ETSI SMG11 **Tdoc SMG11/NS 419R/99**

Kyoto, Japan 6-10 December 1999

Title: **ETSI/AMR NS Selection Experiment 1 - COMSAT Results**

Source*: **COMSAT Laboratories**

Summary

This document presents a summary for ETSI Adaptive Multi Rate (AMR) Noise Selection (NS) Selection Phase Experiment 1 conducted by COMSAT in the Spanish language.

1. Introduction

COMSAT Laboratories in accordance with the AMR NS Selection Phase Experiment 1 performed a listening laboratory evaluation in Spanish. This experiment was designed to identify whether the noise suppression candidate algorithms for the AMR codec operating at 12.2 and 5.9 kbit/s would introduce any artifacts during the initial convergence time for noisy and quiet background. The test design is defined in Section 7 of the AMR NS Selection Subjective Test Plan [1], which used speech material processed under selected experimental conditions from Experiments 2, 4, and 5.

COMSAT performed Experiment 1 using a subset of the Spanish speech material available in the NTT Speech Database. Six subjects fluent in the Spanish language performed as “expert” subjects in the test. The raw data collected was used to derive gender-wise and combined-gender pair comparison scores and standard deviation statistics for each of the three questions posed to the experts.

2. Source Material

Six single-sentence stimuli were selected for two male and two female talkers from the NTT Speech Database. The source material was provided to COMSAT, the designated Host Laboratory, which was responsible for all pre- and post-processing according to the test plan in [2].

3. Experimental Design

The test design followed the specification in the AMR NS Selection Test Plan, as summarized in Tables 1 and 2.

4. Processed Material

The host laboratory provided a CDROM with 960 processed speech files, which corresponds to the processing of 5 files per talker through 96 reference conditions and 96 test conditions. Speech processing details are given in [3]. Test and reference files were concatenated for each condition following the presentation order given in the test plan before presentation to the subjects.

Contact: Simão F. Campos-Neto
COMSAT Labs
22300 COMSAT Drive
Clarksburg, MD 20871
USA

Tel: +1-301-428-4516
Fax: +1-301-428-9287
E-mail: simao.campos@comsat.com

Table 1:
Factors and Conditions for Experiment 1: a subset of conditions in Experiments 2, 4, and 5

Main Codec Conditions	**#**	**Notes**
Noise Suppresser Candidates	6	
Codec	1	AMR
Codec Modes (FR/HR)	HR	5.9 kbit/s rate
	FR	12.2 kbit/s rate
BER	0	Clear channel, no transmission errors
Input level	1	Nominal (-26dBov)
Acoustic Background Noise	0	• Experiment 2: None • Experiments 4A and 5A: Car (6 dB), street (9 dB), and babble (9 dB) • Experiments 4B and 5B: Car (12 dB), street (15 dB), and babble (15 dB)
Transcodings	2	Self tandem condition from Experiment 2
Input Characteristic	1	GSM Filtered
Codec references	**#**	**Notes**
Test vocoders	1	AMR with NS
Reference vocoder	2	AMR@12.2, AMR@5.9
Other references	**#**	**Notes**
None		
Common Conditions	**#**	**Notes**
GSM Channel	0	NO channel model
Number of talkers	4	2 male + 2 female
Number of speech samples	52	6/talker in split-plot mode
Sentences/sample	1	Single sentence stimuli with 2-second silence or noise preamble
Listening Level	1	-18dBPa (76dB SPL) at ERP, binaural
Listeners	6	Expert listeners
Listening instrument		AKG K240DF Studio Monitor Headphone
Randomizations	6	1 expert at a time
Rating Scale	1	Comparison of processed and reference stimuli according to special instructions (Section 7.7 of the test plan).

5. Listening Sessions

COMSAT performed Experiment 1 according to the Test Plan.

5.1 Presentation Sequence Material

COMSAT used the randomization sequences specified in the AMR NS Selection Test Plan for Experiment 1.

5.2 Listeners

The subjective assessment was performed using Six “expert” subjects fluent in the Spanish language. The listener selection criteria were compliant with the AMR NS Selection Test Plan. The pre-test listener orientation used by COMSAT conformed to that specified in the Test Plan.

5.3 Audio Presentation

The processed speech material was presented to the subjects in listening stations contained within an acoustically conditioned sound room meeting the requirements for an NC 20 acoustic facility. Presentation was made binaurally using an AKG Studio Monitor headset, driven by a distribution amplifier set to deliver an active level of -18 dBPa (76 dB SPL) in each ear, using a B&K 4153 Artificial Ear with circumaural headphone adapter, 4134 Microphone element and 2610 Measurement Amplifier.

Table 2:

Experiment 1: Artifacts in initial convergence

Cond.No Exp. 1	Original Cond.No.	Noise	SNR (dB)	Reference	Processed	Tandems
01	4A/15	Car	6	AMR@5.9	AMR@5.9/NS1	1
02	4A/16	Car	6	AMR@5.9	AMR@5.9/NS2	1
03	4A/17	Car	6	AMR@5.9	AMR@5.9/NS3	1
04	4A/18	Car	6	AMR@5.9	AMR@5.9/NS4	1
05	4A/19	Car	6	AMR@5.9	AMR@5.9/NS5	1
06	4A/20	Car	6	AMR@5.9	AMR@5.9/NS6	1
07	4A/21	Street	9	AMR@5.9	AMR@5.9/NS1	1
08	4A/22	Street	9	AMR@5.9	AMR@5.9/NS2	1
09	4A/23	Street	9	AMR@5.9	AMR@5.9/NS3	1
10	4A/24	Street	9	AMR@5.9	AMR@5.9/NS4	1
11	4A/25	Street	9	AMR@5.9	AMR@5.9/NS5	1
12	4A/26	Street	9	AMR@5.9	AMR@5.9/NS6	1
13	4A/27	Babble	9	AMR@5.9	AMR@5.9/NS1	1
14	4A/28	Babble	9	AMR@5.9	AMR@5.9/NS2	1
15	4A/29	Babble	9	AMR@5.9	AMR@5.9/NS3	1
16	4A/30	Babble	9	AMR@5.9	AMR@5.9/NS4	1
17	4A/31	Babble	9	AMR@5.9	AMR@5.9/NS5	1
18	4A/32	Babble	9	AMR@5.9	AMR@5.9/NS6	1
19	4B/15	Car	12	AMR@5.9	AMR@5.9/NS1	1
20	4B/16	Car	12	AMR@5.9	AMR@5.9/NS2	1
21	4B/17	Car	12	AMR@5.9	AMR@5.9/NS3	1
22	4B/18	Car	12	AMR@5.9	AMR@5.9/NS4	1
23	4B/19	Car	12	AMR@5.9	AMR@5.9/NS5	1
24	4B/20	Car	12	AMR@5.9	AMR@5.9/NS6	1
25	4B/21	Street	15	AMR@5.9	AMR@5.9/NS1	1
26	4B/22	Street	15	AMR@5.9	AMR@5.9/NS2	1
27	4B/23	Street	15	AMR@5.9	AMR@5.9/NS3	1
28	4B/24	Street	15	AMR@5.9	AMR@5.9/NS4	1
29	4B/25	Street	15	AMR@5.9	AMR@5.9/NS5	1
30	4B/26	Street	15	AMR@5.9	AMR@5.9/NS6	1
31	4B/27	Babble	15	AMR@5.9	AMR@5.9/NS1	1
32	4B/28	Babble	15	AMR@5.9	AMR@5.9/NS2	1
33	4B/29	Babble	15	AMR@5.9	AMR@5.9/NS3	1
34	4B/30	Babble	15	AMR@5.9	AMR@5.9/NS4	1
35	4B/31	Babble	15	AMR@5.9	AMR@5.9/NS5	1
36	4B/32	Babble	15	AMR@5.9	AMR@5.9/NS6	1
37	5A/15	Car	12	AMR@12.2	AMR@12.2/NS1	1
38	5A/16	Car	12	AMR@12.2	AMR@12.2/NS2	1
39	5A/17	Car	12	AMR@12.2	AMR@12.2/NS3	1
40	5A/18	Car	12	AMR@12.2	AMR@12.2/NS4	1
41	5A/19	Car	12	AMR@12.2	AMR@12.2/NS5	1
42	5A/20	Car	12	AMR@12.2	AMR@12.2/NS6	1
43	5A/21	Street	15	AMR@12.2	AMR@12.2/NS1	1
44	5A/22	Street	15	AMR@12.2	AMR@12.2/NS2	1
45	5A/23	Street	15	AMR@12.2	AMR@12.2/NS3	1
46	5A/24	Street	15	AMR@12.2	AMR@12.2/NS4	1
47	5A/25	Street	15	AMR@12.2	AMR@12.2/NS5	1
48	5A/26	Street	15	AMR@12.2	AMR@12.2/NS6	1
49	5A/27	Babble	15	AMR@12.2	AMR@12.2/NS1	1
50	5A/28	Babble	15	AMR@12.2	AMR@12.2/NS2	1
51	5A/29	Babble	15	AMR@12.2	AMR@12.2/NS3	1
52	5A/30	Babble	15	AMR@12.2	AMR@12.2/NS4	1
53	5A/31	Babble	15	AMR@12.2	AMR@12.2/NS5	1
54	5A/32	Babble	15	AMR@12.2	AMR@12.2/NS6	1
55	5B/15	Car	12	AMR@12.2	AMR@12.2/NS1	1
56	5B/16	Car	12	AMR@12.2	AMR@12.2/NS2	1
57	5B/17	Car	12	AMR@12.2	AMR@12.2/NS3	1
58	5B/18	Car	12	AMR@12.2	AMR@12.2/NS4	1
59	5B/19	Car	12	AMR@12.2	AMR@12.2/NS5	1
60	5B/20	Car	12	AMR@12.2	AMR@12.2/NS6	1
61	5B/21	Street	15	AMR@12.2	AMR@12.2/NS1	1
62	5B/22	Street	15	AMR@12.2	AMR@12.2/NS2	1
63	5B/23	Street	15	AMR@12.2	AMR@12.2/NS3	1
64	5B/24	Street	15	AMR@12.2	AMR@12.2/NS4	1
65	5B/25	Street	15	AMR@12.2	AMR@12.2/NS5	1
66	5B/26	Street	15	AMR@12.2	AMR@12.2/NS6	1
67	5B/27	Babble	15	AMR@12.2	AMR@12.2/NS1	1
68	5B/28	Babble	15	AMR@12.2	AMR@12.2/NS2	1
69	5B/29	Babble	15	AMR@12.2	AMR@12.2/NS3	1
70	5B/30	Babble	15	AMR@12.2	AMR@12.2/NS4	1
71	5B/31	Babble	15	AMR@12.2	AMR@12.2/NS5	1
72	5B/32	Babble	15	AMR@12.2	AMR@12.2/NS6	1
73	2A/05	—	—	AMR@12.2	AMR@12.2/NS1	1
74	2A/06	—	—	AMR@12.2	AMR@12.2/NS2	1
75	2A/07	—	—	AMR@12.2	AMR@12.2/NS3	1
76	2A/08	—	—	AMR@12.2	AMR@12.2/NS4	1
77	2A/09	—	—	AMR@12.2	AMR@12.2/NS5	1
78	2A/10	—	—	AMR@12.2	AMR@12.2/NS6	1
79	2A/17	—	—	AMR@5.9	AMR@5.9/NS1	1
80	2A/18	—	—	AMR@5.9	AMR@5.9/NS2	1
81	2A/19	—	—	AMR@5.9	AMR@5.9/NS3	1
82	2A/20	—	—	AMR@5.9	AMR@5.9/NS4	1
83	2A/21	—	—	AMR@5.9	AMR@5.9/NS5	1
84	2A/22	—	—	AMR@5.9	AMR@5.9/NS6	1
85	2A/11	—	—	AMR@12.2	AMR@12.2/NS1	2
86	2A/12	—	—	AMR@12.2	AMR@12.2/NS2	2
87	2A/13	—	—	AMR@12.2	AMR@12.2/NS3	2
88	2A/14	—	—	AMR@12.2	AMR@12.2/NS4	2
89	2A/15	—	—	AMR@12.2	AMR@12.2/NS5	2
90	2A/16	—	—	AMR@12.2	AMR@12.2/NS6	2
91	2A/23	—	—	AMR@5.9	AMR@5.9/NS1	2
92	2A/24	—	—	AMR@5.9	AMR@5.9/NS2	2
93	2A/25	—	—	AMR@5.9	AMR@5.9/NS3	2
94	2A/26	—	—	AMR@5.9	AMR@5.9/NS4	2
95	2A/27	—	—	AMR@5.9	AMR@5.9/NS5	2
96	2A/28	—	—	AMR@5.9	AMR@5.9/NS6	2

The processed speech files were stored within the main facility computer and presented to the listeners under program control as 16 kHz samples. The signals were passed through a 16-bit D/A converter coupled to the input of the distribution amplifier. The reference and test signals were available to the subjects via an A/B switch box, where the P position corresponded to the Processed signal and the R position corresponded to the Reference signal. The subjects were free to switch between the P and R position as many times as they wanted, before casting a vote for each of the three conditions. On average, each of the subjects took 2 hours to perform the test.

