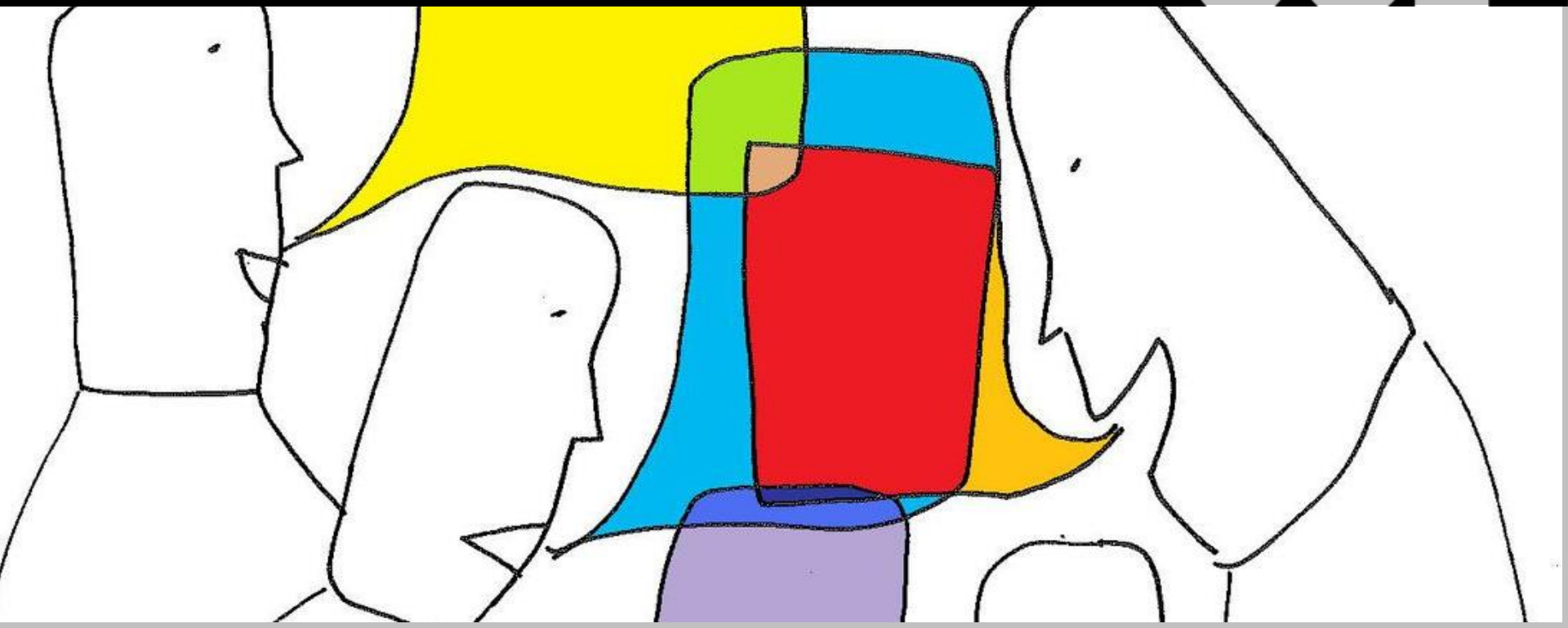
UCL



# Understanding speech in the background of other sounds

It's complicated!


BEEP
I'M SORRY I CAN'T HEAR YOU, WHAT
KNOCK
WHAT DID YOU SAY? SPEAK UP
yeah yeah yeah yeah
TALK TALK TALK
ZAP!
PLOP
swoosh
CLANG
BANG
SMASH

# Why is this interesting?

- Speech is really important …
  - and often occurs in the presence of other sounds
- People of every age vary a lot in how well they can understand speech in the presence of 'noise'
  - Even when they have normal hearing …
  - And impaired hearing makes everything worse!
- Lots of developmental disorders seem to have an impact on this ability
  - Language disorders (DLD & dyslexia)
  - Autism spectrum disorders
  - Auditory processing disorder (APD)

# Essential terminology

- ***signal*** or ***target***
  - what you are trying to listen to
  - typically speech or music or …
- ***'noise'*** or ***masker***
  - what you are trying to ignore
  - can be noise like from a hoover, but also other speech
- ***Signal-to-Noise Ratio*** (SNR)
  - the amount of energy in the signal divided by the amount of energy in the noise
  - Typically expressed in decibels (dB)

TOP DEFINITION

# signal to noise ratio

The ratio of useful information to useless information in any given statement.

by flame August 29, 2003

73 25

SHOP

6 MORE DEFINITIONS

Add your own

20 WORDS RELATED TO SIGNAL TO NOISE RATIO

spam • aol • internet • usenet • alt.binaries.erotica • bikeshedding



TEN WORDS TRENDING NOW

modulations
dual tasking
cognitive load
concurrent task
informational masking
pupillometry
cocktail party
non-native listening
vocal folds
McGurk effect

