Hot Topics in Speech Communication

Peter Assmann

Human Development, Univ of Texas/Dallas

Philip Loizou

Electrical Engineering, Univ of Texas/Dallas

Robustness of speech communication

- Adaptations of the talker
- Separation of competing sounds
- Perceptual compensation for distortion
- Audio-visual speech synthesis

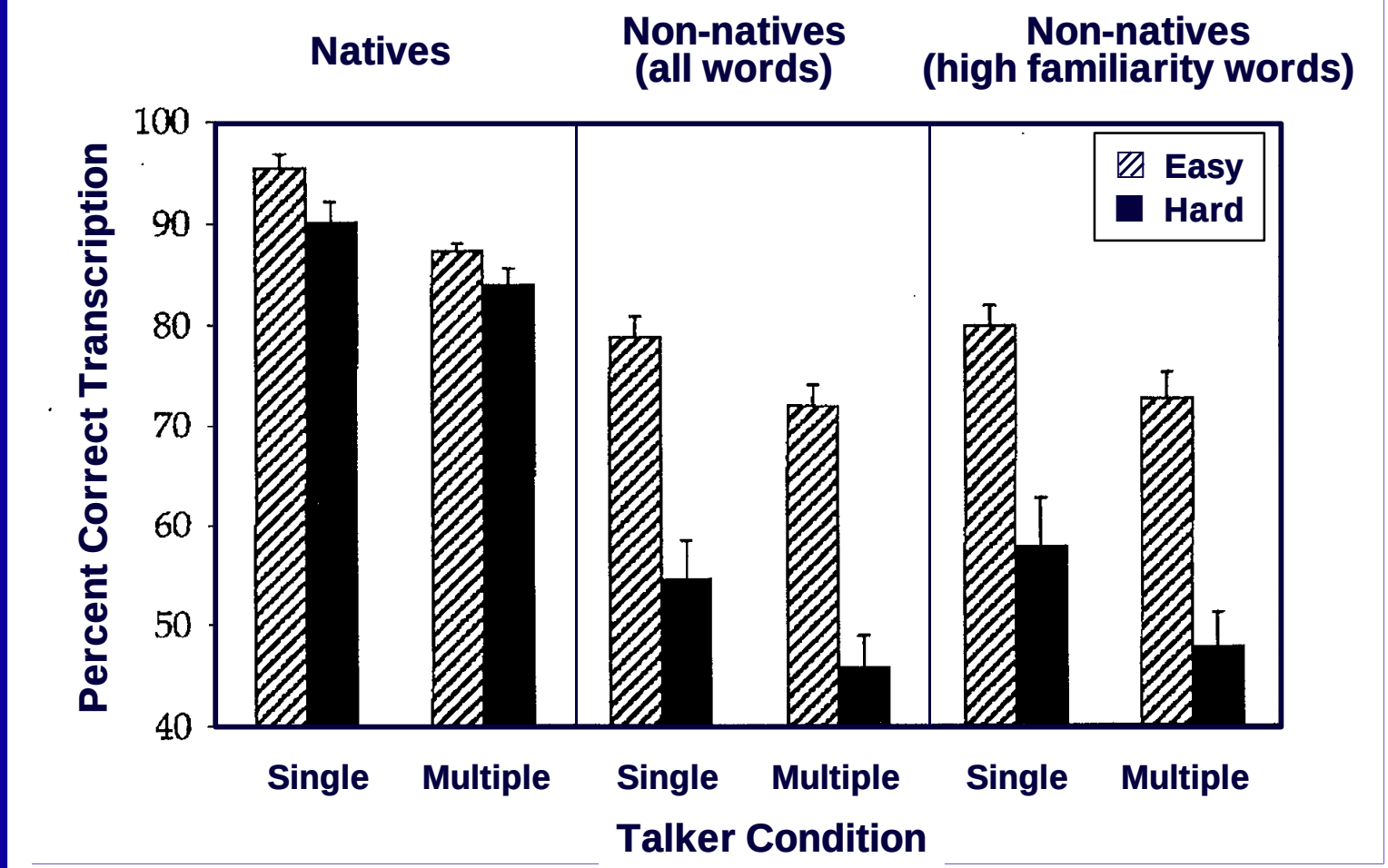
Adaptations of the Talker

Talker variability

- Native and non-native listeners make use of information about the talker in spoken word recognition.

  Bradlow and Pisoni (JASA, Oct 1999)

Bradlow and Pisoni (1999)



Separation of competing sounds

Cocktail party problem

Spatial location	**Level differences**
Fundamental frequency	**Amplitude fluctuations**
Vocal tract size	**Hearing impairment**

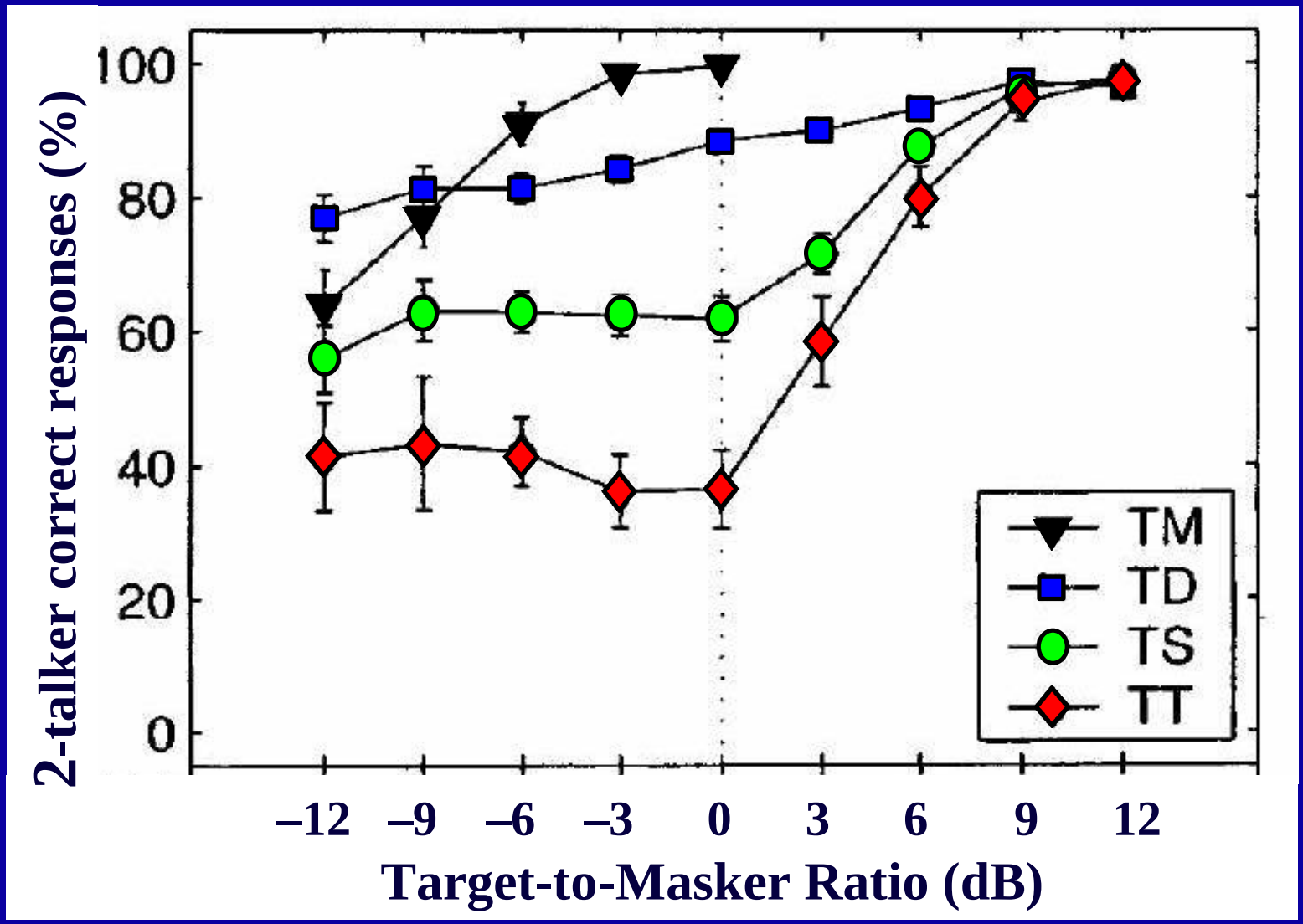
Bronkhorst (*Acta Acustica*, 2000)
Cooke & Ellis (*Speech Communication*, 2001)

Competing talkers

- Performance depends on degree of similarity between target and masker.
- Listeners can attend to the quieter of two competing voices.
- Interference is caused by informational masking, not just energetic masking.

