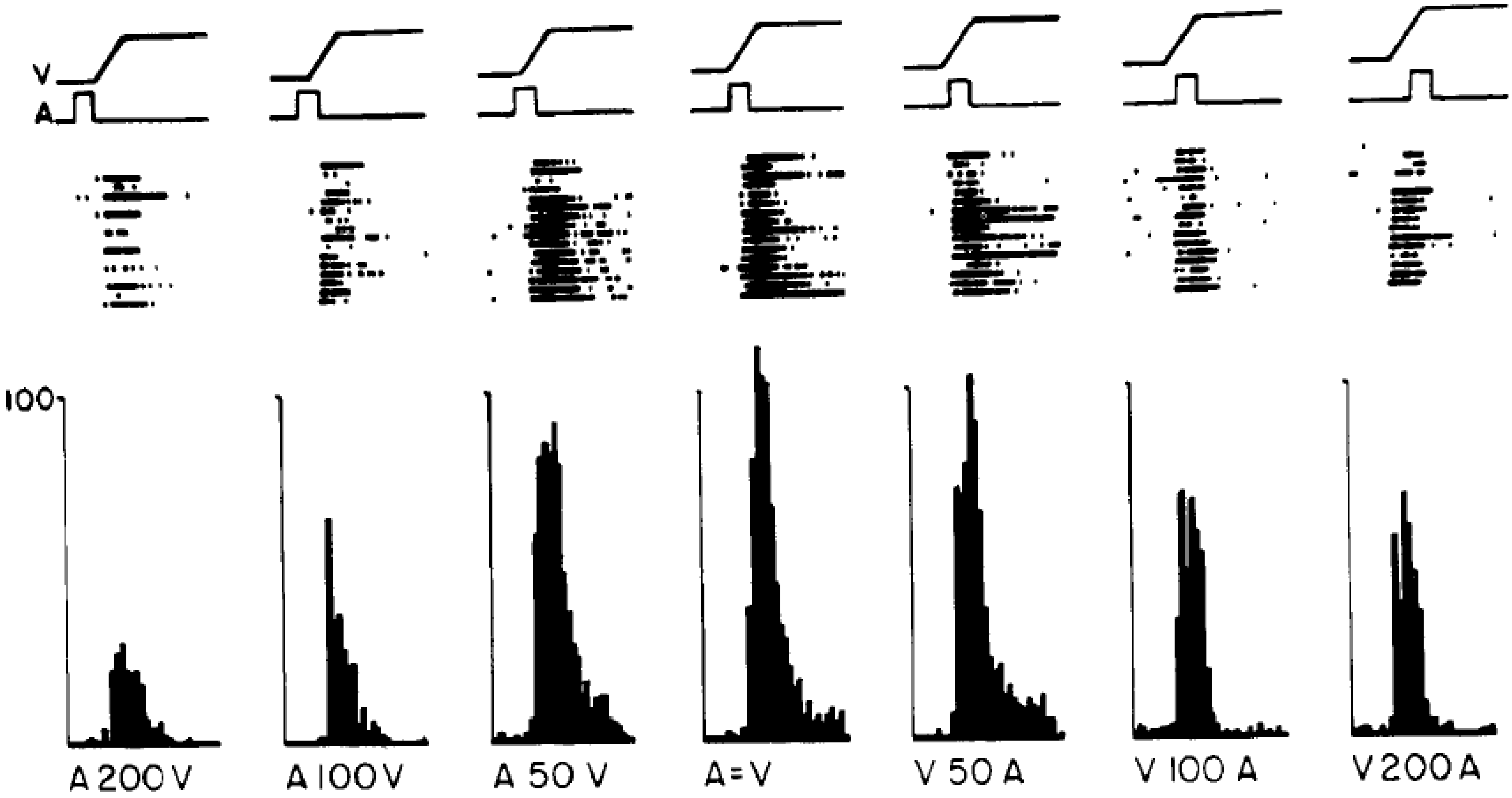
Other cat superior colliculus (deep layers) neurons showed <u>response depression</u>, i.e., a neuron's response to a visual stimulus decreased when it was paired with an auditory stimulus (in particular when the origin of the sound is outside the cell's receptive field).



Meredith and Stein, Science, 1983

Superior Colliculus: Need to line up in time too!

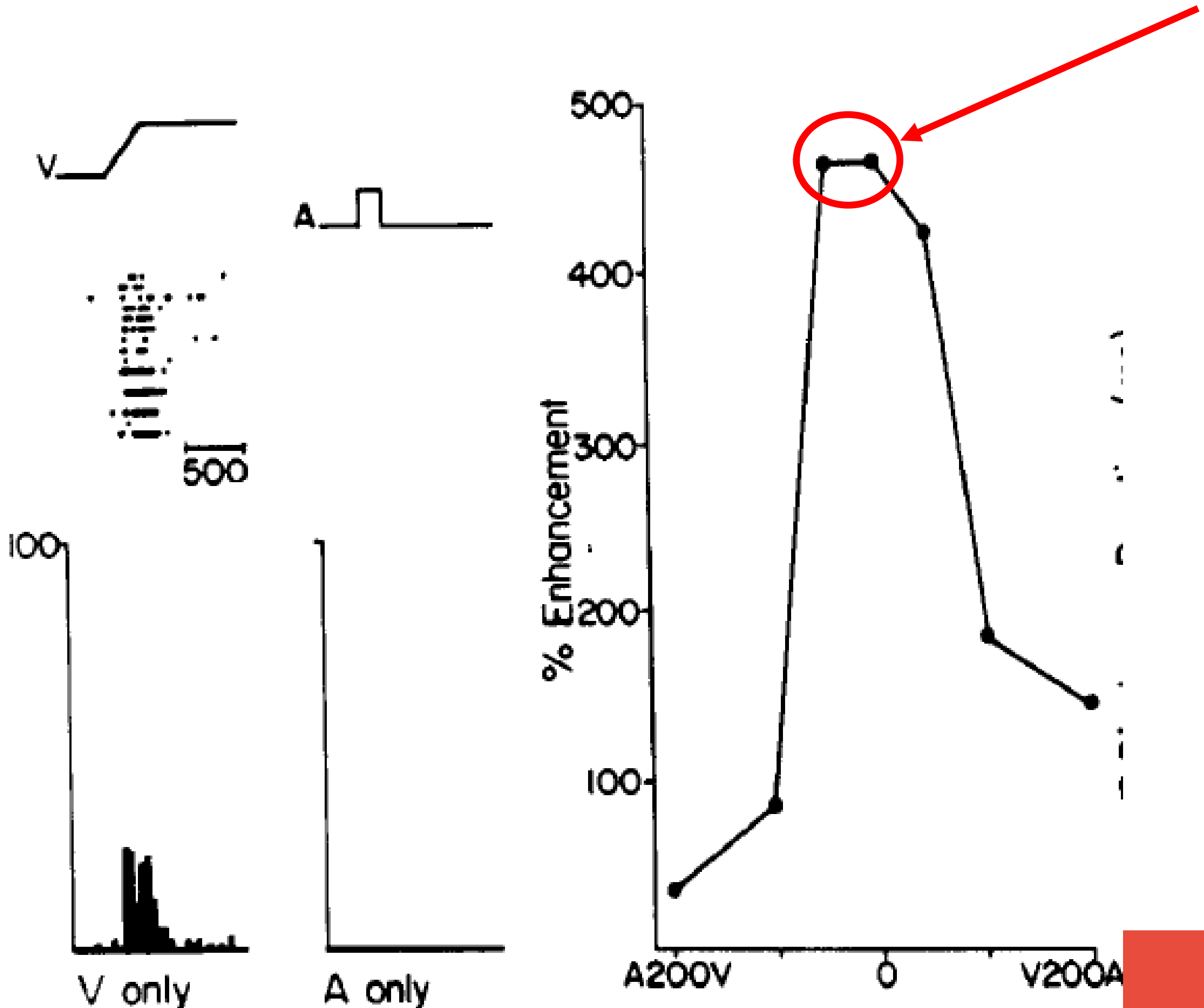
An important factor for audiovisual response enhancement is <u>temporal alignment</u>: that is, when visual and auditory stimuli are presented in synchrony. However, the temporal window is rather broad, possibly taking into account different neural conduction and stimulus propagation times (light is faster than sound).



