Listening tests of Opus at Google

Fall 2011

Jan Skoglund

Google

Introduction

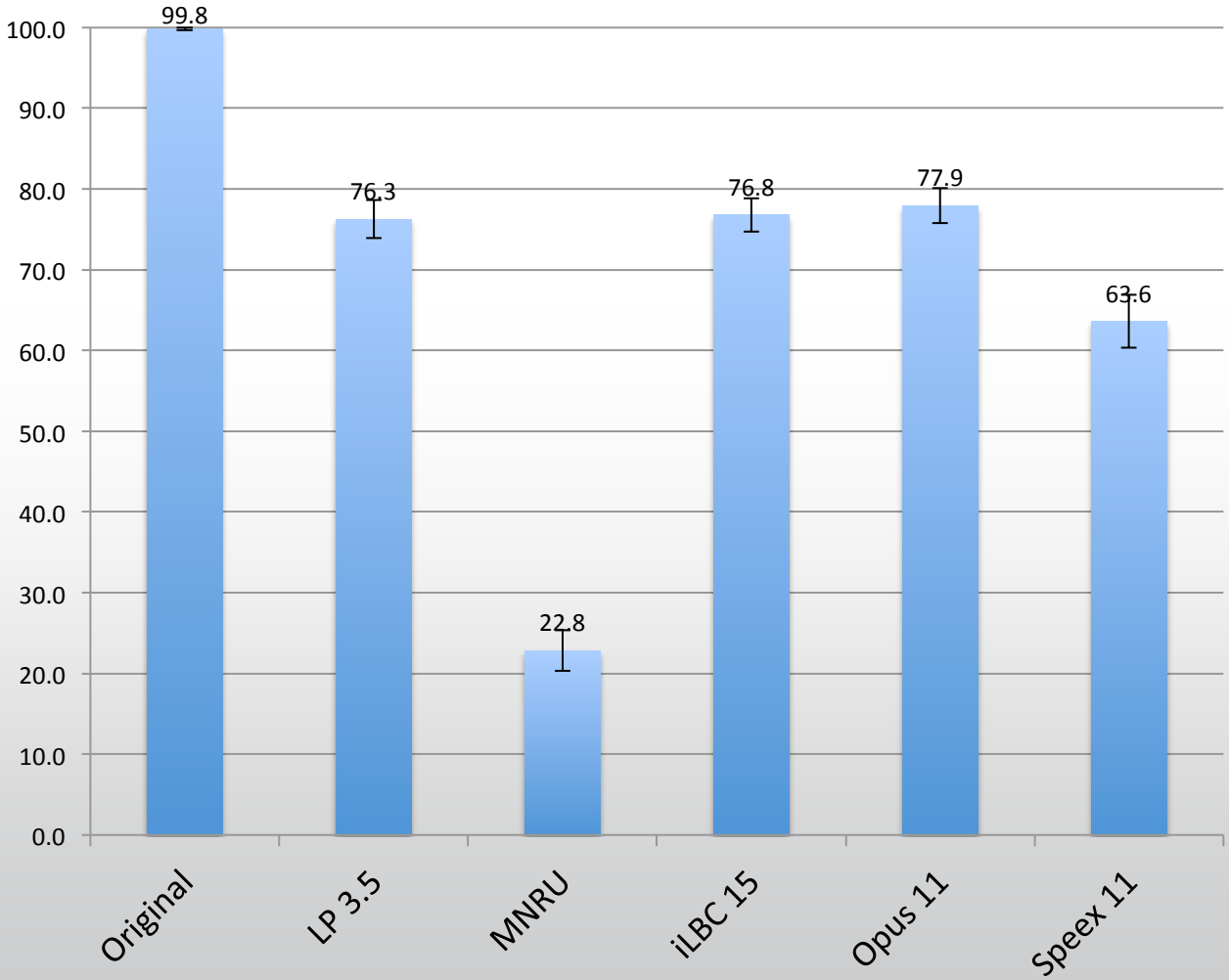
- Four MUSHRA-type tests performed in Aug-Sep 2011 at Google
- Two tests of coding Mandarin speech
- Two tests of transcoding English speech
- Both trained and untrained English-speaking listeners
- Only untrained Mandarin-speaking listeners
- All tests presented on Windows PC with headphones

