

# How Much Better Can Wideband Telephony Be? — Estimating the Necessary R-scale Extension

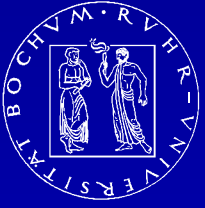
*Alexander Raake*

Institute of Communication Acoustics (Prof. R. Martin, em. Prof. J. Blauert)  
Laboratory for Auditory Quality of Systems (PD. U. Jekosch)  
Ruhr-University Bochum

Workshop on Wideband Speech Quality in Terminals and Networks:  
Assessment and Prediction

8th and 9th June 2004 - Mainz, Germany

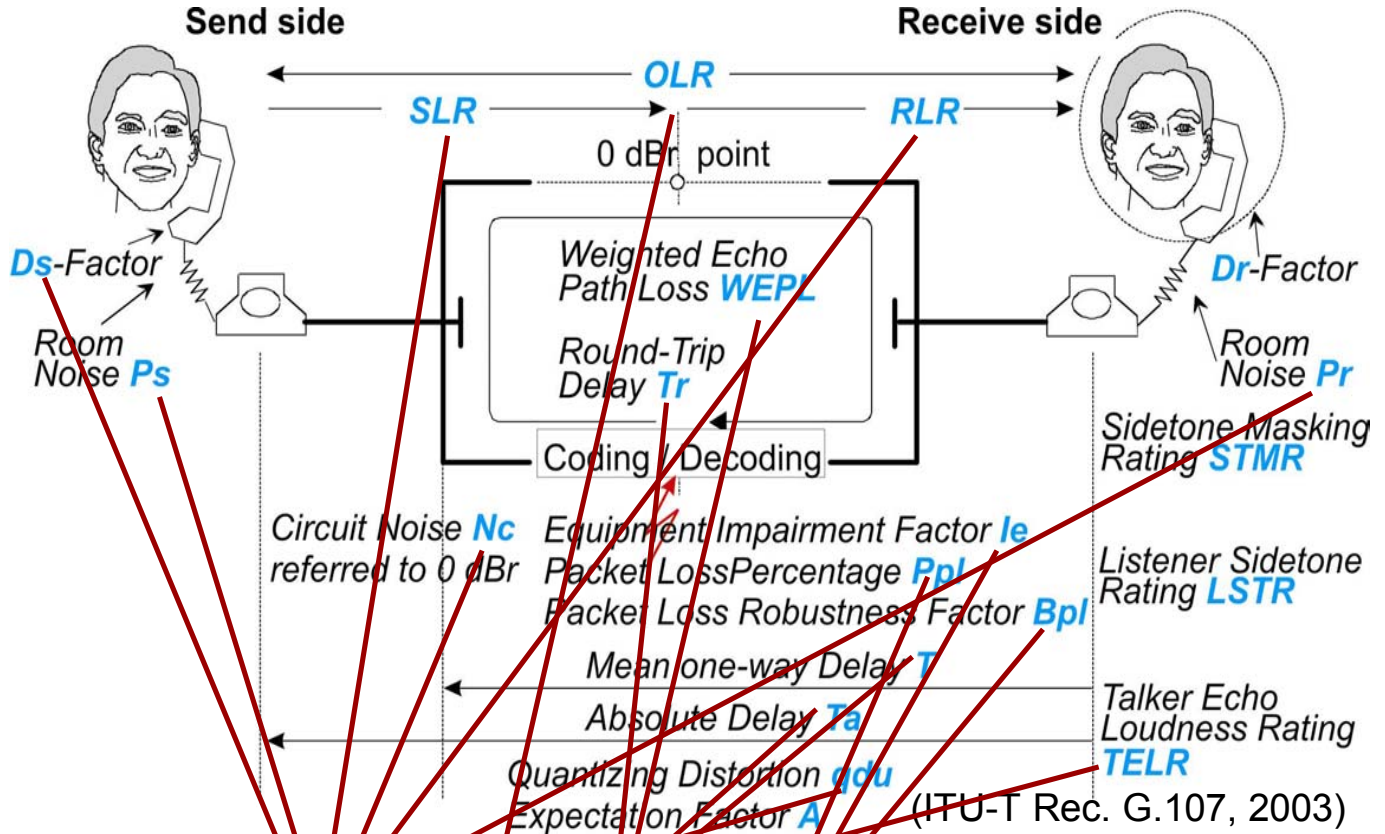
**IKA**



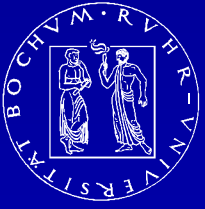
# Introduction

Parametric description

NB E-Model



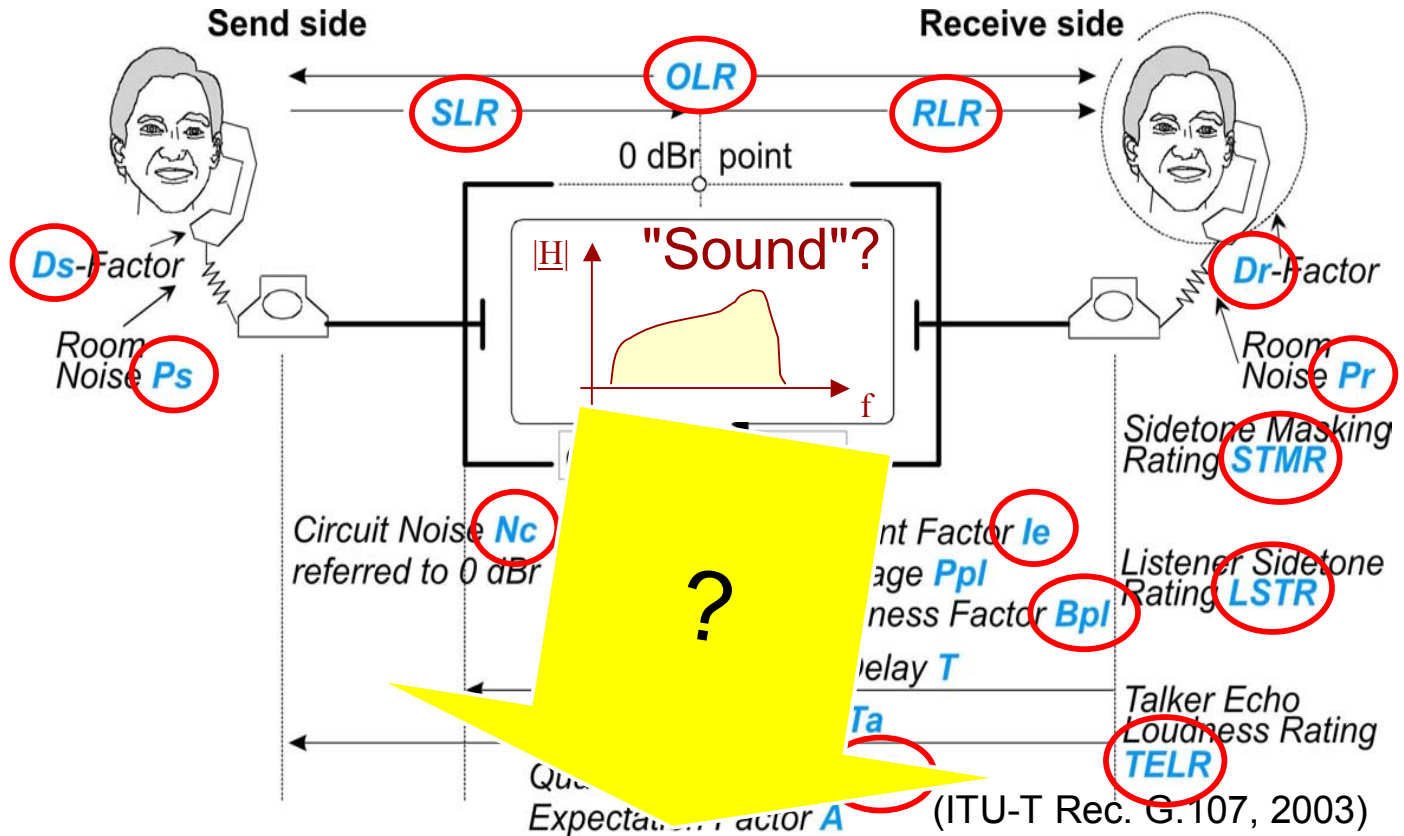
$$R = R_o - I_s - I_d - I_{e,eff} + A$$



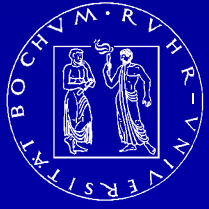
# Introduction

WB-parameters?