Meredith and Stein, Journal of Neuroscience, 1987

Superior Colliculus: Temporal Congruence

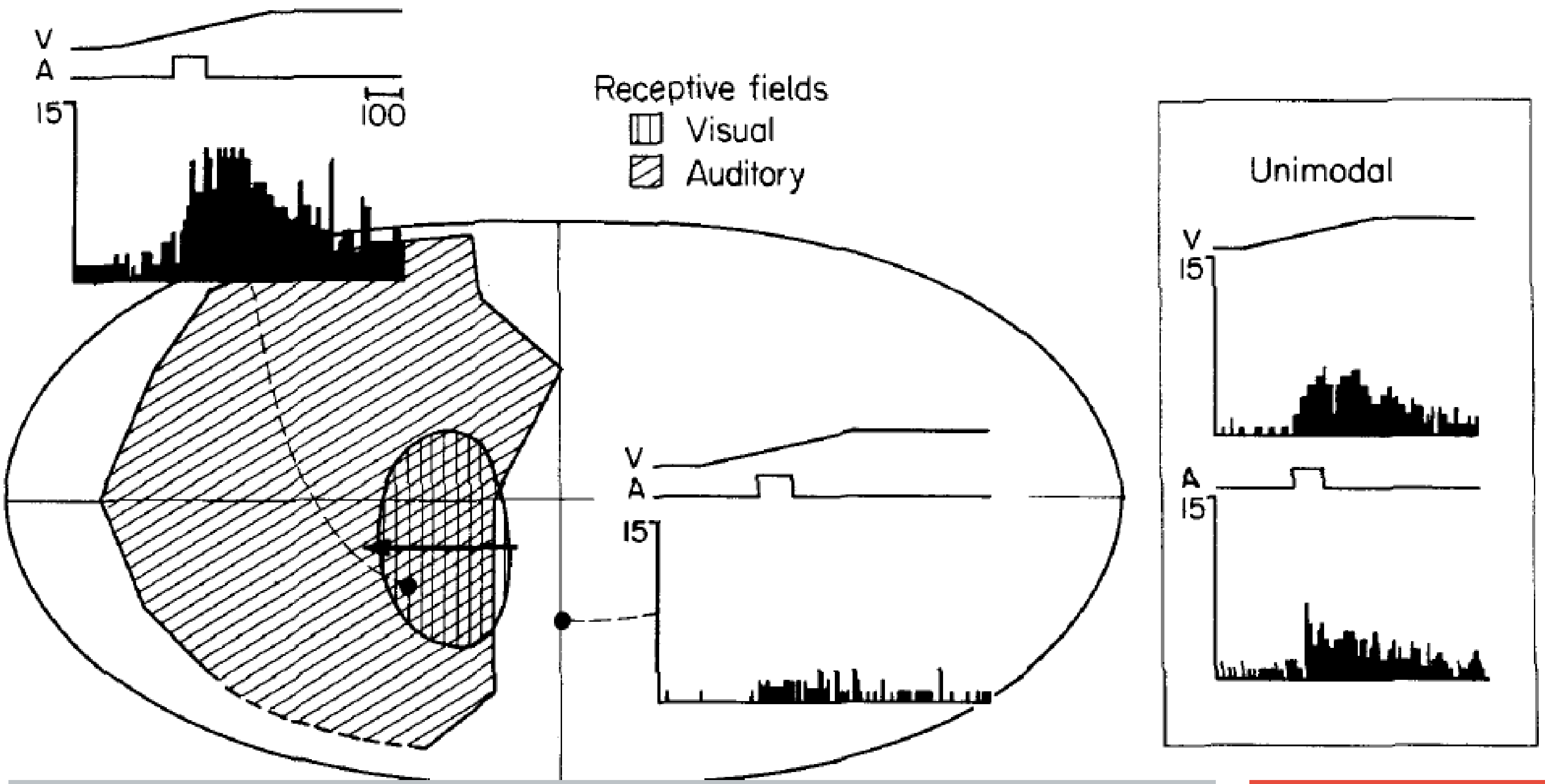
This graph shows maximal response enhancement for asynchronies of -50 to 0 ms.



Meredith and Stein, Journal of Neuroscience, 1987

Superior Colliculus: Spatial Congruence

A second important factor is spatial congruence: for maximum audio-visual response enhancement, spatial location of the stimuli should be in spatial register and within the receptive fields of the neuron. Please note the large auditory receptive field compared to the small visual receptive field.

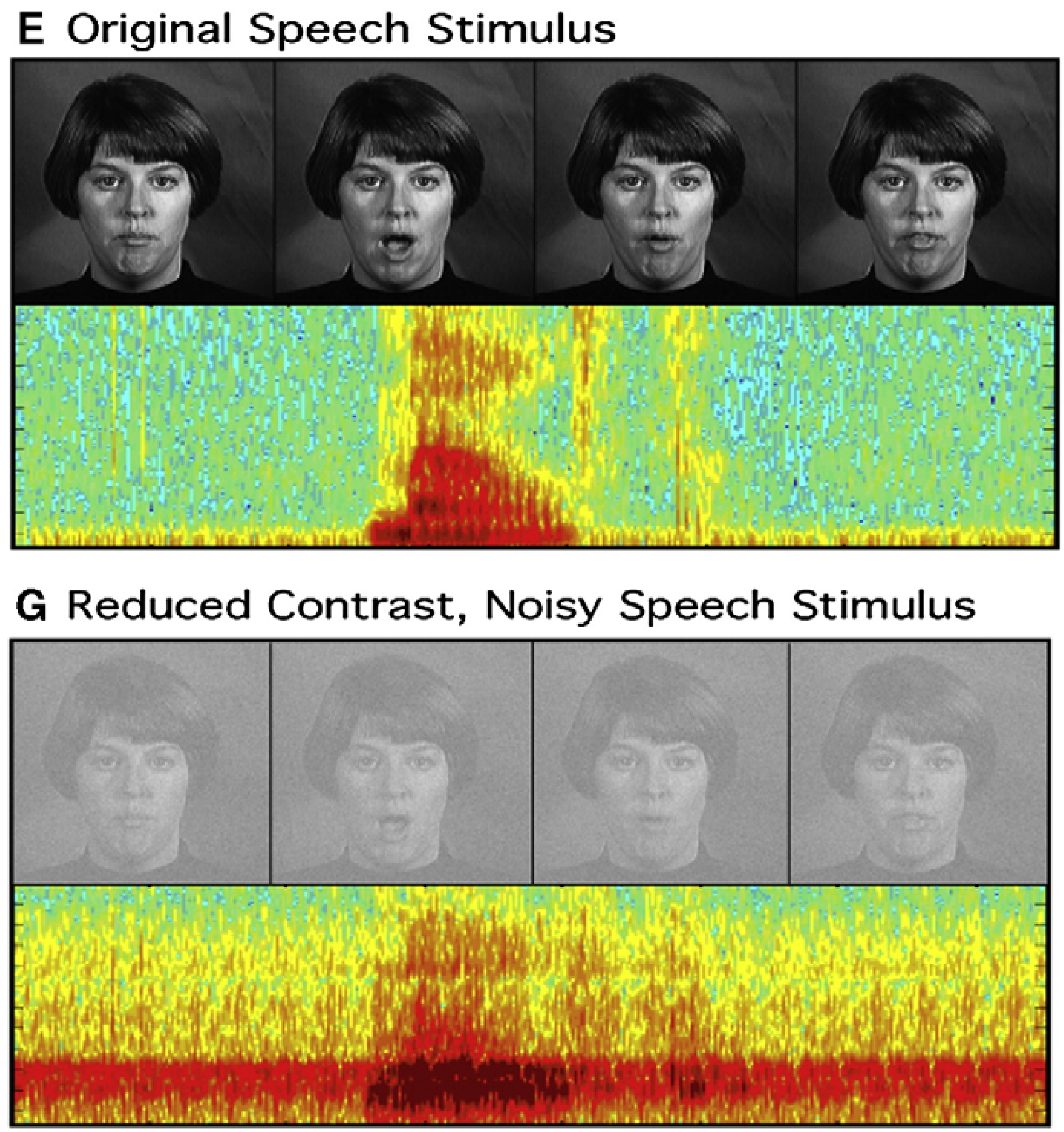


Stein and Meredith, Ann NY Acad Sci, 1990

Multisensory Integration:
The "law" of inverse effectiveness

Inverse effectiveness means that multisensory responses are largest when using weak unisensory stimuli.