5.4 Scoring

The test plan specified that subjects would respond to three questions about each sample. These questions were designed to target the convergence behavior of the noise suppressors and to identify the presence of artifacts in the sample. Subjects were instructed to compare the test and reference samples to determine whether an audible artifact was present in the original sample or added by the noise suppressor. Where artifacts were identified, the subjects are asked to provide a short description.

The questions were:

A. Assess the convergence of the noise suppression.
1. Not audible
2. Audible, but fast
3. Audible and relatively slow

B. Assess the presence of artifacts during the silence at the beginning of the sample:
1. No audible artifacts
2. Audible
3. Clearly audible

C. Assess the presence of artifacts at the onset of the speech and during the speech:
1. No audible artifacts
2. Audible
3. Clearly audible

The responses to these assessment questions yielded quantitative estimates of the number and severity of artifacts heard, as well as a description of those artifacts (see Annex A).

The listener responses were registered on paper for each of the questions and for each of the 96 test conditions after they listened to the processed and reference material to their satisfaction. The scores were later entered into a database for subsequent analysis and presentation.

6. Statistical Analysis

The statistics to be reported for this expert experiment are the average of scores for each of the three questions asked to the subjects. However, since the population sample is too small, standard deviation values are not useful.

The overall figures for each noise suppression candidate are presented in Table 3. In Table 3(a) the percentage of 2 and 3 votes for the NS candidates is presented for each noise type, while Table 3(b) presents the grand mean of the votes for each question for the NS candidates and the overall percentage of votes 2 and 3.

Table 4 presents the basic statistical analysis data produced by COMSAT for AMR NS Selection Experiment 1, and is sorted in test condition order. Each test condition received a total of 6 votes. In the table, *Cond* represents the test condition number in Experiment 1, *SNR/Tandem* is the circuit impairment, *Yconv*, *Ysil*, and *Ysp* are the mean scores for the convergence question A, for the artifacts in silence question B, and for the artifacts in speech question C, respectively. *Sdconv*, *SDsil*, and *SDsp* are the standard deviation for Questions A, B, and C, respectively. The Table also presents the histogram of votes for Questions A (*Convergence*), B (*Silence*), and C (*Speech*)

Table 5 is an alternative presentation of the data in Table 4, whereby the data is grouped by Noise Suppressor and sorted by increasing *Yconv* and *Ysil*.

Table 3

(a) Percentage of votes 2 and 3 for COMSAT Experiment 1 organized by noise type

	Babble			Car			Street			Quiet		
NS	**Yconv**	**Ysil**	**Ysp**	**Yconv**	**Ysil**	**Ysp**	**Yconv**	**Ysil**	**Ysp**	**Yconv**	**Ysil**	**Ysp**
1	4%	17%	33%	46%	58%	54%	21%	25%	33%	0%	0%	33%
2	13%	13%	25%	100%	96%	67%	67%	67%	21%	29%	29%	21%
3	8%	17%	8%	63%	71%	54%	33%	38%	21%	13%	13%	38%
4	8%	8%	13%	100%	100%	71%	50%	50%	38%	4%	4%	29%
5	21%	38%	21%	100%	100%	63%	54%	54%	25%	13%	13%	33%
6	17%	17%	29%	67%	79%	79%	25%	33%	38%	8%	8%	21%

(b) Grand means and percentage of 2&3 votes for COMSAT Experiment 1 for all noise backgrounds confounded

	Grand-means			Percentage 2&3 votes		
NS	**Yconv**	**Ysil**	**Ysp**	**Yconv**	**Ysil**	**Ysp**
1	1.3	1.3	1.5	18%	25%	39%
2	1.6	1.6	1.4	52%	51%	33%
3	1.4	1.4	1.4	29%	34%	30%
4	1.5	1.5	1.5	41%	41%	38%
5	1.8	1.6	1.4	47%	51%	35%
6	1.4	1.4	1.5	29%	34%	42%

It can be seen from Tables 4 and 5 that in general the car noise conditions had the largest number of perceived effects. On the other end, babble and quiet background were the conditions that elicited the least comments. Similarly, the grand-means for each NS algorithm, computed across the 16 impairments, indicate no major difference across the algorithms. However, the proportion of votes 2 and 3 showed a large variance for the different NS candidates, which slightly contradicts the indications of the grand means. When one looks into the scores by noise type, car noise represents the conditions with most noticeable convergence effects.

7. Conclusion

COMSAT performed AMR NS Selection Experiment 1 for the Spanish language in compliance with the test plan. Six expert subjects were used. This experiment was designed to determine any initial convergence artifacts introduced by the noise suppression candidate algorithms for the AMR codec operating at 12.2 and 5.9 kbit/s. It was observed that in general the car noise conditions had the largest number of perceived effects. On the other end, babble and quiet background conditions elicited the least comments. The grand-means for each NS algorithm, computed across the 16 test impairments, indicate no major difference across the algorithms. However, the proportion of votes 2 and 3 showed a large variance for the different NS candidates, which slightly contradicts the indications of the grand means.

References

[1] SMG11SQ, *"Test Plan for the AMR Specification for the AMR-NS Selection Phase"*; Tdoc SMG11 288/99.

[2] SMG11SQ, *"Processing Functions for the GSM AMR Noise Suppression Selection Tests"*; Tdoc SMG11 281R/99

[3] COMSAT Laboratories, *"Host Laboratory Processing for ETSI/AMR Noise Suppression Selection Tests"*, Tdoc SMG11 417/99

Table 4 (Part 1/3):
Experiment 1: Artifacts in initial convergence

											Histograms								
											Convergence			Silence			Speech		
Cond.	**Noise/Tandem**	**Processed**	**Reference**	**N**	**Yconv**	**Sdconv**	**Ysil**	**Sdsil**	**Ysp**	**SDsp**	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
01	Car@6dB/1x	AMR@5.9/NS1	AMR@5.9	6	1.5	0.8	1.3	0.5	1.5	0.5	4	1	1	4	2	0	3	3	0
02	Car@6dB/1x	AMR@5.9/NS2	AMR@5.9	6	2.2	0.4	2.5	0.8	1.8	0.8	0	5	1	1	1	4	2	3	1
03	Car@6dB/1x	AMR@5.9/NS3	AMR@5.9	6	1.3	0.5	1.3	0.5	2.0	0.9	4	2	0	4	2	0	2	2	2
04	Car@6dB/1x	AMR@5.9/NS4	AMR@5.9	6	2.2	0.4	2.5	0.5	2.0	0.9	0	5	1	0	3	3	2	2	2
05	Car@6dB/1x	AMR@5.9/NS5	AMR@5.9	6	2.7	0.5	2.5	0.5	2.0	0.6	0	2	4	0	3	3	1	4	1
06	Car@6dB/1x	AMR@5.9/NS6	AMR@5.9	6	1.7	0.5	2.0	0.9	2.2	0.8	2	4	0	2	2	2	1	3	2
07	Street@9dB/1x	AMR@5.9/NS1	AMR@5.9	6	1.5	0.5	1.8	0.8	1.7	0.8	3	3	0	2	3	1	3	2	1
08	Street@9dB/1x	AMR@5.9/NS2	AMR@5.9	6	1.7	0.8	1.7	0.8	1.5	0.5	3	2	1	3	2	1	3	3	0
09	Street@9dB/1x	AMR@5.9/NS3	AMR@5.9	6	2.0	0.9	1.8	0.8	1.5	0.8	2	2	2	2	3	1	4	1	1
10	Street@9dB/1x	AMR@5.9/NS4	AMR@5.9	6	1.5	0.5	1.5	0.5	1.3	0.5	3	3	0	3	3	0	4	2	0
11	Street@9dB/1x	AMR@5.9/NS5	AMR@5.9	6	1.7	0.8	1.5	0.5	1.2	0.4	3	2	1	3	3	0	5	1	0
12	Street@9dB/1x	AMR@5.9/NS6	AMR@5.9	6	1.2	0.4	1.2	0.4	1.5	0.5	5	1	0	5	1	0	3	3	0
13	Babble@9dB/1x	AMR@5.9/NS1	AMR@5.9	6	1.0	0.0	1.2	0.4	1.2	0.4	6	0	0	5	1	0	5	1	0
14	Babble@9dB/1x	AMR@5.9/NS2	AMR@5.9	6	1.2	0.4	1.2	0.4	1.3	0.5	5	1	0	5	1	0	4	2	0
15	Babble@9dB/1x	AMR@5.9/NS3	AMR@5.9	6	1.0	0.0	1.2	0.4	1.2	0.4	6	0	0	5	1	0	5	1	0
16	Babble@9dB/1x	AMR@5.9/NS4	AMR@5.9	6	1.2	0.4	1.2	0.4	1.2	0.4	5	1	0	5	1	0	5	1	0
17	Babble@9dB/1x	AMR@5.9/NS5	AMR@5.9	6	1.3	0.8	1.3	0.5	1.2	0.4	5	0	1	4	2	0	5	1	0
18	Babble@9dB/1x	AMR@5.9/NS6	AMR@5.9	6	1.0	0.0	1.0	0.0	1.2	0.4	6	0	0	6	0	0	5	1	0
19	Car@12dB/1x	AMR@5.9/NS1	AMR@5.9	6	2.0	1.1	1.8	0.8	1.8	0.8	3	0	3	2	3	1	2	3	1
20	Car@12dB/1x	AMR@5.9/NS2	AMR@5.9	6	2.0	0.0	2.3	0.5	1.7	0.5	0	6	0	0	4	2	2	4	0
21	Car@12dB/1x	AMR@5.9/NS3	AMR@5.9	6	2.0	0.6	2.2	0.4	1.5	0.8	1	4	1	0	5	1	4	1	1
22	Car@12dB/1x	AMR@5.9/NS4	AMR@5.9	6	2.5	0.5	2.2	0.4	1.5	0.8	0	3	3	0	5	1	4	1	1
23	Car@12dB/1x	AMR@5.9/NS5	AMR@5.9	6	2.7	0.5	2.3	0.5	1.7	0.5	0	2	4	0	4	2	2	4	0
24	Car@12dB/1x	AMR@5.9/NS6	AMR@5.9	6	2.0	0.9	1.8	0.4	2.0	0.6	2	2	2	1	5	0	1	4	1
25	Street@15dB/1x	AMR@5.9/NS1	AMR@5.9	6	1.0	0.0	1.0	0.0	1.2	0.4	6	0	0	6	0	0	5	1	0
26	Street@15dB/1x	AMR@5.9/NS2	AMR@5.9	6	1.7	0.5	1.8	0.8	1.0	0.0	2	4	0	2	3	1	6	0	0
27	Street@15dB/1x	AMR@5.9/NS3	AMR@5.9	6	1.2	0.4	1.2	0.4	1.0	0.0	5	1	0	5	1	0	6	0	0
28	Street@15dB/1x	AMR@5.9/NS4	AMR@5.9	6	1.3	0.5	1.3	0.5	1.2	0.4	4	2	0	4	2	0	5	1	0
29	Street@15dB/1x	AMR@5.9/NS5	AMR@5.9	6	1.5	0.5	1.5	0.5	1.5	0.5	3	3	0	3	3	0	3	3	0
30	Street@15dB/1x	AMR@5.9/NS6	AMR@5.9	6	1.0	0.0	1.2	0.4	1.3	0.5	6	0	0	5	1	0	4	2	0
31	Babble@15dB/1x	AMR@5.9/NS1	AMR@5.9	6	1.0	0.0	1.0	0.0	1.7	0.8	6	0	0	6	0	0	3	2	1
32	Babble@15dB/1x	AMR@5.9/NS2	AMR@5.9	6	1.0	0.0	1.0	0.0	1.3	0.5	6	0	0	6	0	0	4	2	0
33	Babble@15dB/1x	AMR@5.9/NS3	AMR@5.9	6	1.2	0.4	1.2	0.4	1.2	0.4	5	1	0	5	1	0	5	1	0
34	Babble@15dB/1x	AMR@5.9/NS4	AMR@5.9	6	1.0	0.0	1.0	0.0	1.0	0.0	6	0	0	6	0	0	6	0	0
35	Babble@15dB/1x	AMR@5.9/NS5	AMR@5.9	6	1.3	0.8	1.3	0.5	1.3	0.8	5	0	1	4	2	0	5	0	1
36	Babble@15dB/1x	AMR@5.9/NS6	AMR@5.9	6	1.3	0.5	1.3	0.5	1.5	0.5	4	2	0	4	2	0	3	3	0

