Limitations of Speech Reception Threshold (SRT) as an outcome measure in hearing-aid research

Graham Naylor

Oticon A/S

Why the fuss?

- Harvey Dillon, IHCON 2006 summing-up
  - *”Look forward to a time when performance differences between HA systems are consistently characterised in dB. That will give us a common basis for comparisons”*
- HA systems becoming increasingly non-linear
  - increasingly important to do the right measurement
- Spread of convenient & standardised SRT procedures
  - increasingly easy to do the (wrong) measurement

Plan

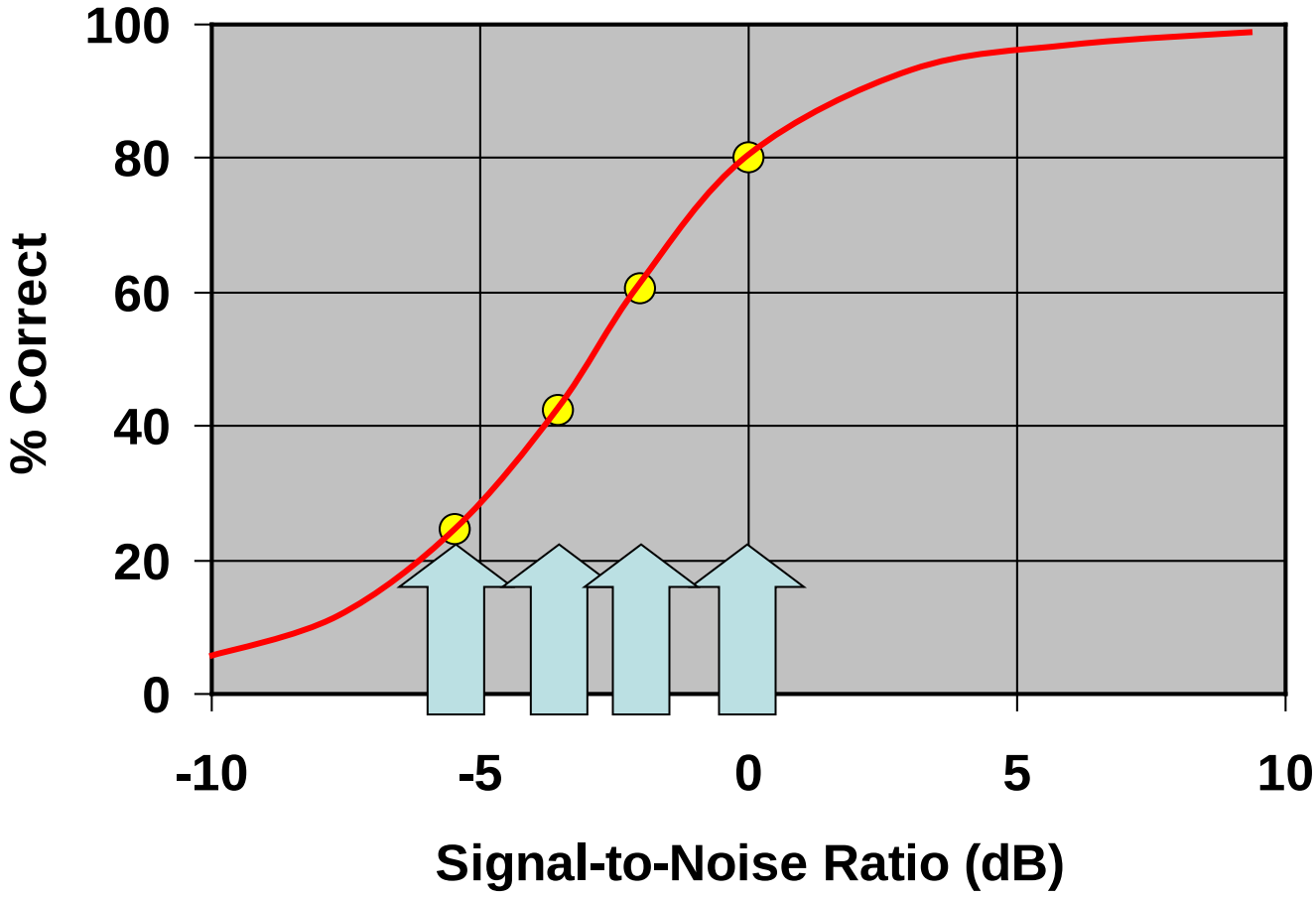
- What ”SRT” am I talking about?
- Factors affecting SRT (unaided)
- Potential consequences for SRT (outcome)
- Non-linear hearing aid (HA) systems
- Why is SRT nevertheless so attractive?
- What to do?

