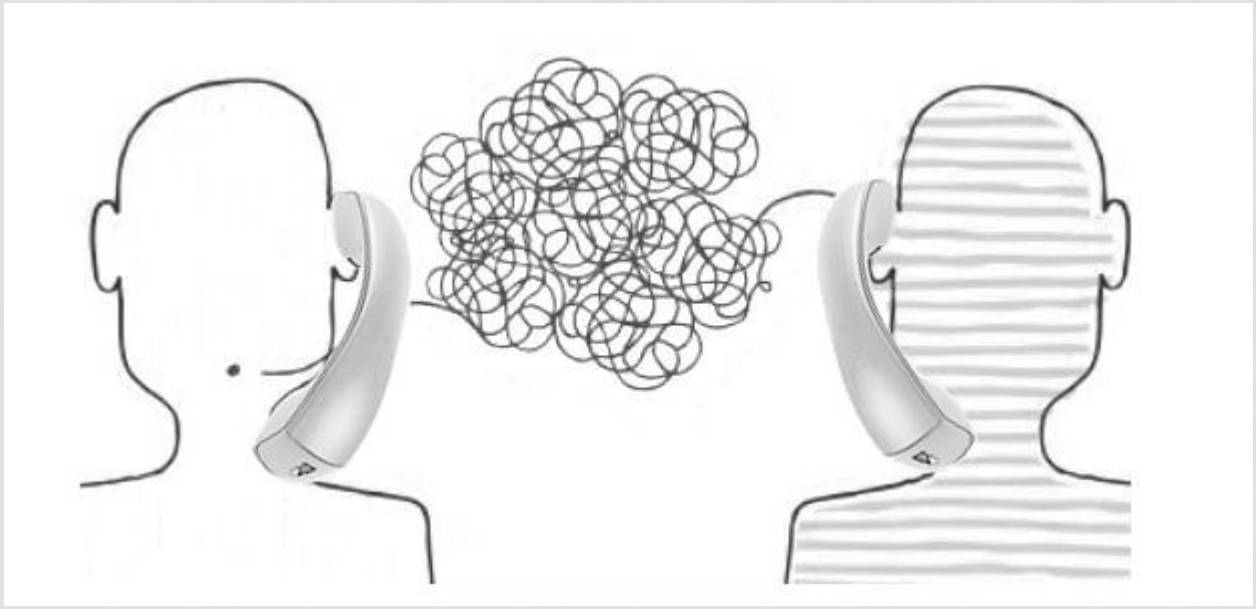
Human Speech Intelligibility Measurements over VoIP Channels