Brungart et al. *(JASA, Nov 2001)*

Brungart et al. (2001)



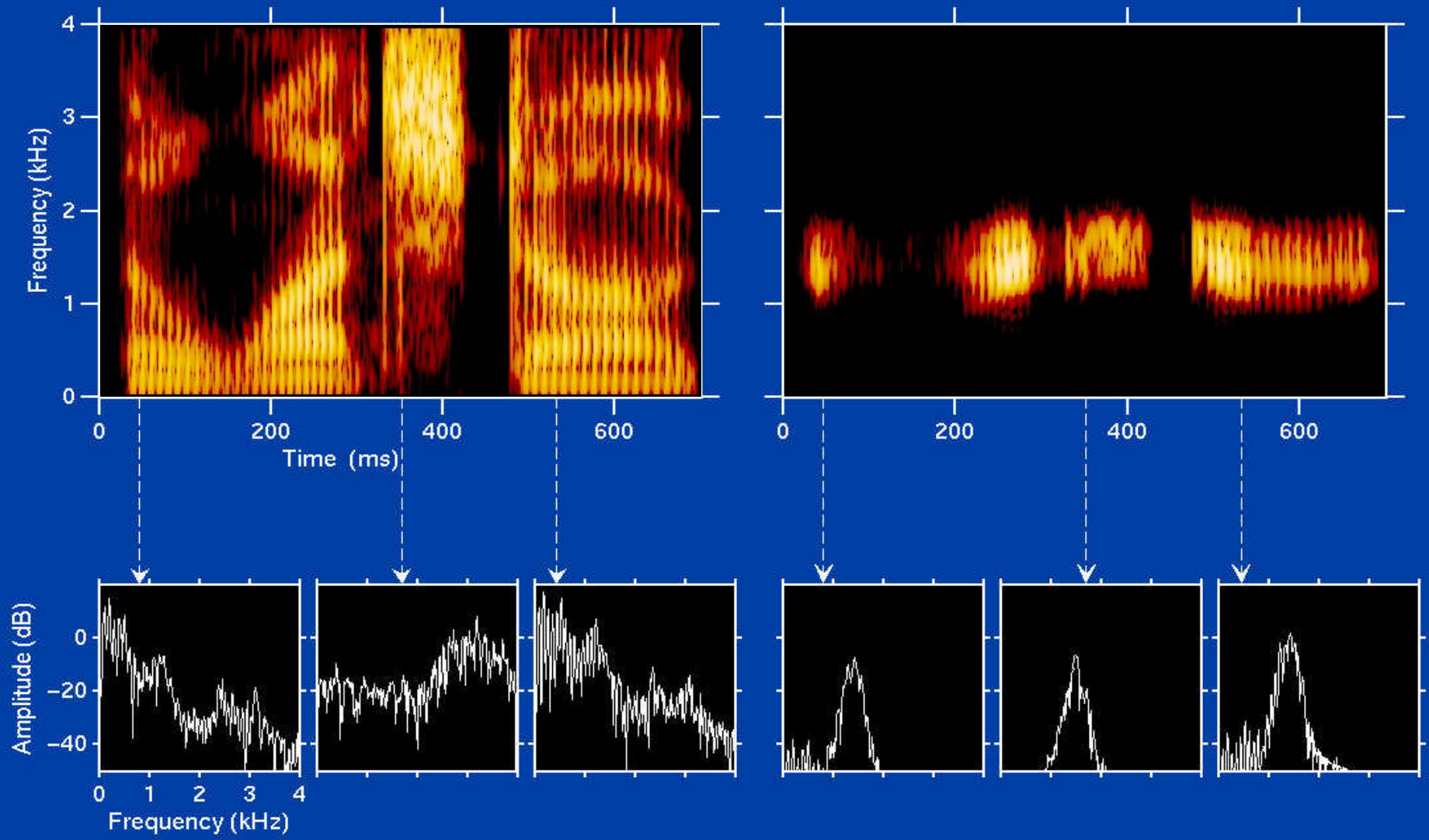
Perceptual compensation for distortion

Perception of filtered speech

- Everyday English sentences filtered using narrow bandpass filters remain highly intelligible (>90% words correct)
  - *one-third octave bandwidth, 1500 Hz center frequency, 100 dB/octave slopes*

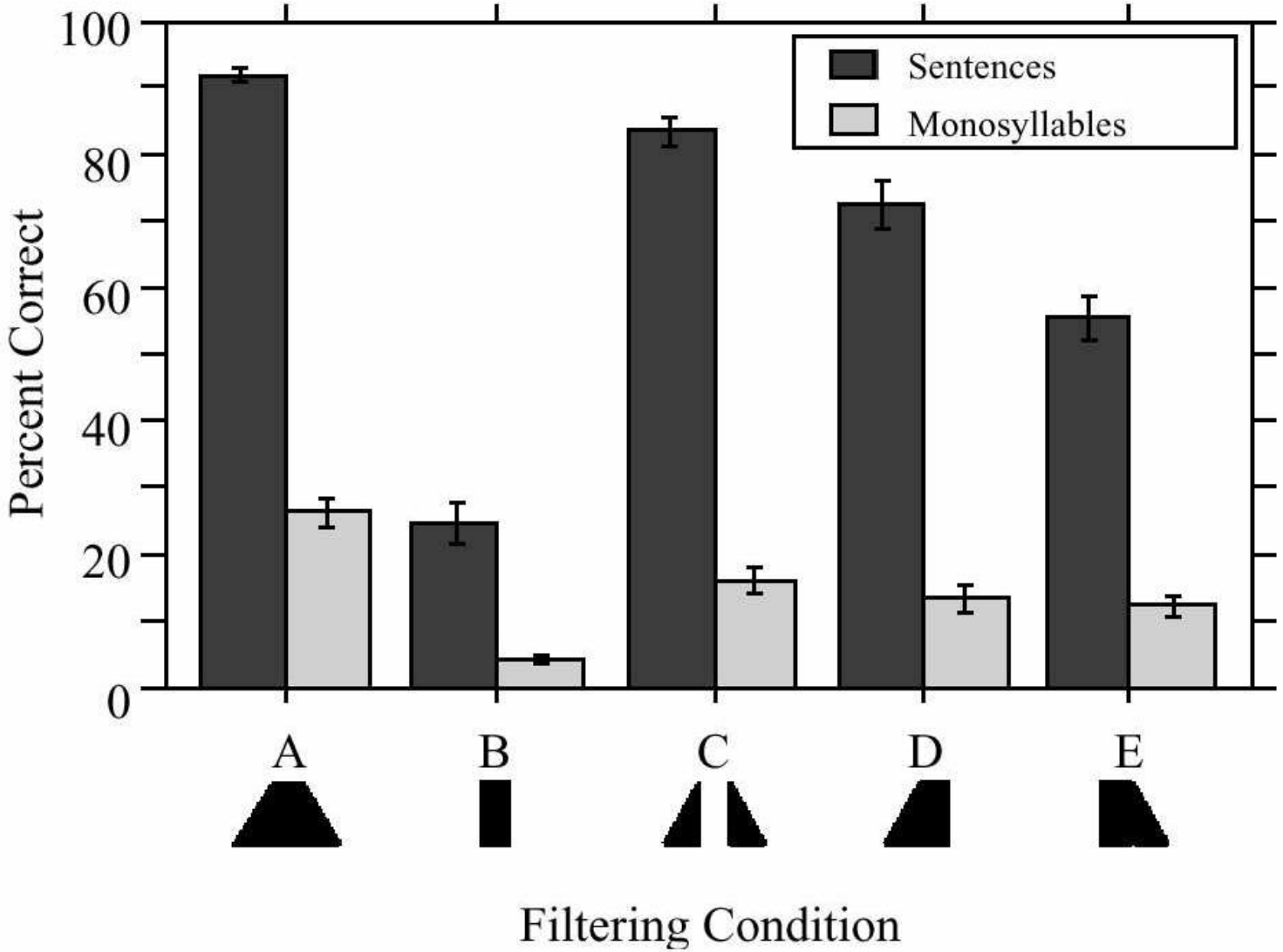
Warren et al. *(Percept Psychophys 1995; JASA 2000)*

Perception of filtered speech



Stickney and Assmann (JASA 2001)

Bashford et al. (2000)



Audio-visual speech synthesis

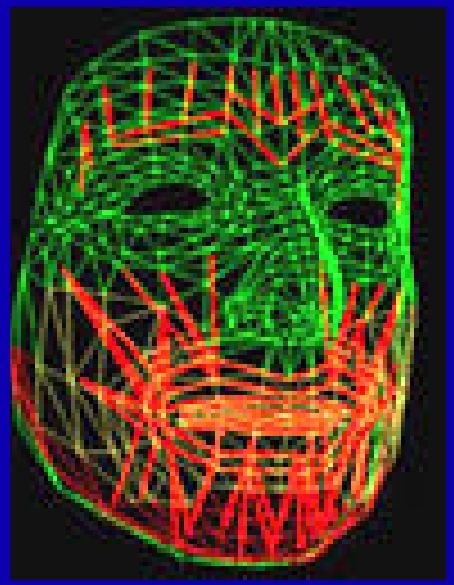
Facial animation

- Dynamic inversion of muscle model
- Kinematic recordings of facial movements

Pitermann & Munhall (*JASA*, Sept 2001)

