

# IS SPEECH A SPECIAL SOUND FOR THE BRAIN?

---

Perception of ambiguous stimuli:  
Audiovisual integration while perceiving  
sinewave speech

# Ambiguous stimulus: Sine-wave speech

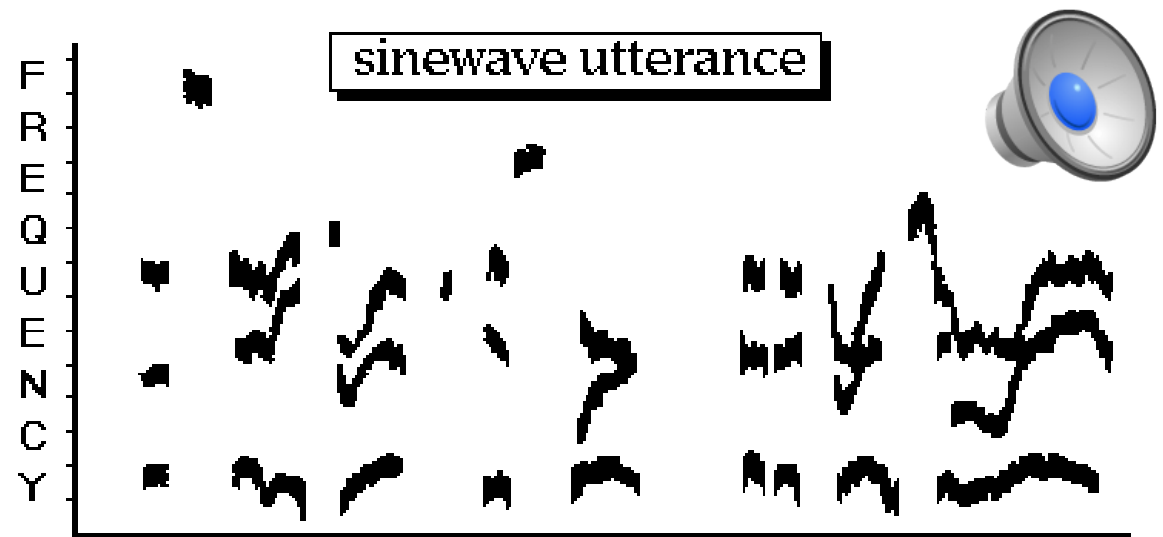
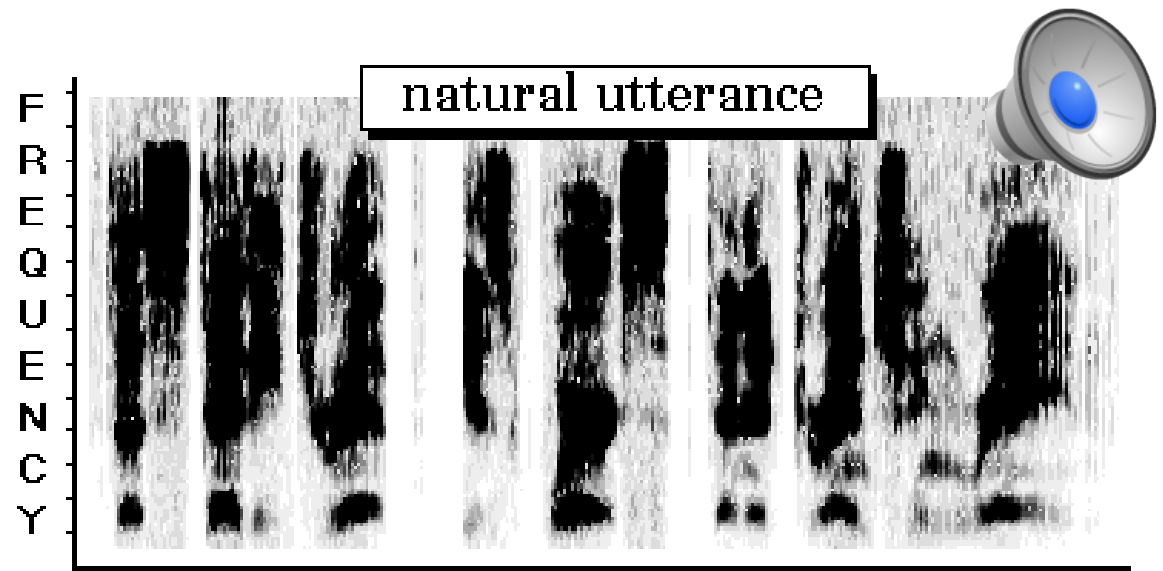


- Remez et al. (1981):
  - Naïve subjects fail to perceive sine-wave stimuli as speech.
  - However, if subjects are instructed about the speech-like nature of sinewave speech, they can easily understand it.
  - Argument for a special “speech mode” of processing.