# Getting to grips with SNRs

SNR = +20 dB 

SNR = +10 dB 

SNR = 0 dB 

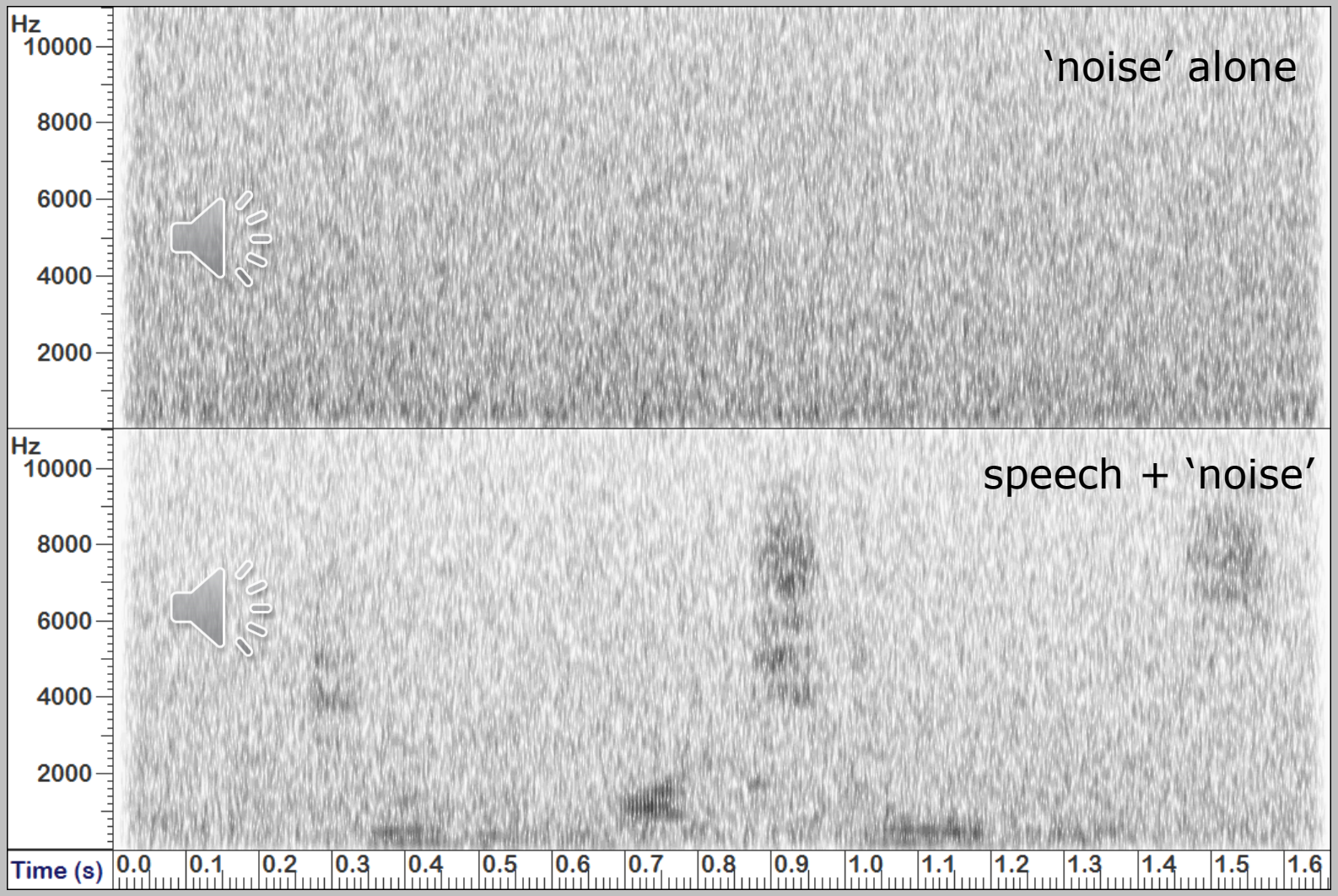
SNR = -5 dB 

SNR = -7 dB 

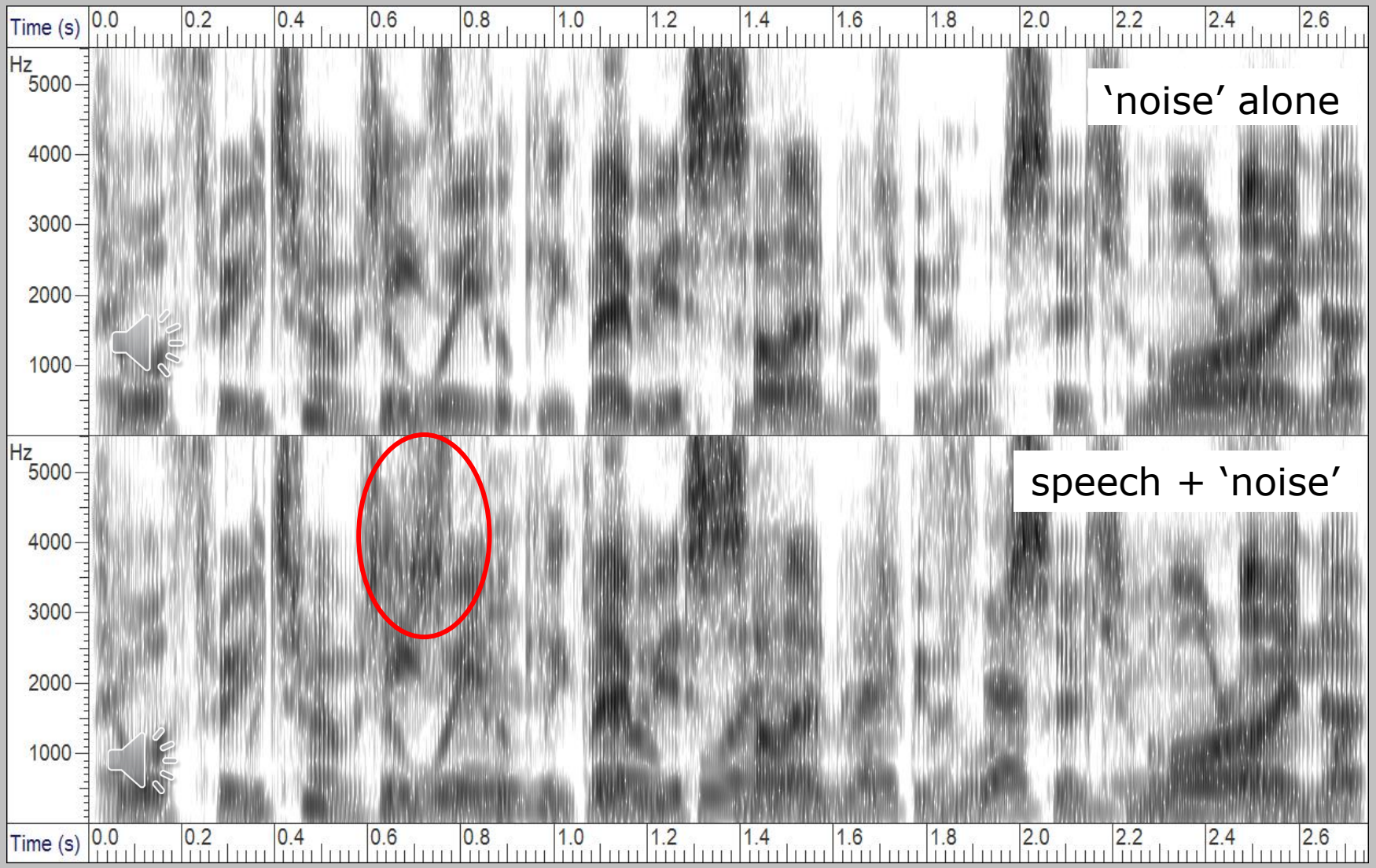
SNR = -10 dB 

SNR = -20 dB 

# There are lots of different kinds of ‘noises’



# Another kind of ‘noise’

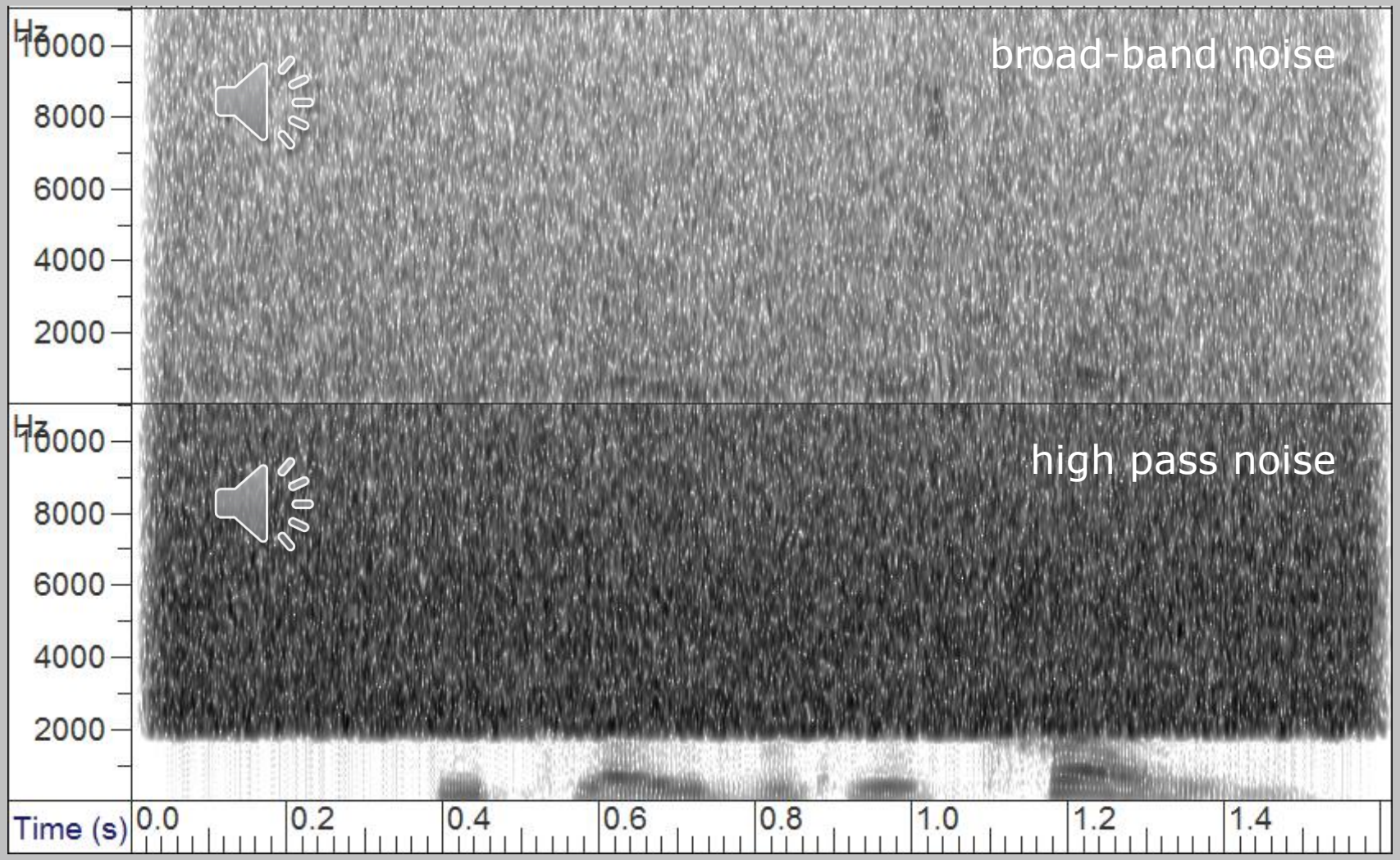



‘show’ starts at t≈0.65 ms

# A long-held distinction between two kinds of masking

- *Energetic masking*
  - maskers interfere with speech to the extent that have energy in the same time/frequency regions
  - primarily reflecting direct interaction of masker and speech in the cochlea
  - making essential information inaudible
  - e.g., a speech-spectrum-shaped noise, like the sound of a boiling kettle
- Two features determine the effectiveness of steady-state energetic maskers