Table 4 (Part 2/3):
Experiment 1: Artifacts in initial convergence

											Histograms								
											Convergence			Silence			Speech		
Cond.	**Noise/Tandem**	**Processed**	**Reference**	**N**	**Yconv**	**Sdconv**	**Ysil**	**Sdsil**	**Ysp**	**SDsp**	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
37	Car@6dB/1x	AMR@12.2/NS1	AMR@12.2	6	2.0	1.1	2.0	0.9	1.7	0.8	3	0	3	2	2	2	3	2	1
38	Car@6dB/1x	AMR@12.2/NS2	AMR@12.2	6	2.0	0.0	2.3	0.5	2.0	0.6	0	6	0	0	4	2	1	4	1
39	Car@6dB/1x	AMR@12.2/NS3	AMR@12.2	6	1.8	1.0	1.8	0.8	1.8	0.8	3	1	2	2	3	1	2	3	1
40	Car@6dB/1x	AMR@12.2/NS4	AMR@12.2	6	2.5	0.5	2.3	0.5	2.3	0.5	0	3	3	0	4	2	0	4	2
41	Car@6dB/1x	AMR@12.2/NS5	AMR@12.2	6	2.8	0.4	2.3	0.5	1.8	0.8	0	1	5	0	4	2	2	3	1
42	Car@6dB/1x	AMR@12.2/NS6	AMR@12.2	6	2.0	0.7	1.8	0.4	2.5	0.5	1	3	1	1	5	0	0	3	3
43	Street@9dB/1x	AMR@12.2/NS1	AMR@12.2	6	1.5	0.8	1.3	0.5	1.5	0.5	4	1	1	4	2	0	3	3	0
44	Street@9dB/1x	AMR@12.2/NS2	AMR@12.2	6	1.8	0.4	2.0	0.6	1.2	0.4	1	5	0	1	4	1	5	1	0
45	Street@9dB/1x	AMR@12.2/NS3	AMR@12.2	6	1.2	0.4	1.3	0.5	1.5	0.8	5	1	0	4	2	0	4	1	1
46	Street@9dB/1x	AMR@12.2/NS4	AMR@12.2	6	1.7	0.5	1.8	0.8	1.7	0.5	2	4	0	2	3	1	2	4	0
47	Street@9dB/1x	AMR@12.2/NS5	AMR@12.2	6	2.0	0.9	1.7	0.5	1.0	0.0	2	2	2	2	4	0	6	0	0
48	Street@9dB/1x	AMR@12.2/NS6	AMR@12.2	6	1.8	1.0	1.7	0.8	1.5	0.8	3	1	2	3	2	1	4	1	1
49	Babble@9dB/1x	AMR@12.2/NS1	AMR@12.2	6	1.2	0.4	1.3	0.5	1.5	0.5	5	1	0	4	2	0	3	3	0
50	Babble@9dB/1x	AMR@12.2/NS2	AMR@12.2	6	1.2	0.4	1.2	0.4	1.3	0.5	5	1	0	5	1	0	4	2	0
51	Babble@9dB/1x	AMR@12.2/NS3	AMR@12.2	6	1.2	0.4	1.2	0.4	1.0	0.0	5	1	0	5	1	0	6	0	0
52	Babble@9dB/1x	AMR@12.2/NS4	AMR@12.2	6	1.2	0.4	1.2	0.4	1.0	0.0	5	1	0	5	1	0	6	0	0
53	Babble@9dB/1x	AMR@12.2/NS5	AMR@12.2	6	1.3	0.8	1.3	0.5	1.2	0.4	5	0	1	4	2	0	5	1	0
54	Babble@9dB/1x	AMR@12.2/NS6	AMR@12.2	6	1.0	0.0	1.0	0.0	1.3	0.5	6	0	0	6	0	0	4	2	0
55	Car@12dB/1x	AMR@12.2/NS1	AMR@12.2	6	1.8	1.0	1.8	0.8	1.7	0.8	3	1	2	2	3	1	3	2	1
56	Car@12dB/1x	AMR@12.2/NS2	AMR@12.2	6	2.2	0.4	2.2	0.4	1.8	1.0	0	5	1	0	5	1	3	1	2
57	Car@12dB/1x	AMR@12.2/NS3	AMR@12.2	6	2.0	0.6	2.0	0.6	1.7	0.8	1	4	1	1	4	1	3	2	1
58	Car@12dB/1x	AMR@12.2/NS4	AMR@12.2	6	2.2	0.4	2.3	0.5	2.2	0.8	0	5	1	0	4	2	1	3	2
59	Car@12dB/1x	AMR@12.2/NS5	AMR@12.2	6	2.3	0.5	2.2	0.4	1.5	0.8	0	4	2	0	5	1	4	1	1
60	Car@12dB/1x	AMR@12.2/NS6	AMR@12.2	6	1.8	0.8	1.8	0.4	1.8	1.0	2	3	1	1	5	0	3	1	2
61	Street@15dB/1x	AMR@12.2/NS1	AMR@12.2	6	1.0	0.0	1.0	0.0	1.2	0.4	6	0	0	6	0	0	5	1	0
62	Street@15dB/1x	AMR@12.2/NS2	AMR@12.2	6	1.7	0.5	1.7	0.5	1.2	0.4	2	4	0	2	4	0	5	1	0
63	Street@15dB/1x	AMR@12.2/NS3	AMR@12.2	6	1.3	0.5	1.3	0.5	1.2	0.4	4	2	0	4	2	0	5	1	0
64	Street@15dB/1x	AMR@12.2/NS4	AMR@12.2	6	1.8	1.0	1.8	1.0	1.5	0.8	3	1	2	3	1	2	4	1	1
65	Street@15dB/1x	AMR@12.2/NS5	AMR@12.2	6	1.8	1.0	1.5	0.5	1.3	0.5	3	1	2	3	3	0	4	2	0
66	Street@15dB/1x	AMR@12.2/NS6	AMR@12.2	6	1.3	0.5	1.5	0.5	1.3	0.5	4	2	0	3	3	0	4	2	0
67	Babble@15dB/1x	AMR@12.2/NS1	AMR@12.2	6	1.0	0.0	1.2	0.4	1.2	0.4	6	0	0	5	1	0	5	1	0
68	Babble@15dB/1x	AMR@12.2/NS2	AMR@12.2	6	1.2	0.4	1.2	0.4	1.0	0.0	5	1	0	5	1	0	6	0	0
69	Babble@15dB/1x	AMR@12.2/NS3	AMR@12.2	6	1.0	0.0	1.2	0.4	1.0	0.0	6	0	0	5	1	0	6	0	0
70	Babble@15dB/1x	AMR@12.2/NS4	AMR@12.2	6	1.0	0.0	1.0	0.0	1.3	0.5	6	0	0	6	0	0	4	2	0
71	Babble@15dB/1x	AMR@12.2/NS5	AMR@12.2	6	1.7	1.0	1.5	0.5	1.3	0.5	4	0	2	3	3	0	4	2	0
72	Babble@15dB/1x	AMR@12.2/NS6	AMR@12.2	6	1.3	0.5	1.3	0.5	1.2	0.4	4	2	0	4	2	0	5	1	0

Table 4 (Part 3/3):
Experiment 1: Artifacts in initial convergence

											Histograms								
											Convergence			Silence			Speech		
Cond.	**Noise/Tandem**	**Processed**	**Reference**	**N**	**Yconv**	**Sdconv**	**Ysil**	**Sdsil**	**Ysp**	**SDsp**	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
73	Quiet/1x	AMR@12.2/NS1	AMR@12.2	6	1.0	0.0	1.0	0.0	1.2	0.4	6	0	0	6	0	0	5	1	0
74	Quiet/1x	AMR@12.2/NS2	AMR@12.2	6	1.7	0.8	1.5	0.5	1.2	0.4	3	2	1	3	3	0	5	1	0
75	Quiet/1x	AMR@12.2/NS3	AMR@12.2	6	1.0	0.0	1.0	0.0	1.2	0.4	6	0	0	6	0	0	5	1	0
76	Quiet/1x	AMR@12.2/NS4	AMR@12.2	6	1.3	0.8	1.2	0.4	1.3	0.5	5	0	1	5	1	0	4	2	0
77	Quiet/1x	AMR@12.2/NS5	AMR@12.2	6	1.3	0.8	1.2	0.4	1.3	0.5	5	0	1	5	1	0	4	2	0
78	Quiet/1x	AMR@12.2/NS6	AMR@12.2	6	1.3	0.8	1.2	0.4	1.3	0.5	5	0	1	5	1	0	4	2	0
79	Quiet/1x	AMR@5.9/NS1	AMR@5.9	6	1.0	0.0	1.0	0.0	1.5	0.8	6	0	0	6	0	0	4	1	1
80	Quiet/1x	AMR@5.9/NS2	AMR@5.9	6	1.2	0.4	1.2	0.4	1.3	0.5	5	1	0	5	1	0	4	2	0
81	Quiet/1x	AMR@5.9/NS3	AMR@5.9	6	1.3	0.8	1.2	0.4	1.3	0.5	5	0	1	5	1	0	4	2	0
82	Quiet/1x	AMR@5.9/NS4	AMR@5.9	6	1.0	0.0	1.0	0.0	1.3	0.5	6	0	0	6	0	0	4	2	0
83	Quiet/1x	AMR@5.9/NS5	AMR@5.9	6	1.3	0.8	1.2	0.4	1.5	0.5	5	0	1	5	1	0	3	3	0
84	Quiet/1x	AMR@5.9/NS6	AMR@5.9	6	1.0	0.0	1.0	0.0	1.2	0.4	6	0	0	6	0	0	5	1	0
85	Quiet/2x	AMR@12.2/NS1	AMR@12.2	6	1.0	0.0	1.0	0.0	1.7	0.8	6	0	0	6	0	0	3	2	1
86	Quiet/2x	AMR@12.2/NS2	AMR@12.2	6	1.3	0.5	1.3	0.5	1.2	0.4	4	2	0	4	2	0	5	1	0
87	Quiet/2x	AMR@12.2/NS3	AMR@12.2	6	1.3	0.8	1.2	0.4	1.3	0.5	5	0	1	5	1	0	4	2	0
88	Quiet/2x	AMR@12.2/NS4	AMR@12.2	6	1.0	0.0	1.0	0.0	1.2	0.4	6	0	0	6	0	0	5	1	0
89	Quiet/2x	AMR@12.2/NS5	AMR@12.2	6	1.3	0.8	1.2	0.4	1.2	0.4	5	0	1	5	1	0	5	1	0
90	Quiet/2x	AMR@12.2/NS6	AMR@12.2	6	1.2	0.4	1.2	0.4	1.2	0.4	5	1	0	5	1	0	5	1	0
91	Quiet/2x	AMR@5.9/NS1	AMR@5.9	6	1.0	0.0	1.0	0.0	1.5	0.8	6	0	0	6	0	0	4	1	1
92	Quiet/2x	AMR@5.9/NS2	AMR@5.9	6	1.2	0.4	1.2	0.4	1.2	0.4	5	1	0	5	1	0	5	1	0
93	Quiet/2x	AMR@5.9/NS3	AMR@5.9	6	1.3	0.8	1.2	0.4	1.7	0.5	5	0	1	5	1	0	2	4	0
94	Quiet/2x	AMR@5.9/NS4	AMR@5.9	6	1.0	0.0	1.0	0.0	1.3	0.5	6	0	0	6	0	0	4	2	0
95	Quiet/2x	AMR@5.9/NS5	AMR@5.9	6	1.0	0.0	1.0	0.0	1.3	0.5	6	0	0	6	0	0	4	2	0
96	Quiet/2x	AMR@5.9/NS6	AMR@5.9	6	1.0	0.0	1.0	0.0	1.2	0.4	6	0	0	6	0	0	5	1	0

Table 5 (Part 1/3):

Experiment 1: Artifacts in initial convergence, grouped by noise suppressor and sorted by Convergence and Silence scores