# Ambiguous stimulus: Sine-wave speech

A spectrogram, or an acoustic "picture" of a speech utterance. Time is represented on the horizontal axis, frequency on the vertical axis. Amplitude corresponds to the darkness.

A sinewave replica of the natural above. All fine-grain acoustic properties of speech are discarded and only the coarse-grain changes in the spectra over time are retained.



## Sine-wave speech

Sentence 1	SWS 	Original 
Sentence 2	SWS 	Original 
Sentence 3	SWS 	Original 
Sentence 4	SWS 	Original 
Sentence 5	SWS 	Original 
Sentence 6	SWS 	Original 

# Approach

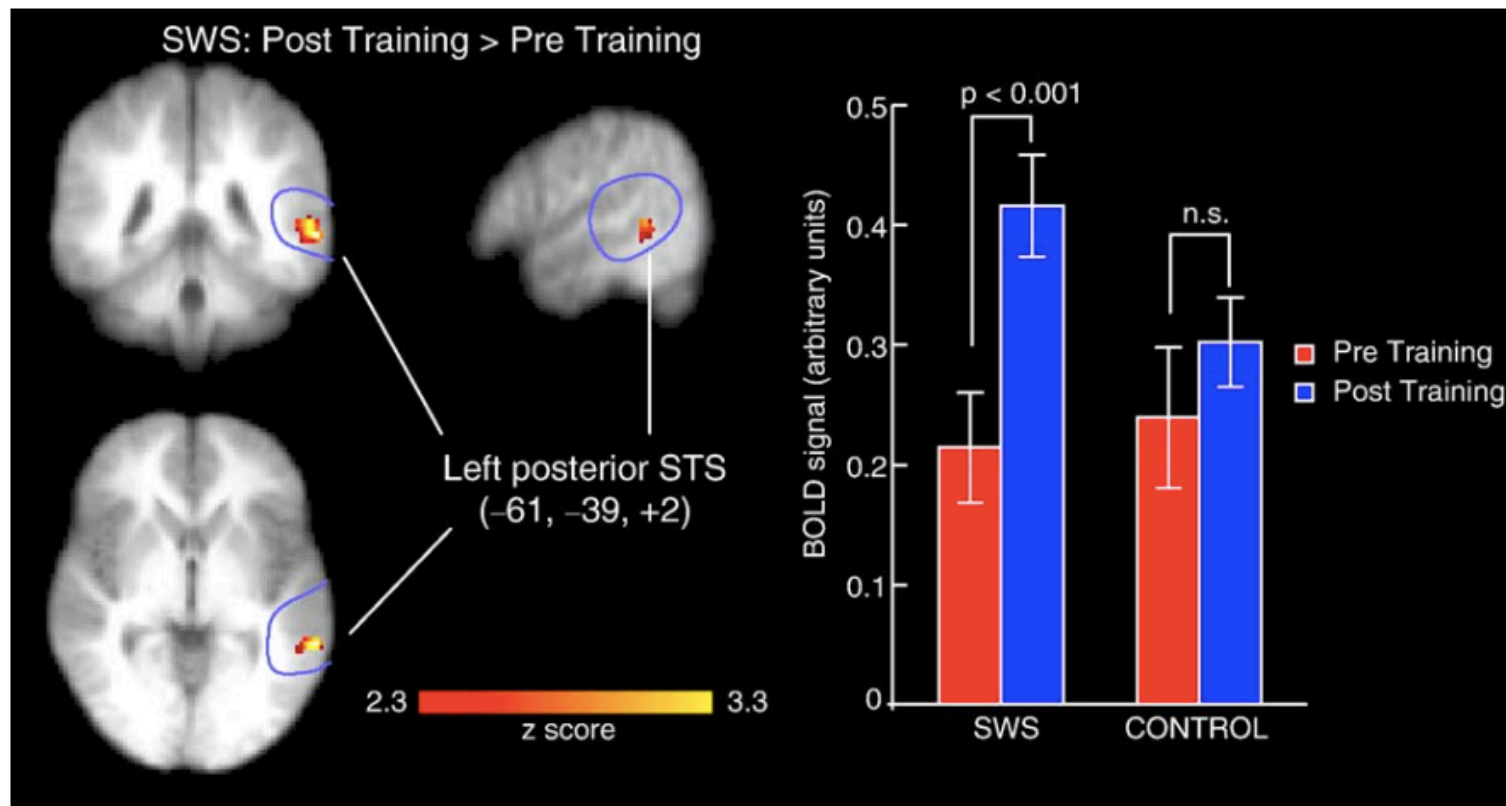
- Have subjects perceive sinewave stimuli either as speech or as non-speech and then study some aspect of processing under those two conditions.