  - the spectral shape of the masker
  - SNR
  - but fluctuations in energy matter a lot too

# Spectral shape is important



# A long-held distinction between two kinds of masking

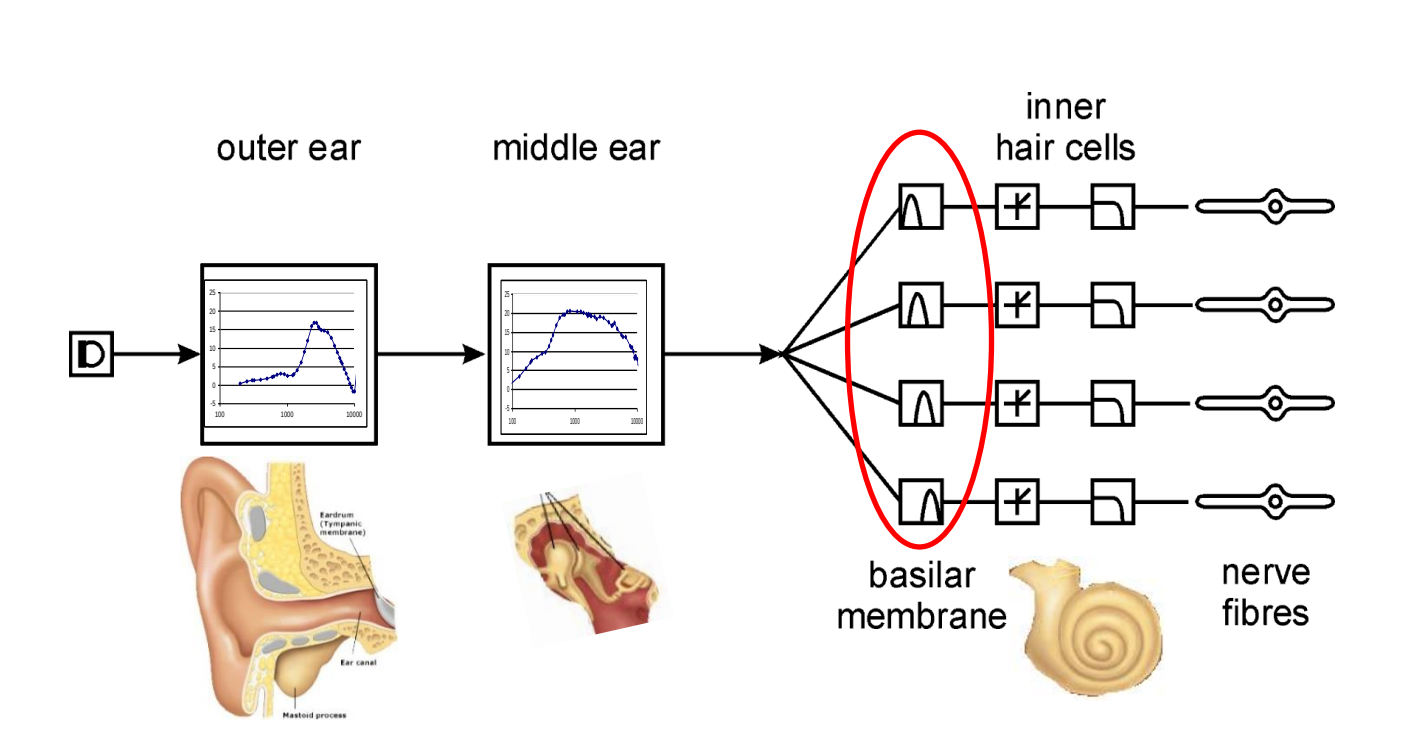
- *Informational masking*
  - everything that is not energetic masking!
  - Something to do with target/masker similarity?
    - signal and masker 'are both audible but the listener is unable to disentangle the elements of the target speech from a similar-sounding distracter'
  - e.g., other talkers

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *J Acoust Soc Am, 110,* 2527-2538.

# Updated notions about Energetic Masking (EM)

- 'Energetic masking' appears to be two different things
  - Genuinely energetic masking (as described before)
  - *Modulation* masking (MM)

