

Speech Quality in Cellular Networks

–

Analysis and Modeling of Structures in Telephony Conversations

*2nd Workshop on Wideband Speech Quality in Terminals and Networks:
Assessment and Prediction*

22nd and 23rd June 2005 - Mainz, Germany

Jens Berger

SwissQual AG, Zuchwil, Switzerland

jens.berger@swissqual.com

- Limitations of today's QoS benchmarking systems
 - Do test calls model human conversations?
 - Statistical analysis of real live calls
 - Modeling conversations and open issues
-
- Focus on modelling conversations and not to the measures to be applied on

Quality Evaluation in Cellular Networks

- Today's Quality Benchmarking Systems are focussed on one-way transmission 'Listening Quality' by establishing Mobile-to-PSTN test calls
 - Networks will be evaluated with clean speech under ideal terminal conditions(*)
 - Length of speech utterances are variable from 3 ... 30s and more
 - 'Nobody' cares about background noises, signal delay and double talk capability
 - Test calls are concatenated test sequences only but not a 'model of a conversation'
- * In EFR equipped networks under good coverage conditions an LQ of >3.9 will be reached normally. Is this real?

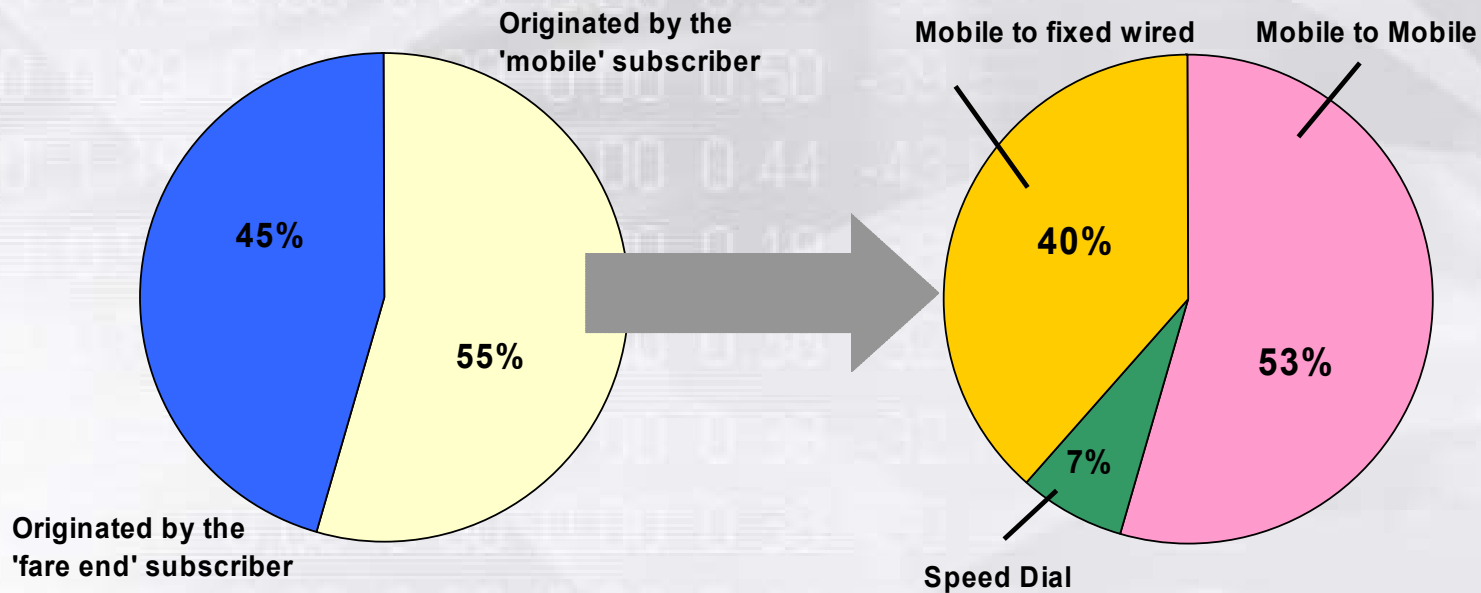
Quality Evaluation in Cellular Networks

- How is the structure of a typical mobile call?
 - Duration
 - Number and length of utterances
- What's about noise and its different handling?
- How is the double talk behaviour?

- 1st step: How to model a call / conversation?
- 2nd step: Which measures can be applied?