Pitermann & Munhall (2001)

- Inverse dynamics approach to face animation



Epidermal mesh

Texture map

References

- Bashford, JA, Warren, RM, Lenz, PW (2000). Relative contributions of passband and filter skirts to the intelligibility of bandpass speech: Some effects of context and amplitude. Acoustics Research Letters Online 1(2): 31-36.
- Bradlow AR, Pisoni DB. (1999). Recognition of spoken words by native and non-native listeners: talker-, listener-, and item-related factors. J Acoust Soc Am. 1999 Oct;106(4 Pt 1):2074-85.
- Bronkhorst AW (2000). The cocktail party phenomena: A review of research on speech intelligibility in multiple-talker conditions. Acustica-acta acustica 86(2000): 117-128.
- Brungart DS. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. J Acoust Soc Am. 2001 Mar;109(3):1101-9.
- Brungart DS. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. J Acoust Soc Am. 2001 Mar;109(3):1101-9.
- Cooke M, Ellis DPW (2001). The auditory organization of speech and other sources in listeners and computational models. Speech Commun 35(2001): 141-177.
- Pitermann M, Munhall KG. (2001). An inverse dynamics approach to face animation. J Acoust Soc Am. 2001 Sep;110(3 Pt 1):1570-80.
- Stickney GS, Assmann PF. (2001). Acoustic and linguistic factors in the perception of bandpass-filtered speech. J Acoust Soc Am. 2001 Mar;109(3):1157-65.
- Warren RM, Bashford JA Jr, Lenz PW. (2000). Intelligibility of bandpass speech: effects of truncation or removal of transition bands. J Acoust Soc Am. 2000 Sep;108(3 Pt 1):1264-8.
- Warren RM, Bashford JA Jr. (1999). Intelligibility of 1/3-octave speech: greater contribution of frequencies outside than inside the nominal passband. J Acoust Soc Am. 1999 Nov;106(5):L47-52.
- Warren RM, Riener KR, Bashford JA Jr, Brubaker BS. (1995). Spectral redundancy: intelligibility of sentences heard through narrow spectral slits. Percept Psychophys. 1995 Feb;57(2):175-82.

Hot topics in cochlear implants

Combining acoustic and electric hearing

Bilateral cochlear implants

Combating channel interaction: Electrode design

Noise reduction strategies

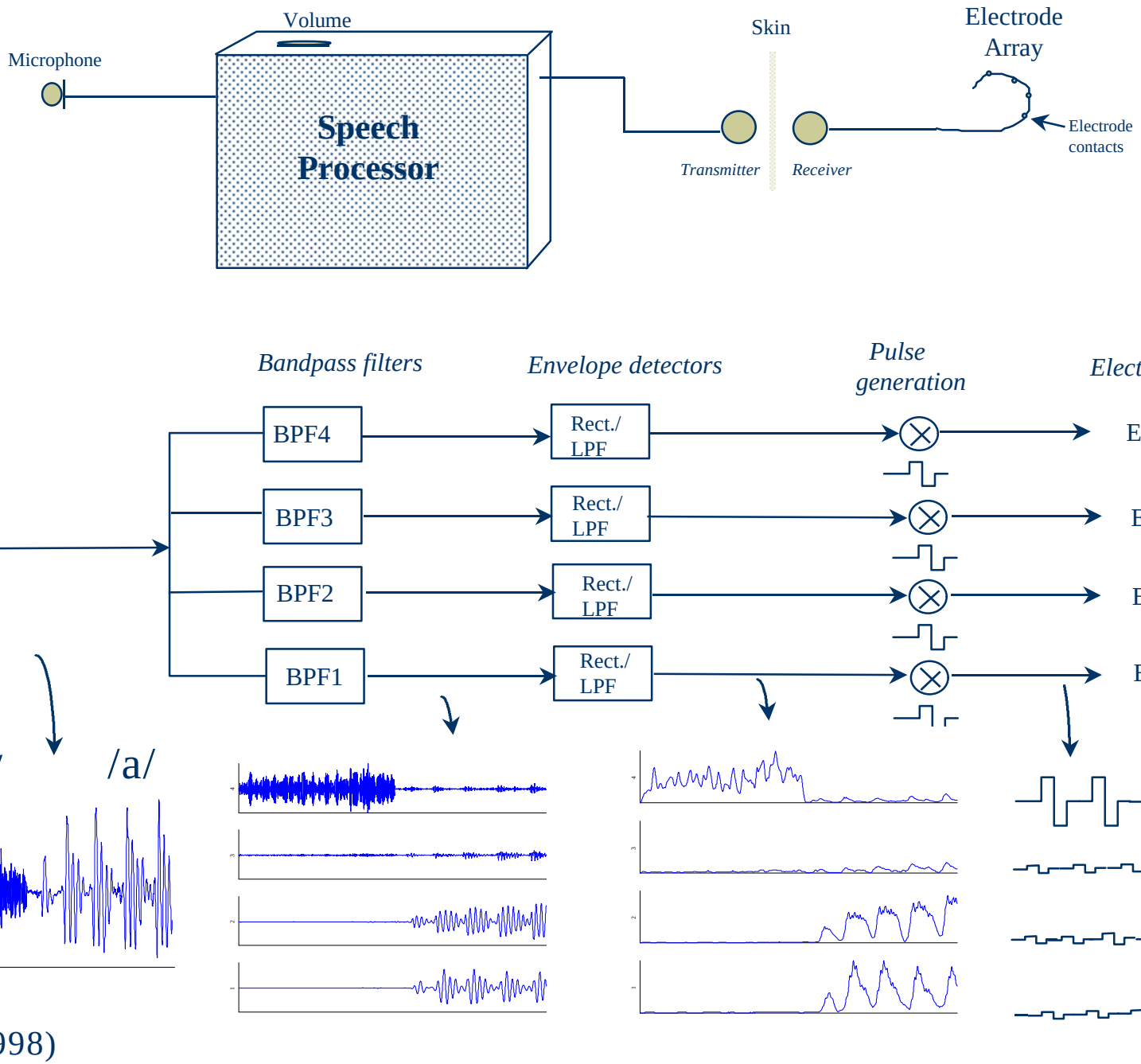
Cochlear Implants

For patients with profound hearing loss.

Bypasses the normal hearing mechanism and electrically stimulates the auditory nerve.

Components

- External (speech processor, mic)
- Internal (RF receiver)