WB  
E-Model?

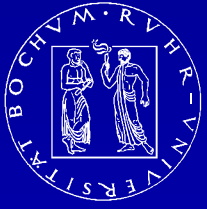


$$R' = Ro' - Is' - Id' - Ie,eff' + A'$$



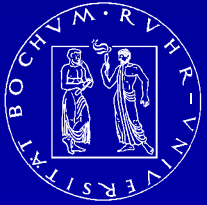
# Overview

- Introduction
- E-model: From NB to WB
- Listening Tests
- R-scale Extension
- Conclusions and Outlook



# E-Model Algorithm, NB

- Parametric speech quality prediction model for NB-network planning
- ITU-T Rec. G.107 (2003)



# E-Model Algorithm, NB

“[...] evaluation of psychological factors on a psychological scale is additive”

$$R = R_0 - I_s - I_d - I_{e,eff} + A$$

*R* Transmission rating factor [0,100]

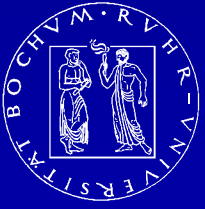
*R<sub>0</sub>* Basic signal-to-noise ratio

*I<sub>s</sub>* Simultaneous impairment factor

*I<sub>d</sub>* Delayed impairment factor

*I<sub>e,eff</sub>* Effective equipment impairment factor

*A* Advantage factor (user expectation)



# E-Model Algorithm: **WB**

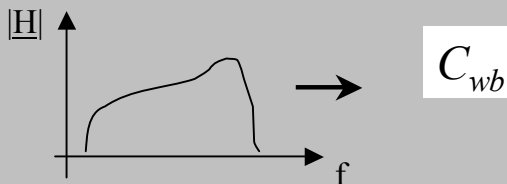
## R-Scale Extension

$$R_{wb} = R_0 - I_s - I_d - I_{e,eff} + A + C_{wb}(?)$$

(wideband:  
Intermediate?)

### Spectral characteristics

- 1st considerations:  
ITU-T SG 12 D. 12, 2001



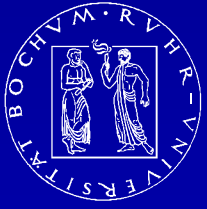
### Equipment Impairment Factor $I_e$

- Next Talk (Möller)



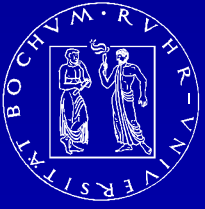
# Input Parameters: **WB**

Parameter	WB coverage	Reference	Remarks
SLR / RLR / OLR	✓ / ✓ / ✓	P.79, Ann. G	–
Nc / Nfor	(✓?) / (✓?)	P.83, Ann. A, O.41	Problem: Psophometric weighting 0.41 → NB
Ps / Pr	✓ / ✓	ISO 10845	–
STMR / LSTR	(✓?) / (✓?)	P.79 ; P.64	–
qdu / Q	– / ✓	– / P.810	–
TELR / WEPL	– / –	– / –	Problem: Rec. like G.122 → NB
Ta / T / Tr	✓ / ✓ / ✓	G.114 / ETR 250	Bandwidth-independent
le	✓	next talk (Möller)	–
Bpl	–	–	So far too few tests
Ppl	✓	–	Bandwidth-independent
"Sound"	-	-	Not at all in the E-model



# Overview

- Introduction
- E-model: From NB to WB
- Listening Tests
- R-scale Extension
- Conclusions and Outlook



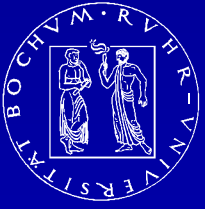
# Bandwidth and User Interfaces (I)

## LOT Headsets vs. Handset Telephone



$$\text{MOS} \approx \max\left(\frac{1}{2} \cdot (S_{10P} - 1) + 1, 1\right)$$

(ITU-T Rec. P.910 /  
IEC 268-13)

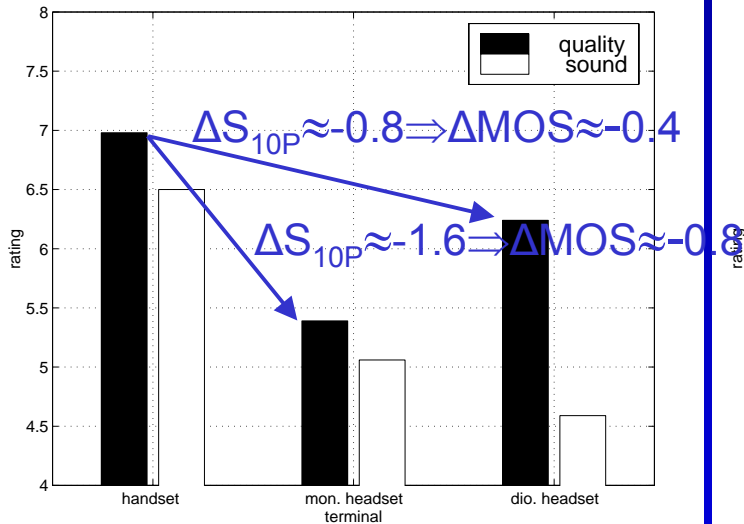


# Bandwidth and User Interfaces (II)

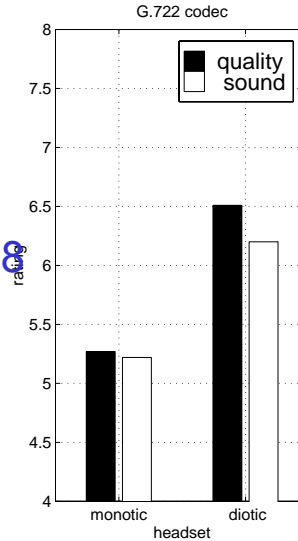
## LOT Headsets vs. Handset Telephone

### Test Results

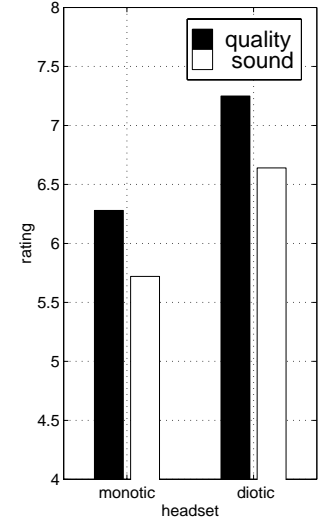
300-3400 Hz



50-7000 Hz



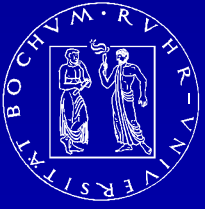
without codec



### Influencing Factors

- Bandwidth
- User Interface

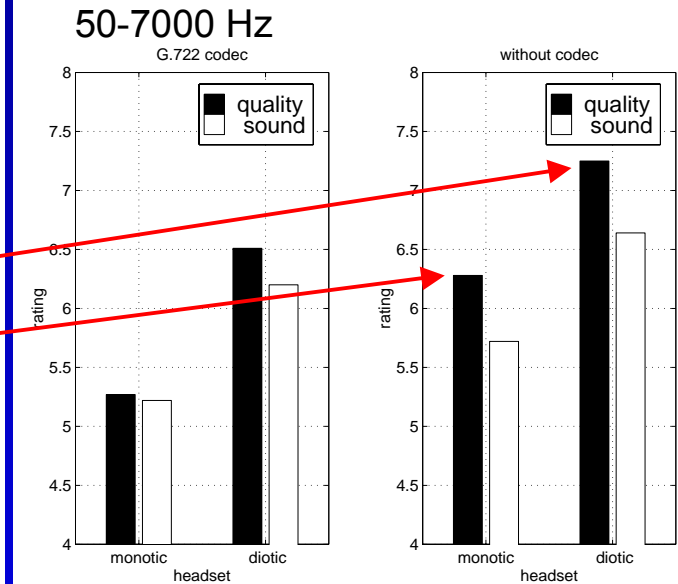
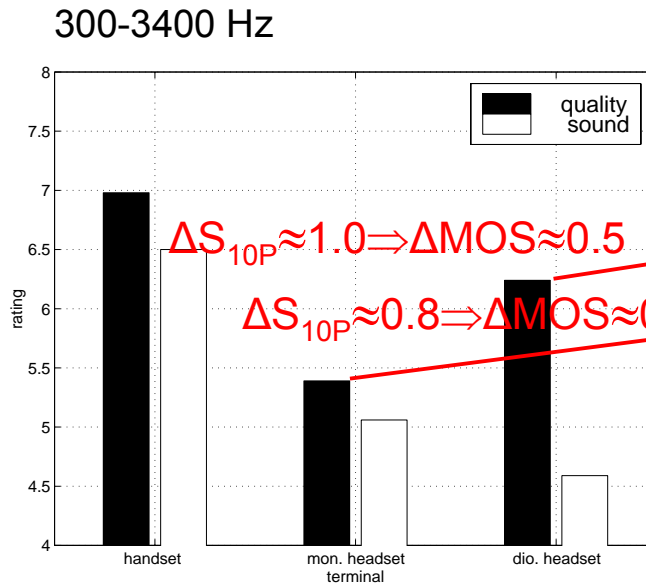
(Raake, 2000)