# The auditory periphery as a signal processor




125 Hz

250 Hz

500 Hz

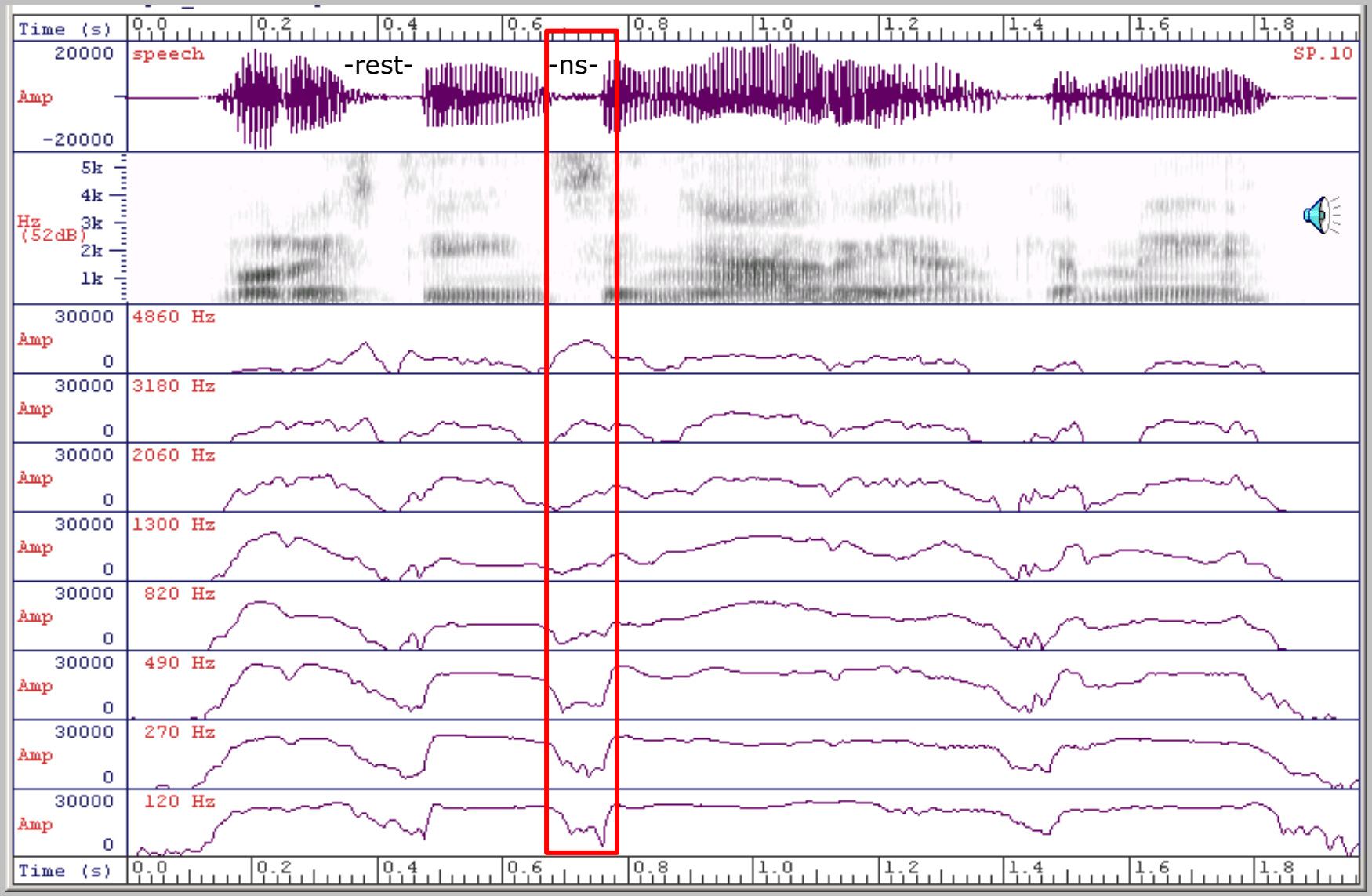
1 kHz

2 kHz

4 kHz

8 kHz

# Speech intelligibility relies on envelope modulations



# Updated notions about Energetic Masking (EM)

- Modulation Masking is the disruptive effect that modulations in the masker have on the modulations in the target
  - So it's not the *energy* in the masker that is so important
  - Similar to EM, in happening at the periphery (needing to be in the same time/frequency region)

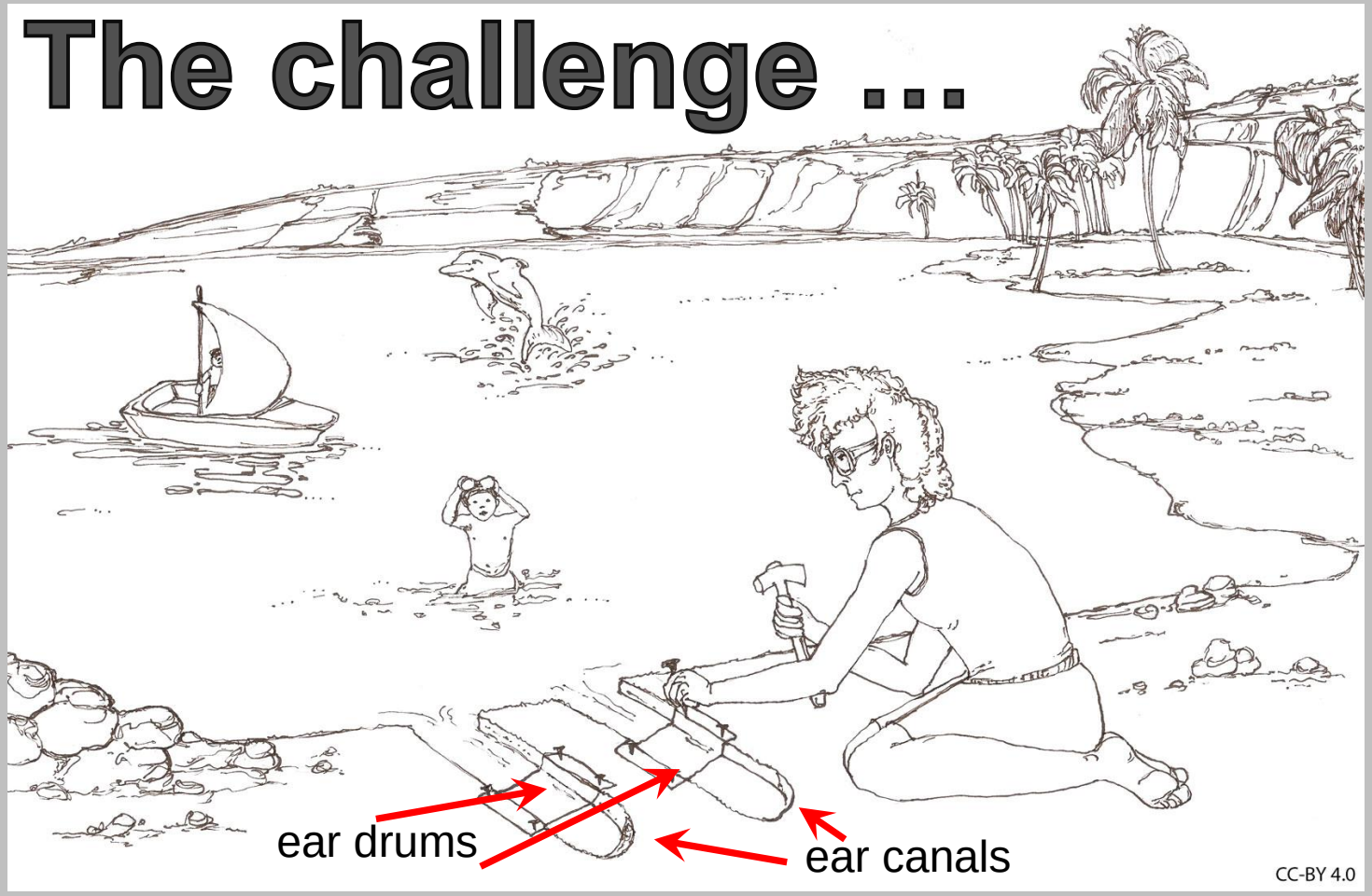
Stone, M. A., Fullgrabe, C., & Moore, B. C. J. (2012). Notionally steady background noise acts primarily as a modulation masker of speech. *J Acoust Soc Am*, 132, 317-326.

Jorgensen, S., Ewert, S. D., & Dau, T. (2013). A multi-resolution envelope-power based model for speech intelligibility. *J Acoust Soc Am*, 134(1), 436-446. doi:10.1121/1.4807563

# Updated notions about Informational Masking (IM)

- 'Informational masking' appears to involve at least two different processes
  - Object formation, or *auditory scene analysis (ASA)*
  - Object selection, related to attention

# The challenge …




*From the work of Al Bregman:* Two narrow channels, with handkerchiefs stretched across, are dug into the side of a busy lake. Objects and events on the lake make waves that cause the two handkerchiefs to move.

What can you say about what is happening on the lake by observing the motion of the handkerchiefs? For example, how many boats are there, where are they, and in what direction are they moving? What else is going on in the lake?

# Updated notions about Informational Masking

- Problems in 'object formation'
  - Related to auditory scene analysis
  - similarities in auditory properties make segregation difficult
    - voice pitch, timbre, rate
- Problems in 'object selection'
  - Related to attention and distraction
  - the masker may distract attention from the target
    - e.g., hearing your name
- And more? e.g., Competition for cognitive resources?

Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends In Cognitive Sciences, 12,* 182-186.