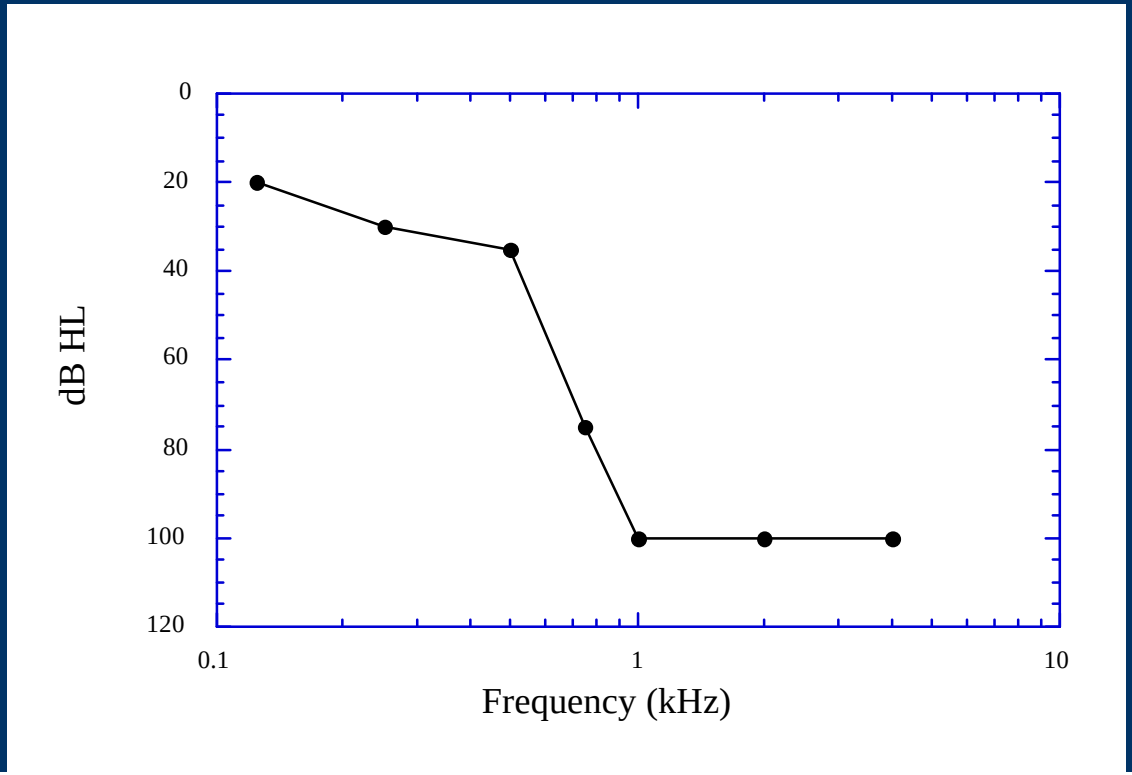

Volume
Microphone
Speech Processor
Skin
Electrode Array
Transmitter
Receiver
Electrode contacts
Bandpass filters
Envelope detectors
Pulse generation
Electrodes
Microphone
BPF4
BPF3
BPF2
BPF1
Rect./ LPF
Rect./ LPF
Rect./ LPF
Rect./ LPF
El-4
El-3
El-2
El-1
/s/
/a/

(Loizou, 1998)

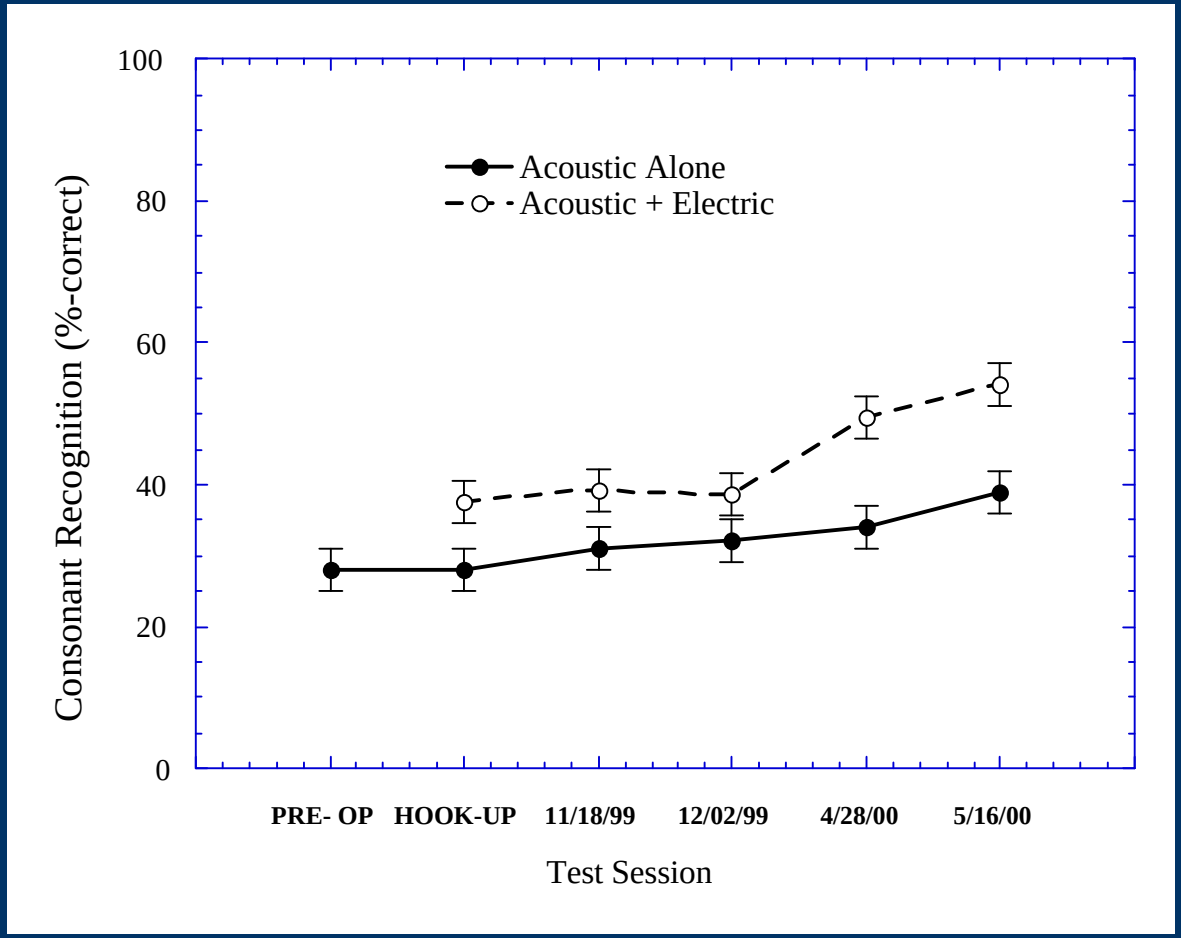
Combining acoustic and electric hearing

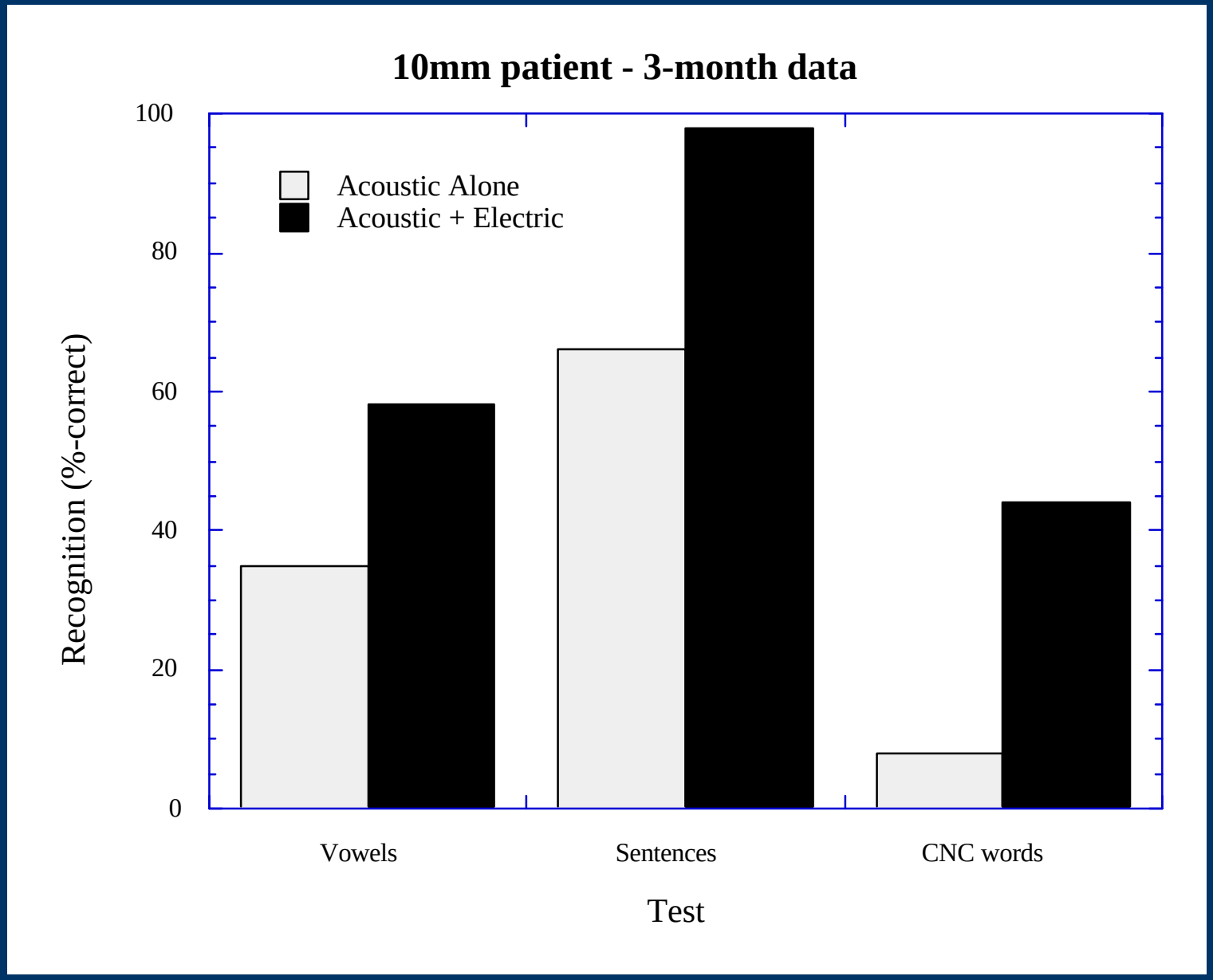
What type of patient?