# Bandwidth and User Interfaces (II)

## LOT Headsets vs. Handset Telephone

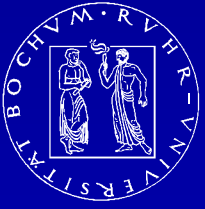
### Test Results



### Influencing Factors

- Bandwidth
- User Interface

(Raake, 2000)

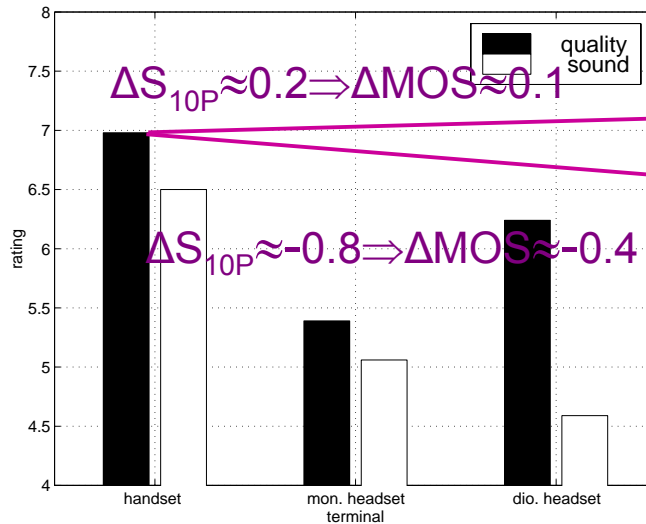


# Bandwidth and User Interfaces (II)

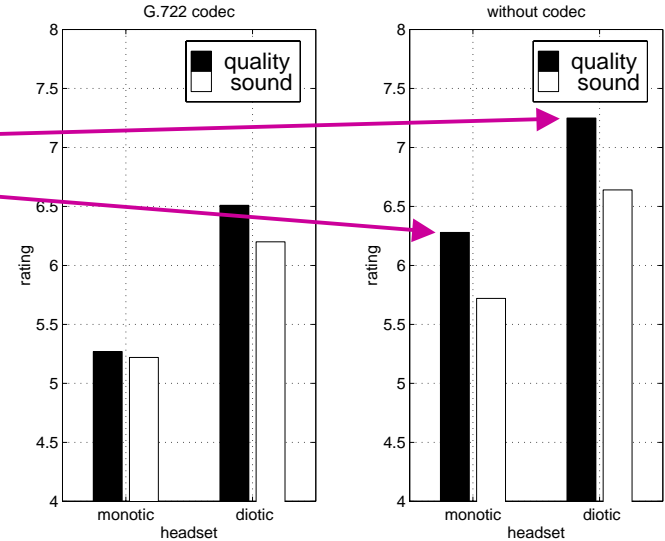
## LOT Headsets vs. Handset Telephone

### Test Results

300-3400 Hz



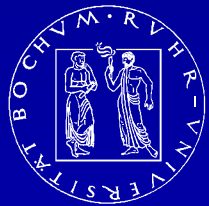
50-7000 Hz



### Influencing Factors

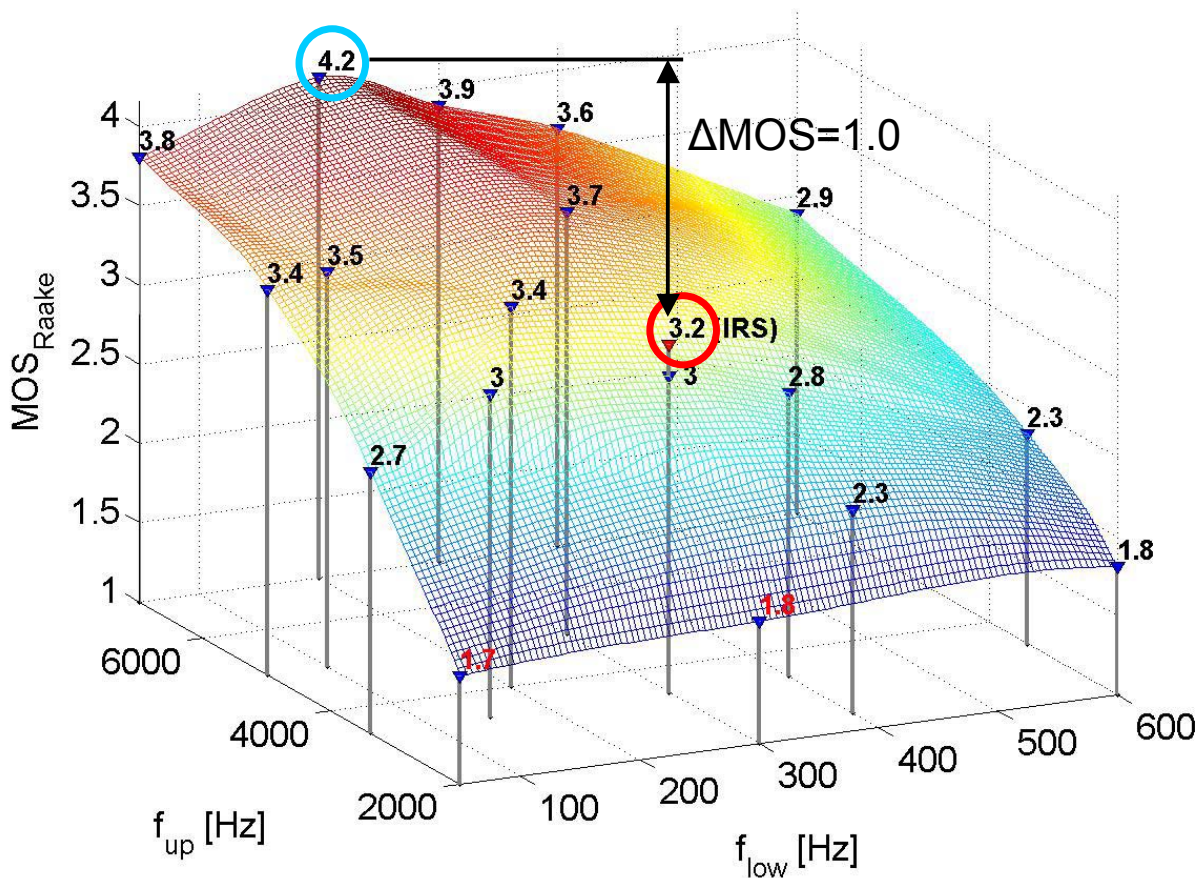
- Bandwidth
- User Interface

(Raake, 2000)