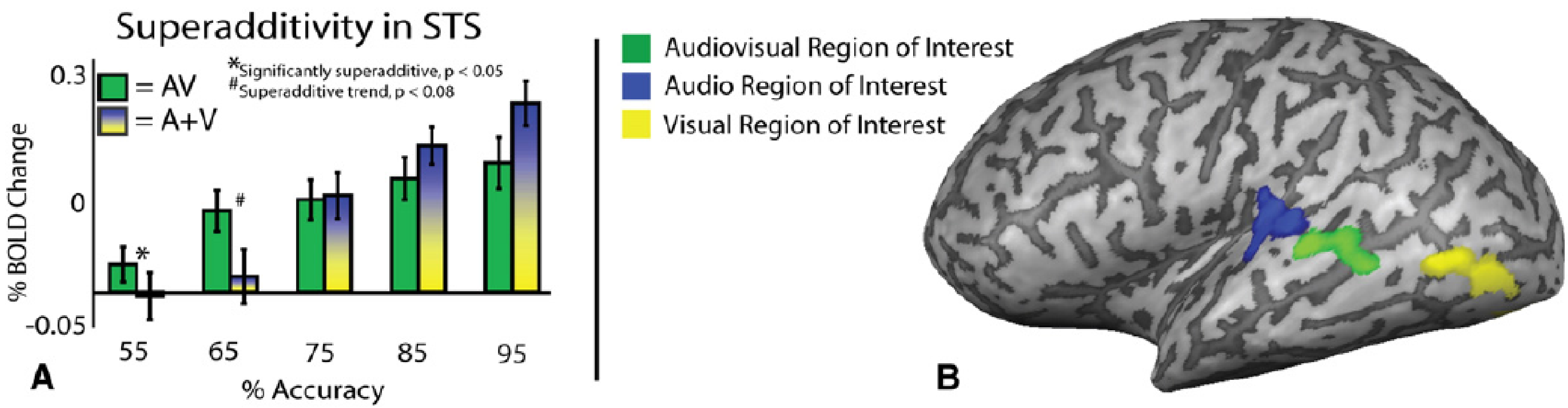
In this case, the lower stimuli were degraded movies with degraded speech and perceptually weaker than those in the upper graph.



Stevenson and James, NeuroImage, 2009

Inverse Effectiveness

Stevenson and James, NeuroImage, 2009



Inverse effectiveness: Multisensory effects are strongest when the input is hard to recognize (e.g., reduced contrast / speech in noise).

In this fMRI study, brain responses in the superior temporal sulcus (STS, green area on the right) showed the largest (AV>A+V) difference for the weak stimuli (55% and 65% accuracy).

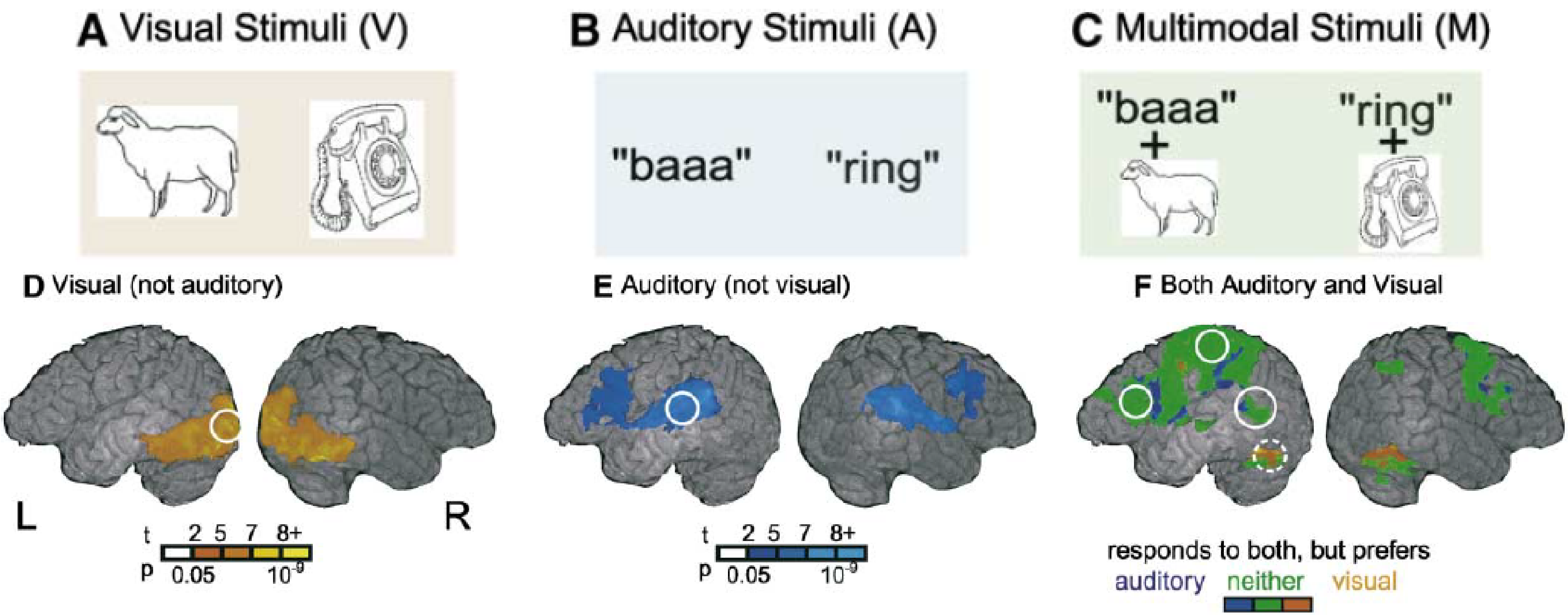
“Principles” of multisensory integration

1) Temporal congruence:
   multisensory responses (AV>A+V) are stronger for temporally congruent AV stimuli.

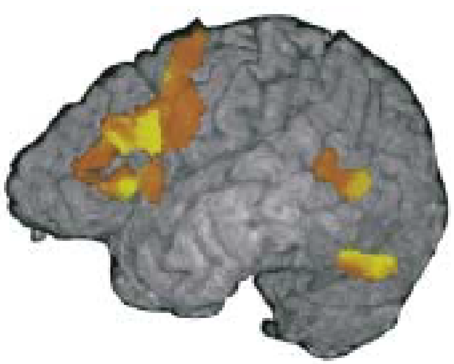
2) Spatial congruence:
   multisensory responses (AV>A+V) are stronger for spatially congruent AV stimuli.

3) Inverse effectiveness:
   multisensory responses (AV>A+V) are often stronger for AV stimuli that consist of weak unisensory (A or V) stimuli.

What about "contents"?



Enhanced Multimodal Response



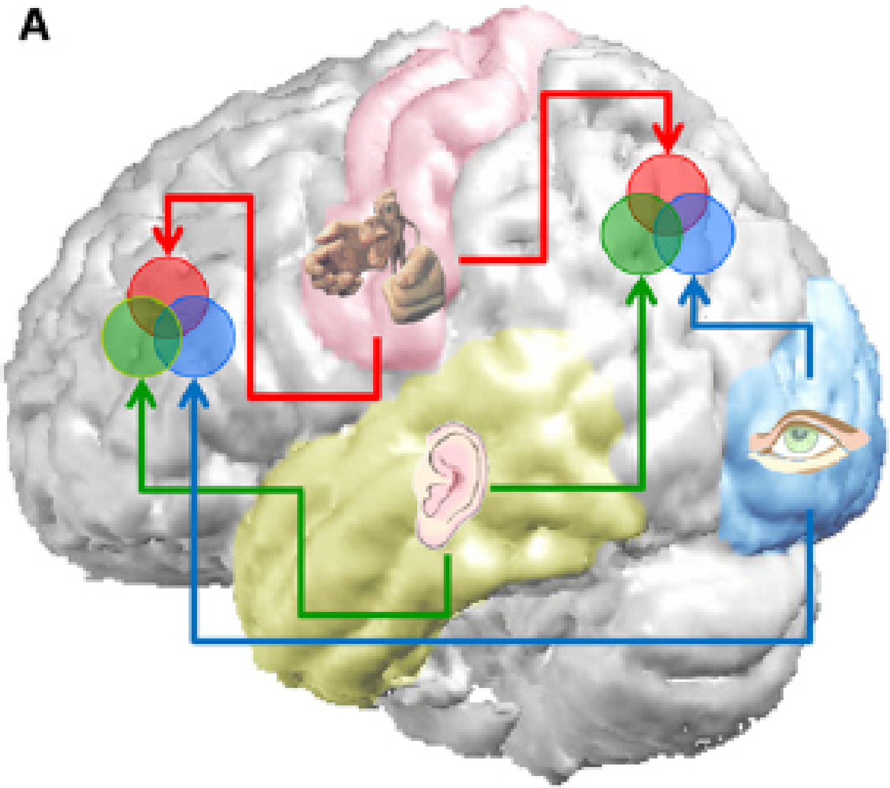
For cortical areas, <u>semantic congruence</u> is also important: areas in dorsolateral prefrontal cortex, ventral temporal cortex, and <u>posterior superior temporal sulcus/middle temporal gyrus</u> respond more for <u>multimodal</u> than unimodal stimuli.