Example: 50% discrimination bilaterally, not happy with amplification



Speech recognition for 6mm electrode patient (Turner et al., 2001)





(Turner et al., 2001)

Bilateral cochlear implants

Are two implants better than one?

Potential benefits of bilateral implants:

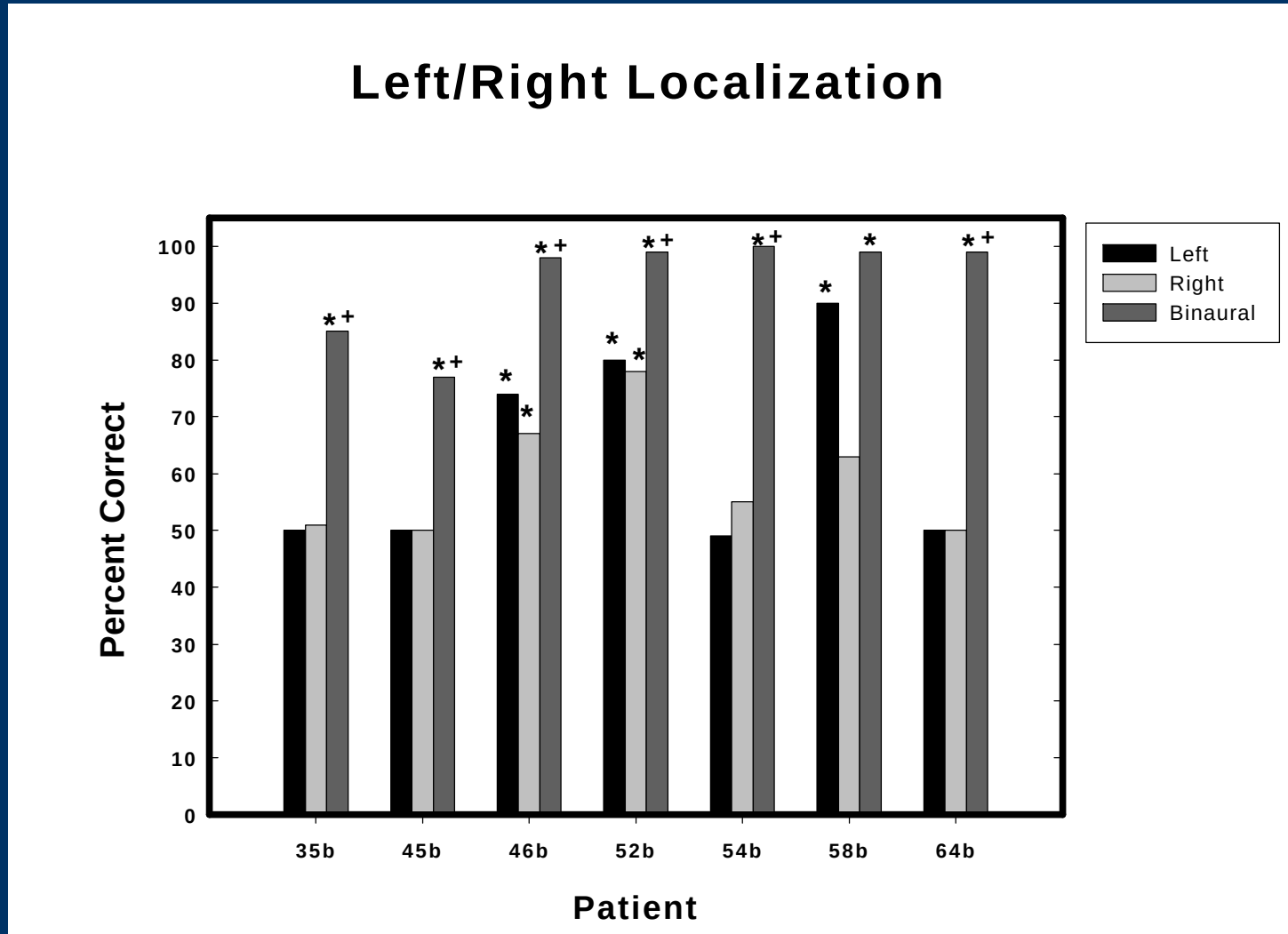
- Better localization
- Better speech recognition in noise

There are two possible advantages of adding a second ear when listening to speech in noise:

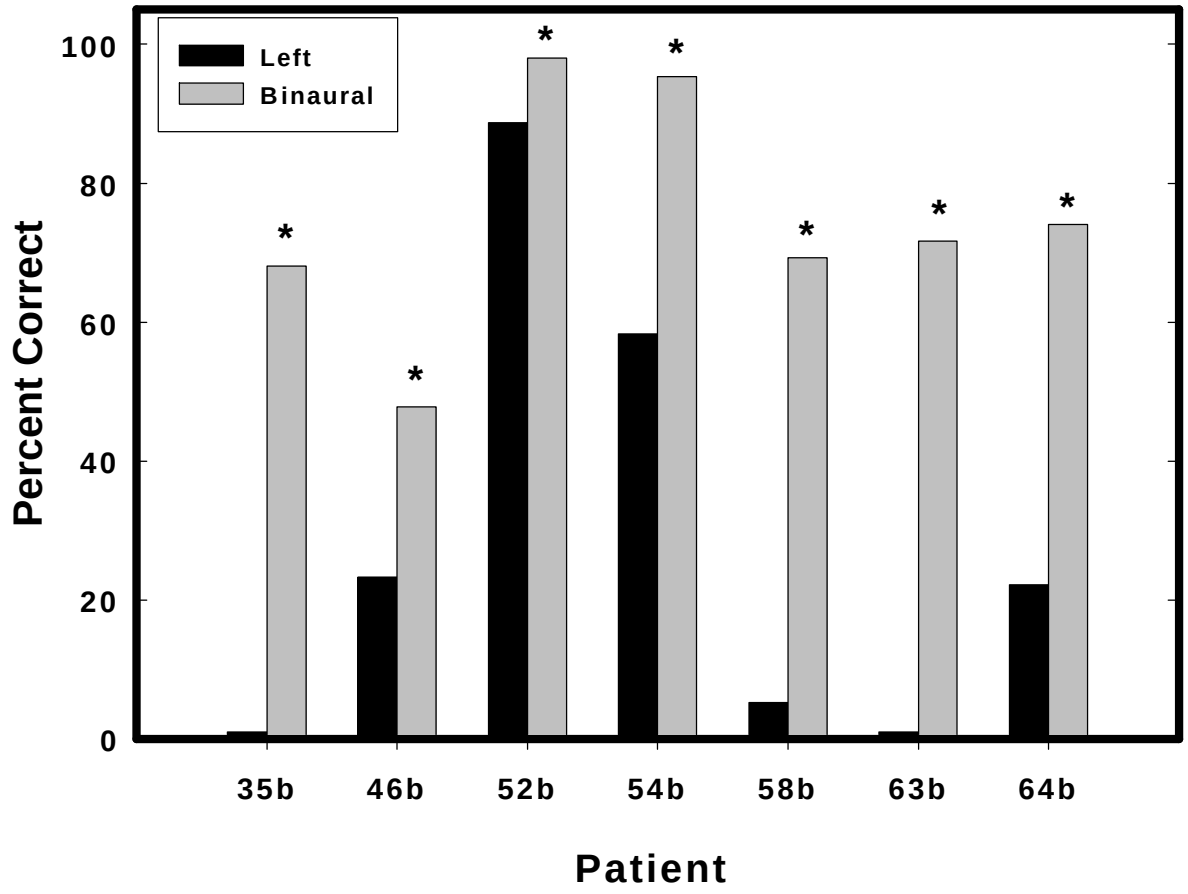
First, the head creates a physical barrier between noise on the contralateral side of one ear. This '*head shadow*' attenuates the noise reaching the ear on the other side of the head, resulting in having an ear with better speech-to-noise ratio.

Second, when the speech and noise are from different locations, the brain may be able to separate them using their distinct localization cues, and thus improve intelligibility. This second potential advantage is referred to as *binaural squelch*.