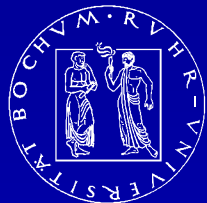


# Bandwidth

## LOT Stax-phone

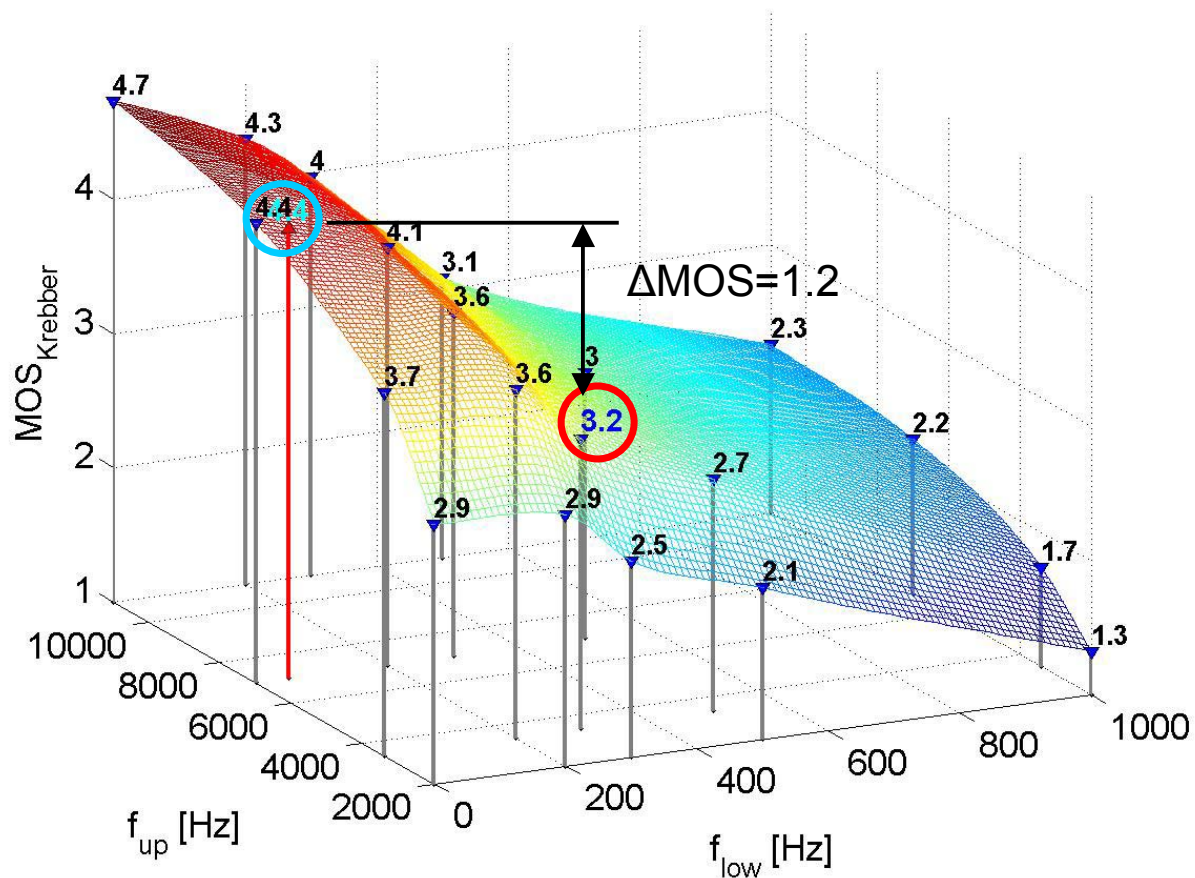


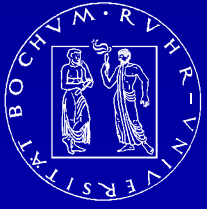
(Raake, 2001)



# Bandwidth

## LOT Headphone

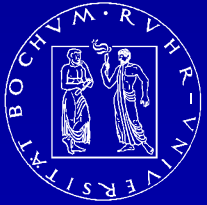




# Conclusions 1

Jekosch, 2000:

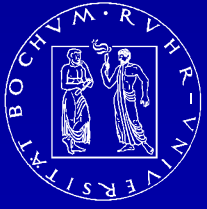
- *Quality*: Result of appraisal of perceived composition of a unit with respect to its desired composition ( $\equiv$  reference).



# Conclusions 1

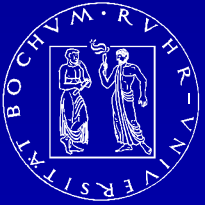
## Influencing Factors

- User Interface
  - Acoustics
  - Expectation  
(non-handset: No established reference for telephony!)
- Bandwidth



# Overview

- Introduction
- E-model: From NB to WB
- Listening Tests
- R-scale Extension
- Conclusions and Outlook



# R-Scale Extension (I) Procedure

## Reference Experiment

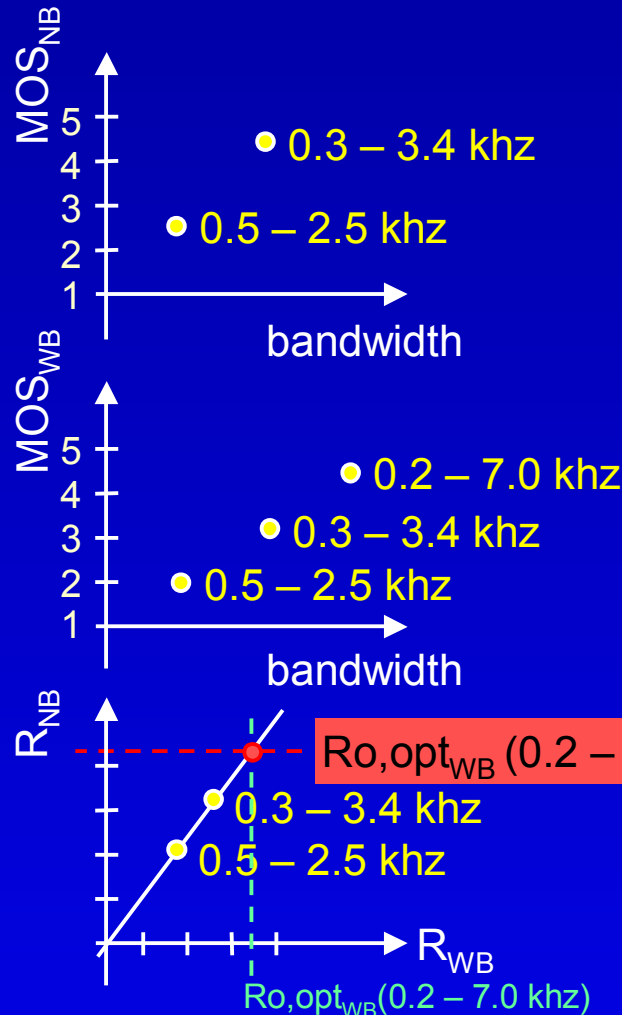
- narrow-band
- different degradation types (bandwidths, MNRU, codecs)

## WB-Tests

- wide-band
- different bandwidths

## R-scale extension

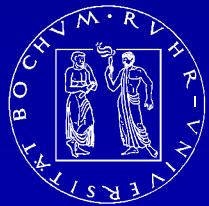
- plot  $R_{NB}$  vs.  $R_{WB}$
- linear interpolation
- extrapolation:  
 $R_{o,opt_{WB}}$