SRT as a diagnostic measure

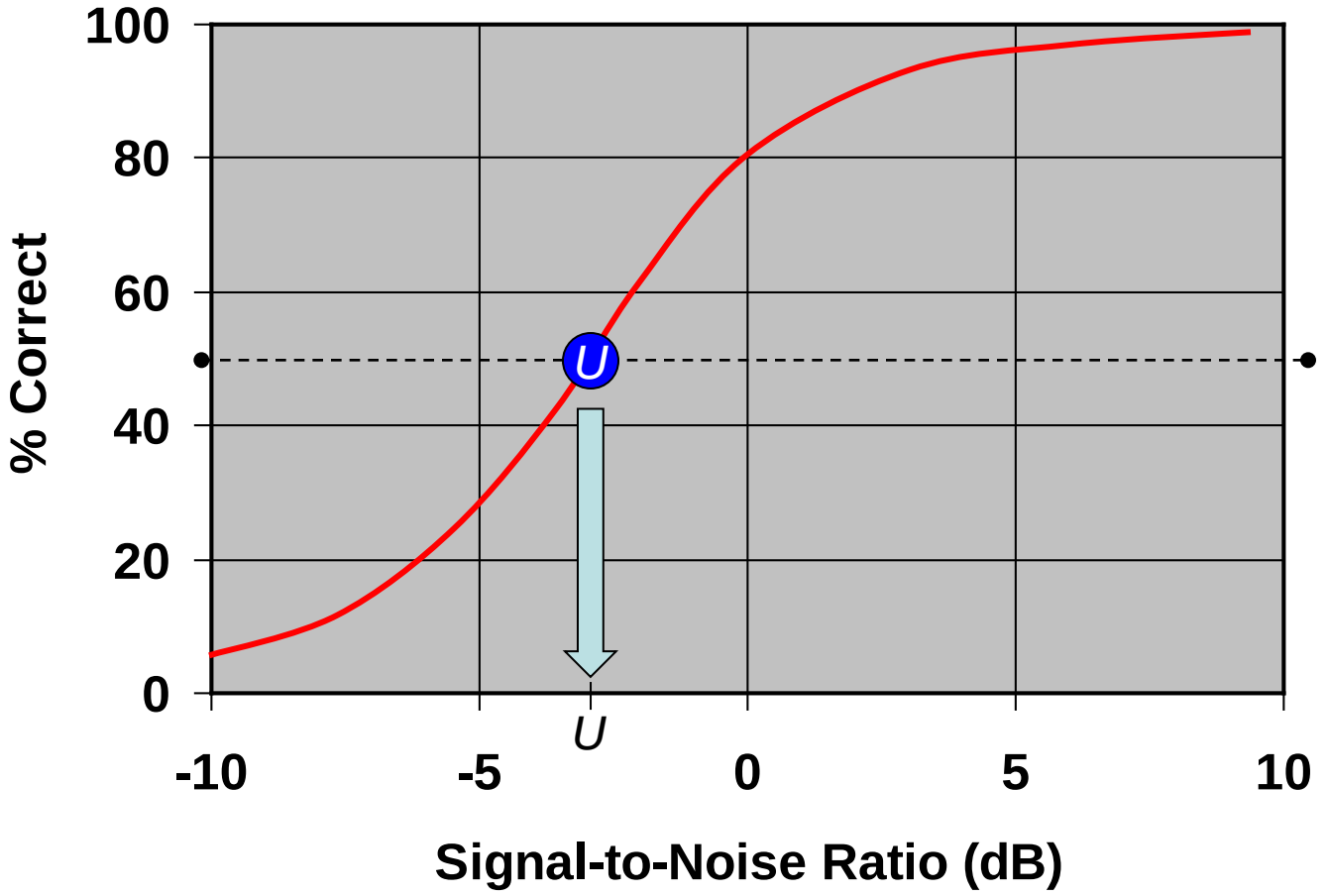
- dB SPL or dB SNR ?

dB SNR

SRT determination (unaided)