Beauchamp et al., Neuron, 2004

Model of multisensory integration

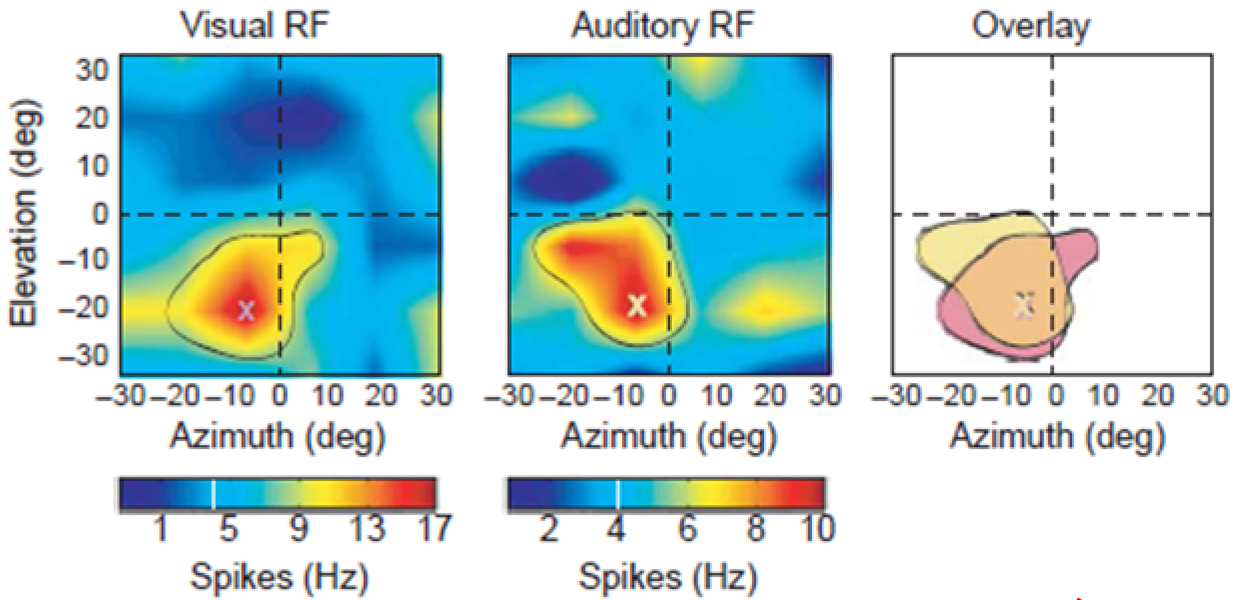
The traditional view of multisensory integration is that it occurs in multimodal association areas (A) where information from unisensory cortices is integrated.

Integration in multisensory association areas:

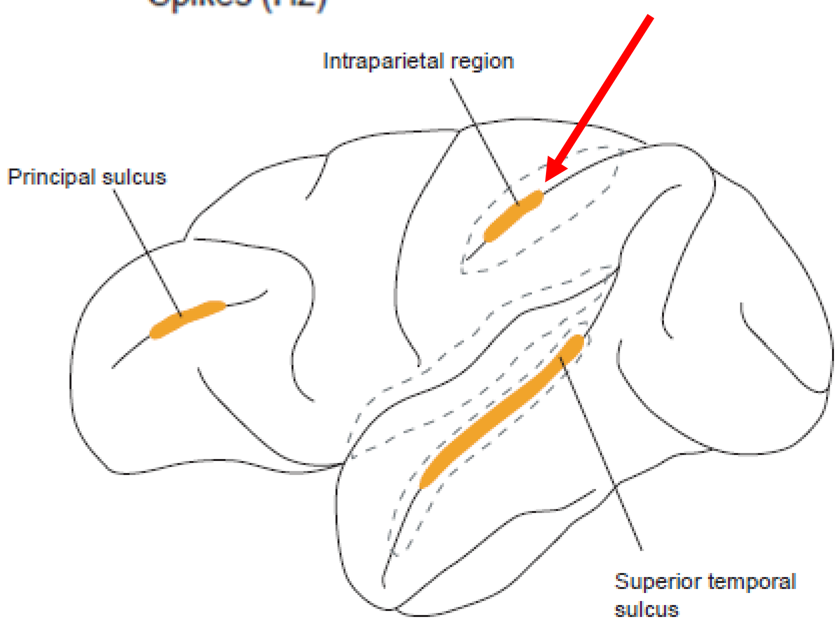


Murray and Spierer, 2011

Association Areas

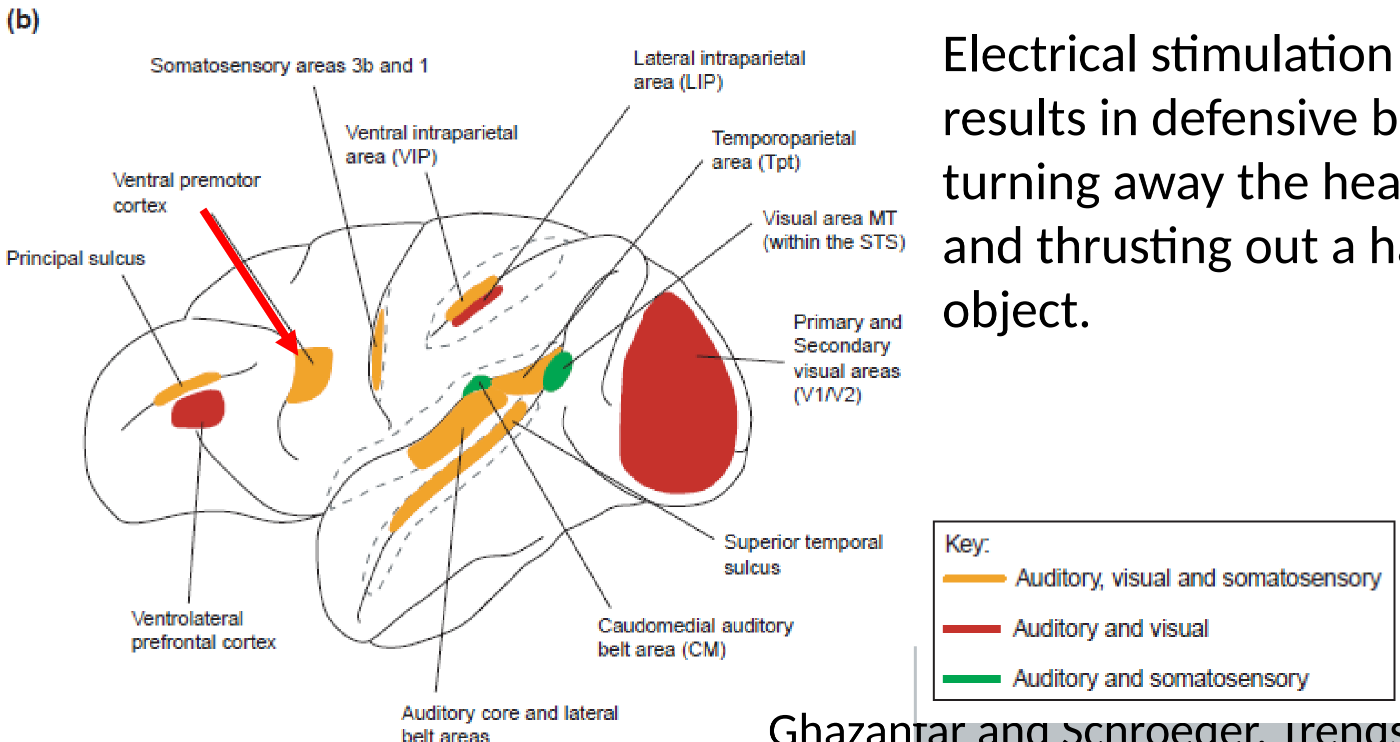


Some neurons in macaque area VIP (ventral intraparietal) have overlapping visual and auditory receptive fields. There is also somatosensory and vestibular input to area VIP which is involved in the multisensory guidance of <u>head and mouth movements</u>.



Ghazanfar and Schroeder, Trends in Cognitive Science, 2006

Association Areas



Multisensory neurons (left image: visual and somatosensory) in premotor cortex may be related to defense behaviors and are very sensitive to approaching or nearby stimuli (auditory, visual, and somatosensory).