Test 1 – Narrowband coding of Mandarin speech

- 4 different male and 4 different female speakers
  - 2 male and 2 female speakers from ITU-T P.501
  - 2 male and 2 female speakers recorded at Google
- Reference files sampled at 48 kHz in low background noise
- 2 anchors
  - Reference file lowpass-filtered at 3.5 kHz
  - Reference file resampled at 8 kHz, with MNRU at 15 dB SNR
- 21 listeners after post-screening
  - No listeners rejected
- 3 narrowband codecs
  - Opus NB at 11 kbps, variable bit rate
  - Speex NB at 11 kbps, variable bit rate
  - iLBC at 15.2 kbps, constant bit rate

Overall result – Narrowband Mandarin speech

- Opus at 11 kbps is comparable to iLBC at 15 kbps
- Opus at 11 kbps is better than Speex at 11 kbps

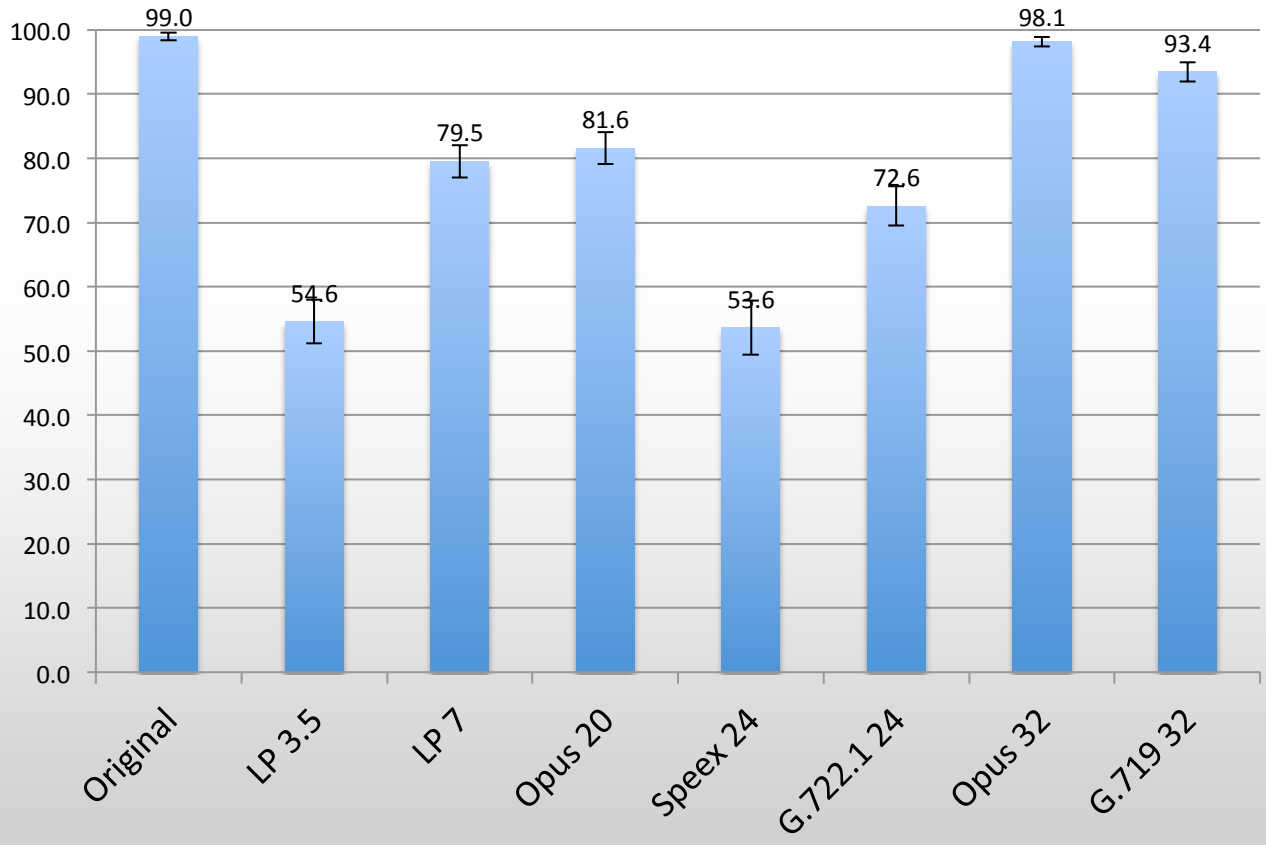


Test 2 – Wideband and fullband coding of Mandarin speech

- 4 different male and 4 different female speakers
  - 2 male and 2 female speakers from ITU-T P.501
  - 2 male and 2 female speakers recorded at Google
- Reference files sampled at 48 kHz in low background noise
- 2 anchors: lowpass-filtered at 3.5 kHz and 7.0 kHz
- 19 listeners after post-screening
  - Rejected 3 listeners having score correlation with the total average lower than 0.8
- 3 wideband codecs
  - Opus WB at 19.85 kbps, variable bit rate
  - Speex WB at 23.8 kbps, constant bit rate
  - G.722.1 at 24 kbps, constant bit rate
- 2 fullband codecs
  - Opus FB at 32 kbps, constant bit rate
  - G.719 at 32 kbps, constant bit rate

Overall result – Wideband and fullband Mandarin speech

- Opus at 32 kbps is better than G.719 at 32 kbps
- Opus at 20 kbps is better than Speex and G.722.1 at 24 kbps