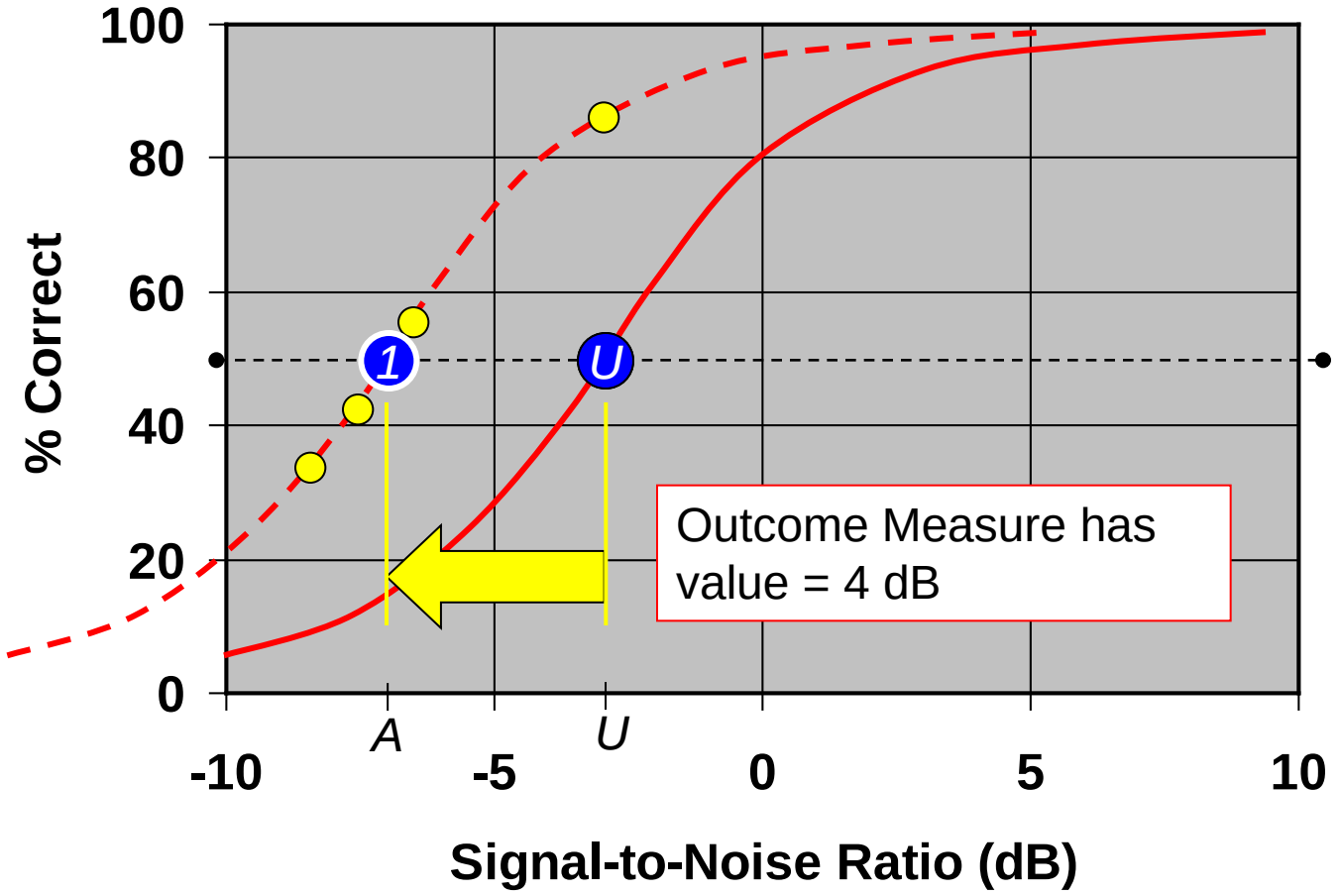


SRT determination (unaided)