Laura Fernández Gallardo

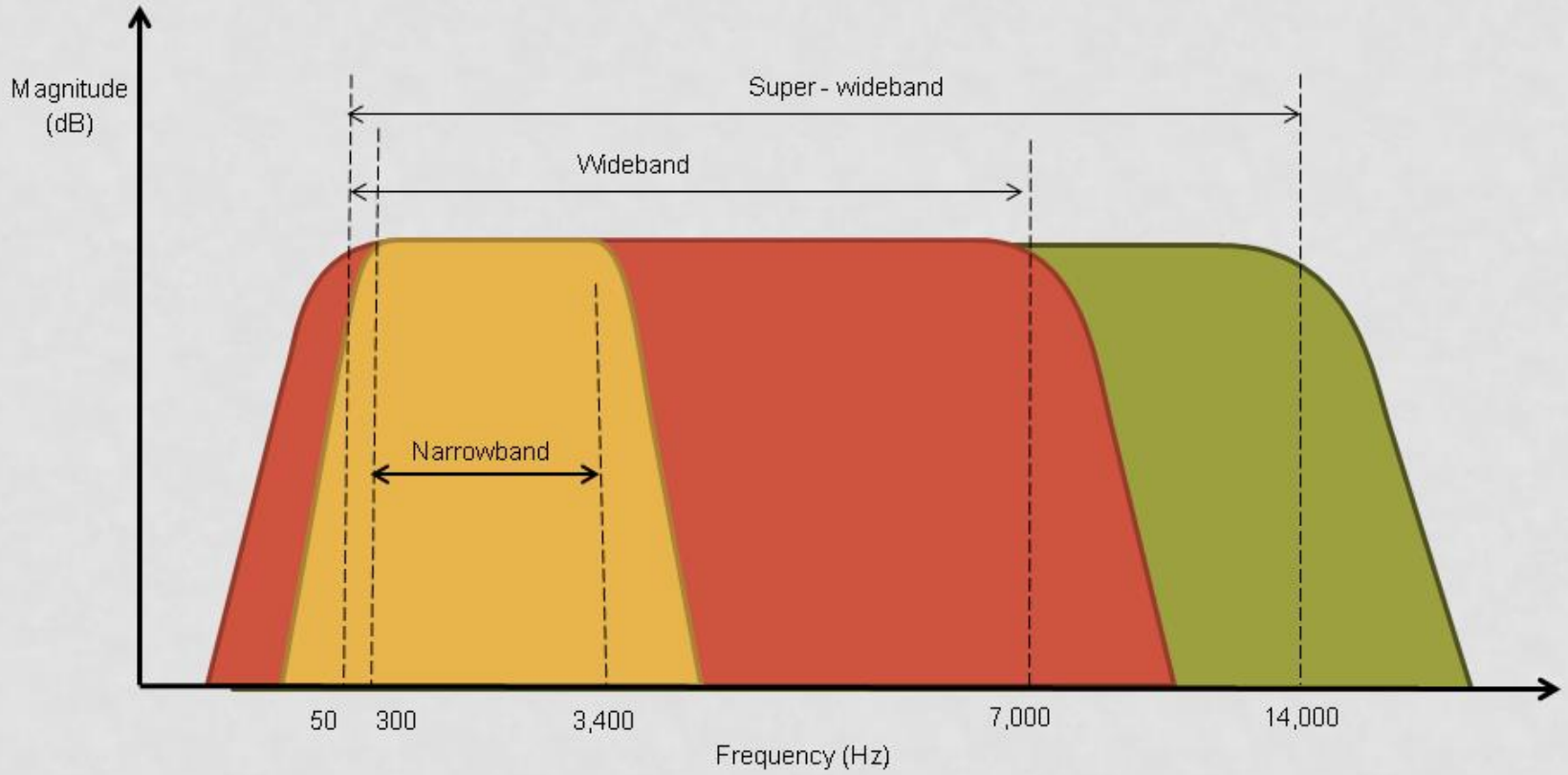
Quality and Usability Lab, Technische Universität Berlin, Germany

43. Jahrestagung für Akustik, DAGA.
March 6-9 Kiel, Germany

Intelligibility of transmitted speech



Intelligibility of transmitted speech



Intelligibility of transmitted speech

- Previous studies
  - Benefits of extended bandwidths shown
  - Intelligibility tests:
    - [Spiegel et al., 1990][Teng and Kubichek, 2006]: NB codecs only
    - [Jokinen et al., 2015]: 16 kHz speech only
    - [Fernández and Möller, 2015]: G.722 (WB) higher intell than AMR-NB
- Contributions of this work
  - Effects of 23 channel degradations: NB, WB, SWB
  - Comparing to objective measures
    - POLQA-intelligibility predictions
    - POLQA MOS

Outline

- Motivation
- Speech material
- Intelligibility test
- Subjective vs. objective intelligibility
- Conclusions

Speech material

- 4 speakers (2m, 2f)
- 8 vowel-consonant-vowel logatomes

"ama"