								Histograms								
								Convergence			Silence			Speech		
Cnd	**Noise/Tandem**	**Processed**	**Reference**	**N**	**Yconv**	**Ysil**	**Ysp**	**1**	**2**	***3***	**1**	**2**	***3***	**1**	**2**	***3***
25	Street@15dB/1x	AMR@5.9/NS 1	AMR@5.9	6	1.0	1.0	1.2	6	0	*0*	6	0	*0*	5	1	*0*
31	Babble@15dB/1x	AMR@5.9/NS 1	AMR@5.9	6	1.0	1.0	1.7	6	0	*0*	6	0	*0*	3	2	*1*
61	Street@15dB/1x	AMR@12.2/NS 1	AMR@12.2	6	1.0	1.0	1.2	6	0	*0*	6	0	*0*	5	1	*0*
73	Quiet/1x	AMR@12.2/NS 1	AMR@12.2	6	1.0	1.0	1.2	6	0	*0*	6	0	*0*	5	1	*0*
79	Quiet/1x	AMR@5.9/NS 1	AMR@5.9	6	1.0	1.0	1.5	6	0	*0*	6	0	*0*	4	1	*1*
85	Quiet/2x	AMR@12.2/NS 1	AMR@12.2	6	1.0	1.0	1.7	6	0	*0*	6	0	*0*	3	2	*1*
91	Quiet/2x	AMR@5.9/NS 1	AMR@5.9	6	1.0	1.0	1.5	6	0	*0*	6	0	*0*	4	1	*1*
13	Babble@9dB/1x	AMR@5.9/NS 1	AMR@5.9	6	1.0	1.2	1.2	6	0	*0*	5	1	*0*	5	1	*0*
67	Babble@15dB/1x	AMR@12.2/NS 1	AMR@12.2	6	1.0	1.2	1.2	6	0	*0*	5	1	*0*	5	1	*0*
49	Babble@9dB/1x	AMR@12.2/NS 1	AMR@12.2	6	1.2	1.3	1.5	5	1	*0*	4	2	*0*	3	3	*0*
1	Car@6dB/1x	AMR@5.9/NS 1	AMR@5.9	6	1.5	1.3	1.5	4	1	*1*	4	2	*0*	3	3	*0*
43	Street@9dB/1x	AMR@12.2/NS 1	AMR@12.2	6	1.5	1.3	1.5	4	1	*1*	4	2	*0*	3	3	*0*
7	Street@9dB/1x	AMR@5.9/NS 1	AMR@5.9	6	1.5	1.8	1.7	3	3	*0*	2	3	*1*	3	2	*1*
55	Car@12dB/1x	AMR@12.2/NS 1	AMR@12.2	6	1.8	1.8	1.7	3	1	*2*	2	3	*1*	3	2	*1*
19	Car@12dB/1x	AMR@5.9/NS 1	AMR@5.9	6	2.0	1.8	1.8	3	0	*3*	2	3	*1*	2	3	*1*
37	Car@6dB/1x	AMR@12.2/NS 1	AMR@12.2	6	2.0	2.0	1.7	3	0	*3*	2	2	*2*	3	2	*1*
32	Babble@15dB/1x	AMR@5.9/NS 2	AMR@5.9	6	1.0	1.0	1.3	6	0	*0*	6	0	*0*	4	2	*0*
14	Babble@9dB/1x	AMR@5.9/NS 2	AMR@5.9	6	1.2	1.2	1.3	5	1	*0*	5	1	*0*	4	2	*0*
50	Babble@9dB/1x	AMR@12.2/NS 2	AMR@12.2	6	1.2	1.2	1.3	5	1	*0*	5	1	*0*	4	2	*0*
68	Babble@15dB/1x	AMR@12.2/NS 2	AMR@12.2	6	1.2	1.2	1.0	5	1	*0*	5	1	*0*	6	0	*0*
80	Quiet/1x	AMR@5.9/NS 2	AMR@5.9	6	1.2	1.2	1.3	5	1	*0*	5	1	*0*	4	2	*0*
92	Quiet/2x	AMR@5.9/NS 2	AMR@5.9	6	1.2	1.2	1.2	5	1	*0*	5	1	*0*	5	1	*0*
86	Quiet/2x	AMR@12.2/NS 2	AMR@12.2	6	1.3	1.3	1.2	4	2	*0*	4	2	*0*	5	1	*0*
62	Street@15dB/1x	AMR@12.2/NS 2	AMR@12.2	6	1.7	1.7	1.2	2	4	*0*	2	4	*0*	5	1	*0*
74	Quiet/1x	AMR@12.2/NS 2	AMR@12.2	6	1.7	1.5	1.2	3	2	*1*	3	3	*0*	5	1	*0*
26	Street@15dB/1x	AMR@5.9/NS 2	AMR@5.9	6	1.7	1.8	1.0	2	4	*0*	2	3	*1*	6	0	*0*
8	Street@9dB/1x	AMR@5.9/NS 2	AMR@5.9	6	1.7	1.7	1.5	3	2	*1*	3	2	*1*	3	3	*0*
44	Street@9dB/1x	AMR@12.2/NS 2	AMR@12.2	6	1.8	2.0	1.2	1	5	*0*	1	4	*1*	5	1	*0*
20	Car@12dB/1x	AMR@5.9/NS 2	AMR@5.9	6	2.0	2.3	1.7	0	6	*0*	0	4	*2*	2	4	*0*
38	Car@6dB/1x	AMR@12.2/NS 2	AMR@12.2	6	2.0	2.3	2.0	0	6	*0*	0	4	*2*	1	4	*1*
56	Car@12dB/1x	AMR@12.2/NS 2	AMR@12.2	6	2.2	2.2	1.8	0	5	*1*	0	5	*1*	3	1	*2*
2	Car@6dB/1x	AMR@5.9/NS 2	AMR@5.9	6	2.2	2.5	1.8	0	5	*1*	1	1	*4*	2	3	*1*

Table 5 (Part 2/3):

Experiment 1: Artifacts in initial convergence, grouped by noise suppressor and sorted by Convergence and Silence scores

								Histograms								
								Convergence			Silence			Speech		
Cnd	**Noise/Tandem**	**Processed**	**Reference**	**N**	**Yconv**	**Ysil**	**Ysp**	**1**	**2**	***3***	**1**	**2**	***3***	**1**	**2**	***3***
75	Quiet/1x	AMR@12.2/NS 3	AMR@12.2	6	1.0	1.0	1.2	6	0	*0*	6	0	*0*	5	1	*0*
15	Babble@9dB/1x	AMR@5.9/NS 3	AMR@5.9	6	1.0	1.2	1.2	6	0	*0*	5	1	*0*	5	1	*0*
69	Babble@15dB/1x	AMR@12.2/NS 3	AMR@12.2	6	1.0	1.2	1.0	6	0	*0*	5	1	*0*	6	0	*0*
27	Street@15dB/1x	AMR@5.9/NS 3	AMR@5.9	6	1.2	1.2	1.0	5	1	*0*	5	1	*0*	6	0	*0*
33	Babble@15dB/1x	AMR@5.9/NS 3	AMR@5.9	6	1.2	1.2	1.2	5	1	*0*	5	1	*0*	5	1	*0*
51	Babble@9dB/1x	AMR@12.2/NS 3	AMR@12.2	6	1.2	1.2	1.0	5	1	*0*	5	1	*0*	6	0	*0*
45	Street@9dB/1x	AMR@12.2/NS 3	AMR@12.2	6	1.2	1.3	1.5	5	1	*0*	4	2	*0*	4	1	*1*
81	Quiet/1x	AMR@5.9/NS 3	AMR@5.9	6	1.3	1.2	1.3	5	0	*1*	5	1	*0*	4	2	*0*
87	Quiet/2x	AMR@12.2/NS 3	AMR@12.2	6	1.3	1.2	1.3	5	0	*1*	5	1	*0*	4	2	*0*
93	Quiet/2x	AMR@5.9/NS 3	AMR@5.9	6	1.3	1.2	1.7	5	0	*1*	5	1	*0*	2	4	*0*
3	Car@6dB/1x	AMR@5.9/NS 3	AMR@5.9	6	1.3	1.3	2.0	4	2	*0*	4	2	*0*	2	2	*2*
63	Street@15dB/1x	AMR@12.2/NS 3	AMR@12.2	6	1.3	1.3	1.2	4	2	*0*	4	2	*0*	5	1	*0*
39	Car@6dB/1x	AMR@12.2/NS 3	AMR@12.2	6	1.8	1.8	1.8	3	1	*2*	2	3	*1*	2	3	*1*
21	Car@12dB/1x	AMR@5.9/NS 3	AMR@5.9	6	2.0	2.2	1.5	1	4	*1*	0	5	*1*	4	1	*1*
57	Car@12dB/1x	AMR@12.2/NS 3	AMR@12.2	6	2.0	2.0	1.7	1	4	*1*	1	4	*1*	3	2	*1*
9	Street@9dB/1x	AMR@5.9/NS 3	AMR@5.9	6	2.0	1.8	1.5	2	2	*2*	2	3	*1*	4	1	*1*
34	Babble@15dB/1x	AMR@5.9/NS 4	AMR@5.9	6	1.0	1.0	1.0	6	0	*0*	6	0	*0*	6	0	*0*
70	Babble@15dB/1x	AMR@12.2/NS 4	AMR@12.2	6	1.0	1.0	1.3	6	0	*0*	6	0	*0*	4	2	*0*
82	Quiet/1x	AMR@5.9/NS 4	AMR@5.9	6	1.0	1.0	1.3	6	0	*0*	6	0	*0*	4	2	*0*
88	Quiet/2x	AMR@12.2/NS 4	AMR@12.2	6	1.0	1.0	1.2	6	0	*0*	6	0	*0*	5	1	*0*
94	Quiet/2x	AMR@5.9/NS 4	AMR@5.9	6	1.0	1.0	1.3	6	0	*0*	6	0	*0*	4	2	*0*
16	Babble@9dB/1x	AMR@5.9/NS 4	AMR@5.9	6	1.2	1.2	1.2	5	1	*0*	5	1	*0*	5	1	*0*
52	Babble@9dB/1x	AMR@12.2/NS 4	AMR@12.2	6	1.2	1.2	1.0	5	1	*0*	5	1	*0*	6	0	*0*
76	Quiet/1x	AMR@12.2/NS 4	AMR@12.2	6	1.3	1.2	1.3	5	0	*1*	5	1	*0*	4	2	*0*
28	Street@15dB/1x	AMR@5.9/NS 4	AMR@5.9	6	1.3	1.3	1.2	4	2	*0*	4	2	*0*	5	1	*0*
10	Street@9dB/1x	AMR@5.9/NS 4	AMR@5.9	6	1.5	1.5	1.3	3	3	*0*	3	3	*0*	4	2	*0*
46	Street@9dB/1x	AMR@12.2/NS 4	AMR@12.2	6	1.7	1.8	1.7	2	4	*0*	2	3	*1*	2	4	*0*
64	Street@15dB/1x	AMR@12.2/NS 4	AMR@12.2	6	1.8	1.8	1.5	3	1	*2*	3	1	*2*	4	1	*1*
58	Car@12dB/1x	AMR@12.2/NS 4	AMR@12.2	6	2.2	2.3	2.2	0	5	*1*	0	4	*2*	1	3	*2*
4	Car@6dB/1x	AMR@5.9/NS 4	AMR@5.9	6	2.2	2.5	2.0	0	5	*1*	0	3	*3*	2	2	*2*
22	Car@12dB/1x	AMR@5.9/NS 4	AMR@5.9	6	2.5	2.2	1.5	0	3	*3*	0	5	*1*	4	1	*1*
40	Car@6dB/1x	AMR@12.2/NS 4	AMR@12.2	6	2.5	2.3	2.3	0	3	*3*	0	4	*2*	0	4	*2*

Table 5 (Part 3/3):

Experiment 1: Artifacts in initial convergence, grouped by noise suppressor and sorted by Convergence and Silence scores