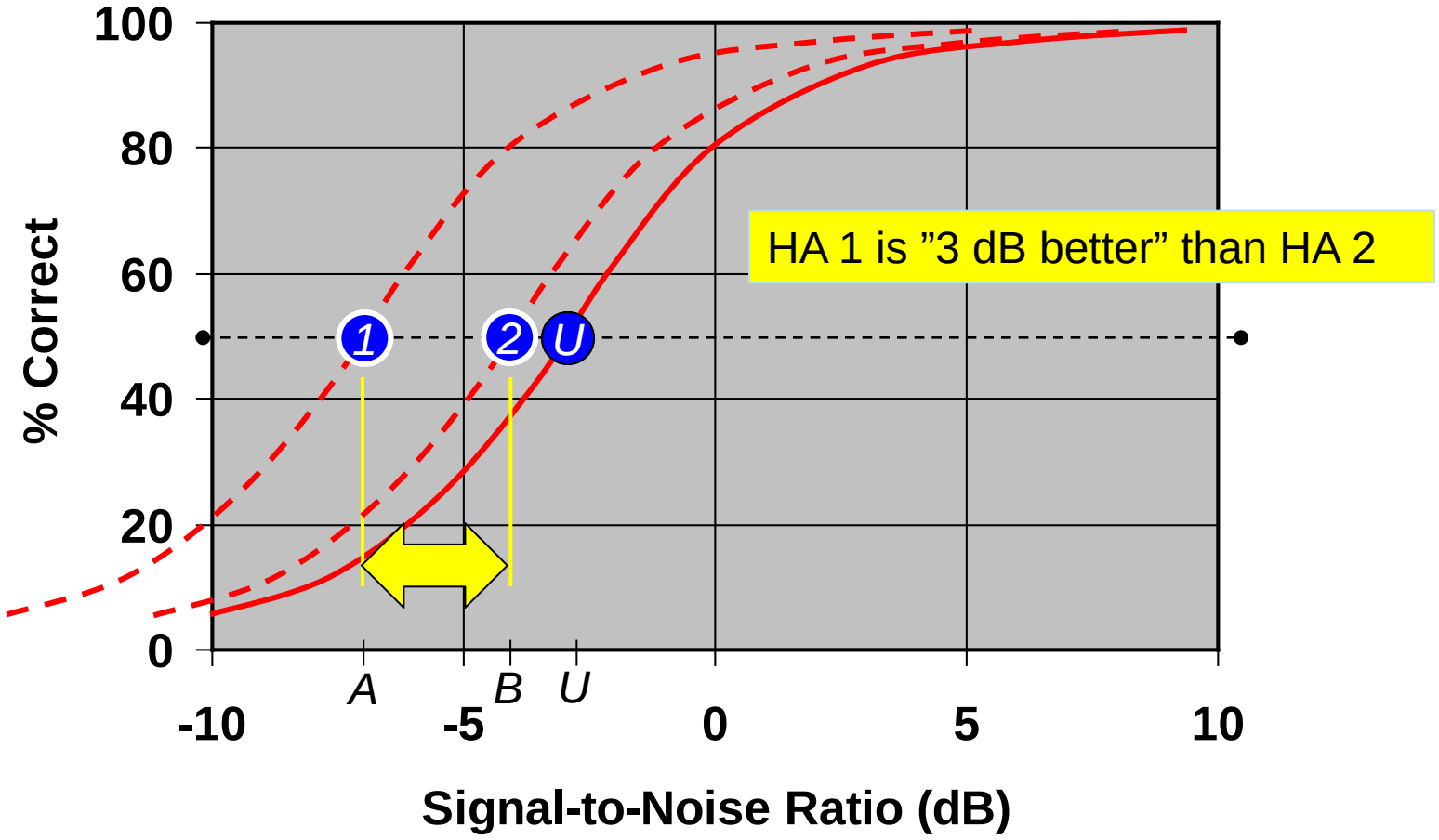


SRT as an outcome measure

The aided SRT outcome for HA 1

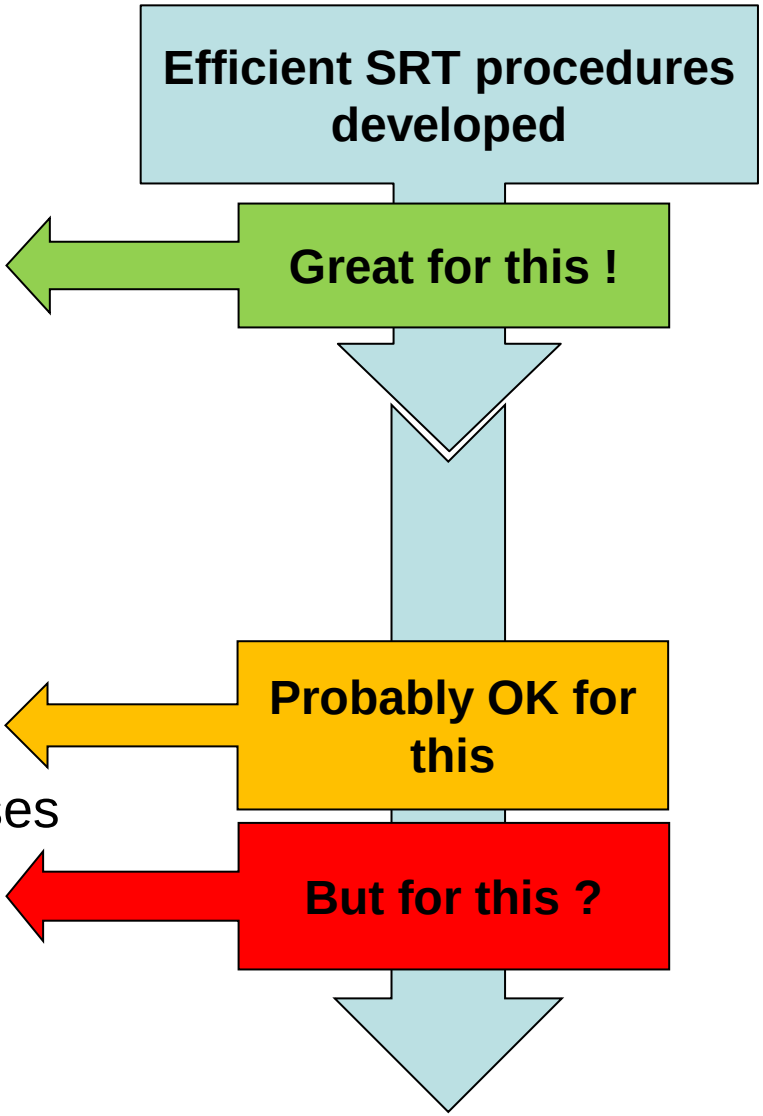


Comparing HA 1 vs. HA 2



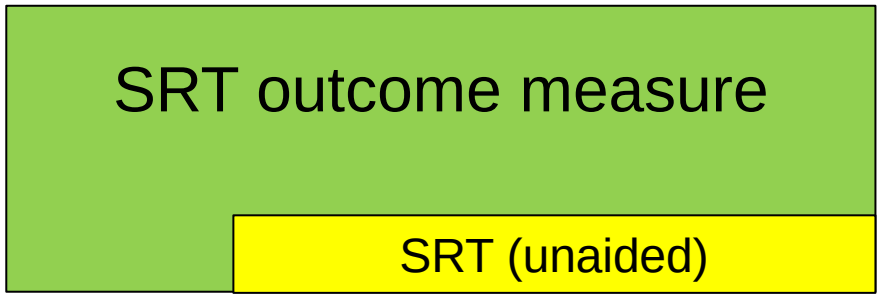
SRT in different arenas

- Unaided (’diagnostic’) SRT
  - Pre-intervention indicator of
    - degree of disability
    - likelihood of benefit with HA
- SRT as outcome measure
  - Aided vs. unaided
    - Quality control of clinical processes
  - Comparison between HAs
    - Obvious idea



Plan

- What ”SRT” am I talking about?
- Factors affecting SRT (unaided)



Effects of some parameters ”orthogonal”

Article X: Stated

”Changing factor Y in the test protocol has the effect of moving the SRT over a range of Z dB”

(Easily interpreted as)

”The underlying psychometric function is shifted by Z dB”.

Potential consequences for SRT (outcome)

- SRT obtained for a given HA system may be
  - determined more by the lab's habitual protocol than by the phenomenon being studied

OR