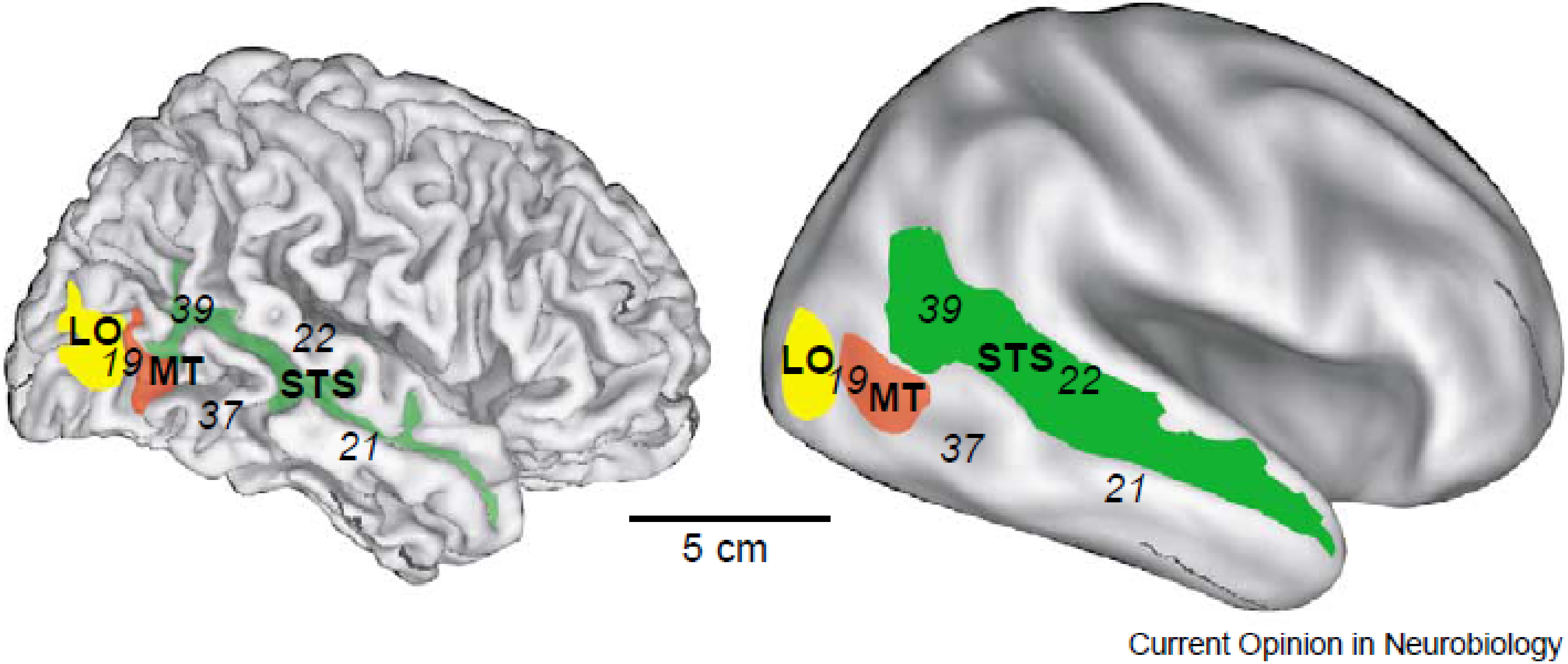
Electrical stimulation of these areas results in defensive behaviors such as turning away the head, closing the eyes, and thrusting out a hand to fend off an object.

Ghazanfar and Schroeder, Trends in Cognitive Science, 2006

In Human (association areas)

Human audiovisual integration areas:

Beauchamp, 2005



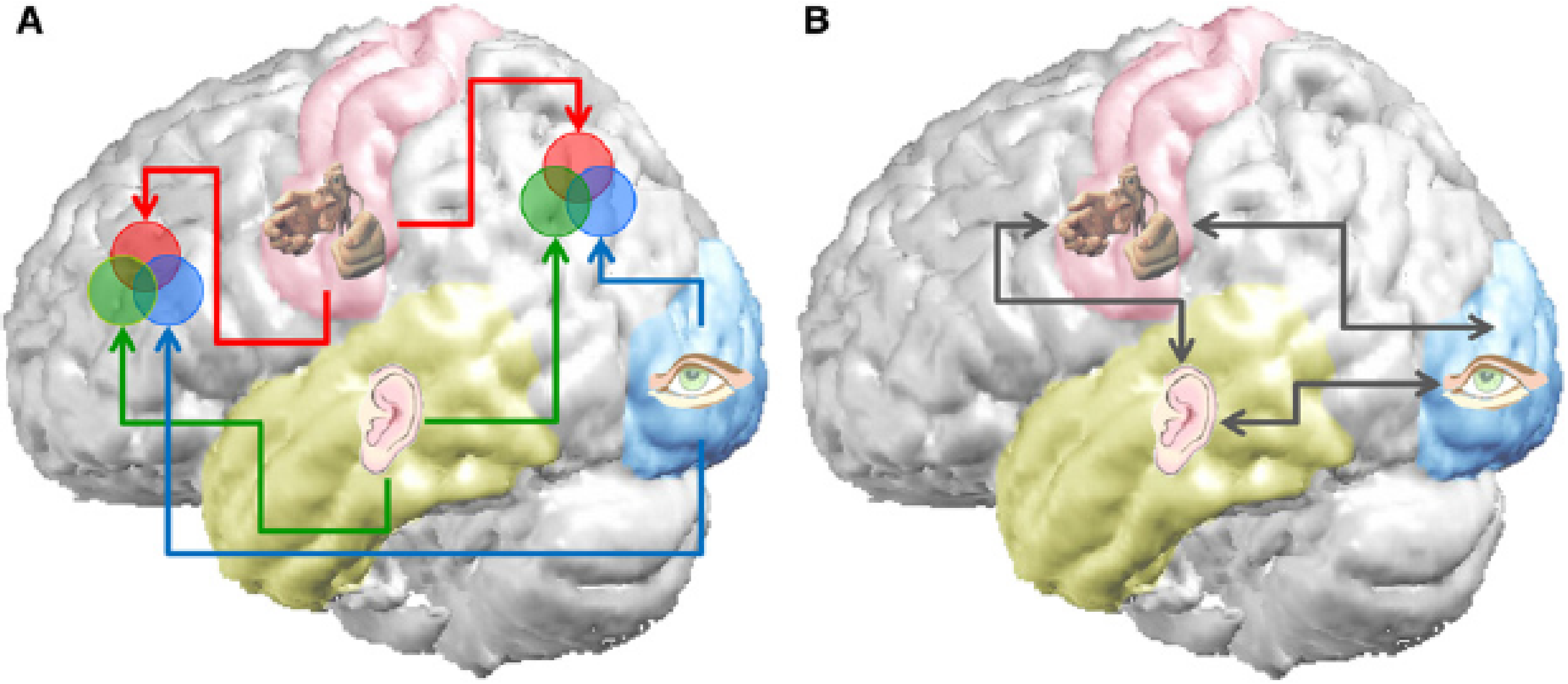
STS: superior temporal sulcus, responds to auditory and visual stimuli, such as:
speech sounds - speaking faces (e.g., McGurk effect)
speech – letters
object sounds - pictures

New models of multisensory integration

Recent studies suggest direct links between unisensory cortices (B) that underlie multisensory integration.

Integration in multisensory association areas:

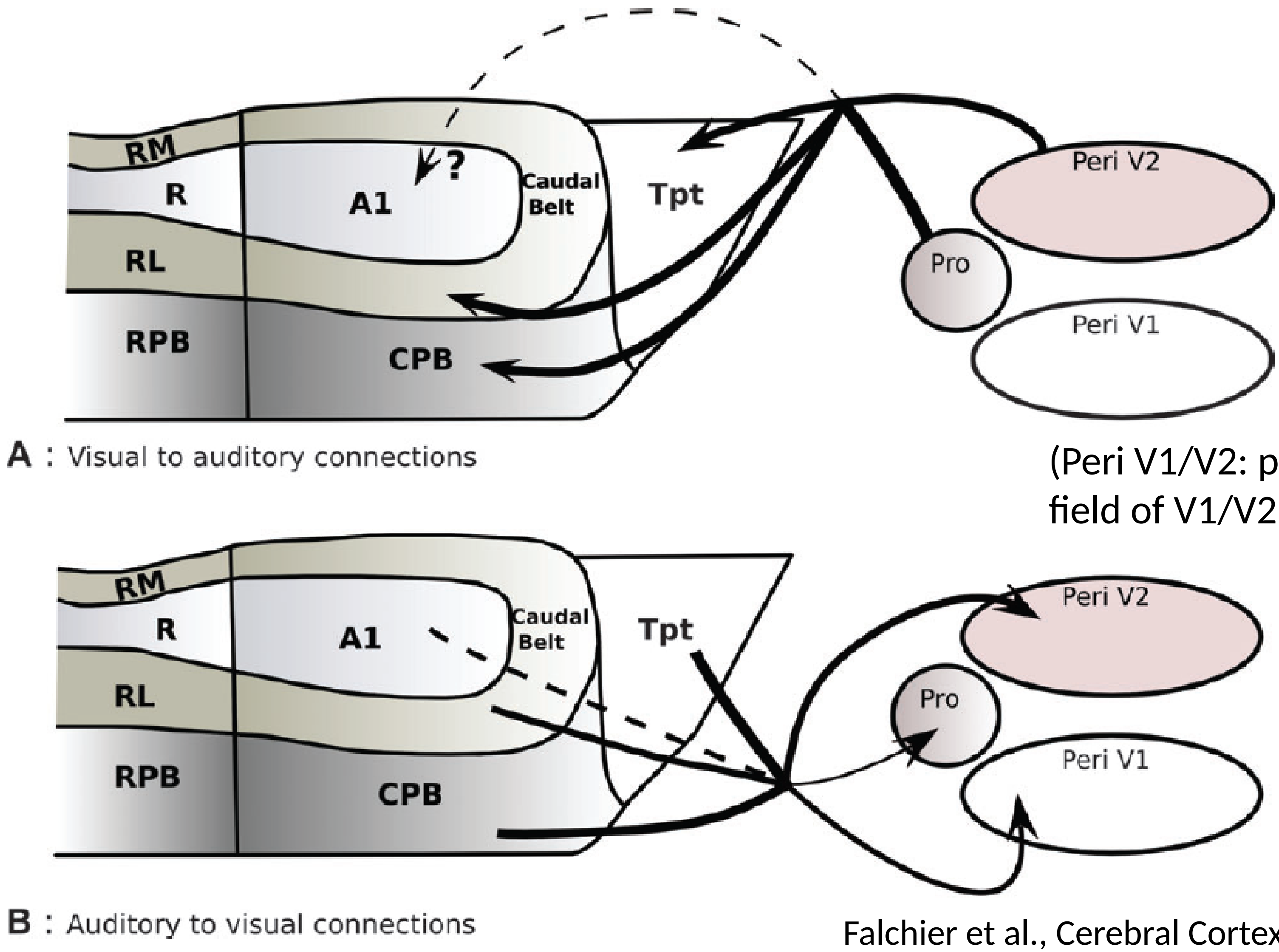
Integration in unisensory association areas:



Murray and Spierer, 2011

Direct sensory cortex connections

Anatomical connections between auditory and visual cortex allow for an early multisensory integration:



(Peri V1/V2: peripheral field of V1/V2)

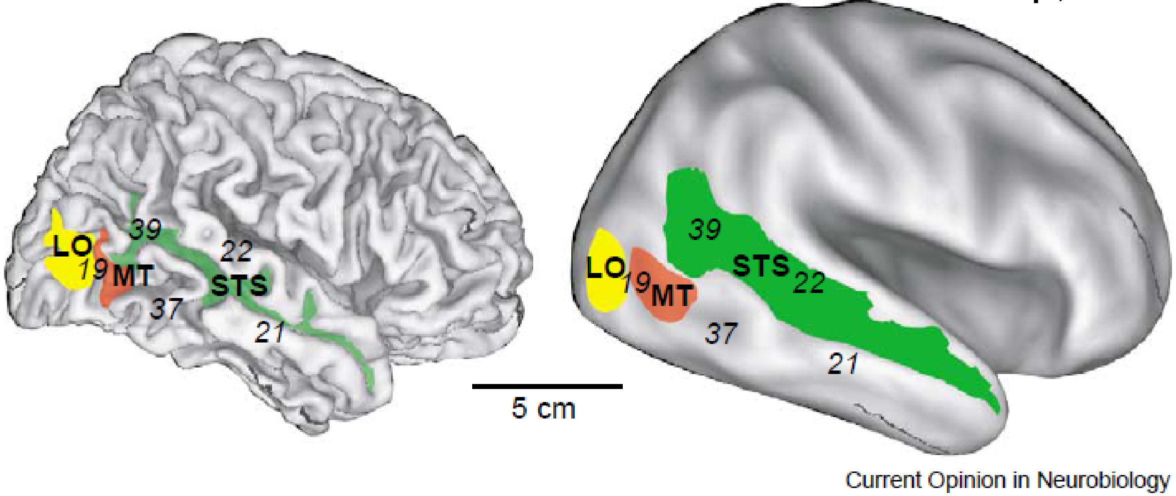
Falchier et al., Cerebral Cortex, 2010

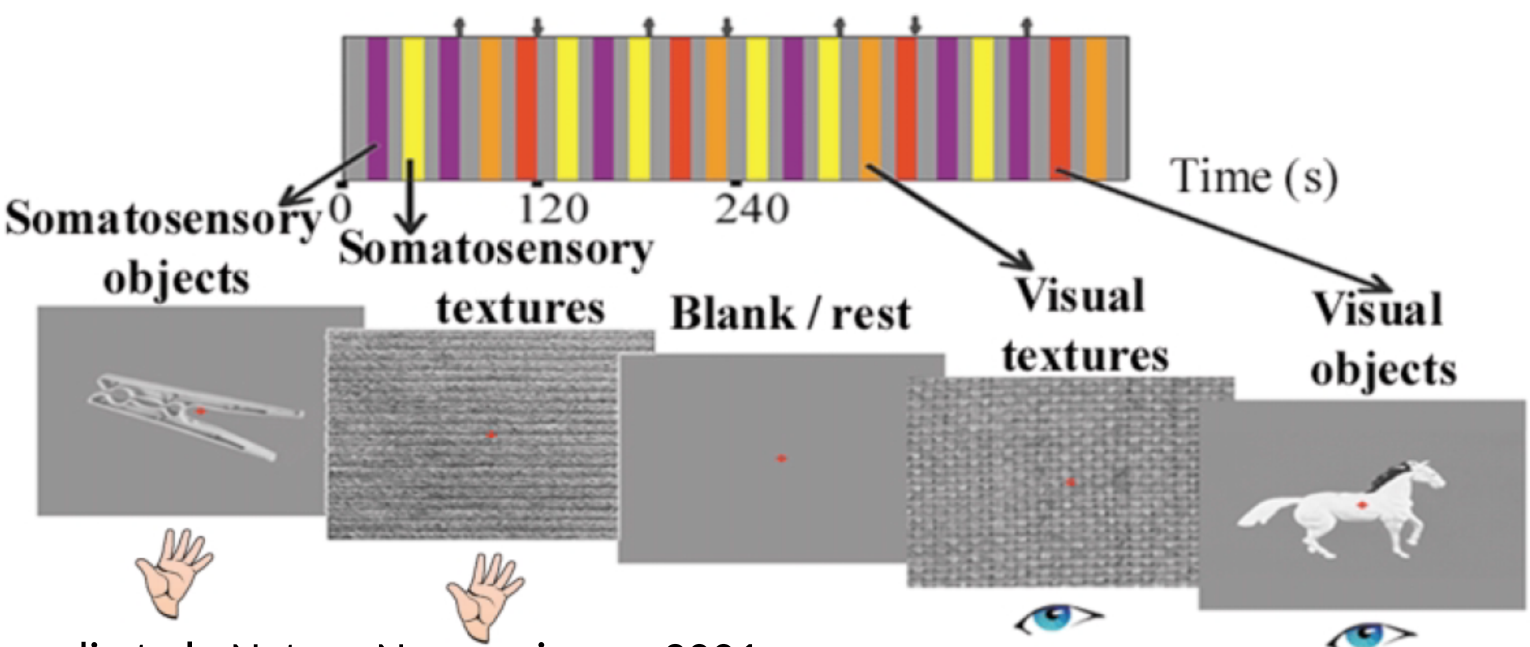
“Higher” visual areas?

Amedi et al. (2001) have shown that a higher order visual area (LO: lateral occipital) responds not only to visual objects, but also to tactile objects.

Control stimuli: visual/tactile textures.