NB-Transformation

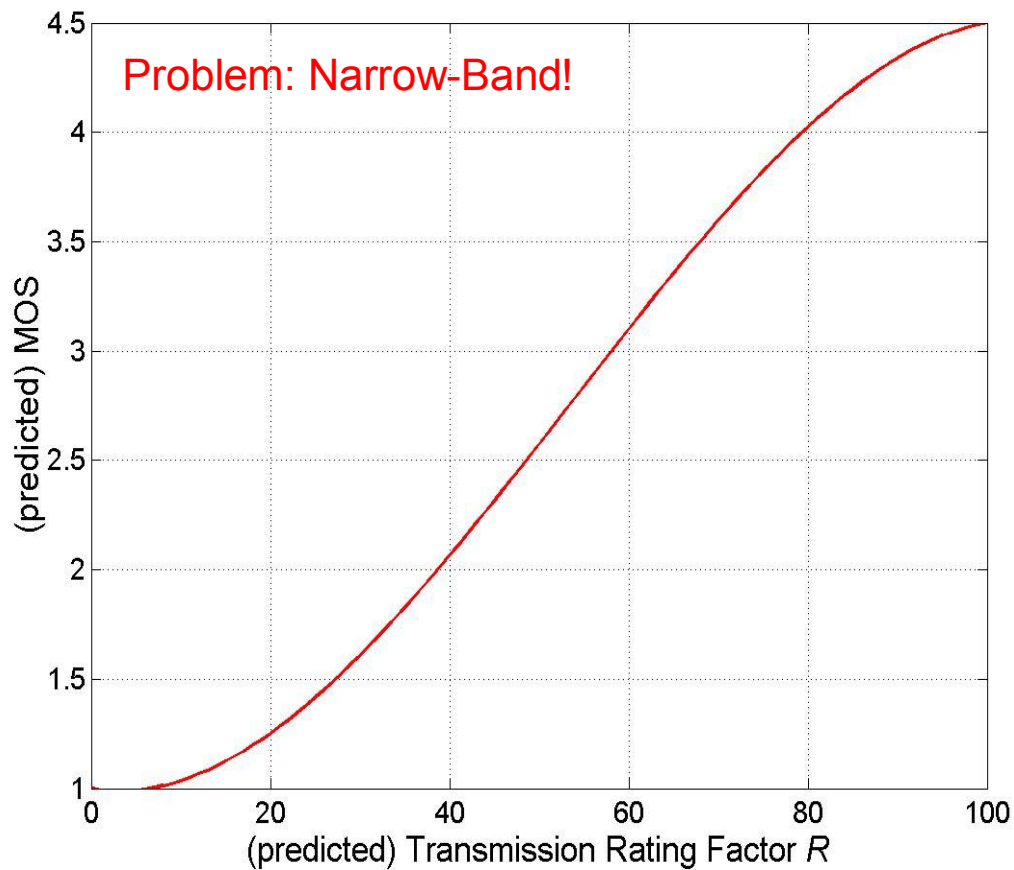
$MOS_{NB} \rightarrow R_{NB}$

$MOS_{WB} \rightarrow R_{WB}$

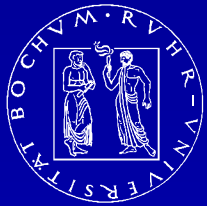


# R-Scale Extension (II)

## Transformation MOS $\rightarrow$ R



(ITU-T Rec. G.107, 2003)

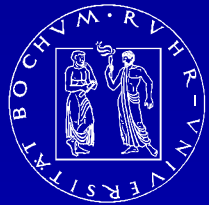


# R-Scale Extension (III)

## Reference LOT (Handset-Telephone)

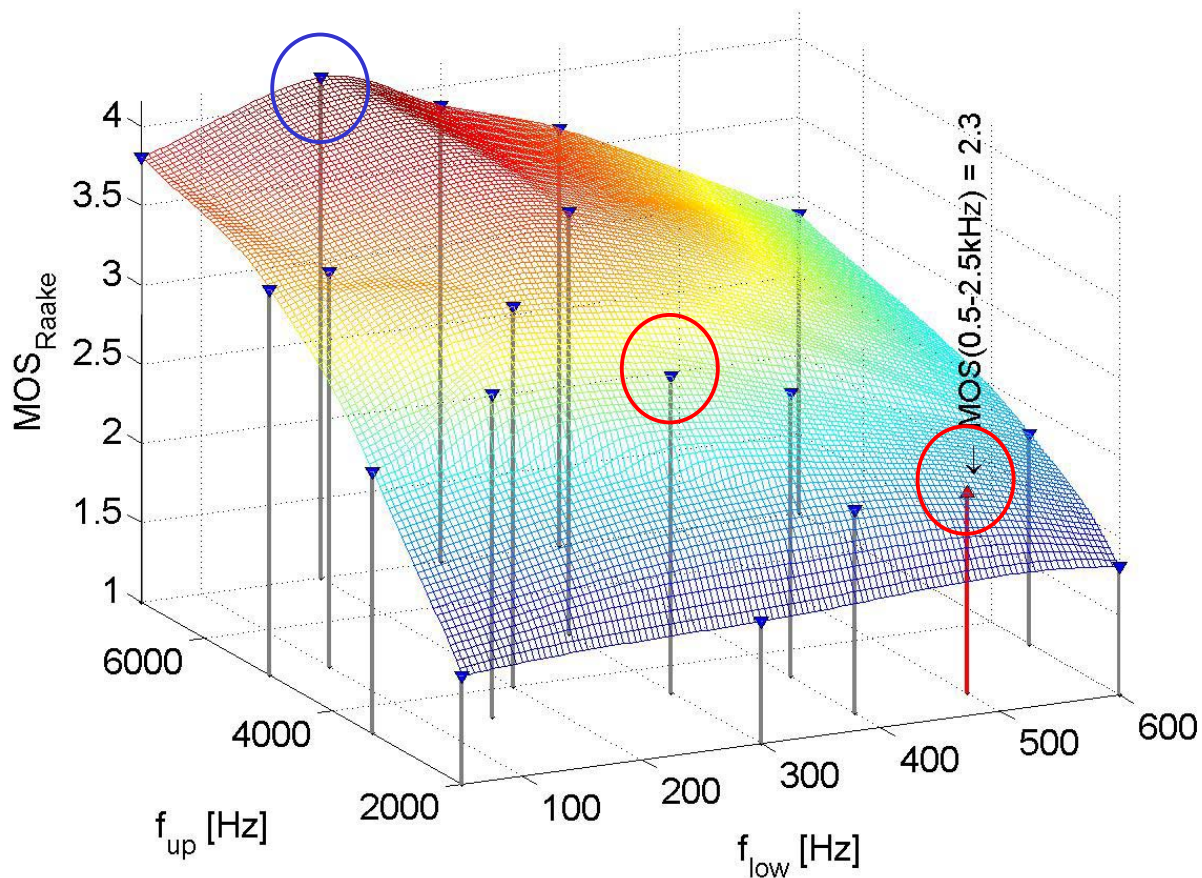
Test Condition						Listening Test
No.	Codec	SLR (dB)	RLR (dB)	SLR-Characteristic	RLR-Characteristic	MOS
1	MNRU 0	8	2	handset	handset	1.07
2	MNRU 5	8	2	handset	handset	1.21
3	MNRU 10	8	2	handset	handset	1.48
4	MNRU 15	8	2	handset	handset	2.25
5	MNRU 20	8	2	handset	handset	3.19
6	MNRU 35	8	2	handset	handset	4.11
7	G.711	8	2	handset	handset	4.14
11	G.711	8	2	BP 0.5-2.5	handset	3.02
12	G.711	8	2	BP 0.8-1.6	handset	1.91

(Möller, 2000)