  - as desired, by adjusting the protocol.

Effects of some parameters

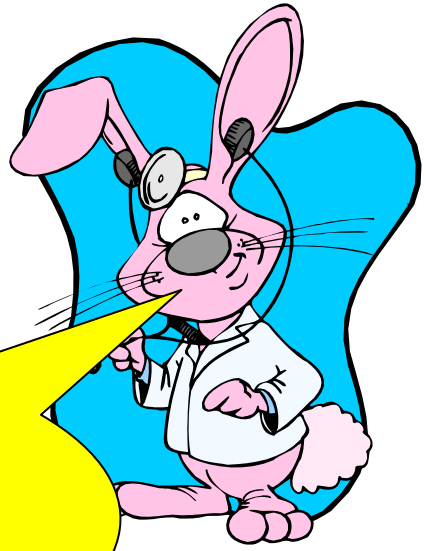
Parameter	Choices				Effect (dB)
Scoring	Sentences	Words	...		3
Criterion %	50 %	80 %	...		4
Masker type	Steady	Mod.	Talker ♀♂	Talker ♀♀	10
Target language	Native	Second			6
Sentence context	High	Low			5
Target talker intellig.	High	Low			1
Room	Dead	Reverb			6
Target/Masker location	Co-located	Separated			7
Response set size	4	32	1000	...	6

You must be joking !

These parameters are ”orthogonal”, so we can add the effects ...

How can that be ?

- Changing parameters is *NOT* equivalent to a parallel shift of an 'underlying PMF'
  - f(a,b,c,...) ≠ g(a) • h(b) • j(c)...
  - 'listening' task changes with listening conditions



Hmm! Well, taking the difference (ΔSRT) to compare devices A & B is still OK, isn't it?

If you're lucky, the SRTs *might* turn out to be in a relevant range

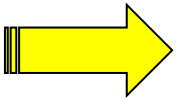
- As we move along the SNR axis ...
  - The listening task changes character

