ITU-T Workshop on "From Speech to Audio: bandwidth extension, binaural perception" Lannion, France, 10-12 September 2008

Quality impact of diotic versus monaural listening on processed speech

> Arnault Nagle R&D Engineer France Telecom

Our Goal

In VoIP audio communication, rendering is usually proposed over either handsets or headphones

two distinct kinds of listening condition:

- Monaural
- Diotic I same content on the two ears

The goal of our study:

 To determine whether listening over the monaural or diotic condition has an impact on the perceived quality of speech processed by VoIP coders

- 1- Protocol Experiment
- 2- Narrowband Results
- 3- Wideband Results
- 4- Conclusion

- 1- Protocol Experiment
- 2- Narrowband Results
- 3- Wideband Results
- 4- Conclusion

Simulation of the transmission between two terminals with one encoding-decoding

Stimuli

- Two sentence-pairs for two male and two female talkers
- Quiet background french speech files:
 - duration = 8 seconds,
 - Fs = 16 kHz,
 - audio bandwidth = [0 8 kHz]

Two tests: one in NB and one in WB

Narrowband conditions







Narrowband & Wideband conditions

Conditions	Coder	% of Packet Loss
1	No coding	0
2		0
3	G.711	3
4	With PLC	6
5		0
6	G.729.1 at 8 kbits/s	3
7	With PLC	6
8		0
9	G.729.1 at 12 kbits/s	3
10	With PLC	6
11		0
12	AMR at 4.75 kbits/s	3
13	With PLC	6
14		0
15	AMR at 12.2 kbits/s	3
16	With PLC	6

Conditions	Coder	% of Packet Loss
1	No coding	0
2		0
3	G.722	3
4	With PLC	6
5		0
6	G.729.1 at 16 kbits/s	3
7	With PLC	6
8		0
9	G.729.1 at 32 kbits/s	3
10	With PLC	6
11		0
12	AMR-WB at 12.65 kbits/s	3
13	With PLC	6
14		0
15	AMR-WB at 23.85 kbits/s	3
16	With PLC	6

Absolute Category Pating test	Excellent	5	
Absolute Category Rating test	Good	4	
Two sessions	Fair	3	
• One with monaural listening	Poor	2	
- One with monadrar listening	Bad	1	
One with diotic listening	MOS Se	cale	
32 subjects by test: 4 groups of 8 l	ister	ners	3
Listening level			
In order to keep the same loudness bet	tween	the	Э
two listenings:			
79 dB SPL for monaural listening			

- 69 dB SPL per channel for diotic listening*
- * G. Reynolds, S. Stevens, Binaural Summation of Loudness, JASA 32 (1960)

* M. Botte, G. Canévet, L. Demany, C. Sorin, Psychoacoustique et Perception auditive (1989) Lannion, France, 10-12 September 2008

MOS

Quality

In one of the two sessions:

- All test stimuli presented monaurally and in randomized order, one order for each of the four listener groups
- Listening done over a Sennheiser HD 25 headset with flat response in the audio-bandwidth: 50Hz-7kHz
- One ear left open

In the other session:

The same stimuli presented to the subjects, in the same randomized order as the other session and over the same headphone, but diotically

In each of the four listener groups:

- Half of the listeners did the monaural session first and next the diotic session,
- The other half did the opposite

I - Protocol Experiment

- 2- Narrowband Results
- 3- Wideband Results
- 4- Conclusion

Effects of the different factors

Factor	Degrees of freedom	F-ratio	Significance
Listening mode	1	157.94	Significant
Coder	5	509.24	Significant
Packet Loss	2	1564.14	Significant
Speaker	3	40.39	Significant
Sample	1	12.99	Significant

Coder ranking without packet loss



Ranking	Monaural listening		Diotic listening	
1	Direct		Direct	
2	G.729.1 12		AMR 12.2	
3	AMR 12.2		G.711	
4	G.729.1 8		G.729.1 12	
5	G.711		G.729.1 8	
6	AMR 4.75		AMR 4.75	

Not the same rankings and not the same equivalences

Lannion, France, 10-12 September 2008

Coder equivalences

Coder ranking with 3% packet loss



Ranking	Monaural listening		Diotic listening	
1	G.729.1 12		G.711	
2	G.711		G.729.1 12	
3	G.729.1 8		AMR 12.2	
4	AMR 12.2		G.729.1 8	
5	AMR 4.75		AMR 4.75	

Not the same rankings and not the same equivalences

Coder ranking with 6% packet loss



- I Protocol Experiment
- 2- Narrowband Results
- 3- Wideband Results
- 4- Conclusion

Effects of the different factors

Factor	Degrees of freedom	F-ratio	Significance
Listening mode	1	94.82	Significant
Coder	5	69.75	Significant
Packet Loss	2	3244.08	Significant
Speaker	3	206.36	Significant
Sample	1	41.08	Significant

Coder ranking without packet loss



Not the same rankings and not the same equivalences

Coder ranking with 3% packet loss



Ranking	Monaural listening		Diotic listen	ing	
1	G.722		G.722		
2	G.729.1 32		G.729.1 32		
3	AMR-WB 23.85		AMR-WB 23.8	35	
4	G.729.1 16	η	G.729.1 16		Ι
5	AMR-WB 12.65		AMR-WB 12.6	5	T

Same ranking but not the same equivalences

Coder ranking with 6% packet loss



Same ranking but not the same equivalences

- I Protocol Experiment
- 2- Narrowband Results
- 3- Wideband Results
- 4- Conclusion

Conclusion

- Listening over the monaural or diotic condition has an impact on the perceived quality of speech processed by VoIP coders
 - For diotic listening, quality is judged more severely when speech is degraded: packet loss or low bit rate
 - Diotic listening seems to help subjects to better discriminate degradations
 - At the opposite, in comparison with monaural listening, diotic listening highlights the benefits of high quality coders

Conclusion

The difference of listening level between the monaural and diotic conditions leads to hide noise defects

=> The potential weight of the listening level for quality evaluation

	Monaural Listening	Diotic Listening (for each ear)
Speech alone	~ 80 dB SPL ~ 79.5 dBA	~ 70 dB SPL ~ 69 dBA
G.711 noise	~ 50 dB SPL ~ 37.5 dBA	~ 50 dB SPL ~ 31 dBA
Ambient noise	~ 49 dB SPL, 28.5 dBA	

Conclusion

In function of the coder and introduced degradations (packet loss or bit rate), the impact can be more or less strong resulting in shifts in coder ranking between the two listening modes

These results suggest that audio coders should be chosen carefully for use cases

Thank you!