# R-Scale Extension (IV)

## LOT Stax-phone

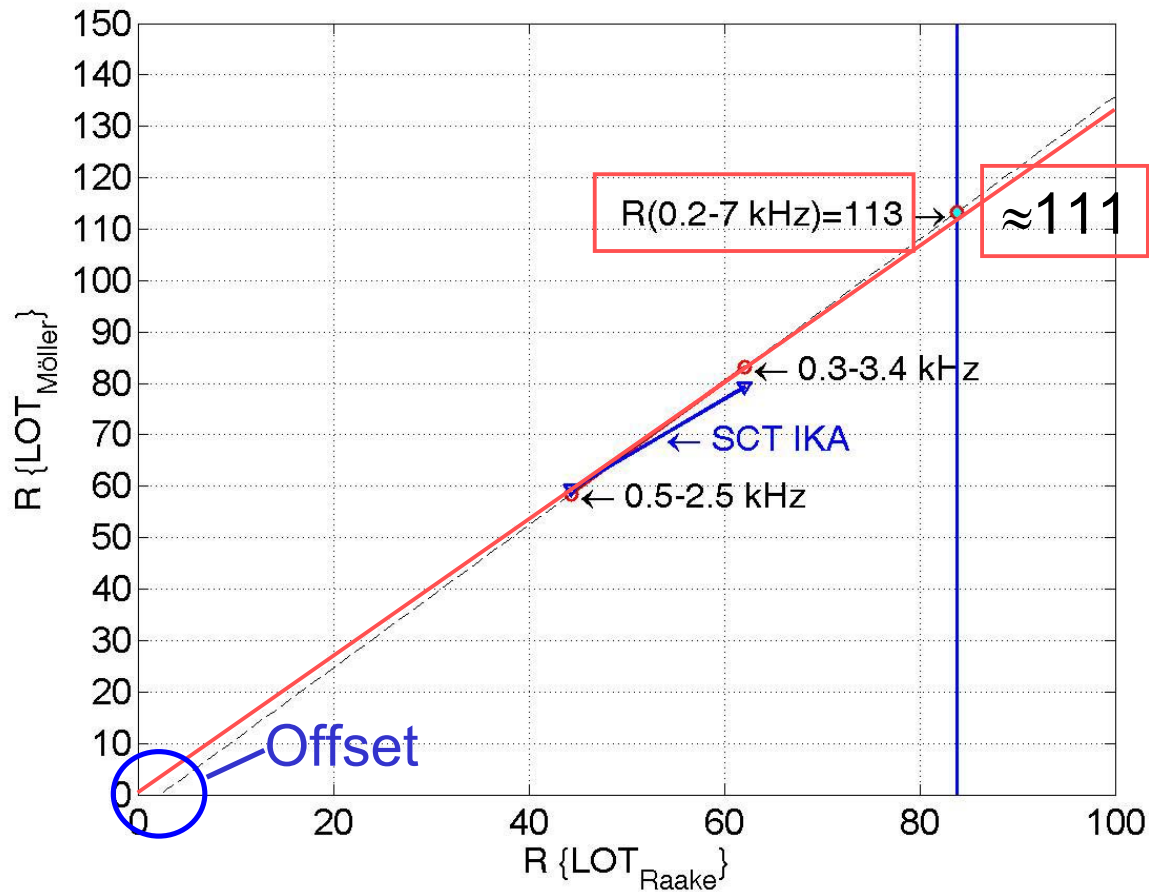


(Raake, 2001)



# R-Scale Extension (IVa)

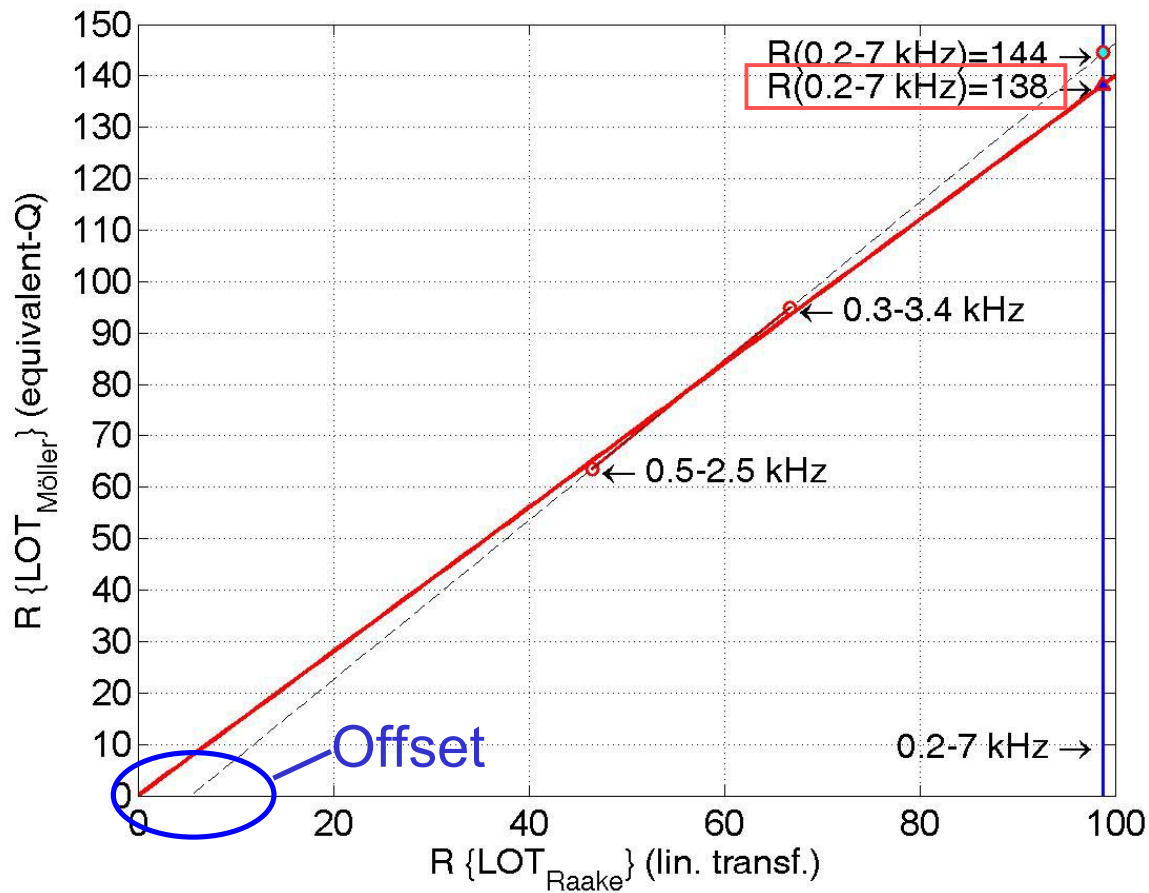
Reference: Raw Data; Stax-Phone: Raw Data

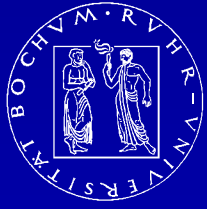




# R-Scale Extension (IVb)

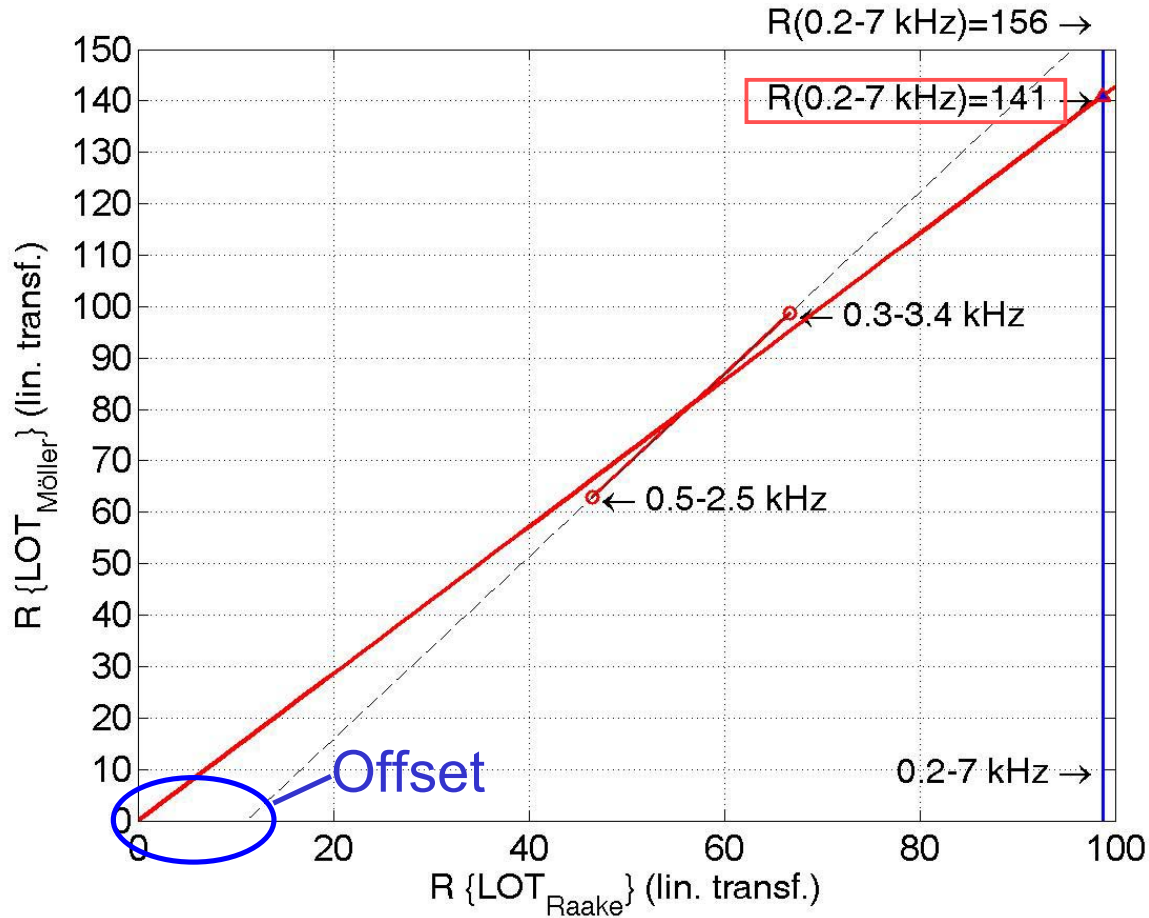
Reference: Equiv.-Q; Stax-Phone: Lin. Transform.