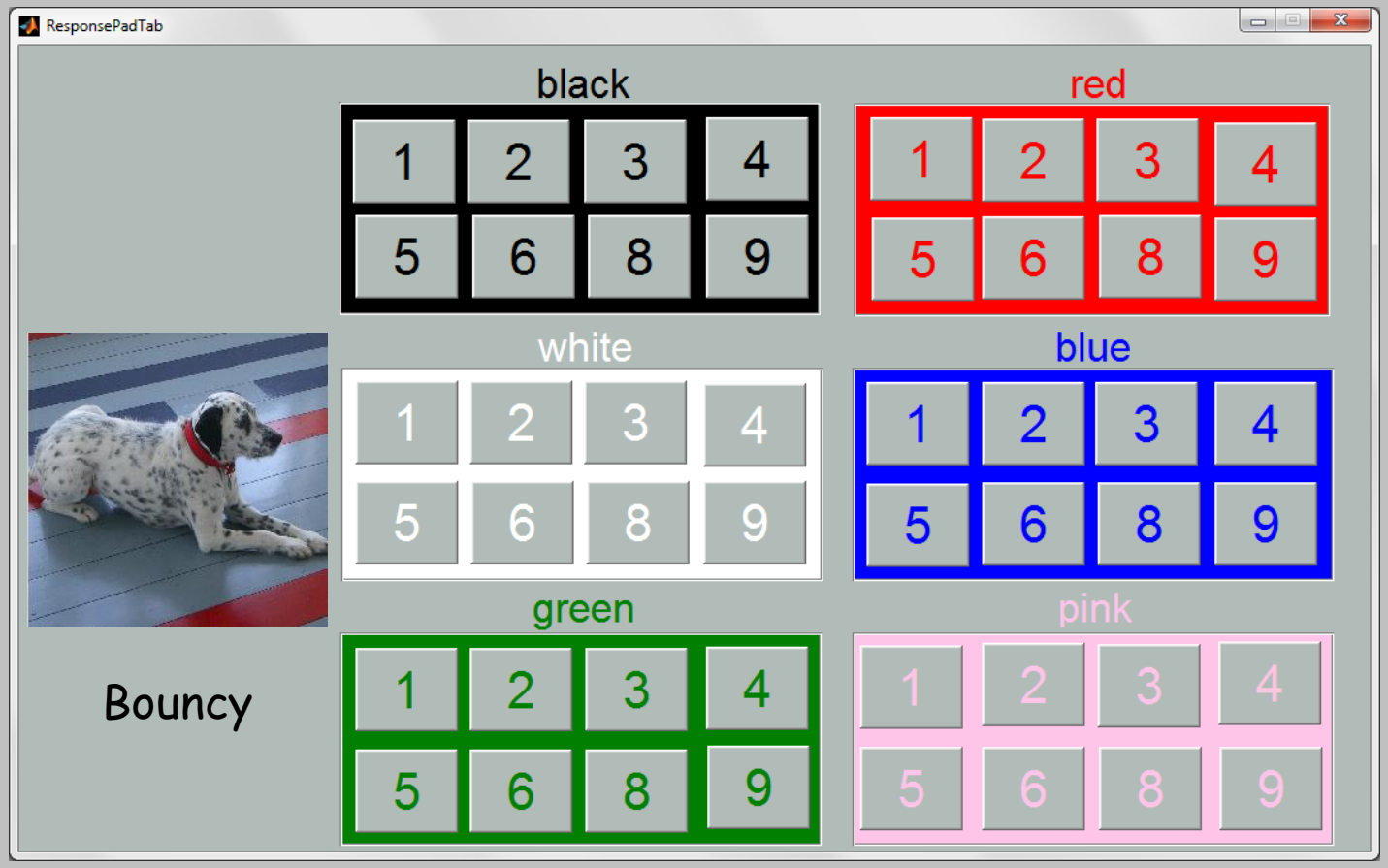
# EM/MM & IM operate at different points in the auditory pathway

- Energetic/Modulation Masking at the periphery, primarily in the cochlea
  - Early developing abilities?
  - Likely to be affected by sensori-neural hearing loss (SNHL)
- Informational Masking at higher centres, in the brain
  - Late developing abilities?
  - Likely affected by various neurological disorders

# What task to use?

- Minimise effects of …
  - vocabulary
  - syntax
  - use of contextual cues
- In short, minimise *language* aspects, to focus on abilities related directly to masking

# The Children's Coordinate Response Measure (CCRM)

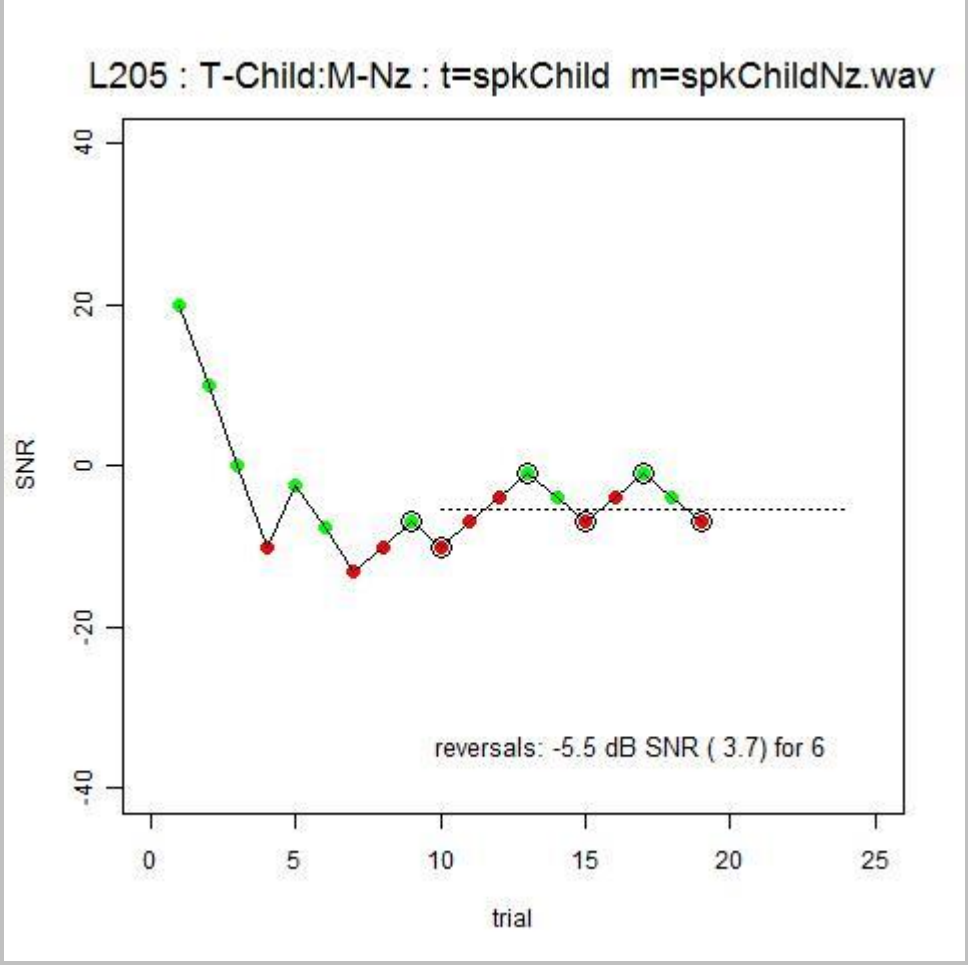



in quiet

in steady noise

against another talker

# Performance measured adaptively




Vary the:

***Signal-to-Noise Ratio* (SNR)**

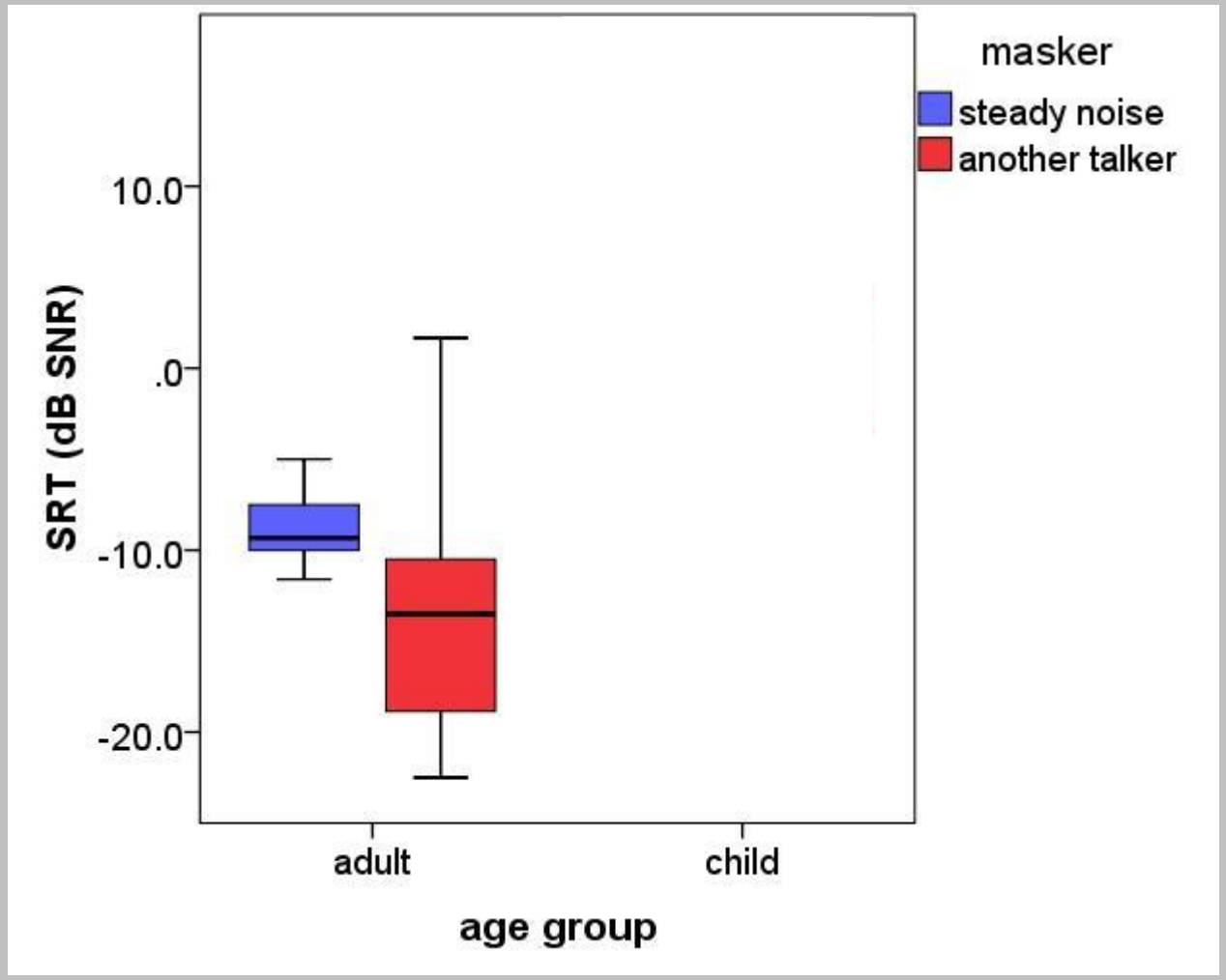
to find the:

***Speech Reception Threshold* (SRT)**

[the SNR which leads to a fixed % correct, typically 50% or 79%]

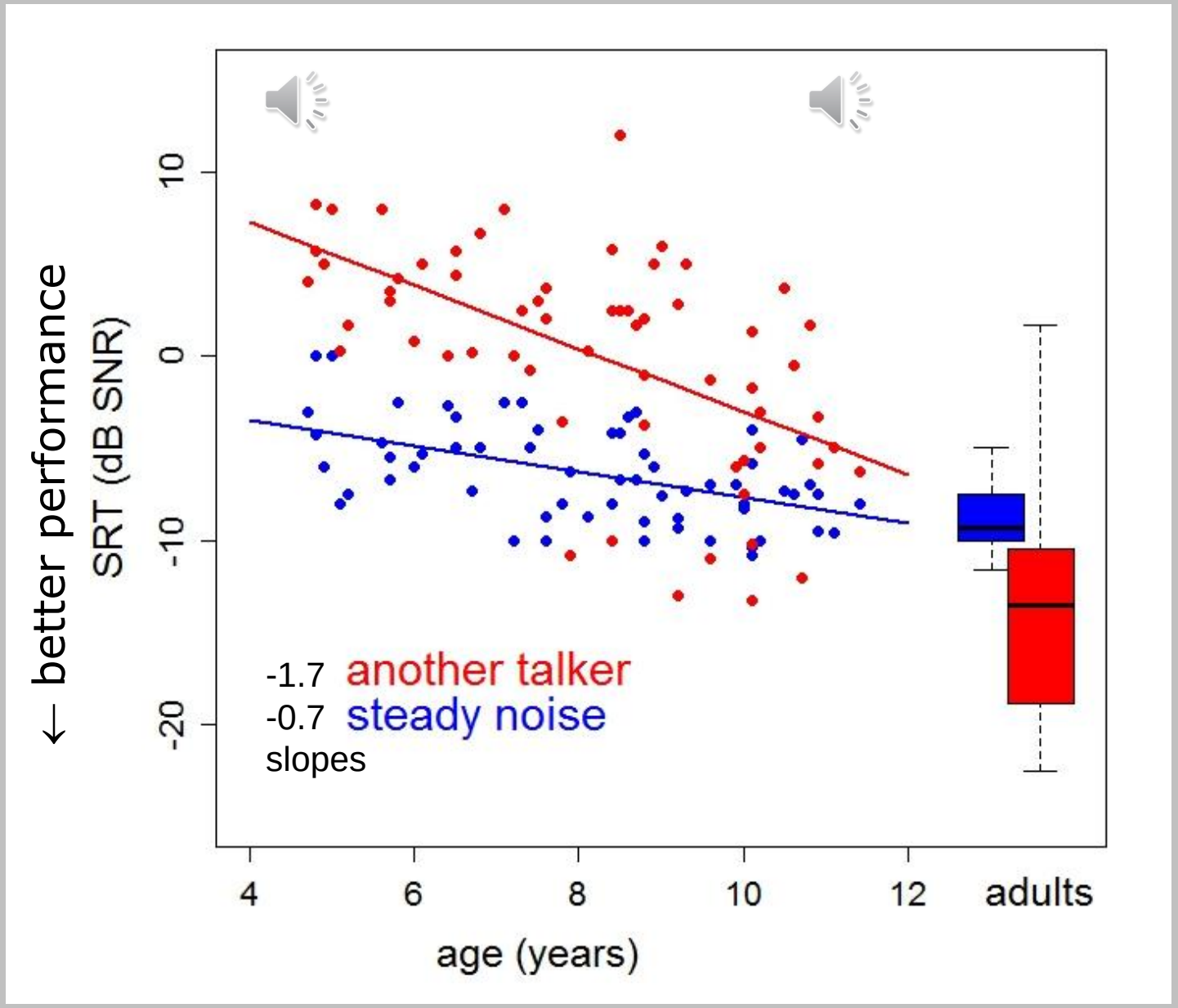
So the lower the better!