Statistical Analysis of Call Origination

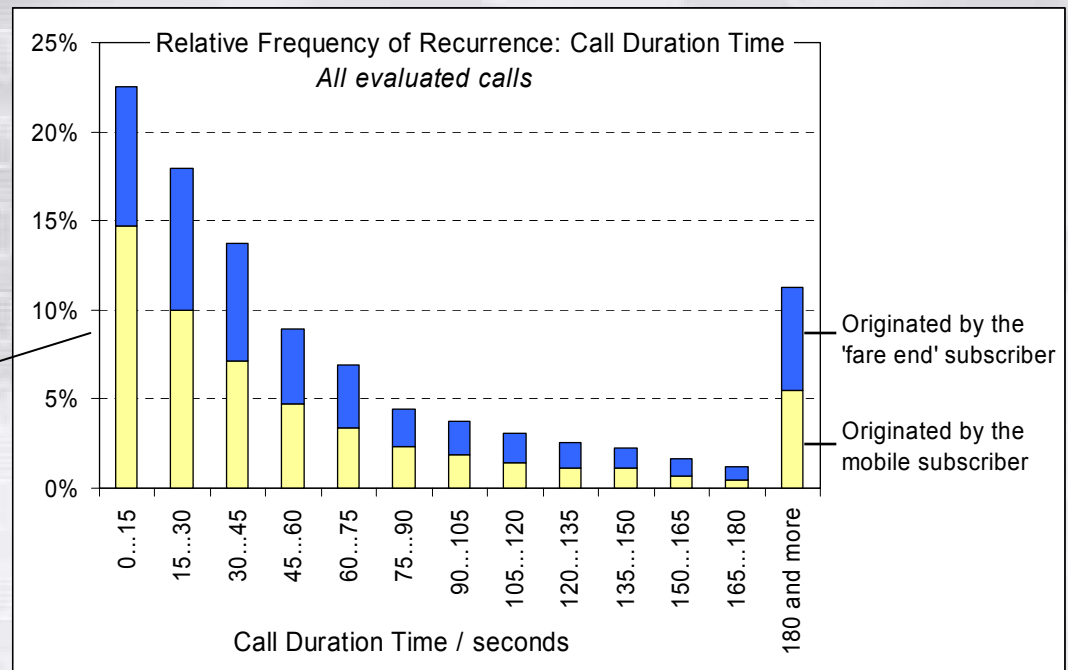
- Investigation was made in 2004
- Evaluation of approx. 5000 calls in European cellular networks
- More than 50% of the calls were originated by the mobile subscriber (MOC)
- More than 50% of them are Mobile to Mobile calls (incl. mobile voice boxes)



Statistical Analysis of Call Duration

- Call Duration Time (CDT) is highly dominated by short calls below 60s
- Average Call Duration Time **90 s**
- Median Call Duration Time **40 s**
- Shorter than 3 min **90%**

Short calls are dominated by MOC



Statistical Analysis of Call Duration

- Short calls are dominated by mobile originated calls (MOC)
- Mainly they contains 'normal conversation' between two human partners...
- ... but they also cover
 - 'speed dials' to the own mobile voice box and requests to the pre-paid account *Average CDT of speed dials: only 21s!*
 - calls to voice boxes 'at the far end'
 - unexpected fax and data modems at the far end
- Calls originated by the far end side contains only 'normal conversation' between two human partners
- Within 'speed dial' and unwanted voice boxes (>70%) shows NO speech activity in uplink!

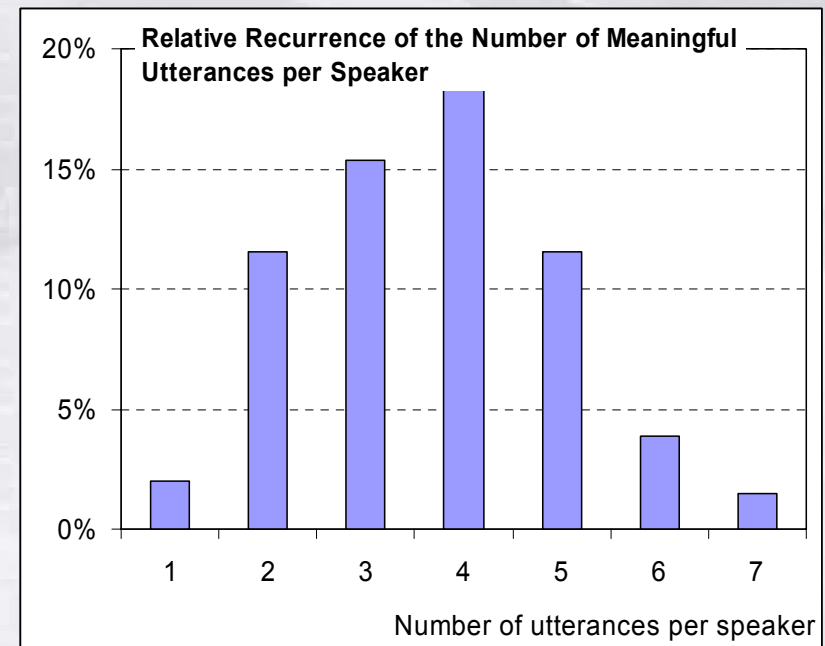
Target: Modelling of real conversations

- Exclusion of speed dials and unwanted voice boxes
- Average Call Duration Time increases to 100 s
- Median Call Duration Time increases to 45 s
- Way: Evaluation of typical human telephone conversation