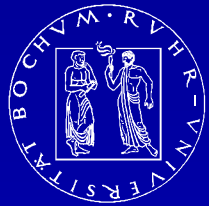




# R-Scale Extension (IVc)

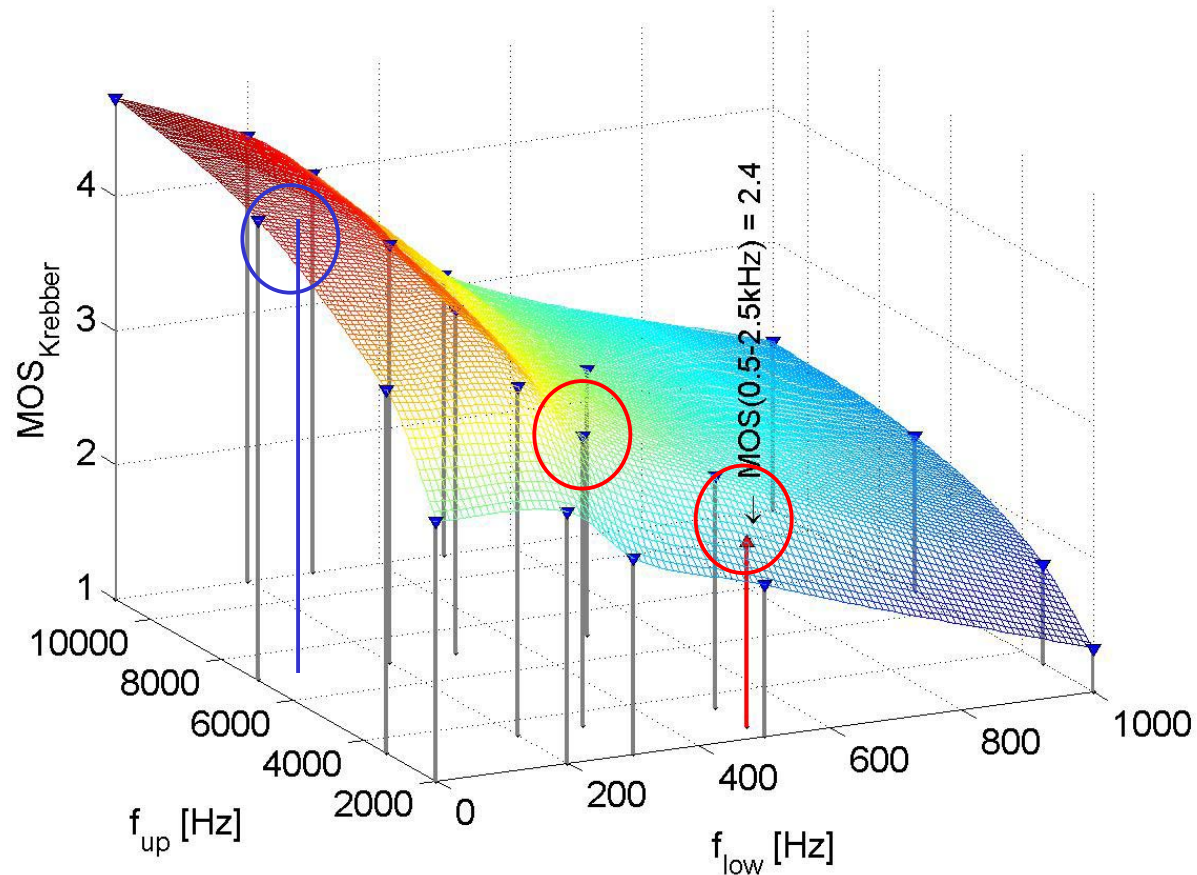
Reference: Lin. Transf.; Stax-Phone: Lin. Transf.