								Histograms								
								Convergence			Silence			Speech		
Cnd	**Noise/Tandem**	**Processed**	**Reference**	**N**	**Yconv**	**Ysil**	**Ysp**	**1**	**2**	***3***	**1**	**2**	***3***	**1**	**2**	***3***
95	Quiet/2x	AMR@5.9/NS 5	AMR@5.9	6	1.0	1.0	1.3	6	0	*0*	6	0	*0*	4	2	*0*
77	Quiet/1x	AMR@12.2/NS 5	AMR@12.2	6	1.3	1.2	1.3	5	0	*1*	5	1	*0*	4	2	*0*
83	Quiet/1x	AMR@5.9/NS 5	AMR@5.9	6	1.3	1.2	1.5	5	0	*1*	5	1	*0*	3	3	*0*
89	Quiet/2x	AMR@12.2/NS 5	AMR@12.2	6	1.3	1.2	1.2	5	0	*1*	5	1	*0*	5	1	*0*
17	Babble@9dB/1x	AMR@5.9/NS 5	AMR@5.9	6	1.3	1.3	1.2	5	0	*1*	4	2	*0*	5	1	*0*
35	Babble@15dB/1x	AMR@5.9/NS 5	AMR@5.9	6	1.3	1.3	1.3	5	0	*1*	4	2	*0*	5	0	*1*
53	Babble@9dB/1x	AMR@12.2/NS 5	AMR@12.2	6	1.3	1.3	1.2	5	0	*1*	4	2	*0*	5	1	*0*
29	Street@15dB/1x	AMR@5.9/NS 5	AMR@5.9	6	1.5	1.5	1.5	3	3	*0*	3	3	*0*	3	3	*0*
11	Street@9dB/1x	AMR@5.9/NS 5	AMR@5.9	6	1.7	1.5	1.2	3	2	*1*	3	3	*0*	5	1	*0*
71	Babble@15dB/1x	AMR@12.2/NS 5	AMR@12.2	6	1.7	1.5	1.3	4	0	*2*	3	3	*0*	4	2	*0*
65	Street@15dB/1x	AMR@12.2/NS 5	AMR@12.2	6	1.8	1.5	1.3	3	1	*2*	3	3	*0*	4	2	*0*
47	Street@9dB/1x	AMR@12.2/NS 5	AMR@12.2	6	2.0	1.7	1.0	2	2	*2*	2	4	*0*	6	0	*0*
59	Car@12dB/1x	AMR@12.2/NS 5	AMR@12.2	6	2.3	2.2	1.5	0	4	*2*	0	5	*1*	4	1	*1*
23	Car@12dB/1x	AMR@5.9/NS 5	AMR@5.9	6	2.7	2.3	1.7	0	2	*4*	0	4	*2*	2	4	*0*
5	Car@6dB/1x	AMR@5.9/NS 5	AMR@5.9	6	2.7	2.5	2.0	0	2	*4*	0	3	*3*	1	4	*1*
41	Car@6dB/1x	AMR@12.2/NS 5	AMR@12.2	6	2.8	2.3	1.8	0	1	*5*	0	4	*2*	2	3	*1*
18	Babble@9dB/1x	AMR@5.9/NS 6	AMR@5.9	6	1.0	1.0	1.2	6	0	*0*	6	0	*0*	5	1	*0*
54	Babble@9dB/1x	AMR@12.2/NS 6	AMR@12.2	6	1.0	1.0	1.3	6	0	*0*	6	0	*0*	4	2	*0*
84	Quiet/1x	AMR@5.9/NS 6	AMR@5.9	6	1.0	1.0	1.2	6	0	*0*	6	0	*0*	5	1	*0*
96	Quiet/2x	AMR@5.9/NS 6	AMR@5.9	6	1.0	1.0	1.2	6	0	*0*	6	0	*0*	5	1	*0*
30	Street@15dB/1x	AMR@5.9/NS 6	AMR@5.9	6	1.0	1.2	1.3	6	0	*0*	5	1	*0*	4	2	*0*
12	Street@9dB/1x	AMR@5.9/NS 6	AMR@5.9	6	1.2	1.2	1.5	5	1	*0*	5	1	*0*	3	3	*0*
90	Quiet/2x	AMR@12.2/NS 6	AMR@12.2	6	1.2	1.2	1.2	5	1	*0*	5	1	*0*	5	1	*0*
78	Quiet/1x	AMR@12.2/NS 6	AMR@12.2	6	1.3	1.2	1.3	5	0	*1*	5	1	*0*	4	2	*0*
36	Babble@15dB/1x	AMR@5.9/NS 6	AMR@5.9	6	1.3	1.3	1.5	4	2	*0*	4	2	*0*	3	3	*0*
72	Babble@15dB/1x	AMR@12.2/NS 6	AMR@12.2	6	1.3	1.3	1.2	4	2	*0*	4	2	*0*	5	1	*0*
66	Street@15dB/1x	AMR@12.2/NS 6	AMR@12.2	6	1.3	1.5	1.3	4	2	*0*	3	3	*0*	4	2	*0*
6	Car@6dB/1x	AMR@5.9/NS 6	AMR@5.9	6	1.7	2.0	2.2	2	4	*0*	2	2	*2*	1	3	*2*
60	Car@12dB/1x	AMR@12.2/NS 6	AMR@12.2	6	1.8	1.8	1.8	2	3	*1*	1	5	*0*	3	1	*2*
48	Street@9dB/1x	AMR@12.2/NS 6	AMR@12.2	6	1.8	1.7	1.5	3	1	*2*	3	2	*1*	4	1	*1*
24	Car@12dB/1x	AMR@5.9/NS 6	AMR@5.9	6	2.0	1.8	2.0	2	2	*2*	1	5	*0*	1	4	*1*
42	Car@6dB/1x	AMR@12.2/NS 6	AMR@12.2	6	2.0	1.8	2.5	1	3	*1*	1	5	*0*	0	3	*3*

Annex A
Subjects comments

Table A.1
Comments for Question B on Artifacts during the leading silence grouped by candidate

					Subjects					
Cnd	**Noise/Tandem**	**NS**	**Processed**	**Reference**	**1**	**2**	**3**	**4**	**5**	**6**
25	Street@15dB/1x	1	AMR@5.9/NS1	AMR@5.9	-	-	-	-	-	-
31	Babble@15dB/1x	1	AMR@5.9/NS1	AMR@5.9	-	-	-	-	-	-
61	Street@15dB/1x	1	AMR@12.2/NS1	AMR@12.2	-	-	-	-	-	-
73	Quiet/1x	1	AMR@12.2/NS1	AMR@12.2	-	-	-	-	-	-
79	Quiet/1x	1	AMR@5.9/NS1	AMR@5.9	-	-	-	-	-	-
85	Quiet/2x	1	AMR@12.2/NS1	AMR@12.2	-	-	-	-	-	-
91	Quiet/2x	1	AMR@5.9/NS1	AMR@5.9	-	-	-	-	-	-
13	Babble@9dB/1x	1	AMR@5.9/NS1	AMR@5.9	-	-	little shush	-	-	-
67	Babble@15dB/1x	1	AMR@12.2/NS1	AMR@12.2	-	-	little shush	-	-	-
49	Babble@9dB/1x	1	AMR@12.2/NS1	AMR@12.2	-	pause	tiny shush	-	-	-
1	Car@6dB/1x	1	AMR@5.9/NS1	AMR@5.9	click	-	-	-	-	whosh
43	Street@9dB/1x	1	AMR@12.2/NS1	AMR@12.2	-	pause	shush, click	-	-	-
7	Street@9dB/1x	1	AMR@5.9/NS1	AMR@5.9	voice like musical tones	-	shush	whosh	-	whosh
55	Car@12dB/1x	1	AMR@12.2/NS1	AMR@12.2	shush	pause	shush	-	shush	-
19	Car@12dB/1x	1	AMR@5.9/NS1	AMR@5.9	shush	wavey	little shush	-	-	-
37	Car@6dB/1x	1	AMR@12.2/NS1	AMR@12.2	shush	wavey, pause	shush	-	shush	-
32	Babble@15dB/1x	2	AMR@5.9/NS2	AMR@5.9	-	-	-	-	-	-
14	Babble@9dB/1x	2	AMR@5.9/NS2	AMR@5.9	-	-	-	whosh	-	-
50	Babble@9dB/1x	2	AMR@12.2/NS2	AMR@12.2	-	pause	-	-	-	-
68	Babble@15dB/1x	2	AMR@12.2/NS2	AMR@12.2	-	ping, pause	-	-	-	-
80	Quiet/1x	2	AMR@5.9/NS2	AMR@5.9	shush	-	-	-	-	-
92	Quiet/2x	2	AMR@5.9/NS2	AMR@5.9	fast shush	-	-	-	-	-
86	Quiet/2x	2	AMR@12.2/NS2	AMR@12.2	-	-	-	noise pumping	-	-
74	Quiet/1x	2	AMR@12.2/NS2	AMR@12.2	shush	-	-	noise pumping	-	-
8	Street@9dB/1x	2	AMR@5.9/NS2	AMR@5.9	scratch and clicking	-	shush	-	-	musical tones
62	Street@15dB/1x	2	AMR@12.2/NS2	AMR@12.2	shush	pause	shush	fast whosh	-	-
26	Street@15dB/1x	2	AMR@5.9/NS2	AMR@5.9	shush	whosh	fast shush	whosh	-	-
44	Street@9dB/1x	2	AMR@12.2/NS2	AMR@12.2	shush	pause	shush	noise fades out	shush	-
20	Car@12dB/1x	2	AMR@5.9/NS2	AMR@5.9	shush	wavey	shush	whosh	-	whosh
38	Car@6dB/1x	2	AMR@12.2/NS2	AMR@12.2	shush	whosh	shush	whosh	shush, broken noise	whosh
56	Car@12dB/1x	2	AMR@12.2/NS2	AMR@12.2	shush	wavey,pop	shush	whosh	shush	whosh
2	Car@6dB/1x	2	AMR@5.9/NS2	AMR@5.9	-	-	-	whosh	shush	whosh
75	Quiet/1x	3	AMR@12.2/NS3	AMR@12.2	-	-	-	-	-	-
15	Babble@9dB/1x	3	AMR@5.9/NS3	AMR@5.9	shush, pumping	-	-	-	-	-
69	Babble@15dB/1x	3	AMR@12.2/NS3	AMR@12.2	-	-	little shush	-	-	-
27	Street@15dB/1x	3	AMR@5.9/NS3	AMR@5.9	-	-	little shush	-	-	-
33	Babble@15dB/1x	3	AMR@5.9/NS3	AMR@5.9	long shush	-	-	-	-	-
51	Babble@9dB/1x	3	AMR@12.2/NS3	AMR@12.2	fast shush	-	-	-	-	-
45	Street@9dB/1x	3	AMR@12.2/NS3	AMR@12.2	-	pause	shush	-	-	-
81	Quiet/1x	3	AMR@5.9/NS3	AMR@5.9	-	-	-	-	shush	-
87	Quiet/2x	3	AMR@12.2/NS3	AMR@12.2	-	-	-	-	-	-
93	Quiet/2x	3	AMR@5.9/NS3	AMR@5.9	-	-	-	-	-	-
3	Car@6dB/1x	3	AMR@5.9/NS3	AMR@5.9	shush	-	-	-	-	-

63	Street@15dB/1x	3	AMR@12.2/NS3	AMR@12.2	super fast shush	bang	-	-	-	-
39	Car@6dB/1x	3	AMR@12.2/NS3	AMR@12.2	shush	whosh	shush	-	shush	-
9	Street@9dB/1x	3	AMR@5.9/NS3	AMR@5.9	-	bumpy	shush	-	shush	whosh
57	Car@12dB/1x	3	AMR@12.2/NS3	AMR@12.2	shush	wavey	shush	whosh	-	-
21	Car@12dB/1x	3	AMR@5.9/NS3	AMR@5.9	shush at the start	wavey	shush	burst at beginning	shush	whosh
34	Babble@15dB/1x	4	AMR@5.9/NS4	AMR@5.9	-	-	-	-	-	-
70	Babble@15dB/1x	4	AMR@12.2/NS4	AMR@12.2	-	-	-	-	-	-
82	Quiet/1x	4	AMR@5.9/NS4	AMR@5.9	-	-	-	-	-	-
88	Quiet/2x	4	AMR@12.2/NS4	AMR@12.2	-	-	-	-	-	-
94	Quiet/2x	4	AMR@5.9/NS4	AMR@5.9	-	-	-	-	-	-
16	Babble@9dB/1x	4	AMR@5.9/NS4	AMR@5.9	-	popping	-	-	-	-
52	Babble@9dB/1x	4	AMR@12.2/NS4	AMR@12.2	-	pause	-	-	-	-
76	Quiet/1x	4	AMR@12.2/NS4	AMR@12.2	-	-	-	-	-	-
28	Street@15dB/1x	4	AMR@5.9/NS4	AMR@5.9	shush	whosh	-	-	-	-
10	Street@9dB/1x	4	AMR@5.9/NS4	AMR@5.9	shush	-	shush	noise burst	-	-
46	Street@9dB/1x	4	AMR@12.2/NS4	AMR@12.2	shush	whosh	shush	noise fades out	-	-
64	Street@15dB/1x	4	AMR@12.2/NS4	AMR@12.2	shush	wavey	-	-	shush	-
58	Car@12dB/1x	4	AMR@12.2/NS4	AMR@12.2	shush	wavey	shush	noise fades out	shush, ocean noise	whosh
4	Car@6dB/1x	4	AMR@5.9/NS4	AMR@5.9	-	wavey	shush	whosh	ocean noise	whosh
22	Car@12dB/1x	4	AMR@5.9/NS4	AMR@5.9	shush	wavey	shush	noise fades out	-	whosh
40	Car@6dB/1x	4	AMR@12.2/NS4	AMR@12.2	shush	wavey	shush	whosh	shush	whosh
95	Quiet/2x	5	AMR@5.9/NS5	AMR@5.9	-	-	-	-	-	-
77	Quiet/1x	5	AMR@12.2/NS5	AMR@12.2	-	-	-	-	shush	-
83	Quiet/1x	5	AMR@5.9/NS5	AMR@5.9	-	-	-	-	shush	-
89	Quiet/2x	5	AMR@12.2/NS5	AMR@12.2	shush	-	-	-	-	-
17	Babble@9dB/1x	5	AMR@5.9/NS5	AMR@5.9	shush	-	little shush	-	-	-
35	Babble@15dB/1x	5	AMR@5.9/NS5	AMR@5.9	shush	-	little shush	-	-	-
53	Babble@9dB/1x	5	AMR@12.2/NS5	AMR@12.2	shush	-	little shush	-	-	-
29	Street@15dB/1x	5	AMR@5.9/NS5	AMR@5.9	shush	wavey	shush	-	-	-
11	Street@9dB/1x	5	AMR@5.9/NS5	AMR@5.9	shush	wavey	shush	-	-	-
71	Babble@15dB/1x	5	AMR@12.2/NS5	AMR@12.2	shush	-	little shush	-	shush	-
65	Street@15dB/1x	5	AMR@12.2/NS5	AMR@12.2	shush	wavey	little shush	-	-	-
47	Street@9dB/1x	5	AMR@12.2/NS5	AMR@12.2	shush	whosh	shush	whosh	-	-
59	Car@12dB/1x	5	AMR@12.2/NS5	AMR@12.2	shush	wavey,pause	shush	whosh	shush	-
23	Car@12dB/1x	5	AMR@5.9/NS5	AMR@5.9	shush	wavey	shush	noise fades out	ocean noise	whosh
5	Car@6dB/1x	5	AMR@5.9/NS5	AMR@5.9	shush	wavey	shush	whosh	-	click
41	Car@6dB/1x	5	AMR@12.2/NS5	AMR@12.2	shush	wavey	shush	whosh	shush	whosh
18	Babble@9dB/1x	6	AMR@5.9/NS6	AMR@5.9	-	-	-	-	-	-
54	Babble@9dB/1x	6	AMR@12.2/NS6	AMR@12.2	-	-	-	-	-	-
84	Quiet/1x	6	AMR@5.9/NS6	AMR@5.9	-	-	-	-	-	-
96	Quiet/2x	6	AMR@5.9/NS6	AMR@5.9	-	-	-	-	-	-
30	Street@15dB/1x	6	AMR@5.9/NS6	AMR@5.9	-	-	little shush	-	-	-
12	Street@9dB/1x	6	AMR@5.9/NS6	AMR@5.9	-	bumpy	shush	-	-	-
90	Quiet/2x	6	AMR@12.2/NS6	AMR@12.2	-	-	-	noise pumping	-	-
78	Quiet/1x	6	AMR@12.2/NS6	AMR@12.2	-	-	-	-	shush	-
36	Babble@15dB/1x	6	AMR@5.9/NS6	AMR@5.9	-	-	little shush	-	-	-
72	Babble@15dB/1x	6	AMR@12.2/NS6	AMR@12.2	fast shush	ping	-	-	-	-
66	Street@15dB/1x	6	AMR@12.2/NS6	AMR@12.2	fast shush	pause	little shush	-	-	-
6	Car@6dB/1x	6	AMR@5.9/NS6	AMR@5.9	shush	-	shush	whosh	-	whosh
48	Street@9dB/1x	6	AMR@12.2/NS6	AMR@12.2	very fast shush	ping	-	-	-	-
60	Car@12dB/1x	6	AMR@12.2/NS6	AMR@12.2	shush	ping	little shush	fast whosh	shush	-
24	Car@12dB/1x	6	AMR@5.9/NS6	AMR@5.9	shush	popping	little shush	noise fades out	-	whosh
42	Car@6dB/1x	6	AMR@12.2/NS6	AMR@12.2	shush	popping	shush	whosh	shush	whosh