"aba“

"afa“

"ana“

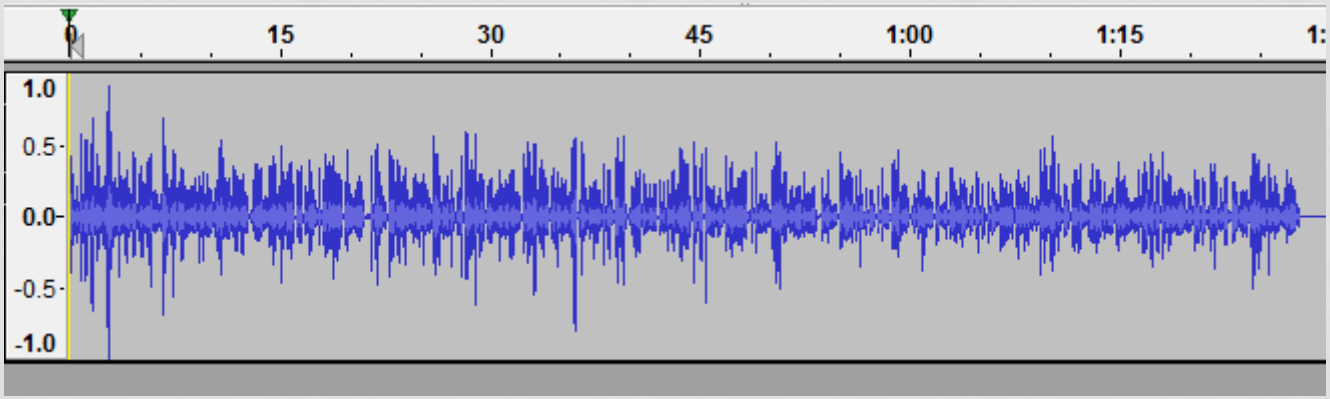
"apa“

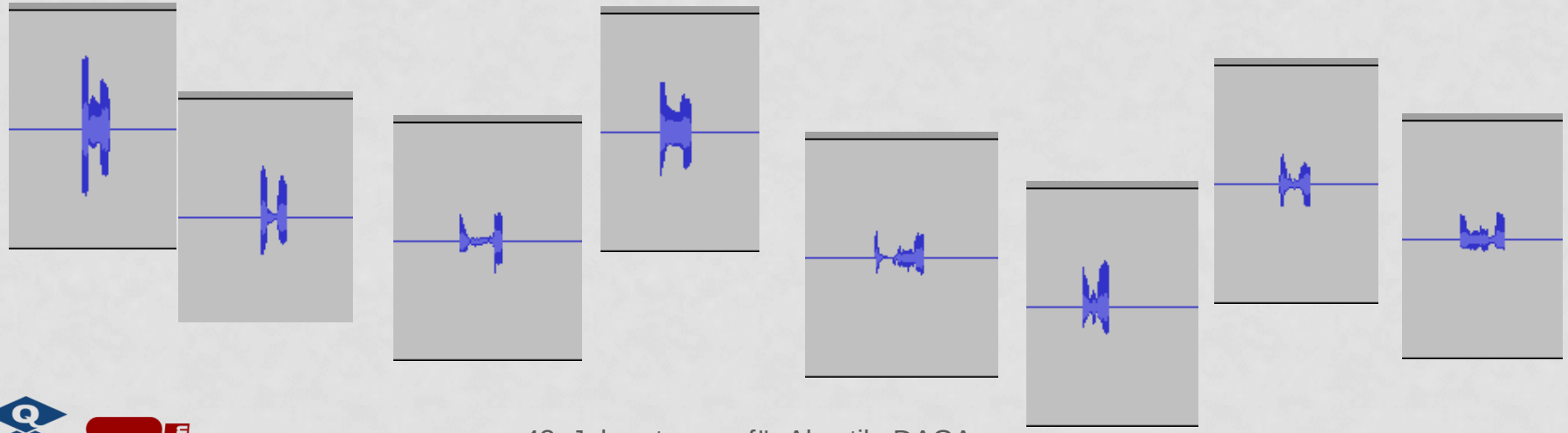
"asa“

"awa“

"ascha"

Speech material





Speech material

Bandwidth	Codec
Direct (fs=8 kHz)	No
NB	G.712 filter
NB	G711 A-law (64)
NB	G711 A-law (64) + IRS
NB	G723.1 (6.3)
NB	GSMEFR (12.2)
NB	AMRNB (4.75)
NB	AMRNB (12.2)
NB	Speex (2.15)
NB	Speex (11)
NB	Speex (24.6)
Direct (fs=16 kHz)	No
WB	P.341 filter
WB	G722 (64)
WB	AMRWB (12.65)
WB	AMRWB (23.05)
WB	Speex (3.95)
WB	Speex (23.8)
WB	Speex (42.2)
Direct (fs=32 kHz)	No
SWB	14KBP filter
SWB	G722.1C (24)
SWB	G722.1C (48)
SWB	EVS (24.4)
SWB	EVS (48)
SWB	Opus (160)
Direct (fs=48 kHz)	No