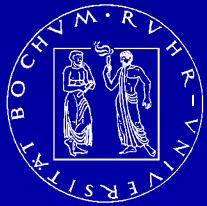




# R-Scale Extension (V)

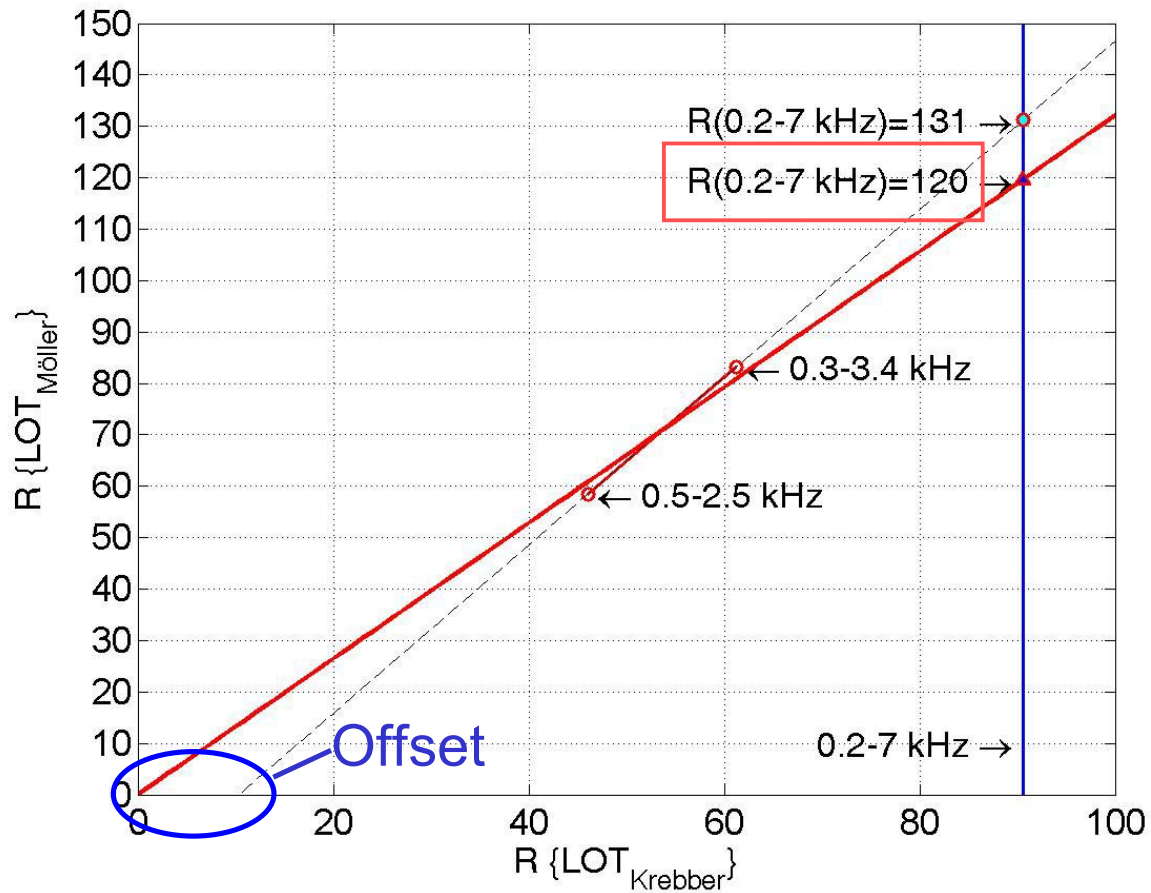
## LOT Headphone

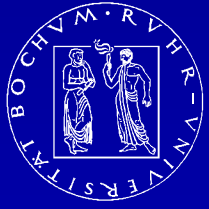




# R-Scale Extension (Va)

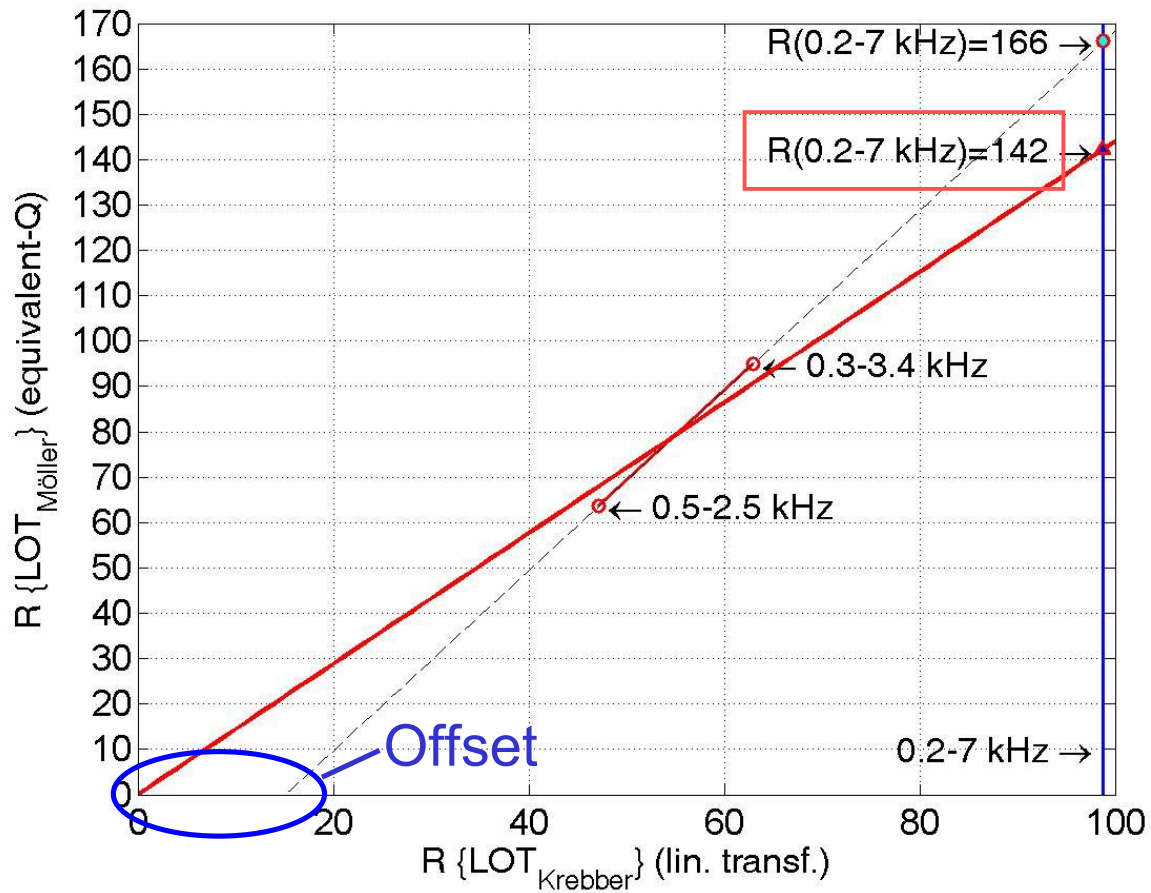
Reference: Raw Data; Headphone: Raw Data





# R-Scale Extension (Vb)

Reference: Equiv.-Q; Headphone: Lin. Transform.



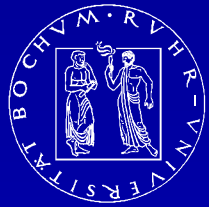


# R-Scale Extension (VI)

## Summary

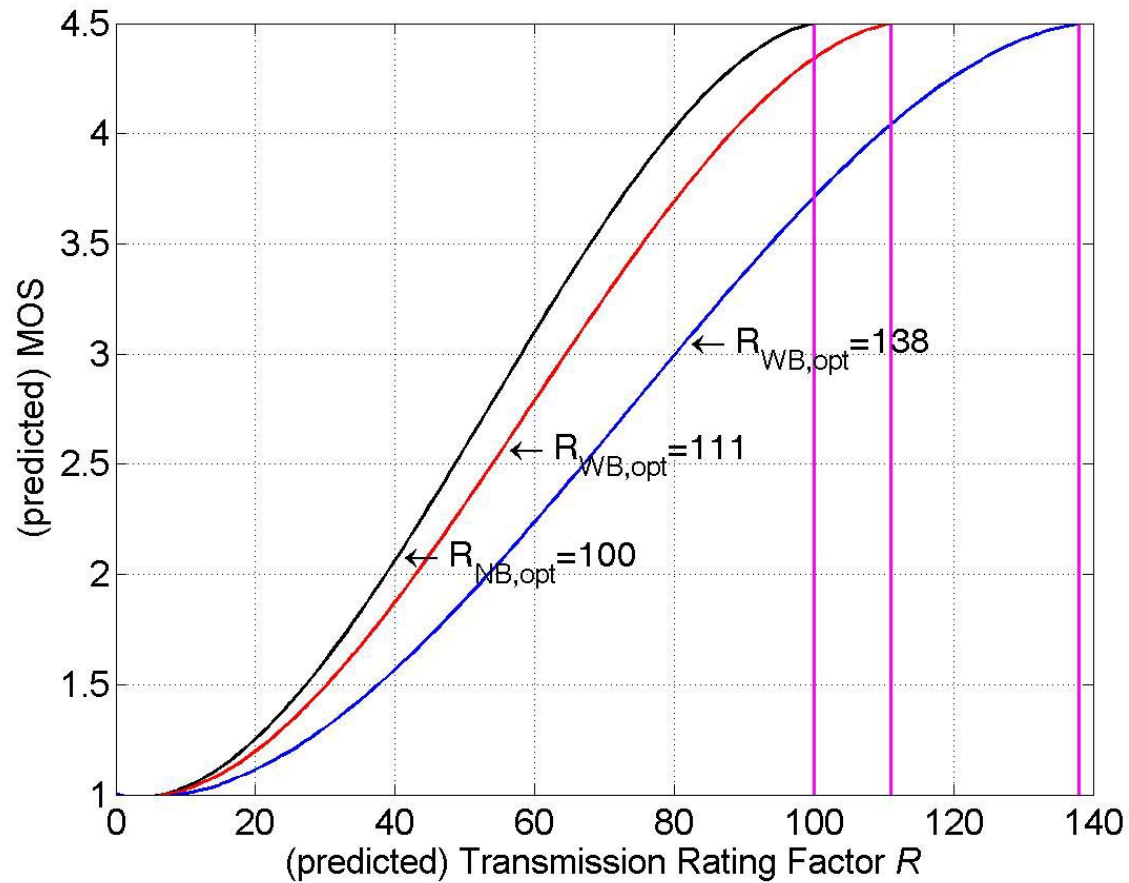
Test	Reference	$R_{o_{WB}}$	$R_{o_{WB}'}$ (offset = 0)	$C_{WB}$
Raake LOT Raw	Möller LOT Raw	113	111	11
Raake LOT Lin. Transform.	Möller LOT Equiv. Q	144	138	38
Krebber LOT Lin. Transform.	Möller LOT Raw	131	120	20
Krebber LOT Lin. Transform.	Möller LOT Equiv. Q	166	141	41

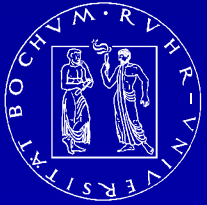
$$(R_{wb} = R_0 - I_s - I_d - I_{e,eff} + A + C_{wb})$$



# R-Scale Extension (VII)

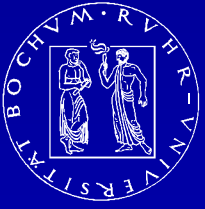
## R to MOS conversion





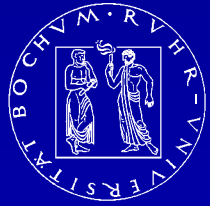
# Conclusions

- **Limitations of the approach**
  - NB-transformation MOS  $\rightarrow$  R
  - Only one reference experiment readily available
    - Others (e.g.: Voran, 1997; Gabrielsson, 1985; Gleiss, 1989)
      - Only WB
      - WB and NB, but no direct case
      - Different scales
      - No conditions anchored vs. NB-E-model
  - Linear transformation probably not ideal
- **Even with very conservative estimate  $>10\%$  extension**
  - Significant improvement over NB-case!



# Outlook

- **Refine RoWB-estimates**
  - Compile data presented here and in other tests presented at this workshop
- **New Cts and LOTs necessary for directly determining the R-scale extension**
  - Scale to be used?
- **Extensive CTs and LOTs necessary for WB + other degradations (noise, echo, delay, etc.)**



Thank you!

**IKA**