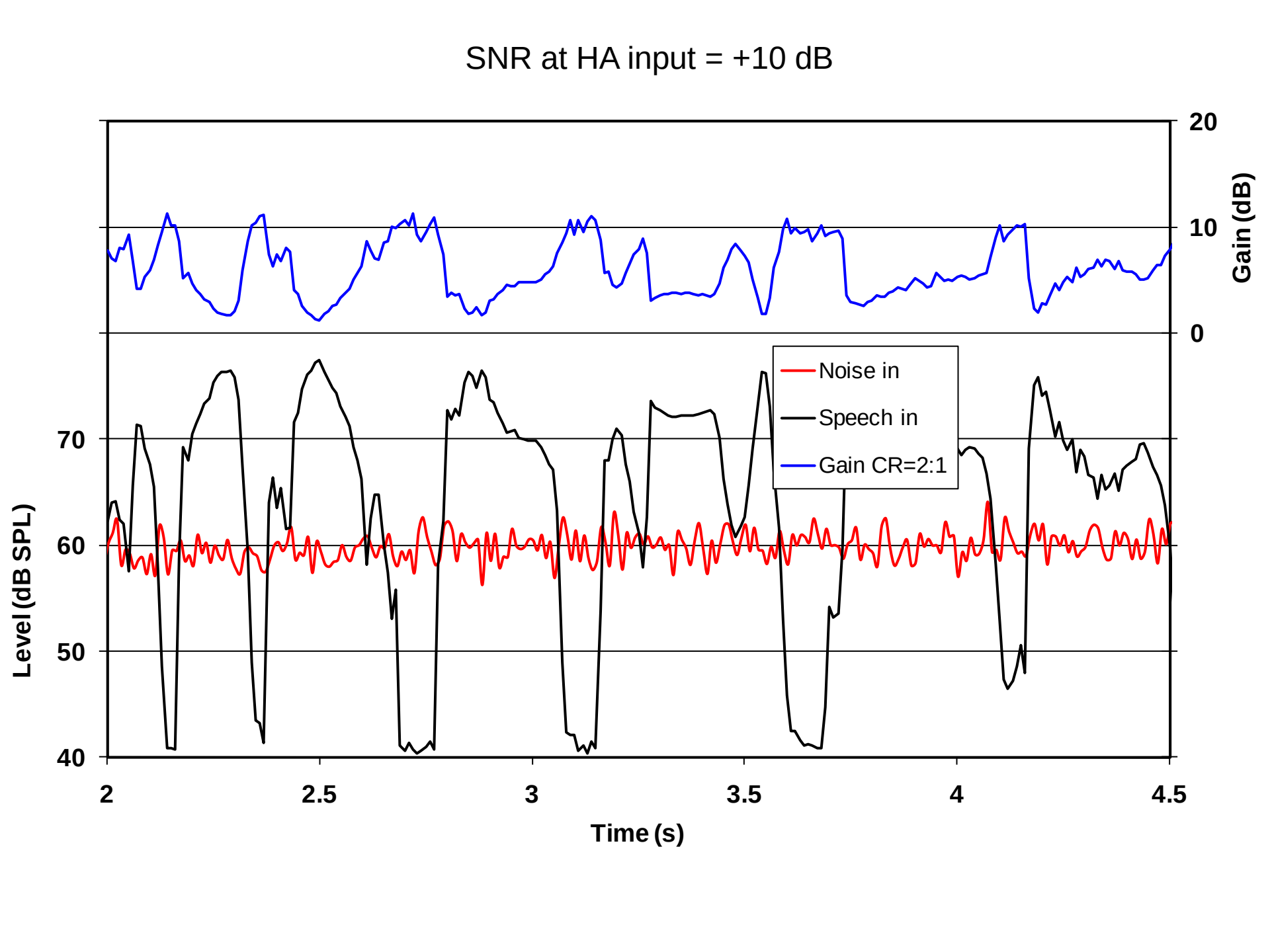
Plan

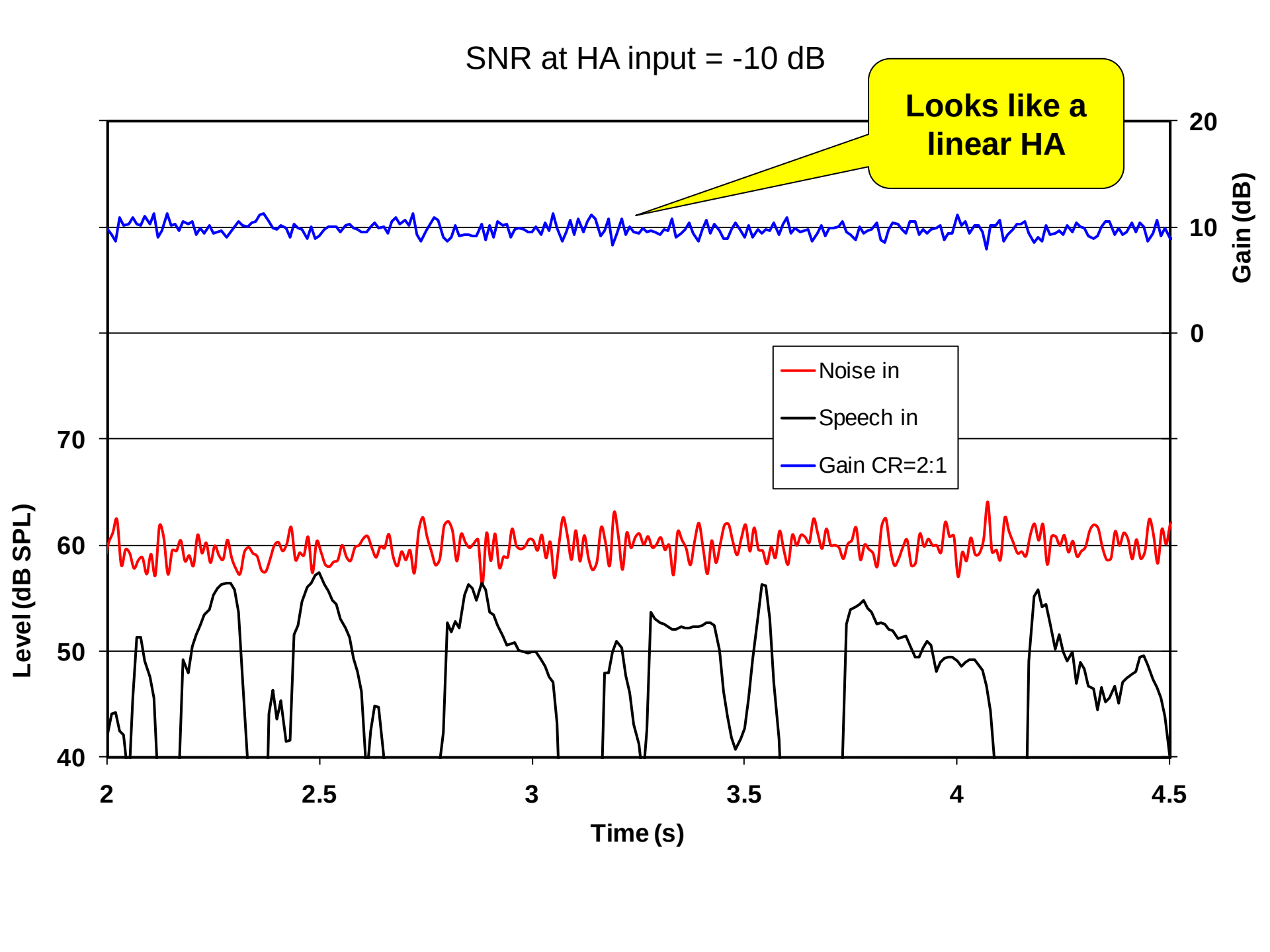
- What ”SRT” am I talking about?
- Factors affecting SRT (unaided)
- Potential consequences for SRT (outcome)
- Non-linear hearing aid (HA) systems

Non-linear hearing aids

- Behave differently depending on input signal
  - thus behaviour depends on SNR at input
- Types of nonlinearity (examples)
  - Noise Reduction
    - controlled by modulation-based estimate of Voice-to-Noise Ratio
  - Steering of directional processing
    - based on Direction-Of-Arrival estimation
  - Binary Masks
    - based on SNR estimates in Time x Freq cells
  - Dynamic range compression
    - Example: fast-acting WDRC at +10 and -10 dB input SNR




SNR at HA input = +10 dB
Noise in
Speech in
Gain CR=2:1
Level (dB SPL)
Gain (dB)
Time (s)
20
10
0
70
60
50
40
2