Table A.2

Comments for Question C on Artifacts during speech grouped by candidate

					Subjects					
Cnd	**Noise/Tandem**	**NS**	**Processed**	**Reference**	**1**	**2**	**3**	**4**	**5**	**6**
25	Street@15dB/1x	1	AMR@5.9/NS1	AMR@5.9	-	-	-	-	-	muffling
31	Babble@15dB/1x	1	AMR@5.9/NS1	AMR@5.9	metallic	-	-	-	pumping	-
61	Street@15dB/1x	1	AMR@12.2/NS1	AMR@12.2	-	-	-	-	clear voice	pumping
73	Quiet/1x	1	AMR@12.2/NS1	AMR@12.2	-	-	-	pumping	-	-
79	Quiet/1x	1	AMR@5.9/NS1	AMR@5.9	metallic	-	-	-	pumping	-
85	Quiet/2x	1	AMR@12.2/NS1	AMR@12.2	the voice is scratched	-	-	-	pumping	modulation
91	Quiet/2x	1	AMR@5.9/NS1	AMR@5.9	metallic	more bass	-	-	pumping	-
13	Babble@9dB/1x	1	AMR@5.9/NS1	AMR@5.9	-	-	-	-	pumping	-
67	Babble@15dB/1x	1	AMR@12.2/NS1	AMR@12.2	modulation like a robot	-	-	-	-	-
49	Babble@9dB/1x	1	AMR@12.2/NS1	AMR@12.2	-	more bass	-	-	pumping	pumping
1	Car@6dB/1x	1	AMR@5.9/NS1	AMR@5.9	pumping	-	modulation	-	-	pumping
43	Street@9dB/1x	1	AMR@12.2/NS1	AMR@12.2	-	-	shush in bkgd.	cracking voice	-	-
7	Street@9dB/1x	1	AMR@5.9/NS1	AMR@5.9	void like a robot	-	-	granular	crackling	-
55	Car@12dB/1x	1	AMR@12.2/NS1	AMR@12.2	-	more clearer	little shush in the bkgd.	-	pumping	-
19	Car@12dB/1x	1	AMR@5.9/NS1	AMR@5.9	metallic	-	-	muffling	pumping	muffling
37	Car@6dB/1x	1	AMR@12.2/NS1	AMR@12.2	hits within the voice	more bass	-	-	crackling	-
32	Babble@15dB/1x	2	AMR@5.9/NS2	AMR@5.9	-	-	little scratching	-	pumping	-
14	Babble@9dB/1x	2	AMR@5.9/NS2	AMR@5.9	modul.,shush & scratched	-	-	-	-	modulation
50	Babble@9dB/1x	2	AMR@12.2/NS2	AMR@12.2	-	-	-	envelope distortion	pumping	-
68	Babble@15dB/1x	2	AMR@12.2/NS2	AMR@12.2	-	-	-	-	-	-
80	Quiet/1x	2	AMR@5.9/NS2	AMR@5.9	metallic	-	little shush in the bkgd.	-	-	-
92	Quiet/2x	2	AMR@5.9/NS2	AMR@5.9	a little bit pumped	-	-	-	pumping, like in a tunel	-
86	Quiet/2x	2	AMR@12.2/NS2	AMR@12.2	-	-	little shush in the bkgd.	-	pumping	-
74	Quiet/1x	2	AMR@12.2/NS2	AMR@12.2	-	-	little shush in the bkgd.	-	-	-
8	Street@9dB/1x	2	AMR@5.9/NS2	AMR@5.9	-	-	-	cracking voice	modulation	-
62	Street@15dB/1x	2	AMR@12.2/NS2	AMR@12.2	-	-	-	-	-	modulation
26	Street@15dB/1x	2	AMR@5.9/NS2	AMR@5.9	-	-	-	-	-	-
44	Street@9dB/1x	2	AMR@12.2/NS2	AMR@12.2	-	-	-	-	tremble on the voice	-
20	Car@12dB/1x	2	AMR@5.9/NS2	AMR@5.9	end of sentence is a echo	-	little change in clarity	-	modulation	pumping
38	Car@6dB/1x	2	AMR@12.2/NS2	AMR@12.2	metallic	-	shush in bkgd.	muffling	pumping	-
56	Car@12dB/1x	2	AMR@12.2/NS2	AMR@12.2	-	more clearer	modulation	-	pumping, modulation	-
2	Car@6dB/1x	2	AMR@5.9/NS2	AMR@5.9	voice is scratched	-	-	muffling	tunel	-
75	Quiet/1x	3	AMR@12.2/NS3	AMR@12.2	-	-	little shush in the bkgd.	-	-	-
15	Babble@9dB/1x	3	AMR@5.9/NS3	AMR@5.9	-	more bass	-	-	-	-
69	Babble@15dB/1x	3	AMR@12.2/NS3	AMR@12.2	-	-	-	-	-	-
27	Street@15dB/1x	3	AMR@5.9/NS3	AMR@5.9	-	-	-	-	-	-
33	Babble@15dB/1x	3	AMR@5.9/NS3	AMR@5.9	voice hit with something	-	-	-	-	-
51	Babble@9dB/1x	3	AMR@12.2/NS3	AMR@12.2	-	-	-	-	shortness in the voice	-
45	Street@9dB/1x	3	AMR@12.2/NS3	AMR@12.2	-	more volume	shush in bkgd.	-	-	-
81	Quiet/1x	3	AMR@5.9/NS3	AMR@5.9	-	-	little shush in the bkgd.	-	modulation	-
87	Quiet/2x	3	AMR@12.2/NS3	AMR@12.2	-	-	little shush in the bkgd.	-	pumping	-
93	Quiet/2x	3	AMR@5.9/NS3	AMR@5.9	voice got pumped	modulation	little shush in the bkgd.	-	pumping	-
3	Car@6dB/1x	3	AMR@5.9/NS3	AMR@5.9	modulation	-	-	-	pumping	modulation
63	Street@15dB/1x	3	AMR@12.2/NS3	AMR@12.2	-	more bass	-	-	-	-
39	Car@6dB/1x	3	AMR@12.2/NS3	AMR@12.2	-	-	shush in bkgd.	-	pumping	muffling
9	Street@9dB/1x	3	AMR@5.9/NS3	AMR@5.9	-	-	-	-	pumping	modulation
57	Car@12dB/1x	3	AMR@12.2/NS3	AMR@12.2	-	-	voice clarity unchanged	tonal distortion	pumping	-
21	Car@12dB/1x	3	AMR@5.9/NS3	AMR@5.9	-	-	-	slightly muffing	pumping	-

34	Babble@15dB/1x	4	AMR@5.9/NS4	AMR@5.9	-	-	-	-	crackling	-
70	Babble@15dB/1x	4	AMR@12.2/NS4	AMR@12.2	voice echo	-	-	coding artifacts	pumping	-
82	Quiet/1x	4	AMR@5.9/NS4	AMR@5.9	-	-	little shush in the bkgd.	coding artifacts	broken	-
88	Quiet/2x	4	AMR@12.2/NS4	AMR@12.2	-	-	little shush in the bkgd.	-	pumping	-
94	Quiet/2x	4	AMR@5.9/NS4	AMR@5.9	echo	-	little shush in the bkgd.	-	no changes	-
16	Babble@9dB/1x	4	AMR@5.9/NS4	AMR@5.9	-	-	-	-	pumping	pumping
52	Babble@9dB/1x	4	AMR@12.2/NS4	AMR@12.2	-	-	-	-	clear voice	-
76	Quiet/1x	4	AMR@12.2/NS4	AMR@12.2	-	-	little shush in the bkgd.	-	modulation	-
28	Street@15dB/1x	4	AMR@5.9/NS4	AMR@5.9	-	-	-	-	pumping	-
10	Street@9dB/1x	4	AMR@5.9/NS4	AMR@5.9	-	-	-	muffling	pumping	-
46	Street@9dB/1x	4	AMR@12.2/NS4	AMR@12.2	voice fades out at end	-	shush in bkgd.	-	pumping	modulation
64	Street@15dB/1x	4	AMR@12.2/NS4	AMR@12.2	-	-	-	-	pumping, broken voice	-
58	Car@12dB/1x	4	AMR@12.2/NS4	AMR@12.2	pumping voice	clearer	voice clarity unchanged	coding artifacts	pumping	-
4	Car@6dB/1x	4	AMR@5.9/NS4	AMR@5.9	distorted voice	-	modulation	-	crackling	pumping
22	Car@12dB/1x	4	AMR@5.9/NS4	AMR@5.9	-	stronger	-	-	modulation	-
40	Car@6dB/1x	4	AMR@12.2/NS4	AMR@12.2	voice clipping at the end	more bass	modulation	envelope distortion	modulation	modulation
95	Quiet/2x	5	AMR@5.9/NS5	AMR@5.9	voice scratched	-	little shush in the bkgd.	-	pumping	-
77	Quiet/1x	5	AMR@12.2/NS5	AMR@12.2	-	-	little shush in the bkgd.	-	-	-
83	Quiet/1x	5	AMR@5.9/NS5	AMR@5.9	-	-	little shush in the bkgd.	-	modulation	-
89	Quiet/2x	5	AMR@12.2/NS5	AMR@12.2	-	-	little shush in the bkgd.	-	pumping	-
17	Babble@9dB/1x	5	AMR@5.9/NS5	AMR@5.9	-	-	-	-	-	pumping
35	Babble@15dB/1x	5	AMR@5.9/NS5	AMR@5.9	-	more high end	-	-	-	-
53	Babble@9dB/1x	5	AMR@12.2/NS5	AMR@12.2	-	-	-	-	pumping	-
29	Street@15dB/1x	5	AMR@5.9/NS5	AMR@5.9	sounds like a empty space	more bass	little shush at end	-	-	-
11	Street@9dB/1x	5	AMR@5.9/NS5	AMR@5.9	-	-	-	slightly muffing	-	-
71	Babble@15dB/1x	5	AMR@12.2/NS5	AMR@12.2	get like a flut horn	-	-	-	pumping	-
65	Street@15dB/1x	5	AMR@12.2/NS5	AMR@12.2	-	more bass	little shush in the bkgd.	-	-	-
47	Street@9dB/1x	5	AMR@12.2/NS5	AMR@12.2	-	-	-	-	-	-
59	Car@12dB/1x	5	AMR@12.2/NS5	AMR@12.2	-	-	clear voice	-	pumping, broken voice	-
23	Car@12dB/1x	5	AMR@5.9/NS5	AMR@5.9	pumping at the end	more bass	-	low pass muffling	pumping	-
5	Car@6dB/1x	5	AMR@5.9/NS5	AMR@5.9	distorted voice	-	like being a small room	envelope distortion	pumping	-
41	Car@6dB/1x	5	AMR@12.2/NS5	AMR@12.2	voice echo	more bass	shush in bkgd.	-	pumping, cracking	-
18	Babble@9dB/1x	6	AMR@5.9/NS6	AMR@5.9	-	-	-	-	pumping	-
54	Babble@9dB/1x	6	AMR@12.2/NS6	AMR@12.2	metallic	-	-	-	pumping	modulation
84	Quiet/1x	6	AMR@5.9/NS6	AMR@5.9	-	-	little shush in the bkgd.	-	-	-
96	Quiet/2x	6	AMR@5.9/NS6	AMR@5.9	-	-	little shush in the bkgd.	-	pumping	-
30	Street@15dB/1x	6	AMR@5.9/NS6	AMR@5.9	-	-	modulation	-	-	muffling
12	Street@9dB/1x	6	AMR@5.9/NS6	AMR@5.9	-	-	-	muffling	modulation	-
90	Quiet/2x	6	AMR@12.2/NS6	AMR@12.2	-	-	little shush in the bkgd.	-	pumping	-
78	Quiet/1x	6	AMR@12.2/NS6	AMR@12.2	-	-	shush in bkgd.	-	modulation	-
36	Babble@15dB/1x	6	AMR@5.9/NS6	AMR@5.9	shushing	-	-	muffling	-	pumping
72	Babble@15dB/1x	6	AMR@12.2/NS6	AMR@12.2	modulation gets slower	-	-	-	-	-
66	Street@15dB/1x	6	AMR@12.2/NS6	AMR@12.2	voice echo	-	-	envelope distortion	-	-
6	Car@6dB/1x	6	AMR@5.9/NS6	AMR@5.9	voice is echoing	echo	-	metalic voice	pumping	modulation
48	Street@9dB/1x	6	AMR@12.2/NS6	AMR@12.2	-	-	little modulation	-	pumping, trembled voice	-
60	Car@12dB/1x	6	AMR@12.2/NS6	AMR@12.2	-	deeper	little modulation	-	modulation	-
24	Car@12dB/1x	6	AMR@5.9/NS6	AMR@5.9	metallic	more bass	-	envelope distortion	broken	modulation
42	Car@6dB/1x	6	AMR@12.2/NS6	AMR@12.2	metallic	more bass	shush in bkgd.	tonal distortion	pumping	noise close effect