**Example 1:**  
**Neural effect of perceiving a signal as speech**

## Neural effect of perceiving sine-wave speech as speech vs. non-speech

- Möttönen et al. (2006, *NeuroImage*):
  - Sinewave **nonwords** are first discriminated as non-speech.
  - Then the subject is told that the stimuli can be understood as tokens of speech.
  - Perception as speech increases activity in left superior temporal cortex.





**Example 2:**  
**Audiovisual integration while perceiving  
speech vs. non-speech**



# McGurk effect



## Role of articulatory gestures on speech perception

- When the visual image of a person saying “ga” is combined with an audio recording of “ba”, the percept is “da” (or for some people “ga”).
- Incongruent articulatory gestures can *change* the auditory percept even when **the signal is clear**.

## Audio-visual integration is NOT specific to speech

- Saldana and Rosenblum (1993):
  - Audio–visual integration of the “plucks” and “bows” of cello playing.
  - Not only speech, but also other ecologically valid combinations of auditory and visual stimuli can integrate in a complex manner.

## But:

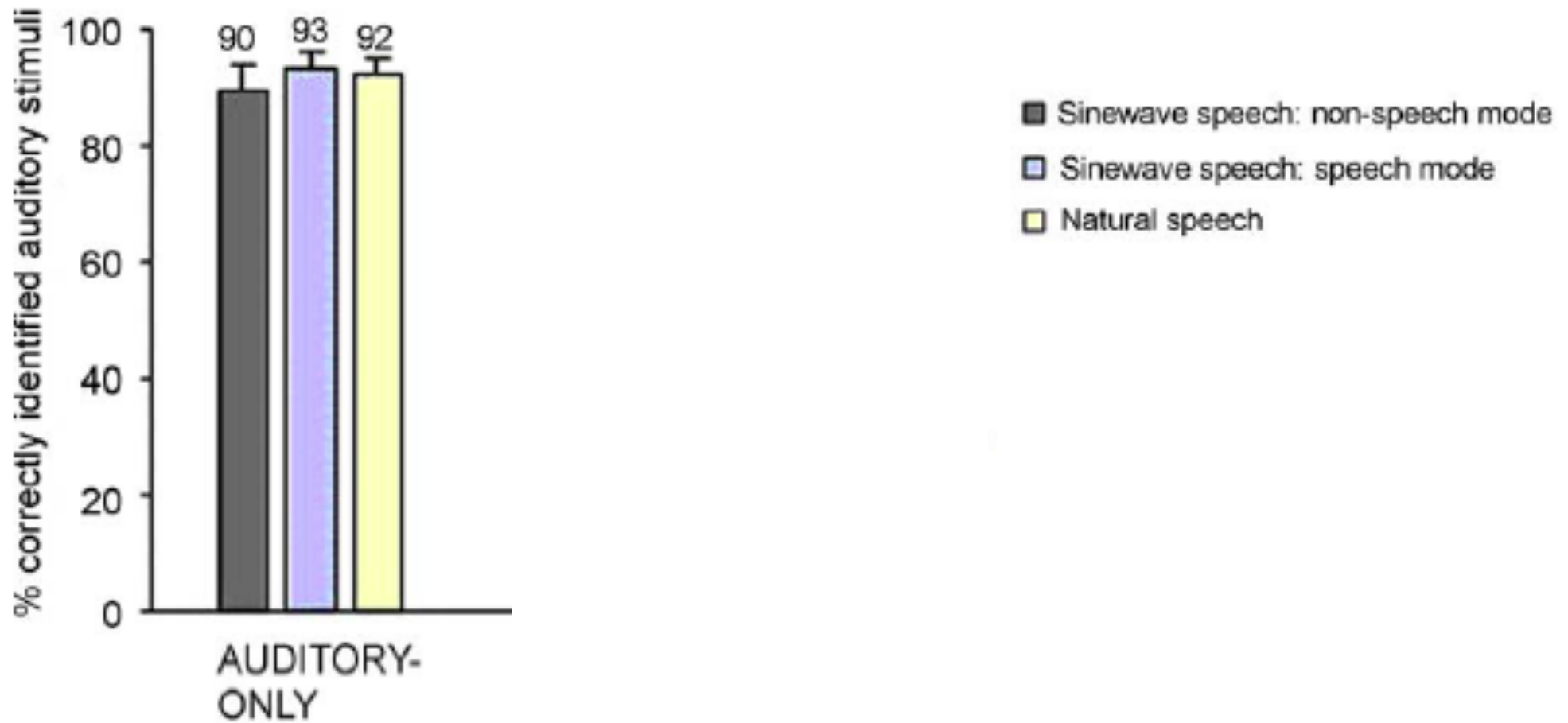
- Audio-visual integration **is used in speech much more than in** the perception of **non-speech** sounds.
- Evidence comes from the effect of incongruent articulatory gestures on sinewave speech, perceived either as speech or not (Tuomainen et al. 2004, *Cognition*)