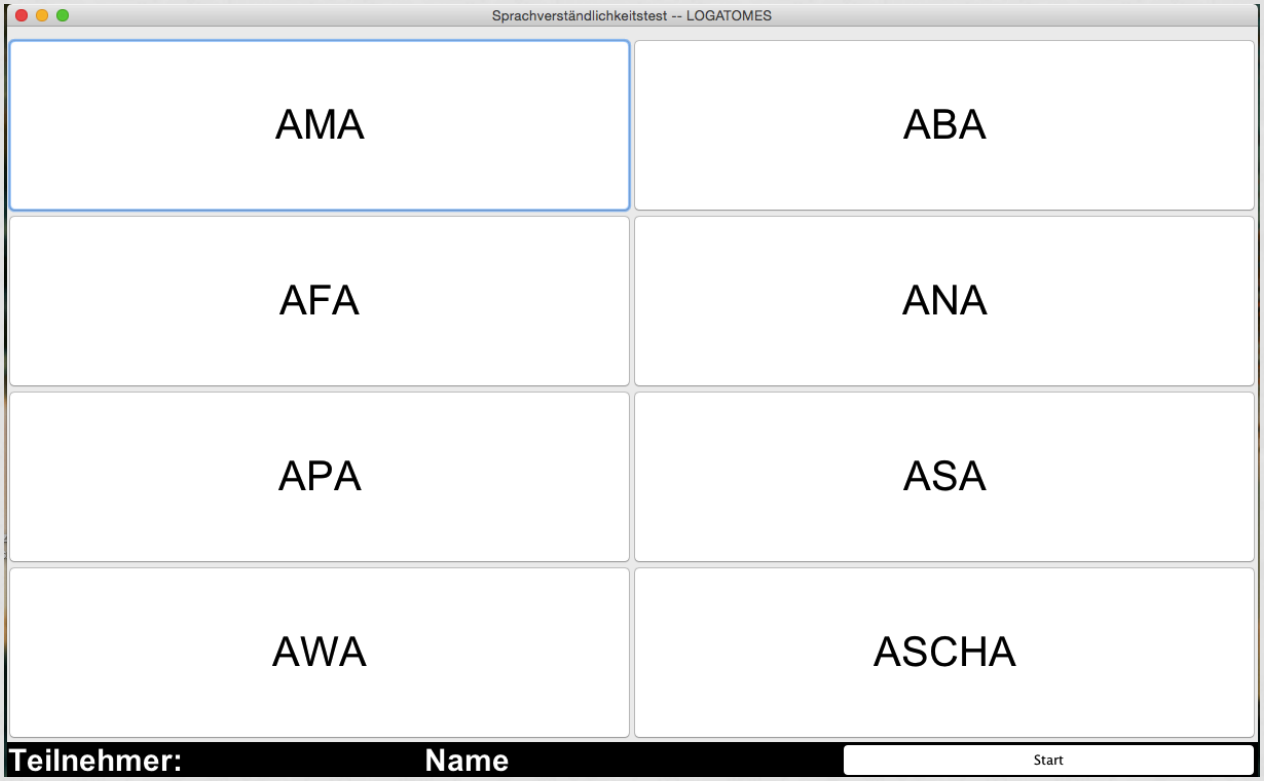
Outline

- Motivation
- Speech material
- Intelligibility test
- Subjective vs. objective intelligibility
- Conclusions