ETSI SMG11 **Tdoc SMG11/NS 420R/99**

Kyoto, Japan 6-10 December 1999

Title: **ETSI/AMR NS Selection Experiment 2 - COMSAT Results**

Source*: **COMSAT Laboratories**

Summary

This document presents a summary for ETSI Adaptive Multi Rate (AMR) Noise Selection (NS) Selection Phase Experiment 2 conducted by COMSAT in the French language.

1. Introduction

COMSAT Laboratories performed a listening laboratory evaluation in French in accordance with the AMR NS Selection Phase Experiment 2. This experiment was designed to assess whether the noise suppression candidate algorithms for the AMR codec operating at 12.2 and 5.9 kbit/s would introduce any degradation for a quiet background for one and two transcodings. The test design is defined in Section 8 of the AMR NS Selection Subjective Test Plan [1]. COMSAT performed Experiment 2 using a subset of the French speech material available in the NTT Speech Database. Forty-eight native speakers of the French language (instead of 24, due to listening session duration constraints) performed as subjects in the test, which was nominally balanced for gender. The raw data collected was used to derive gender-wise and combined-gender pair comparison scores and standard deviation statistics.

2. Source Material

Thirteen single-sentence stimuli were selected for two male and two female talkers from the NTT Speech Database, for a total of 52 different source speech stimuli. Twelve single-sentences were allocated per talker for the main assessment sessions, and one single-sentence was allocated per talker for the practice session. All files had an exact duration of 4 seconds. The source material was provided to COMSAT, the designated Host Laboratory, which was responsible for all pre- and post-processing according to [2].

3. Experimental Design

The test design followed the specification in the AMR NS Selection Test Plan, as summarized in Tables 1 and 2.

4. Processed Material

The host laboratory provided a CDROM with 2720 processed speech files, which corresponds to the processing of (12+1) single-sentence per talker for four talkers through 28 test conditions. See [3] for details on the source speech processing.

Contact: Simão F. Campos-Neto
COMSAT Labs
22300 COMSAT Drive
Clarksburg, MD 20871
USA

Tel: +1-301-428-4516
Fax: +1-301-428-9287
E-mail: simao.campos@comsat.com

Table 1:
Factors and Conditions for Experiment 2.

Main Codec Conditions	**#**	**Notes**
Noise Suppresser Candidates	6	
Codec	1	AMR
Codec Modes (FR/HR)	HR	5.9 kbit/s rate
	FR	12.2 kbit/s rate
BER	0	Clear channel, no transmission errors
Input level	1	Nominal (-26dBov)
Acoustic Background Noise	0	None
Tandeming	1	Self tandem condition
Input Characteristic	1	GSM Filtered
Codec references	**#**	**Notes**
Test vocoders	1	AMR with NS
Reference vocoder	2	AMR@12.2, AMR@5.9
Other references	**#**	**Notes**
Direct		Nominal level, GSM Filtered
MNRU	0	None, but used in preliminaries
Ideal Noise Suppression	0	None
Common Conditions	**#**	**Notes**
GSM Channel	0	NO channel model
Number of talkers	4	2 male + 2 female
Number of speech samples	52	12/talker + 1 practice/talker
Sentences/sample	1	Single sentence stimuli
Listening Instrument		Modified IRS telephone handset
Listening Level	1	-15dBPa (79dB SPL) at ERP
Listeners	48	Naive Listeners
Randomizations	6	6 groups of 4 listeners
Rating Scale	1	PC Instructions
Replications	2	Original presentation order for the first 24 subjects and the modified presentation order for the subsequent 24 subjects.

5. Listening Sessions

COMSAT performed Experiment 2 according to the Test Plan, with the exception that instead of 24 subjects listening to two replications of the processed material, twice as many subjects were used. This was done because the procedure defined in the Test Plan would result in sessions longer than two hours, which COMSAT did not deem appropriate. The option to summon the same subjects at a later date was not a viable solution, in particular because of the tight time schedule proposed for the activity. The most appropriate solution that would still accomplish the objectives of the test design was to use twice the number of subjects, as described below.

5.1 Presentation Sequence Material

COMSAT used the grouping and randomization sequences specified in the AMR NS Selection Test Plan for Experiment 2.

Table 2:
Experiment 2: Degradation in Clean Speech

Cond.	Reference Codec	Processed Codec	Trans-codings
1	AMR@12.2	AMR@12.2	1
2	AMR@12.2	AMR@12.2	2
3	AMR@5.9	AMR@5.9	1
4	AMR@5.9	AMR@5.9	2
5	AMR@12.2	AMR/NS1@12.2	1
6	AMR@12.2	AMR/NS2@12.2	1
7	AMR@12.2	AMR/NS3@12.2	1
8	AMR@12.2	AMR/NS4@12.2	1
9	AMR@12.2	AMR/NS5@12.2	1
10	AMR@12.2	AMR/NS6@12.2	1
11	AMR@12.2	AMR/NS1@12.2	2
12	AMR@12.2	AMR/NS2@12.2	2
13	AMR@12.2	AMR/NS3@12.2	2
14	AMR@12.2	AMR/NS4@12.2	2
15	AMR@12.2	AMR/NS5@12.2	2
16	AMR@12.2	AMR/NS6@12.2	2
17	AMR@5.9	AMR/NS1@5.9	1
18	AMR@5.9	AMR/NS2@5.9	1
19	AMR@5.9	AMR/NS3@5.9	1
20	AMR@5.9	AMR/NS4@5.9	1
21	AMR@5.9	AMR/NS5@5.9	1
22	AMR@5.9	AMR/NS6@5.9	1
23	AMR@5.9	AMR/NS1@5.9	2
24	AMR@5.9	AMR/NS2@5.9	2
25	AMR@5.9	AMR/NS3@5.9	2
26	AMR@5.9	AMR/NS4@5.9	2
27	AMR@5.9	AMR/NS5@5.9	2
28	AMR@5.9	AMR/NS6@5.9	2

5.2 Listeners

The subjective assessment was performed using 48 listeners (nominally balanced between male and female), divided in two rounds of six groups of four listeners each. The first 24 subjects listened to the presentation order defined in Annex F.4 of the Test Plan, and the subsequent 24 subjects listened to the presentation order modified as per the instructions in Annex F.4.

The listener selection criteria were compliant with the AMR NS Selection Test Plan, noting that Audiometric testing was not performed on the listeners, for legal reasons. Test subjects were selected from an existing pool of native French language listeners for which past assessment performance data exists indicating their general hearing integrity. Subject performance within this experiment was compared to the overall performance of all listeners used in each experiment as a check on the hearing integrity of each listener at the time of testing. The pre-test listener orientation used by COMSAT conformed to that specified in the Test Plan.

5.3 Audio Presentation

The processed speech material was presented to groups of listeners, seated at separate, visually screened listening stations contained within an acoustically conditioned sound room meeting the requirements for an NC 20 acoustic facility. Presentation was made monaurally using a telephone handset, driven by a distribution amplifier set to deliver monophonic speech to the listener's preferred listening ear at an active level of -15 dBPa (79 dB SPL), using a B&K 4153 Artificial Ear with circumaural headphone adapter, 4134 Microphone element and 2610 Measurement Amplifier.

The processed speech files were stored within the main facility computer and presented to the listeners under program control as 16 kHz samples through a 16-bit, D/A coupled to the input of the distribution amplifier through a Frequency Devices 9002 Eight-pole Elliptic Filter, set for a bandpass of 200-Hz to 3.4-kHz. Auxiliary filtering was performed to achieve an overall modified-IRS receive characteristic.

The listener responses were registered on auxiliary computers. One of these voting terminals is contained within each voting station. Voting was permitted following the completed presentation of each voting stimulus (in this experiment, a pair of associated single-sentences). All seated listeners were required to register responses prior to the subsequent presentation of a new stimulus. Once a group of listeners was conditioned to the dynamics of the voting procedure, the voting response time for each presented stimulus was nominally three seconds for each presented stimulus.

5.4 Scoring

Within experiments using a pair comparison method of assessment, the presented single-sentences were scored by the listeners using a pair comparison scale either as the first stimulus being preferred over the second (encoded by the computer as a vote 1), or otherwise (encoded by the computer as a vote 0). Subjects were forced to choose one of the options, as equal-preference votes are not allowed in standard pair comparison tests. The semantic designations were presented on the screen of the voting terminals and selected through the use of a pointing device. The voting screen was rendered neutral during the presentation of each new stimulus.

As all seated listeners completed their voting, the votes for all stations were transferred to the main facility computer prior to the presentation of subsequent new material. The votes of each group of listeners for each presentation set of speech material were stored as ASCII files within the main facility computer for subsequent analysis and presentation.

Upon completion of the listening sessions, all raw data were de-scrambled and consolidated into a single ASCII file, which was used for the statistical analysis.

6. Statistical Analysis

The statistics to be reported for this pair-comparison experiment [4] are the proportion *P* of subjects preferring the test stimulus over the reference stimulus (as defined in Table 2) for a total of *N* votes per condition, the standard deviation *s*:

$$s = \sqrt{\frac{P \cdot (1-P)}{N}} \qquad \textbf{(Eq.1)}$$

and the upper and lower confidence limits, as calculated by:

$$CI_{1-\alpha} = \frac{N}{N + z_{1-\alpha/2}^2} \cdot \left(P + \frac{z_{1-\alpha/2}^2}{2N} \pm z_{1-\alpha/2} \sqrt{\frac{P \cdot (1-P)}{N} + \frac{z_{1-\alpha/2}^2}{4N^2}} \right) \qquad \textbf{(Eq.2)}$$

where $z_{1-\alpha/2}$ is the standardized score for a normal distribution cutting off the lower $\alpha/2$ proportion of cases.