- 1. Selection of 50 'typical' 35s ... 60s calls**
- 2. Number of interactions between partners**
- 3. Duration time of meaningful utterances**
- 4. Double talk situations**

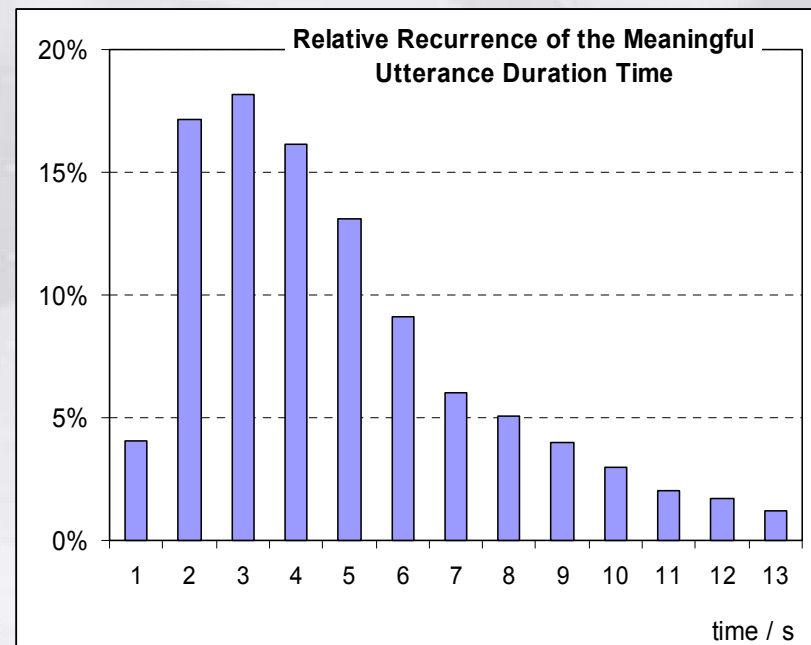
Interactions between partners (35 ... 60s calls)

- Initial 'Hello' and final 'Good Bye' phrases
 - *Mean duration time per partner is nearly exact 1.0s*
- Number of meaningful utterances (interactions) per partner (except initial / final phrases)
 - *Average: 4 utterances per partner*



Duration of meaningful utterances (35 ... 60s calls)