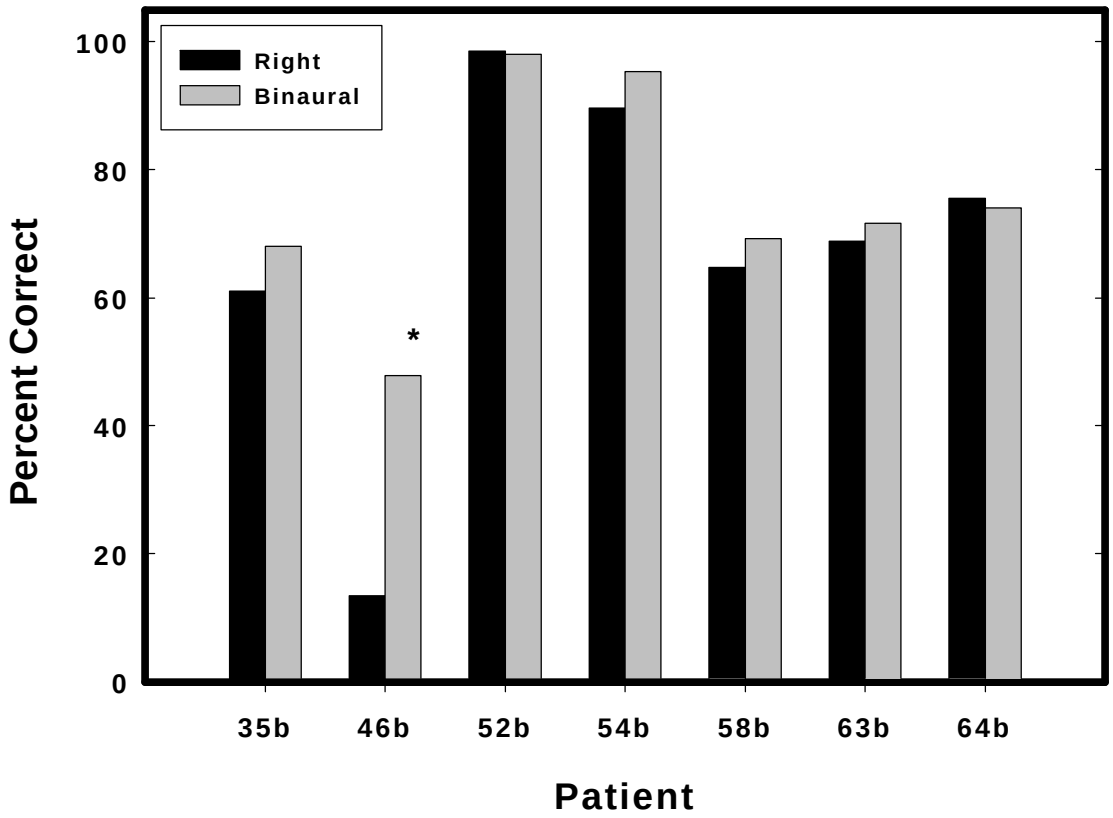
Sound Localization (Tyler et al., 2002)



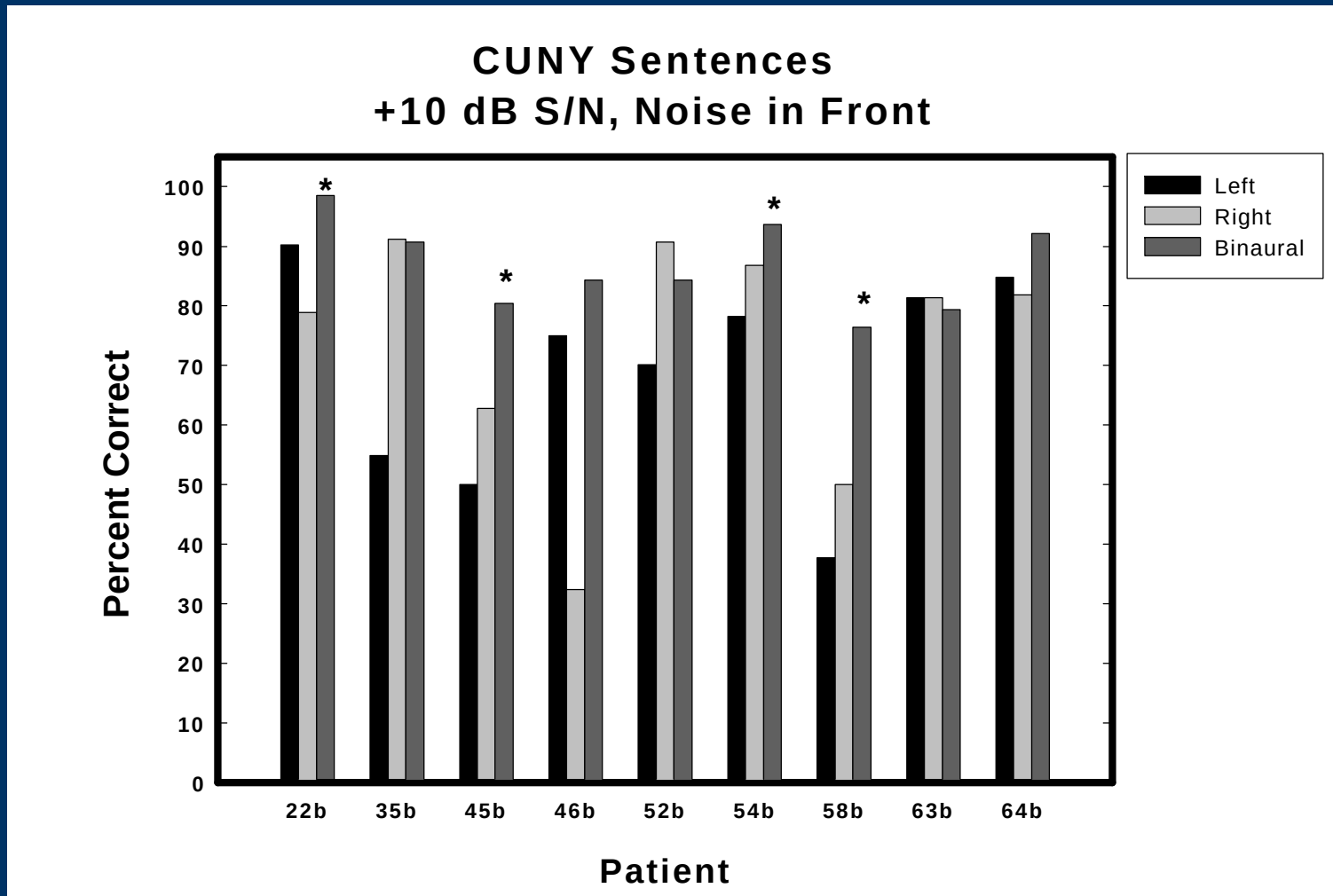
Noise Ipsilateral to Monaural Implant
Noise on Left
CUNY Sentences



(Tyler et al., 2002)



(Tyler et al., 2002)



(Tyler et al., 2002)

Electrode Design

Channel Interaction

Cochlea is filled with conductive salt solutions, which allows current to spread.

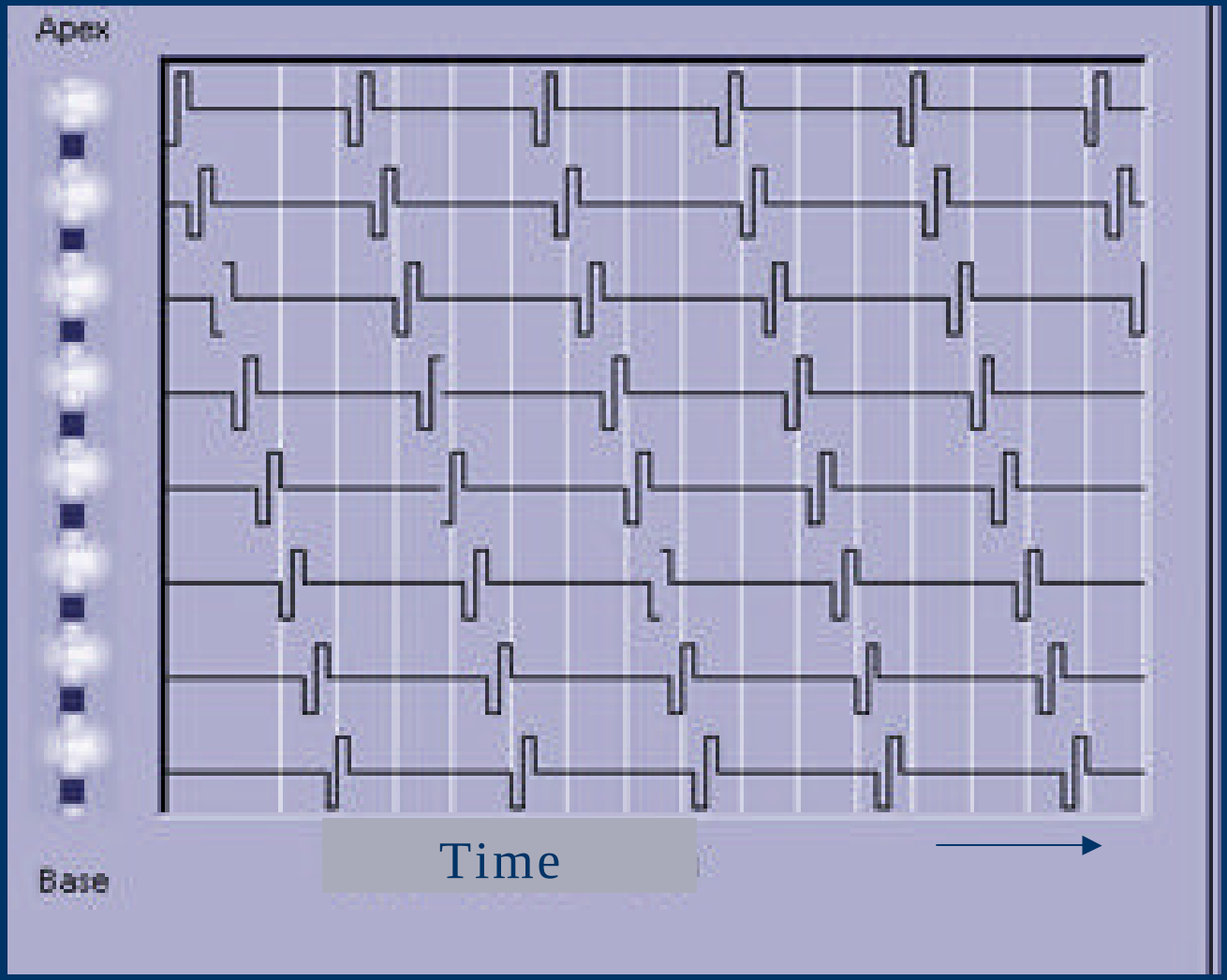
Current spread is detrimental to speech recognition when current fields from different electrodes interact.