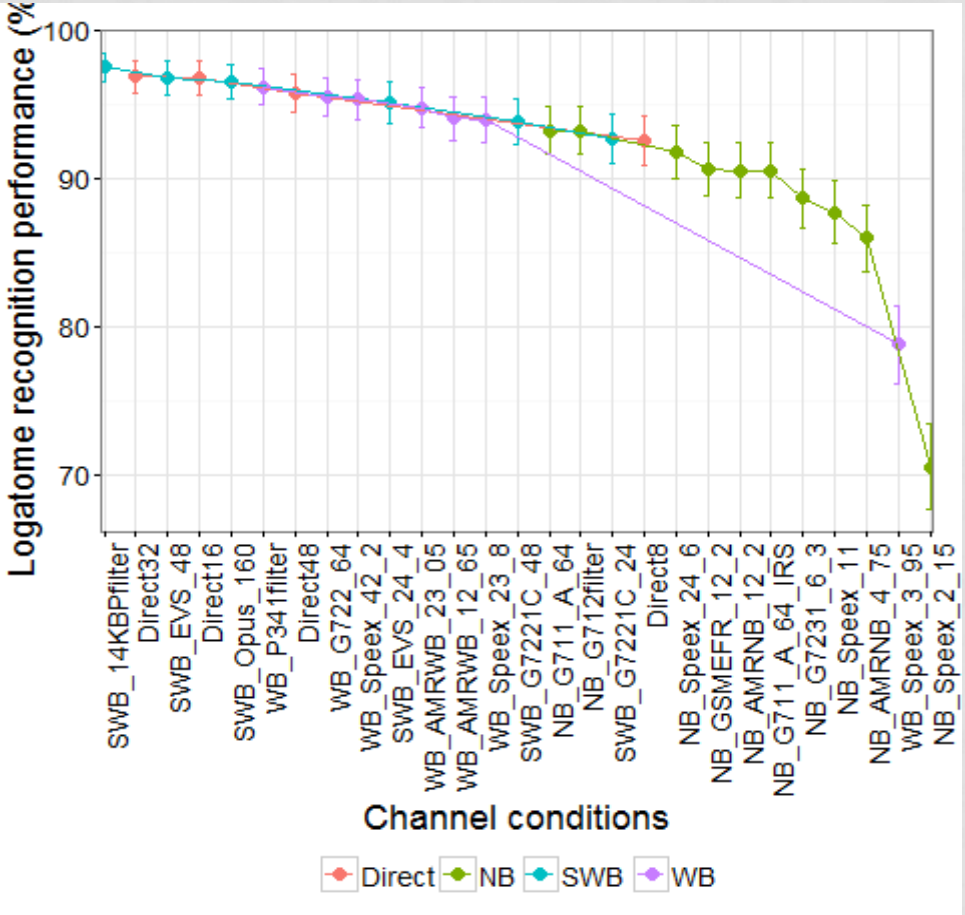
Intelligibility test

- 30 listeners (15m, 15f), normal-hearing, German
- 54m^2 acoustically-treated room
- Shure SHR240 headphones
  - freq.: 20-20,000 Hz
  - diotic listening
- 864 stimuli
- Duration: one hour

Intelligibility test



Intelligibility test



Intelligibility test

- Better performance with higher bandwidth and bitrate
- WB significant over NB?
  - only for WB codecs at high bitrate and NB codecs at low bitrate
  - Speex-WB@42.2, G.722@64, P.341 filter better than NB@ ≤11
- SWB significant over WB?
  - No, except for Speex-WB@3.95
- SWB significant over NB?
  - Not always
  - G.722.1C@24 and @48 only better than Speex-NB@2.15
  - 14 KPB filter same as G.711@64 and G.712 filter

Intelligibility test

- Logatome confusions
  - “aba”-”awa” -> prevail in SWB and in “Direct” with fs ≥ 16 kHz
  - “awa”-”aba” -> reduced from NB to WB
  - “afa”-”asa” -> reduced from NB to WB