- Mean Duration Time of meaningful utterances is 5.4 s

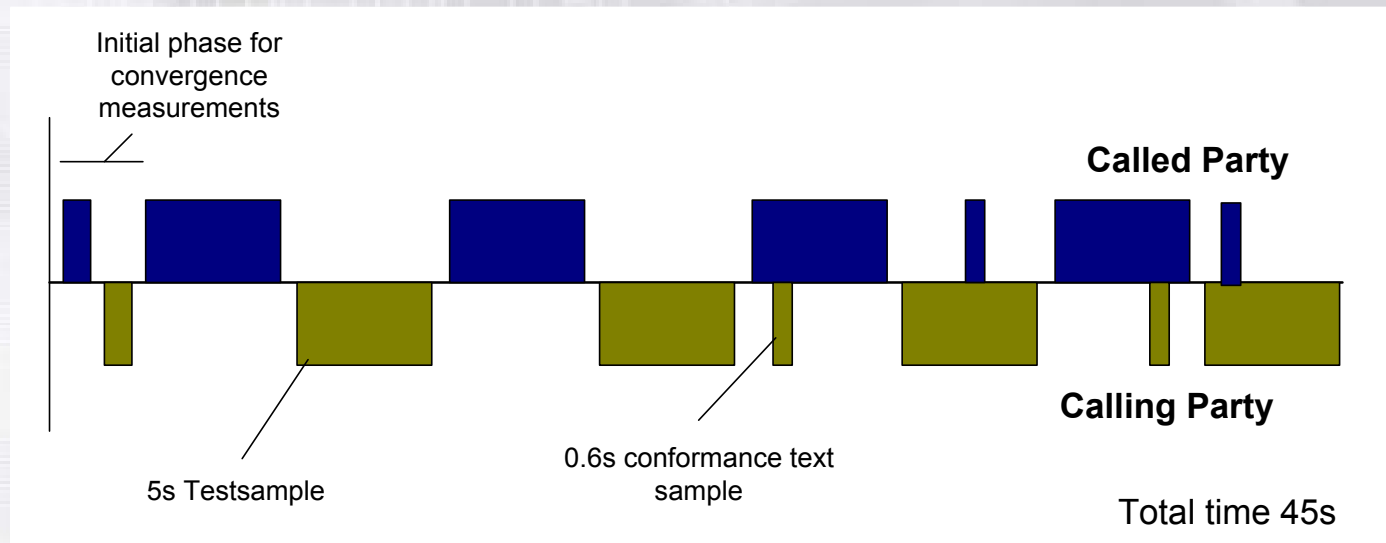


Occurrences of Double Talk

- Double talk is mostly occurred by short confirmation phrases or requests (mean duration time: 0.6s)
- Short confirmations or requests are more typical in case of longer utterances of the partner
- In case of an utterance duration between 4 ... 6s the probability of a short confirmation or request by the partner is approx. 60%
- The double talk ratio is between 5 ... 10 % typically

How to model conversations?

- Initial phrases for convergence measurements
- Alternating transmission of 5s speech samples
- No consideration of final phrases
- Double talk is occurring at different points in the test sample

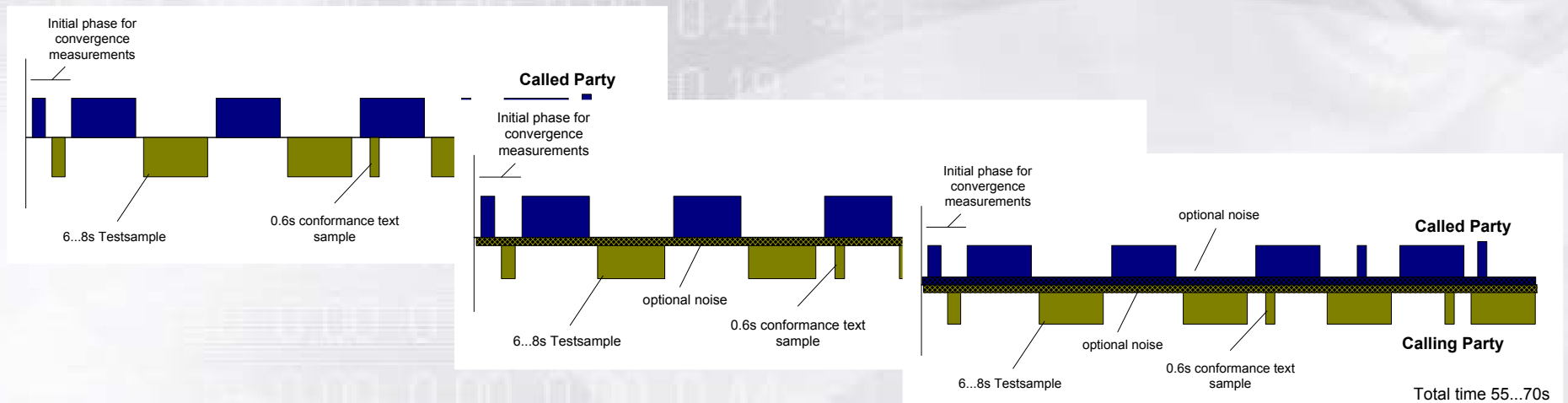


Open issues

- 5 s samples are in conflict with draft P.862.2
Compromise in Call Length?
- What's about background noise?
- How to deal with the received double talk?
- How to score 'Call Quality'? More than simple average.
- What's about other dimensions
Signal delay (e.g. in VideoCalls $\geq 300\text{ms}$, Mobile-to-PSTN $\approx 100\text{ms}$)
Echo Cancelling Performance

Compromise for narrow-band telephony?

- Usage of 6...8s samples as recommended in P.862.2
number of interactions remains four
- Optional insertion of background noise
 - No noise for ideal clean connection
 - Continuous noise at one side (typical for mobile-to-PSTN calls)
 - Continuous noise at both sides (typical for mobile-to-mobile calls)



Wideband Telephony (assumptions)

- Structure will remain unchanged
- Call duration might be increasing at all since people will become more familiar with cellular phones
- Increasing of background noise caused by 'un-restricted' usage of mobile phones

Video Telephony (assumptions)

- Call duration might be increasing since high ratio of business calls
- Increasing of background noise caused by hands-free application (camera is in front of the user)