Solutions:

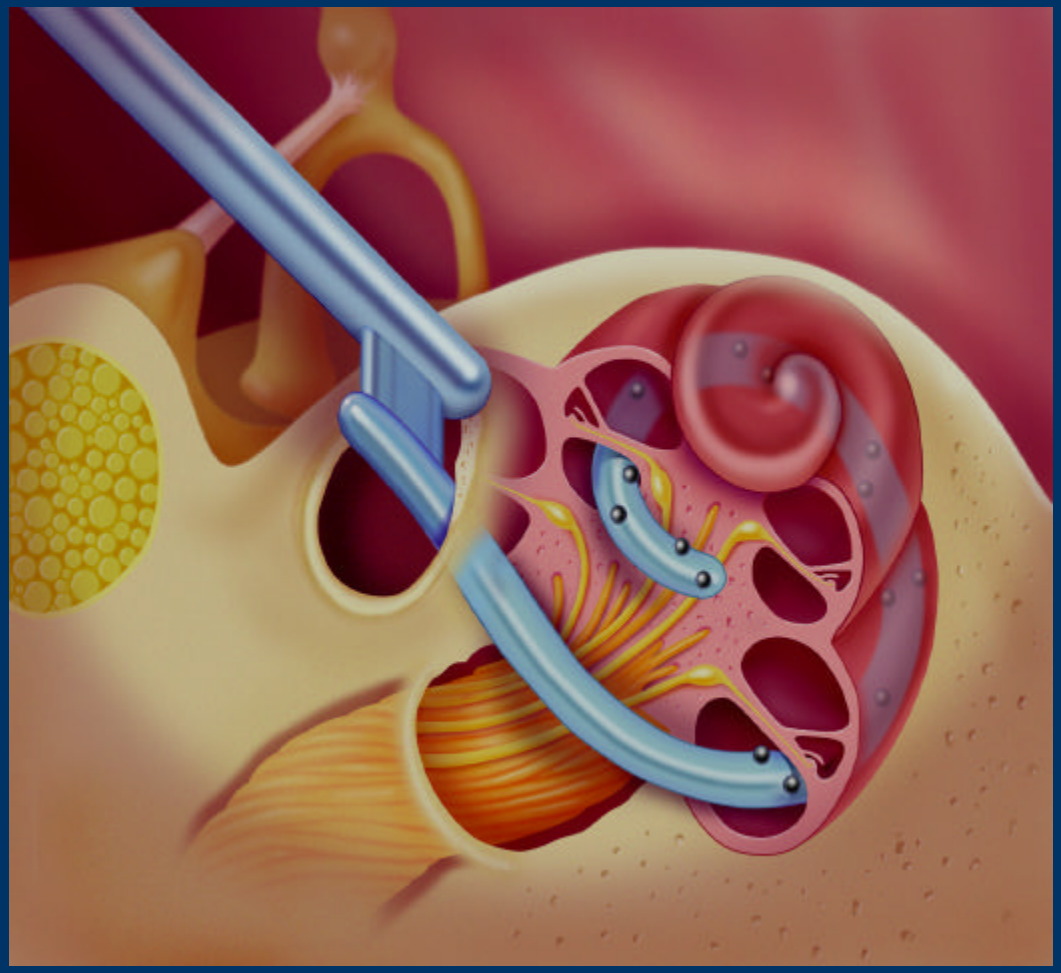
- Speech processing strategies (e.g., CIS) which stimulate the electrodes sequentially rather than simultaneously.
- Better electrode design:
  - Better positioning – closer to the target cells
  - Inserting buffers between electrodes for better localization