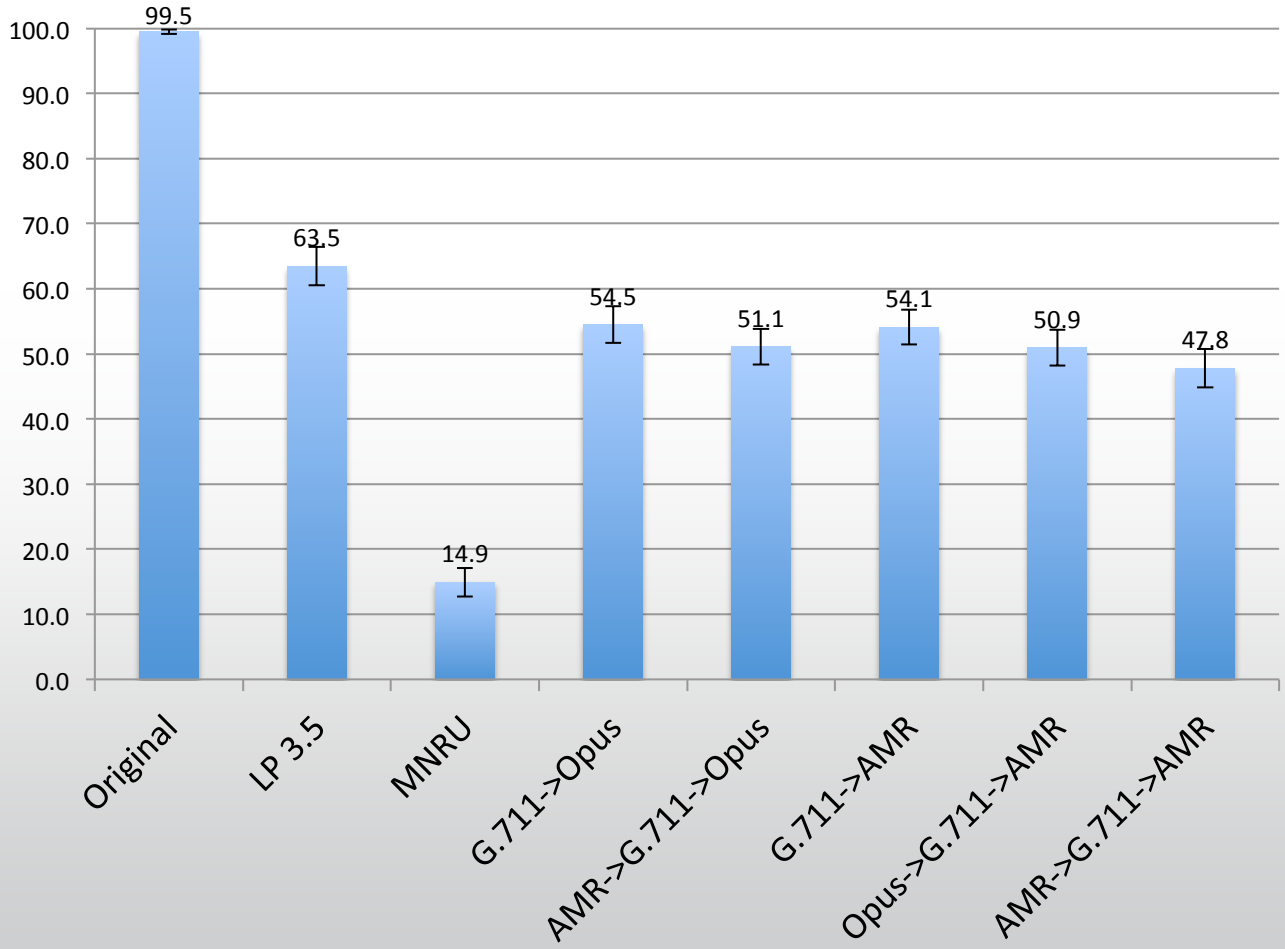
Google

Test 3 – Narrowband transcoding of English speech

- 4 different male and 4 different female speakers
  - 2 male and 2 female speakers from ITU-T P.501
  - 2 male and 2 female speakers from McGill database
- Reference files sampled at 48 kHz in low background noise
- 2 anchors
  - Reference file lowpass-filtered at 3.5 kHz
  - Reference file resampled at 8 kHz, with MNRU at 15 dB SNR
- 19 listeners after post-screening
  - No listeners rejected
- 5 narrowband transcoding scenarios
  - G.711 at 64 kbps -> Opus NB at 12.2 kbps, variable bit rate
  - G.711 at 64 kbps -> AMR NB at 12.2 kbps, constant bit rate
  - AMR NB at 12.2 kbps -> G.711 at 64 kbps -> Opus NB at 12.2 kbps
  - Opus NB at 12.2 kbps -> G.711 at 64 kbps -> AMR NB at 12.2 kbps
  - AMR NB at 12.2 kbps -> G.711 at 64 kbps -> AMR NB at 12.2 kbps

Overall result – Narrowband transcoding

- Opus NB pre-coded with G.711 is comparable to AMR NB pre-coded with G.711
- Opus NB transcoded to AMR NB via G.711 is better than AMR NB tandem-coded via G.711



Test 4 – Wideband transcoding of English speech

- 4 different male and 4 different female speakers
  - 2 male and 2 female speakers from ITU-T P.501
  - 2 male and 2 female speakers from McGill database
- Reference files sampled at 48 kHz in low background noise
- 2 anchors: lowpass-filtered at 3.5 kHz and 7 kHz
- 18 listeners after post-screening
  - No listeners rejected
- 4 wideband single coding and transcoding scenarios
  - Opus WB at 19.85 kbps, variable bit rate
  - AMR WB at 19.85 kbps, constant bit rate
  - AMR WB at 19.85 kbps -> Opus WB at 19.85 kbps
  - Opus WB at 19.85 kbps -> AMR WB at 19.85 kbps

Overall result – Wideband transcoding

- Single-coded Opus WB is better than single-coded AMR WB
- Single-coded AMR WB is slightly better than transcoding AMR WB -> Opus WB and Opus WB -> AMR WB (statistically significant)