# Children find it hard to ignore another talker

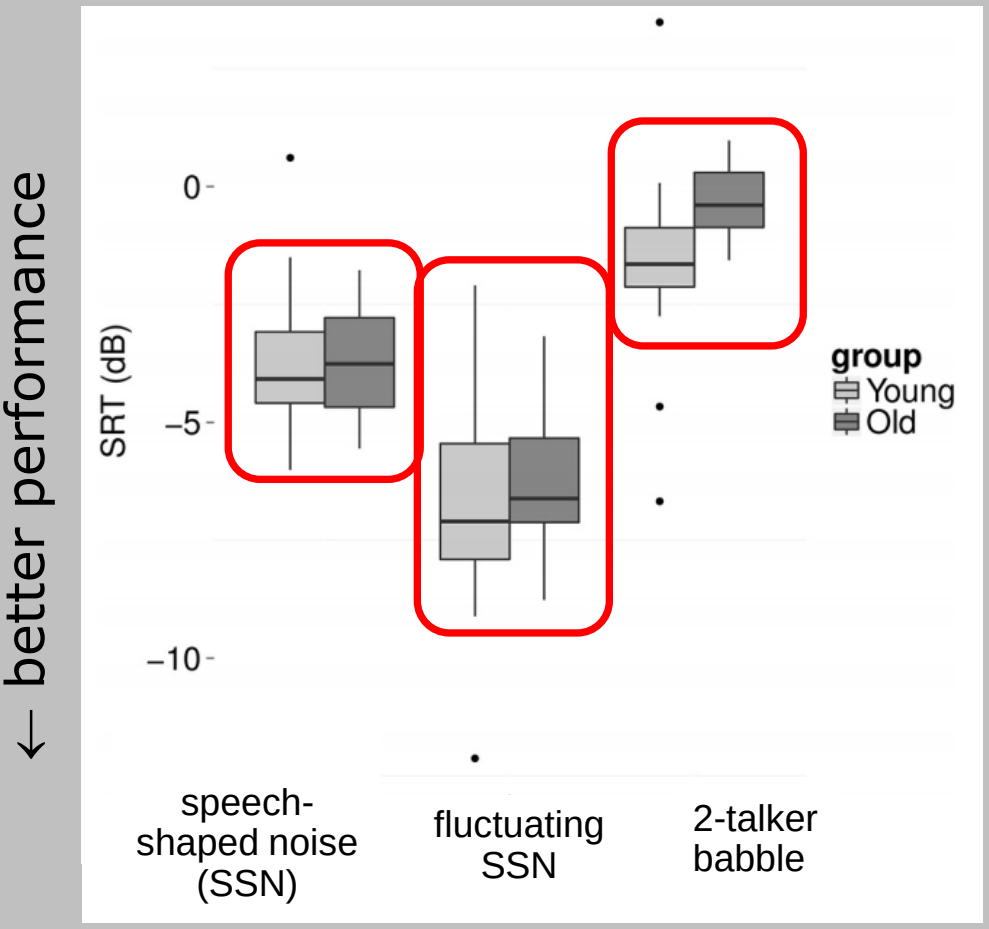



← better performance

# Children can’t ignore another talker



# Increased IM in older listeners



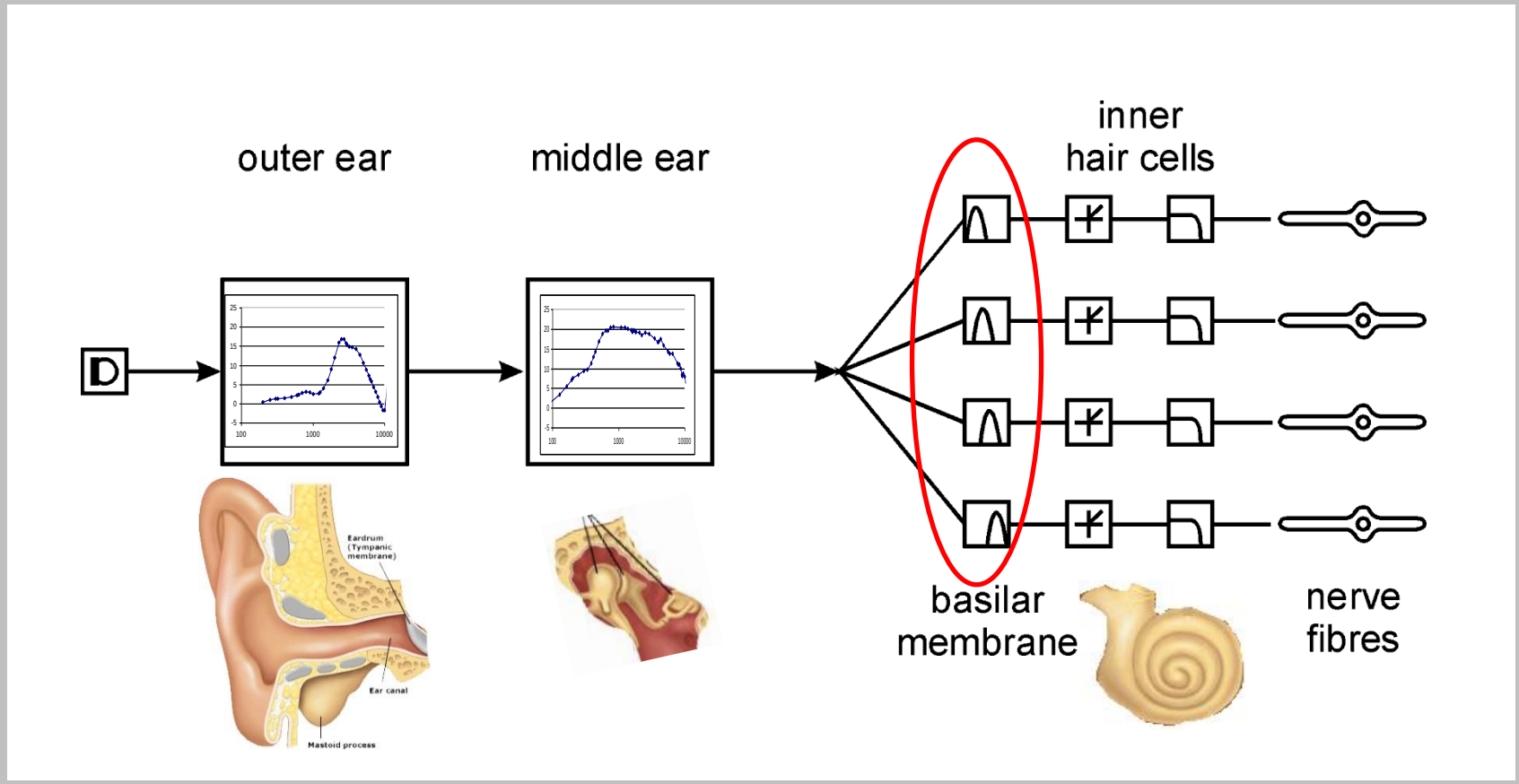

Schoof T & Rosen S (2014) The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners. *Front Aging Neurosci, 6*. doi:10.3389/fnagi.2014.00307

# Interim Summary

- Adult-like performance for Informational Maskers is not reached before age 12.
  - And evidence that development takes even longer when spatial aspects are involved
- Children appear to be much more vulnerable to informational masking than adults …
  - probably because they cannot ignore other distracting speech
- Similar specific deficits have been found in ageing, and a number of cognitive and neurological disorders

# So, what about hearing impairment (HI)?

# Better frequency analysis keeps noise ‘in its place’



# HI affects Energetic & Modulation Masking

- Remember: EM/MM reflects the direct interaction of masker and speech in the cochlea
- Hearing impaired listeners are more affected by noise …
  - because they typically have poorer frequency analysis
  - so noise and speech get more ‘mixed up’