Non-simultaneous stimulation

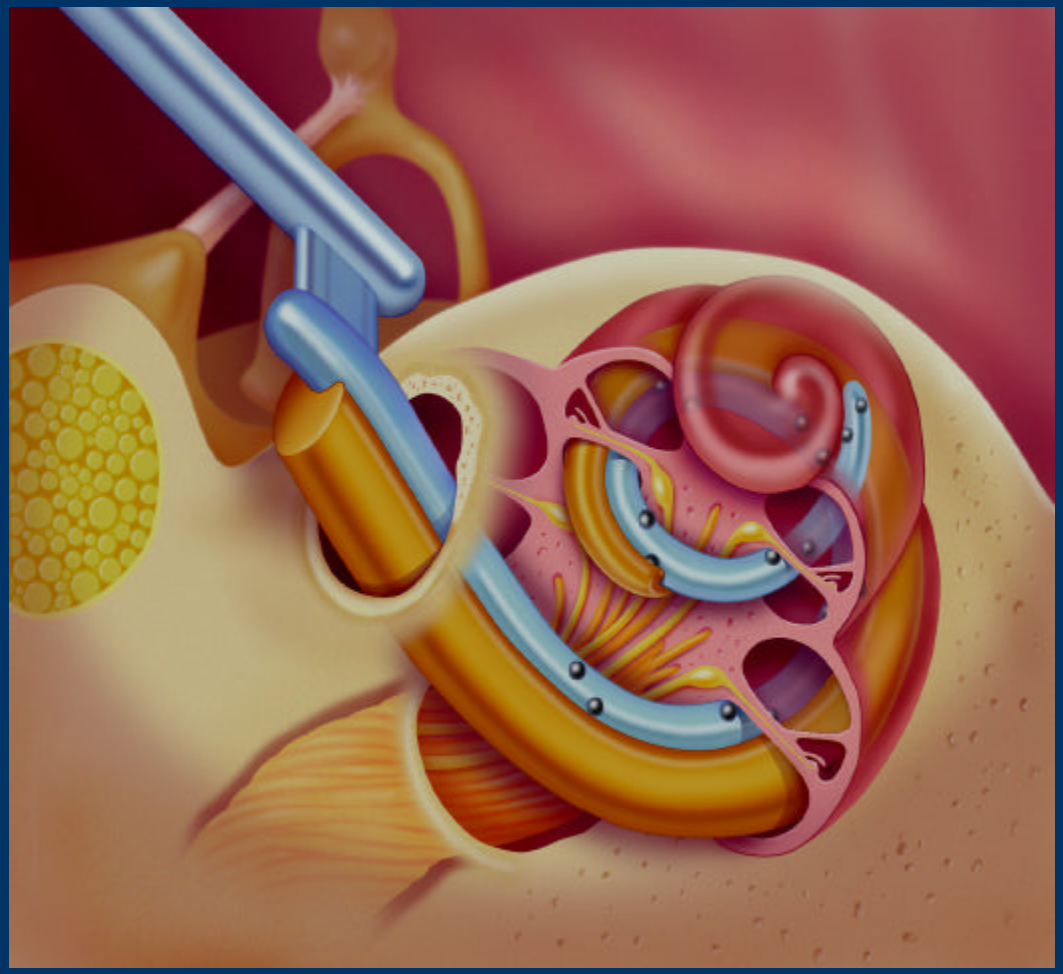
CIS speech processing strategy



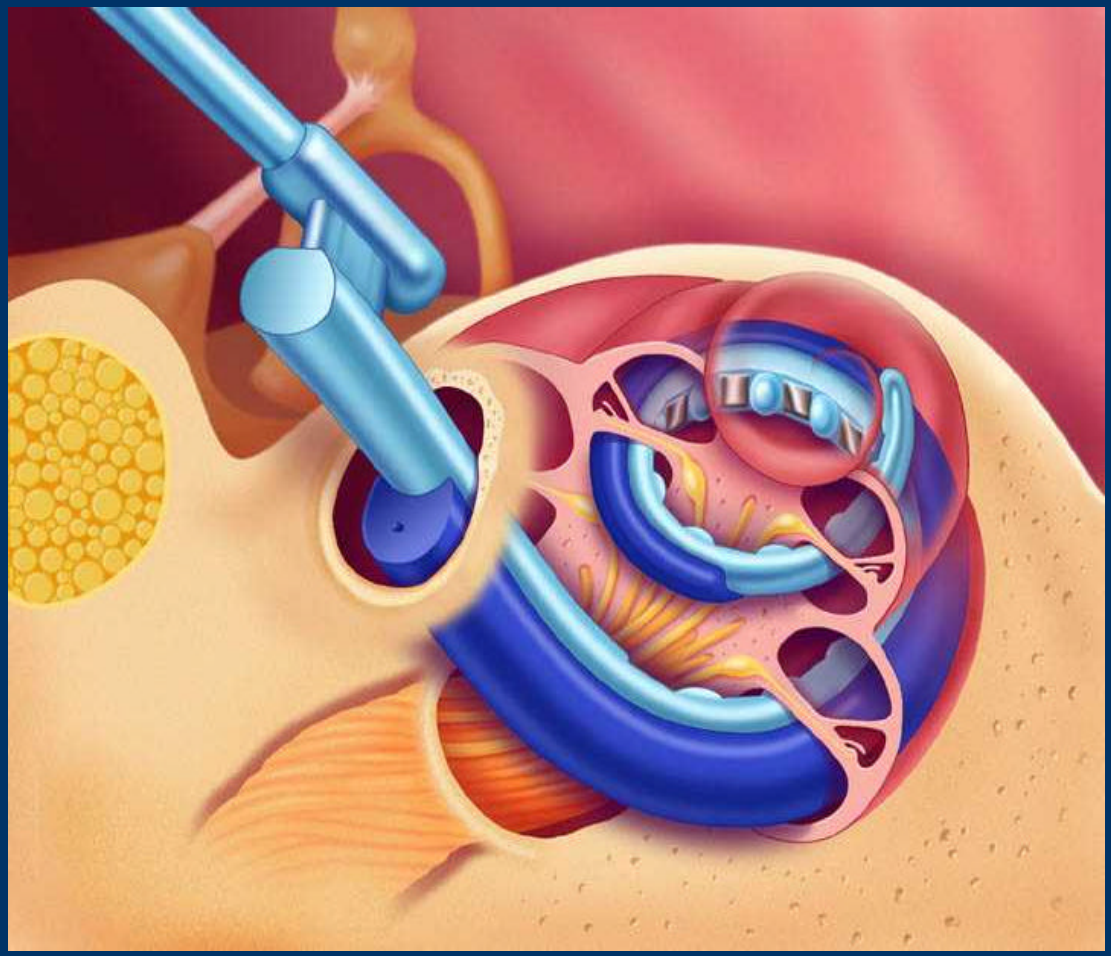
Standard Electrode



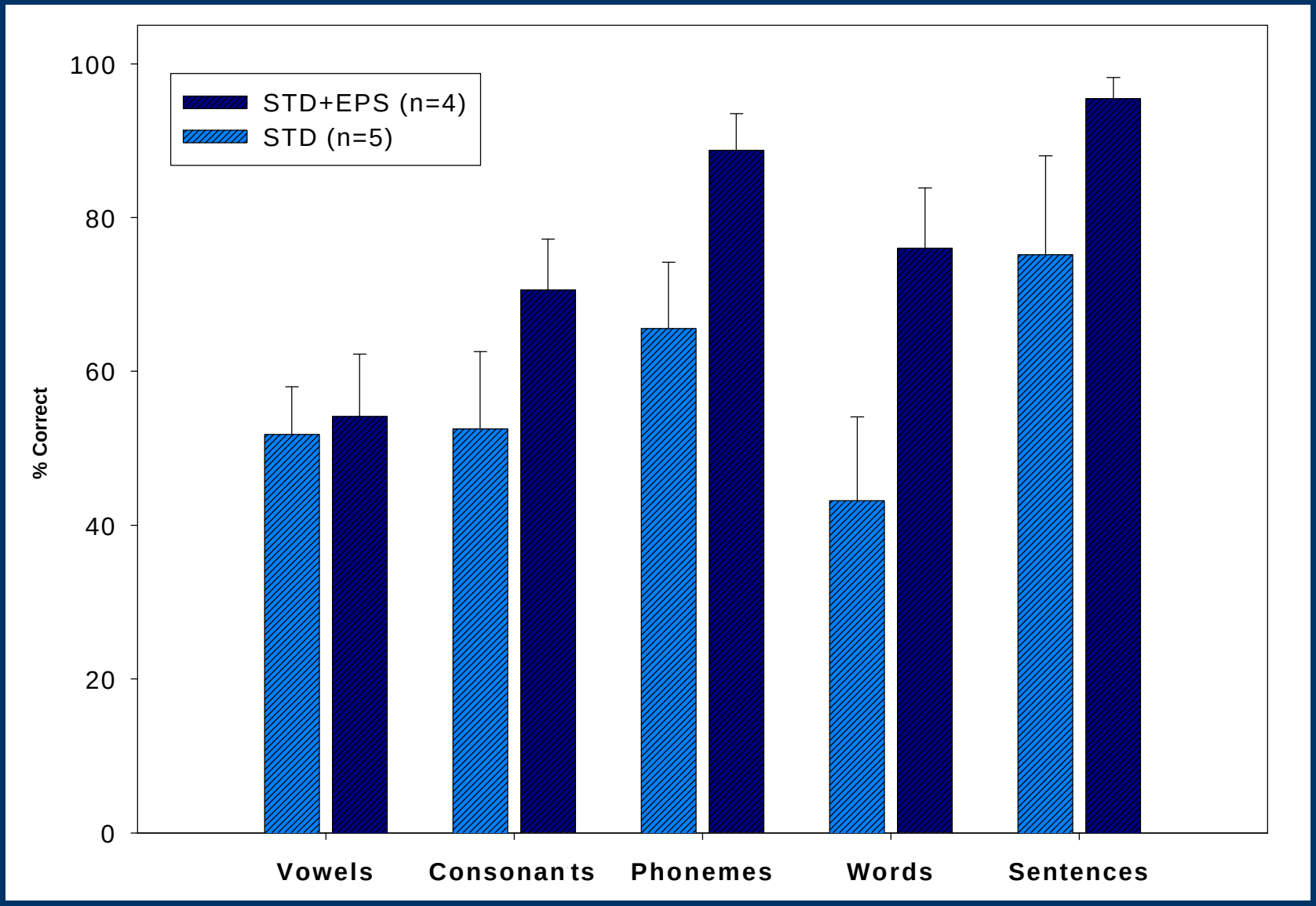
Standard Electrode w/ EPS



Hi-Focus Electrode w/ EPS



Speech Recognition Performance (Stickney et al., 2000)



More hot topics

- Stochastic resonance
  - Increasing the stochastic properties of electrical stimulation for reduced fiber synchrony.
  - High-rate pulse trains (> 5000 pulses/sec) may create random spike patterns in auditory nerve fibers (Rubinstein *et al.*, 1999)
- Noise reduction strategies
  - Pre-processing the noisy signal with a standard noise-reduction technique (e.g., spectral subtraction) may benefit some subjects in noise (Loizou *et al.*, submitted)

References

Loizou, P. (1998). "Mimicking the human ear: An overview of signal processing strategies for cochlear implants," IEEE Signal Processing Magazine,15(5), 101-130.

Loizou, P., Hu, Y. and Mukhul, B. (submitted). “Noise reduction strategies for cochlear impla Journal of Acoustical Society of America.

Rubinstein, J., Wilson, B., Finley, C. and Abbas, P. (1999). “Pseudospontaneous activity: Stochastic independence of auditory nerve fibers with electrical stimulation,” Hear. Res., 127, 108-118.

Stickney, G., Loizou, P., Mishra, M., Assmann, P., R. Shannon, and J. Opie. (2000). "Channel interaction and speech processing strategies for cochlear implants," Journal of Acoustical Society of America , 108(5), Pt. 2, 2601.

Turner, C., Gantz, B. et al. (2001). “Combining acoustic and electric hearing: Th 2001 Conference on Implantable Auditory Prostheses, Asilomar, Monterey, CA.

Tyler, R., Gantz, B. *et al*. (2002). “Three-month results with bilateral cochlear implants,” Ear Hear., in press.