Beauchamp, 2005





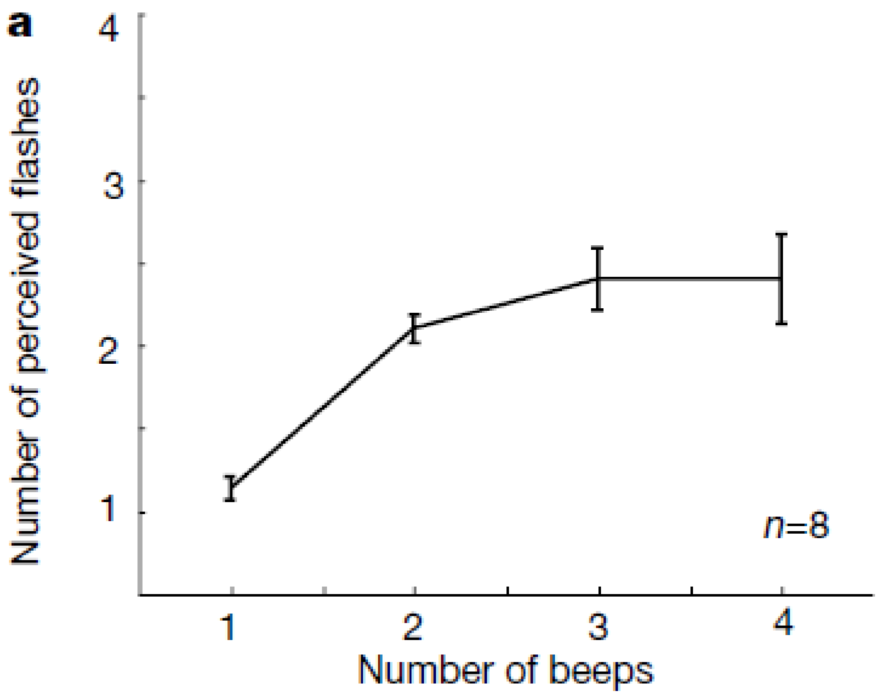
Amedi et al., Nature Neuroscience, 2001

Ventriloquism: Visual Capture

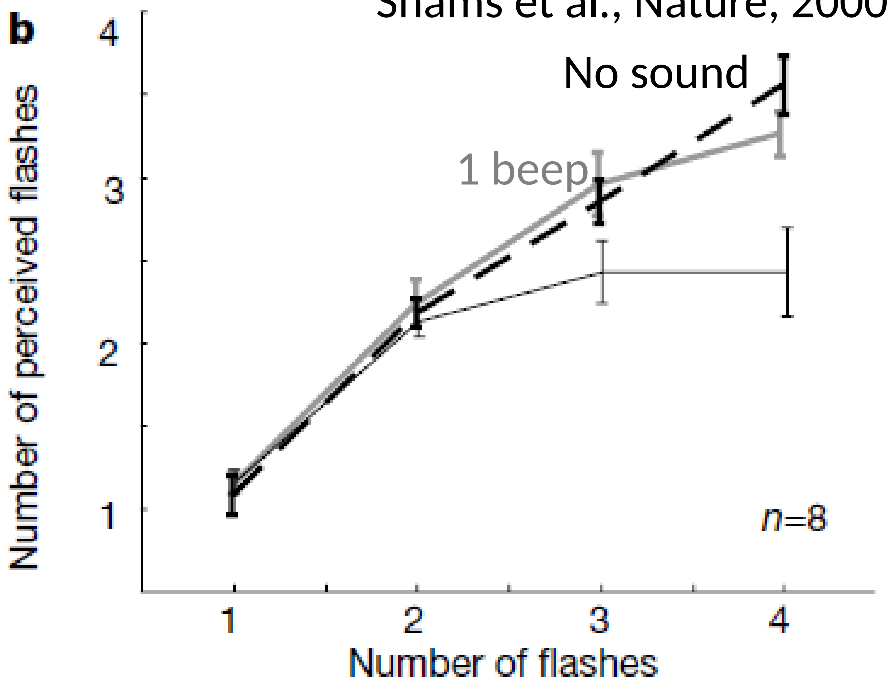
Sound Source appears shifted

<u>Visual cues</u> for source location can shift the perceived <u>location</u> of a sound source: ventriloquism effect. Higher reliability of the visual cue for spatial location has been proposed to underlie this effect.

Flashes and Beeps



Shams et al., Nature, 2000



The number of perceived visual flashes (which is always 1) is influenced by the number of beeps.

Auditory cues can shift the perceived timing of a visual flash. Higher reliability of the auditory cue for temporal information has been proposed to underlie this effect.

Synesthesia

Synesthesia is a condition in which stimulation of one sensory modality causes unusual experiences in a second, unstimulated modality (Hubbard and Ramachandran, Neuron, 2005).

e.g.,

Grapheme -> Color

Music -> Color

Numbers -> Objects

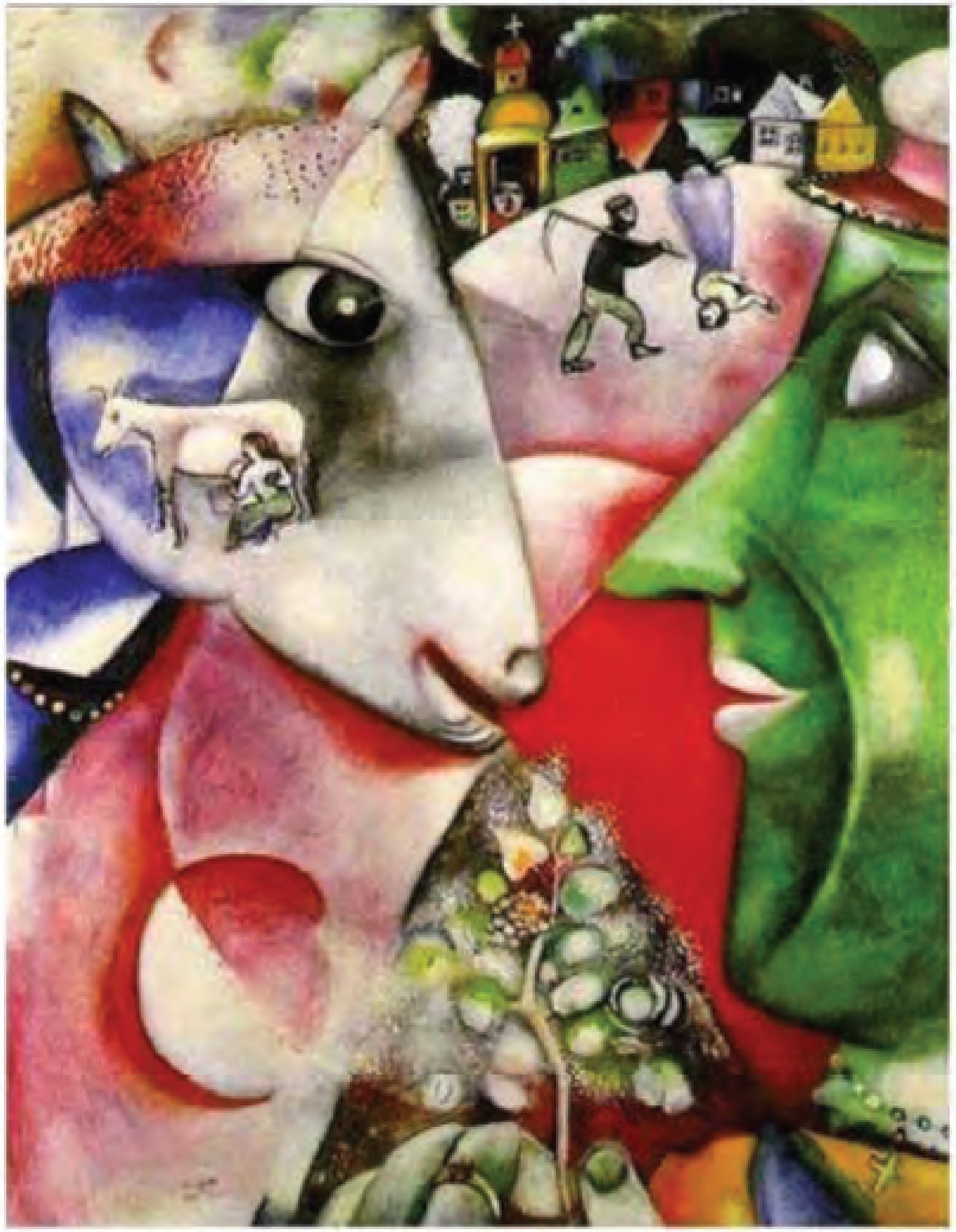
“Trumpet Sound”

Safran and Sanda, Current Opinion in Neurology, 2015

Synesthesia

Marc Chagall
I and the Village, 1911

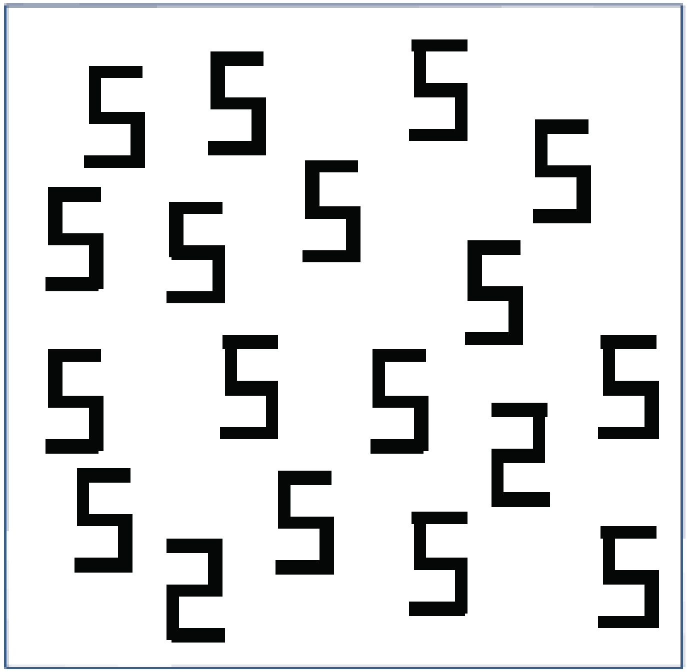
Personality-Color
Synesthesia?