Outline

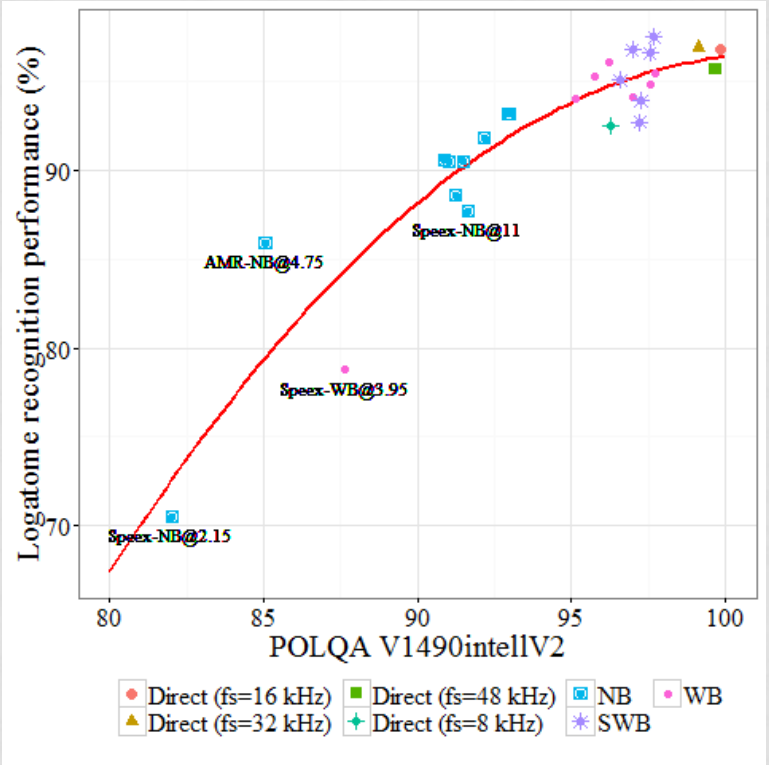
- Motivation
- Speech material
- Intelligibility test
- Subjective vs. objective intelligibility
- Conclusions

Objective intelligibility

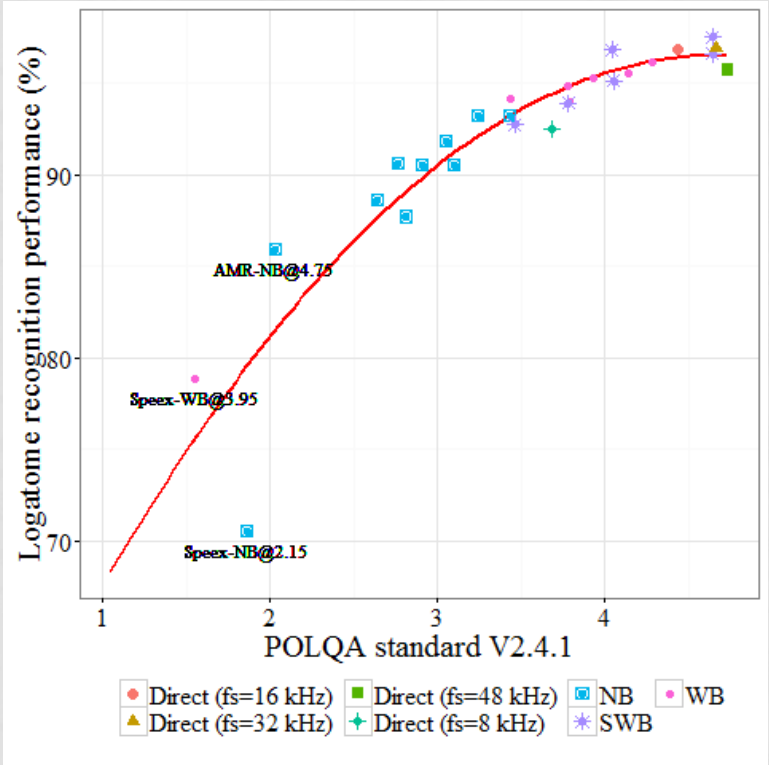
- POLQA-intelligibility model (V1490intellV2)
- POLQA standard (V2.4.1)

- 2nd-order curves can be fitted to subjective vs. objective intelligibility
  - slightly better fit with POLQA-intelligibility as predictor

2nd-order fits



R2 = 0.870, RMSE = 2.10



R2 = 0.858, RMSE = 2.20

Outline

- Motivation
- Speech material
- Intelligibility test
- Subjective vs. objective intelligibility
- Conclusions

Conclusions

- Intelligibility test
  - closed-response with 8 VCV logatomes
  - 23 channel degradations: NB, WB, SWB
    - + “Direct” with fs = 8, 16, 32, and 48 kHz

- Greater gain in intelligibility for the transition NB → WB
- No statistical differences between WB and SWB (except for Speex-WB@3.95)

- Strong quadratic correspondences between subjective and objective intelligibility

Thank you for your attention!

Laura Fernández Gallardo, Ph.D.
laura.fernandezgallardo@tu-berlin.de

Questions?

Laura Fernández Gallardo, Ph.D.
laura.fernandezgallardo@tu-berlin.de