## Basic idea

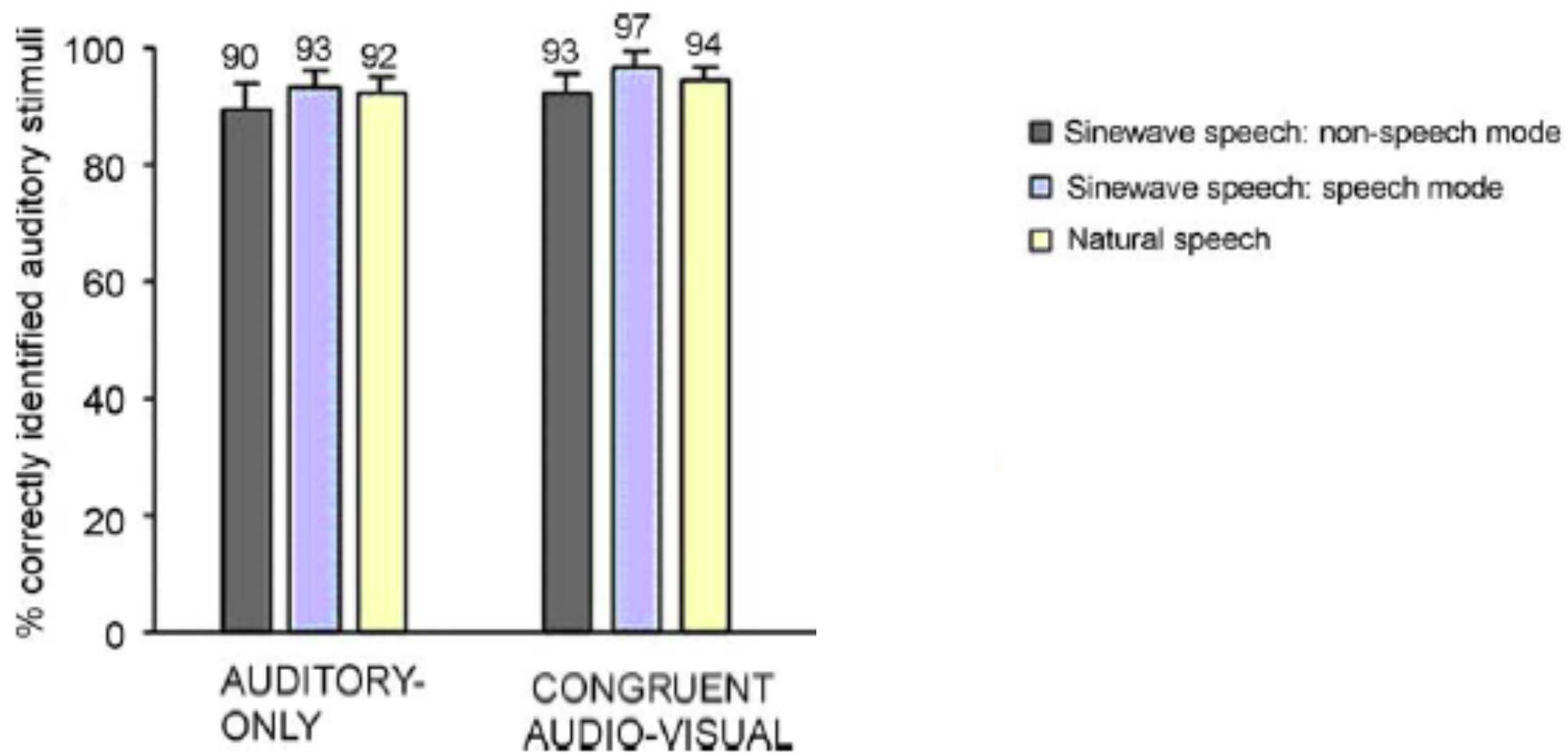
- Subjects watch videos of a mouthing head and the auditory signal either (A) matches the mouth movement or (B) mismatches the mouth movement.
  - If the mismatching mouth movement changes the perception, then that's the McGurk effect.
- The auditory stimuli are sinewave speech.
  - Sinewave replicas of the Finnish nonwords /omso/ and /onso/
  - Task: to indicate which one you heard.
- Subjects either think the stimuli are speech or not.
- Will the McGurk effect be larger when subjects think the stimuli are speech?
- The physical properties of the auditory stimuli are not changing!
- Natural speech is used as an extra control (straightforward McGurk effect is expected)