Objective test for synesthesia

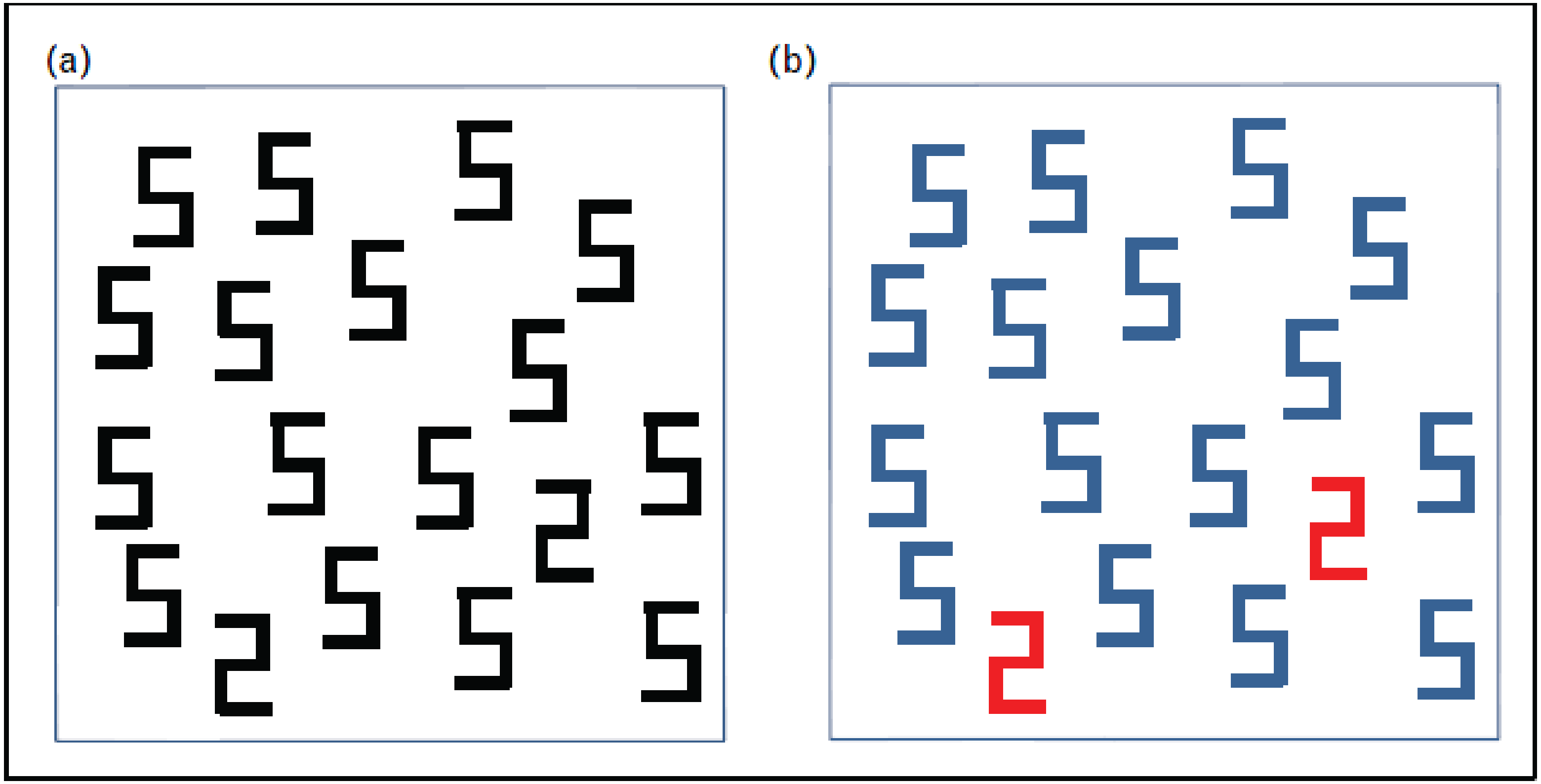
Where are the 2s?

(a)

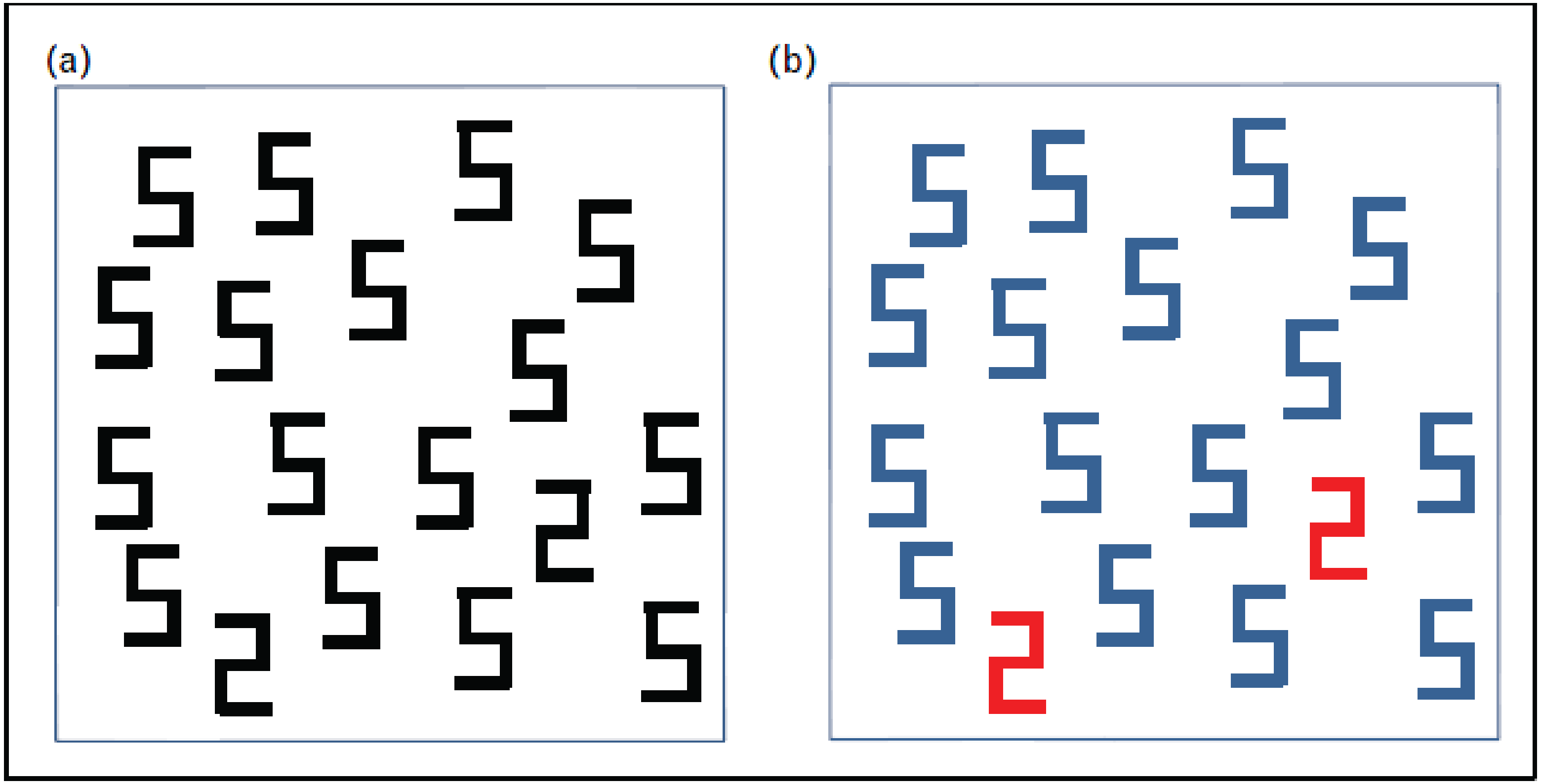


Objective test for synestheisa

Where are the 2s?



Objective test for synesthesia



The task here is to find the "2" in "5" (pop-out)
a) is difficult for non-synesthetes, but easy for synesthetes, because the different perceived colors make it similar to b).

Summary (Multisensory Integration)

- Multisensory integration can occur as response enhancement or depression. It is maximal when multimodal stimuli have spatial and temporal congruence.

- Inverse effectiveness refers to the observation that multisensory integration is most effective/important when unisensory information is weak.

- In cortical areas, semantic congruence (e.g., seeing a cat and hearing ‘meow’) plays an important role for multisensory integration.

- Recent studies suggest that in cortex, multisensory integration not only occurs in multisensory association areas (superior temporal sulcus, intraparietal sulcus, prefrontal cortex), but also in ‘unisensory’ areas (such as visual or auditory cortex).

- Synesthesia is a special case of multisensory integration.