2.5
3
3.5
4
4.5


SNR at HA input = -10 dB
Looks like a linear HA
20
10
0
Gain (dB)
Noise in
Speech in
Gain CR=2:1
70
60
50
40
Level (dB SPL)
2
2.5
3
3.5
4
4.5
Time (s)

So far, so ...

- Problems, problems
  - Untenable assumption of ’underlying PMF’
  - Non-linearity of HAs
- It matters what SNR we test at
- This is not really news to anyone ...
  - but we still go on using SRT without constraining the SNR.

Why is SRT so attractive, despite these problems?

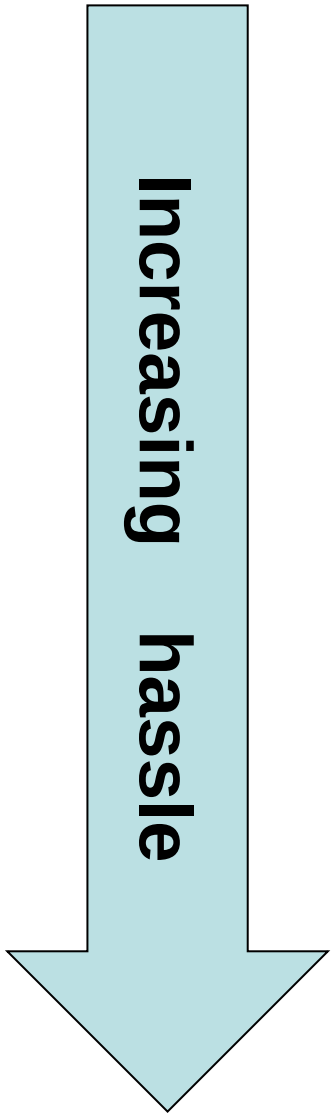
- Always gives a result with convenient statistical properties
- Can be made relatively fast for a given repeatibility
- Nobody has told us what SNRs we should be designing HA systems to work in
  - so guilty conscience does not kick in
- No catastrophic wrong conclusions documented so far
  - but there may be some candidates ...