(Tuomainen et al. 2004, Cognition)

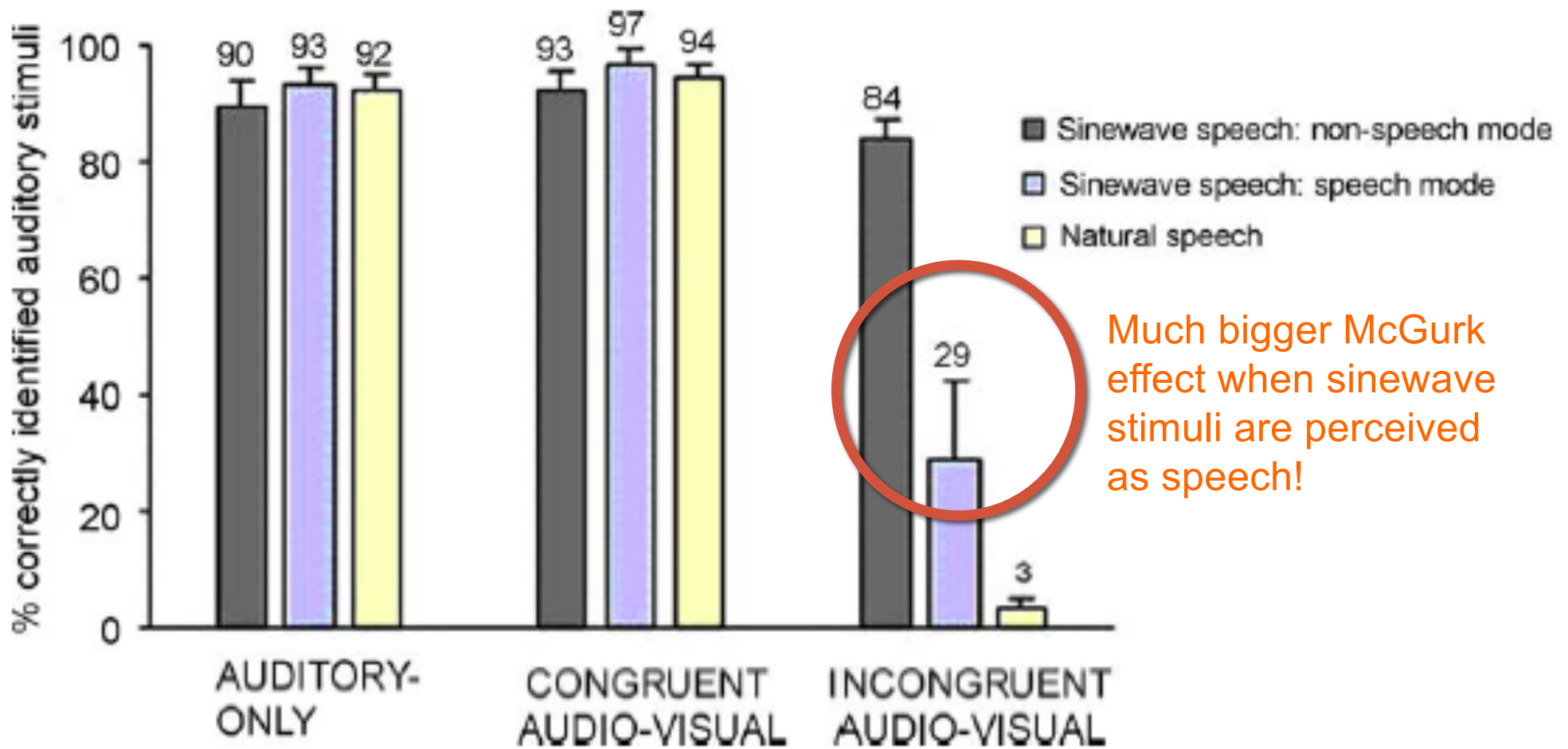
## Results



## Results



# Results





# Bottomline

- Whether speech is a special kind of sound for our brains is a really hard question that we still don't have a clear answer to.
- Controlling the physical properties of the signal is always the beast when addressing this question.
- Clever experimental design is needed.
- Audiovisual integration is very robust for speech, perhaps in a way that may be unique.