# Final remarks

Different kinds of background sounds affect speech intelligibility in different ways …

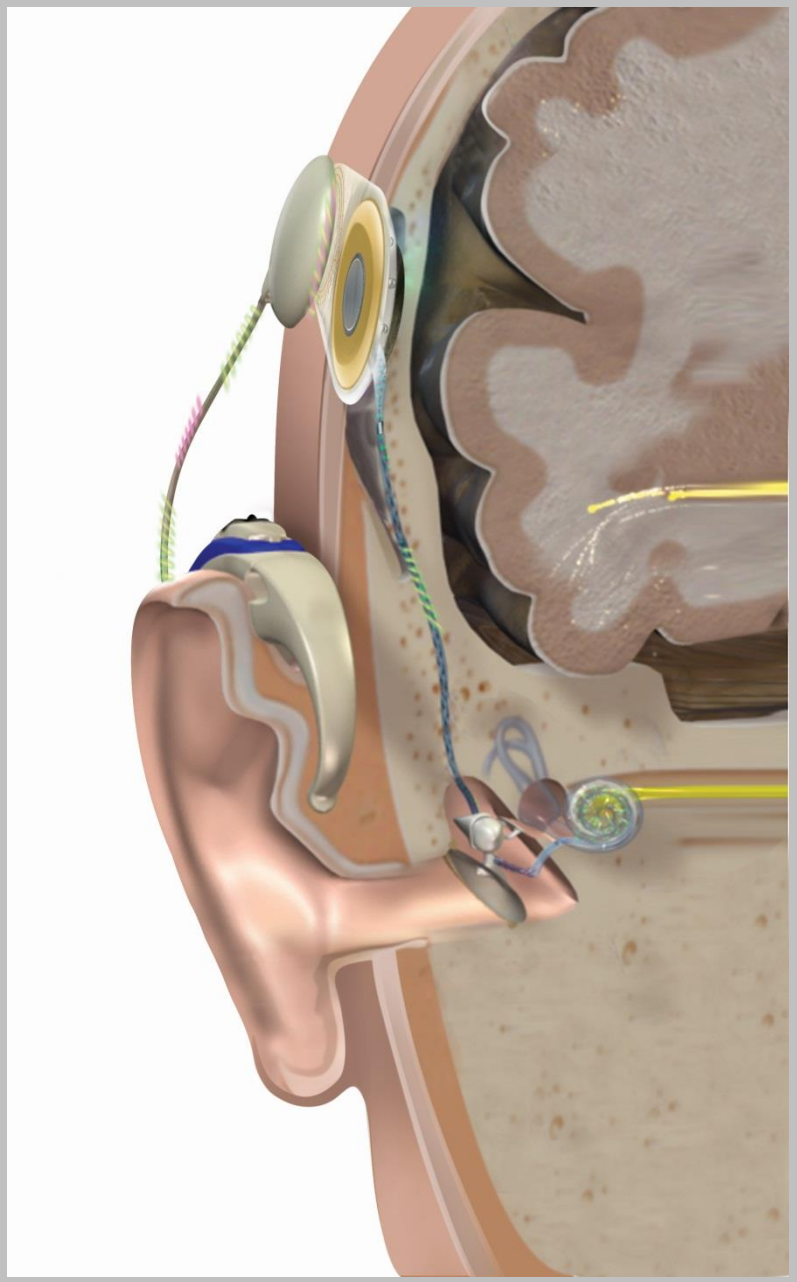
# Even people with normal hearing can have difficulty hearing speech in noisy backgrounds

- young children
- older people with normal hearing
- children with language impairment
- adults with dyslexia
- people with High Functioning Autism

# And the situation is worse with impaired hearing because of …

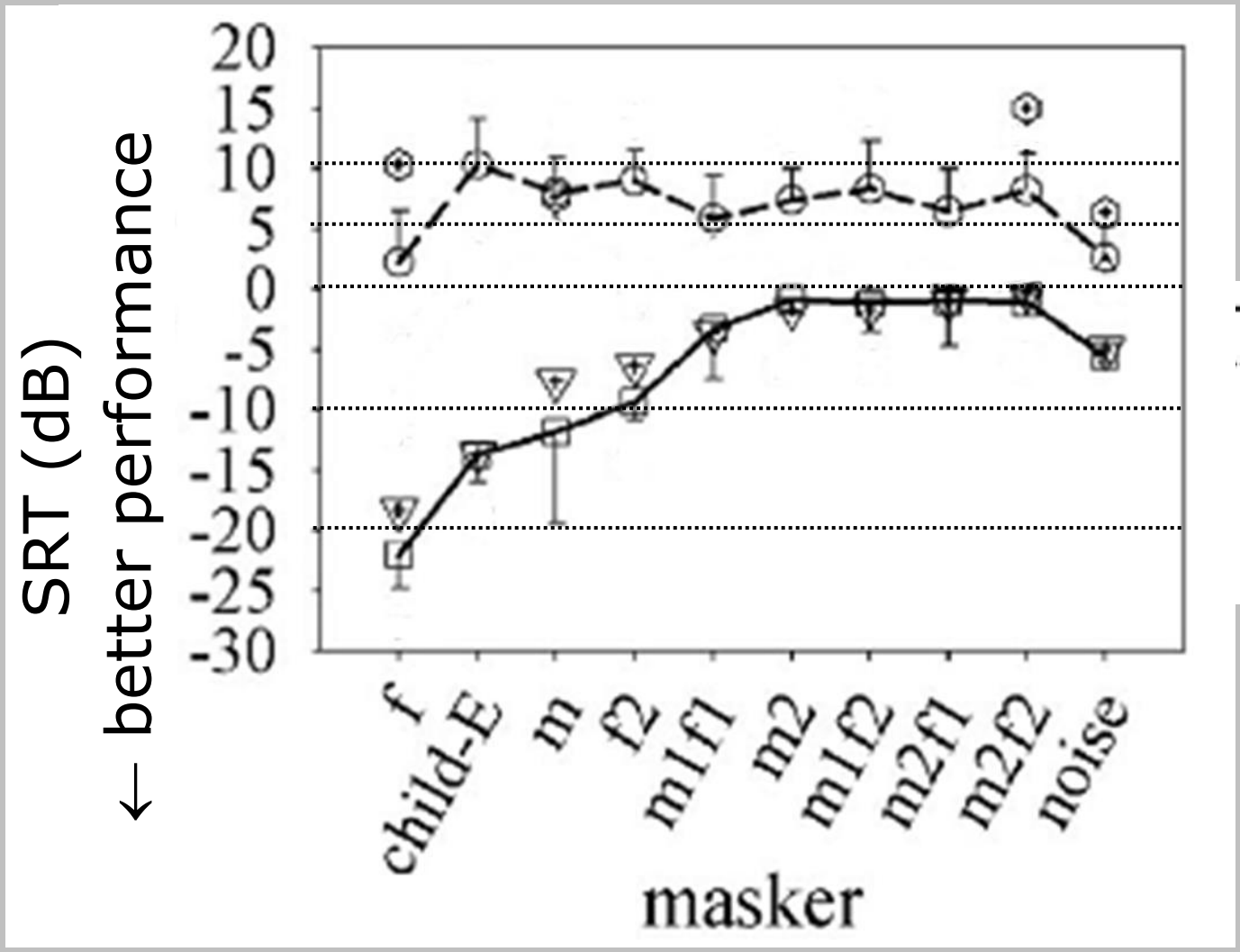
- increased energetic and modulation masking
- which also interferes with auditory scene analysis

# Thank you!



# A special kind of auditory prosthesis: A cochlear implant

# CI users are greatly affected by noise




male target sentences

Cullington & Zeng (2008)