Plan

- What ”SRT” am I talking about?
- Factors affecting SRT (unaided)
- Potential consequences for SRT (outcome)
- Non-linear hearing aid (HA) systems
- Why is SRT nevertheless so attractive?
- What to do?

What to do?

- Carry on using SRT ... with care
  - Appropriate respect for the SNR confound



What to do?

- Carry on using SRT ... with care
- Find out what SNRs really occur in what HA use situations, and use this to
  - (a) specify HA feature operation during design and
  - (b) prescribe SNR ranges for testing the feature
    - Pearsons et al 1977 NOT adequate basis

What to do?

- Carry on using SRT ... with care
- Find out what SNRs really occur in what HA use situations
- “Constrained SRT testing”
  - manipulate protocol to ensure operation at appropriate SNR for reference condition: then ΔSRT is probably an OK measure for a comparison
    - (maybe) do it for each individual listener
    - manipulations of protocol must not be of a sort which will affect the HA system operation, e.g.
      - Response set size
      - Repetitions of stimulus
      - Scoring rules, etc. etc.

What to do?

- Carry on using SRT ... with care
- Find out what SNRs really occur in what HA use situations
- “Constrained SRT testing”
- Accept or deal with ceiling/floor effects
  - manipulate protocol to ensure operation at appropriate SNR for reference condition: then measure Δ%-correct for test condition instead of ΔSRT
    - 100% is 100%
    - Two HA systems yielding 100% (or 0%) *are equally good* within the domain of the speech test being used.

What to do?

- Carry on using SRT ... with care
- Find out what SNRs really occur in what HA use situations
- “Constrained SRT testing”
- Accept or deal with ceiling/floor effects
- Measure complete PMFs (per individual listener)
  - Overkill ?

Plan

- What ”SRT” am I talking about?
- Factors affecting SRT (unaided)
- Potential consequences for SRT (outcome)
- Non-linear hearing aid (HA) systems
- Why is SRT nevertheless so attractive?
- What to do?

Summing up

- Good methods for ’diagnostic SRT’
  - may be risky to use as Outcome Measures
- Assumption of // underlying PMF is faulty
  - absurd range of apparently possible SRTs
- Nonlinear HAs
  - behave differently at different input SNRs
- We need to be deciding what SNR range to test at ***before testing***
  - some suggestions (no ready solutions)
- Work needed to chart real-life SNRs.

End

Effect of %-age criterion

- Realistically we wouldn’t vary over whole 0 <-> 100
  - conservative, 50 .. 80 %
  - Pichora-Fuller 1995: 5 dB
  - Wagener & Brand 2005: 2 dB
  - Say 3.5 dB

Sentence context

- changing from high-context to low-context sentences = 5 dB (Pichora-Fuller 1995)

Scoring words vs. sentences

- Low-context sentences (e.g. Oldenburg)
  - Ps = Pw ** 3.96 (Bronkhorst W & B 2002, Boothroyd & Nittrouer 1988)
  - 50% words = 6% sentences (!!)
  - 50% sentences = 84% words
  - 50% words -> 50% sentences requires 4 dB
    - for corpus with steep pmf (Wagener, Jovassen etal 2003)
- Less for higher-context sentences
- Say 3
- ?? what about scoring phonemes

Changing talker

- highly-intelligible to a less-intelligible talker
  - ca. 20% (Cox et.al. 1987a)
  - = 20/12 = 1.7 dB (Cox et.al. 1987b)

Noise type

- Steady-state vs. modulated
  - (Wagener & Brand 2005, many others ???refs),
  - Controversy
    - ”no effect” -> ”??? dB” vs. ”depends on SNR”
    - take your pick! Bernstein & Grant 2009
    - Conservative = 0
- ... vs. N talkers
  - Gender (7..11 dB, Festen & Plomp 1990)

Language

- Bilinguals
  - in s/s noise,
    - 9 dB (Stuart et al 2010)
    - 3 dB (Nilsson et al 1992)
  - babble, 6 dB (Mayo et al 1997)

Acoustics of setup

- Monotic vs. diotic vs. dichotic
- Co-located S & N vs. spatially separated
  - Neher ISAAR 2007 NH 10 dB, HI 7 dB (LR) sp-on-sp
  - Bronkhorst & Plomp 1990 6.5 dB s/s noise 1 masker
  - Marrone ISAAR 2007 NH 10 dB, HI 4 dB (LR) sp-on-sp
  - say 7
- Reverberation
  - Arweiler ISAAR 2007 ca. 4 dB
  - Plomp 1976 6-10 dB
  - say 6

Response set size

- 32 vs. 1000 ca. 6 dB (Bernstein pers. comm.)