Additionally, a hypothesis to test was whether the preference for the noise reduction-enabled AMR codec was statistically different from the ideal proportion π=0.5, i.e. that the AMR with noise suppression is equally preferred to AMR without noise suppression (for quiet background). In other words,

$$H_0: \quad \pi = 0.5$$
$$H_1: \quad \pi \neq 0.5$$

The null hypothesis Ho is tested using a *z* test where:

$$z = \frac{P - \pi}{\sqrt{\frac{\pi(1-\pi)}{N}}} = \frac{P - 0.5}{0.5/\sqrt{384}} = 2\sqrt{384} \cdot (P - 0.5) = 39.192 \cdot (P - 0.5) \qquad \textbf{(Eq.3)}$$

Hence, the null hypothesis is rejected if

$$|z| \geq z_{1-\alpha/2}$$

Or accepted if:

$$0.5 - \frac{z_{1-\alpha/2}}{39.19} < P < 0.5 + \frac{z_{1-\alpha/2}}{39.19} \qquad \textbf{(Eq.4)}$$

For a 95% confidence level, Equations 2 and 4 are reduced to ($z_{1-\alpha/2} = 1.96$, N=384):

$$CI_{95\%} = \frac{N}{N+3.84}\left[P + \frac{1.92}{N} \pm 1.96\sqrt{\frac{P\cdot(1-P)}{N} + \frac{0.96}{N^2}}\right] \approx 0.99\left[P + 0.005 \pm 0.1\sqrt{P(1-P)}\right] \qquad \textbf{(Eq.5)}$$

$$0.45 < P < 0.55 \qquad \textbf{(Eq.6)}$$

Table 3 presents the basic statistical analysis data produced by COMSAT for AMR NS Selection Experiment 2, similar to the data provided to the Global Analysis Laboratory. Each test condition received a total of 96 votes. In the table, *Cnd* represents the test condition number, *P* is the proportion of subjects preferring the Test Codec stimulus, *s* is the standard deviation, +95% and -95% represent the upper and lower (1-α)=95% confidence limits (as calculated by Eq.5), and the *Ho* column identifies whether the null hypothesis is accepted (A) or rejected (R), as calculated by Eq.6.

It can be seen from Table 3 that, at the 95% confidence level, NS 1 failed the null hypothesis for one transcoding at 12.2 kbit/s, NS 2 failed the null hypothesis for two transcodings at 12.2 kbit/s, and NS 4 failed the null hypothesis for both 12.2 kbit/s and 5.9 kbit/s two-transcoding cases.

7. Conclusion

COMSAT performed AMR NS Selection Experiment 2 for the French language in compliance with the test plan, with the exception that 48 subjects were used, instead of 24, to achieve the desired test power while keeping the listening assessment sessions within the time limit usually used by COMSAT. This experiment was designed to assess whether the noise suppression candidate algorithms for the AMR codec operating at 12.2 and 5.9 kbit/s would introduce any degradation for a quiet background. It was observed that the basic hypothesis that the subjects would equally prefer the AMR codec with noise suppression over the AMR codec without noise suppression in a quiet background was not valid at the 95% confidence level for NS candidate 1 for one transcoding at 12.2 kbit/s, for NS candidate 2 for two transcodings at 12.2 kbit/s, and for NS candidate 4 at 12.2 kbit/s and 5.9 kbit/s in the two-transcoding case. In all other cases, the subjects equally preferred the AMR codec with or without noise suppression for a quiet background.

References

[1] SMG11SQ, *"Test Plan for the AMR Specification for the AMR-NS Selection Phase"*; Tdoc SMG11 288/99.

[2] SMG11SQ, *"Processing Functions for the GSM AMR Noise Suppression Selection Tests"*; Tdoc SMG11 281R/99

[3] COMSAT Laboratories, *"Host Laboratory Processing for ETSI/AMR Noise Suppression Selection Tests"*, Tdoc SMG11 417/99

[4] Glass, G.V; Hopkins, K.D., "Statistical Methods in Education and Psychology"; 3rd edition; Allyn & Bacon, Needham Heights, MA 02194, 1995; pp.319-330.

Table 3
MOS, standard deviation, and 95% confidence interval for COMSAT's Experiment 2

Cnd	Reference (Reference Codec)	Processed (Test Codec)	Tandem	N	P	s	+95%	-95%	H_0 ?
1	AMR@12.2	AMR@12.2	1	384	0.4661	0.0255	0.5158	0.4170	A
2	AMR@12.2	AMR@12.2	2	384	0.5000	0.0255	0.5495	0.4505	A
3	AMR@5.9	AMR@5.9	1	384	0.5260	0.0255	0.5752	0.4763	A
4	AMR@5.9	AMR@5.9	2	384	0.5234	0.0255	0.5726	0.4737	A
5	AMR@12.2	AMR/NS1@12.2	1	384	0.4167	0.0252	0.4663	0.3686	**R**
6	AMR@12.2	AMR/NS2@12.2	1	384	0.4688	0.0255	0.5184	0.4196	A
7	AMR@12.2	AMR/NS3@12.2	1	384	0.4974	0.0255	0.5469	0.4479	A
8	AMR@12.2	AMR/NS4@12.2	1	384	0.5052	0.0255	0.5546	0.4556	A
9	AMR@12.2	AMR/NS5@12.2	1	384	0.4922	0.0255	0.5417	0.4427	A
10	AMR@12.2	AMR/NS6@12.2	1	384	0.5000	0.0255	0.5495	0.4505	A
11	AMR@12.2	AMR/NS1@12.2	2	384	0.5130	0.0255	0.5623	0.4634	A
12	AMR@12.2	AMR/NS2@12.2	2	384	0.4115	0.0251	0.4610	0.3636	**R**
13	AMR@12.2	AMR/NS3@12.2	2	384	0.5026	0.0255	0.5520	0.4530	A
14	AMR@12.2	AMR/NS4@12.2	2	384	0.4349	0.0253	0.4846	0.3864	**R**
15	AMR@12.2	AMR/NS5@12.2	2	384	0.5156	0.0255	0.5649	0.4659	A
16	AMR@12.2	AMR/NS6@12.2	2	384	0.4635	0.0254	0.5132	0.4145	A
17	AMR@5.9	AMR/NS1@5.9	1	384	0.4870	0.0255	0.5365	0.4376	A
18	AMR@5.9	AMR/NS2@5.9	1	384	0.4948	0.0255	0.5443	0.4453	A
19	AMR@5.9	AMR/NS3@5.9	1	384	0.5052	0.0255	0.5546	0.4556	A
20	AMR@5.9	AMR/NS4@5.9	1	384	0.4922	0.0255	0.5417	0.4427	A
21	AMR@5.9	AMR/NS5@5.9	1	384	0.5182	0.0255	0.5675	0.4685	A
22	AMR@5.9	AMR/NS6@5.9	1	384	0.4557	0.0254	0.5054	0.4068	A
23	AMR@5.9	AMR/NS1@5.9	2	384	0.5026	0.0255	0.5520	0.4530	A
24	AMR@5.9	AMR/NS2@5.9	2	384	0.4948	0.0255	0.5443	0.4453	A
25	AMR@5.9	AMR/NS3@5.9	2	384	0.5026	0.0255	0.5520	0.4530	A
26	AMR@5.9	AMR/NS4@5.9	2	384	0.4427	0.0253	0.4924	0.3941	**R**
27	AMR@5.9	AMR/NS5@5.9	2	384	0.5156	0.0255	0.5649	0.4659	A
28	AMR@5.9	AMR/NS6@5.9	2	384	0.5156	0.0255	0.5649	0.4659	A

ETSI SMG11 **Tdoc SMG11/S4 421R/99**

Kyoto, Japan 6-10 December 1999

Title: **ETSI/AMR NS Selection Experiments 4&5 - COMSAT Results**

Source*: **COMSAT Laboratories**

Summary

This document presents a summary for ETSI/AMR Noise Suppression (NS) Selection Phase Experiments 4&5 conducted by COMSAT in the Spanish language. This revised version contains the corrections published in the Addendum to SMG11/S4 421, as presented in the SMG11/SA4 Kyoto meeting.

1. Introduction

COMSAT Laboratories performed listening assessments in Spanish for the AMR Noise Selection Experiments 4 and 5. These experiments were designed to evaluate the performance of the six candidate noise suppression algorithms in the presence of three different noise types within the same experiment (car, street, and babble noise) for different signal-to-noise ratios (SNR) and two fixed bit rates (5.9 and 12.2 kbit/s). The six noise suppression candidate algorithms are presented in this contribution using blinded designations NS1 through NS6. A total of four experiments were designed by ETSI, as follows:

- Experiment 4A: Background noise performance for AMR NS at 5.9 kbit/s with low SNR values
- Experiment 4B: Background noise performance for AMR NS at 5.9 kbit/s with high SNR values
- Experiment 5A: Background noise performance for AMR NS at 12.2 kbit/s with low SNR values
- Experiment 5B: Background noise performance for AMR NS at 12.2 kbit/s with high SNR values

As specified in the AMR test plan, the assessment method was a modified version of the Category Comparison Rating (CCR) method defined in [3], in which the reference signal, instead of being the unprocessed ("Direct") speech, was the speech processed through the AMR codec without noise suppression. This was done because the main parameter of interest in this experiment was the NS performance, not the coding quality per se. The only exceptions were the MNRU conditions, whose Reference Signal was the Direct speech. In this report, the reference signal is denoted as stimulus "A", and the test signal is denoted as stimulus "B". In the data reported, a positive CMOS value indicates preference towards the test sample, and a negative value indicates preference towards the reference sample.

After the performance of the subjective sessions, COMSAT discovered that up to version 2.0 of the Test Plan, there was a flaw in the presentation sequences for Experiments 4 & 5. This flaw caused condition 7 (7 dB Ideal noise suppression for car noise) to be systematically paired with the incorrect reference (condition 38, instead of 39) when the B/A direction was played. The problem was discovered in time for the English-language listening laboratory (Nortel) to perform the experiments using corrected presentation sequences. However, half of the votes (96) had to be discarded for condition 7 from COMSAT's analysis in each of the four experiments in their first realization.

Due to time constraints, it was not possible for COMSAT to repeat all the experiments. However, the experiments that could be repeated used corrected presentation sequences, and contain the full 192 votes for condition 7. Additional measures were taken to compensate the CCR presentation direction bias for condition 7 in the experiments that could not be repeated.

Contact: Simão F. Campos-Neto
COMSAT Labs
22300 COMSAT Drive
Clarksburg, MD 20871
USA

Tel: +1-301-428-4516
Fax: +1-301-428-9287
E-mail: simao.campos@comsat.com

2. Source Material

COMSAT performed Experiments 4&5 using a subset of the Spanish speech material available in the NTT Speech Database, which was pre-processed in compliance with the specifications in the AMR NS Selection Processing Test Plan [2]. Twenty-four distinct native speakers of the Spanish language performed as subjects for each of the four sub-experiments, which were nominally balanced for gender. In total, 96 subjects were used. The raw data collected was used to derive gender-wise and combined-gender Comparison Mean Opinion Scores (CMOS) and standard deviation statistics for each sub-experiment. Additionally, rank-order analyses were performed for the different impairments in each sub-experiment.

Six sentence-pairs were selected for two male and two female Spanish-speaking talkers from the NTT Speech Database for each of the four experiments. This database contains quiet background speech sampled at 16 kHz, for a bandwidth of 8 kHz. The speech material was selected such that each experiment used a totally distinct set of speech material in the main part of the test.

The quiet background sentence pairs were provided to the Host Laboratory, also performed by COMSAT. The Host Laboratory was responsible for the pre- and post-processing of the speech material, including addition of the background noise.

3. Experiment Design

The experiment designs for Experiments 4 and 5 are defined in Section 10 of the AMR NS Selection Subjective Test Plan [1], and are summarized in Tables 1 and 2. From the tables, it can be seen that the test designs of Experiments 4A, 4B, 5A, and 5B were identical, with the exception of the specific SNR and bit rate values.

4. Processed Material

The host laboratory provided a CDROM for each of the four experiments. For each experiment, a total of 1568 files were produced (382 MB), which corresponds to the blocked processing of the source sentence-pairs through the 32 test conditions and the 32 reference conditions, plus 32 practice session files. See [4] for details on the source speech processing.

5. Listening Sessions

5.1 Presentation Sequence Material

COMSAT used the grouping and randomization sequences specified in the AMR NS Selection Test Plan for Experiments 4&5. All four experiments used the same sequences. After the performance of the subjective sessions, COMSAT discovered that up to version 2.0 of the Test Plan, there was a flaw in the presentation sequences. This flaw caused condition 7 to be systematically paired with the incorrect reference (condition 38, instead of 39) when the B/A direction was played. The problem was corrected in version 2.1 of the test plan.

5.2 Listeners

Each of the four subjective assessments was performed using 24 listeners (nominally balanced between male and female subjects), divided into six groups of four listeners each. In total, 96 different fluent speakers of the Spanish language performed as test subjects.

The listener selection criterion was compliant with the AMR NS Selection Test Plan, noting that audiometric testing was not performed on the listeners, for legal reasons. Test subjects were selected from an existing pool of native Spanish language listeners for which past assessment performance data exists indicating their general hearing integrity. Subject performance within this experiment was compared to the overall performance of all listeners used in each experiment as a check on the hearing integrity of each listener at the time of testing. The pre-test listener orientation used by COMSAT conformed to that specified in the Test Plan.

Table 1:
Factors and Conditions for Experiments 4&5.

Main Codec Conditions	**#**	**Notes**
Noise Suppresser Candidates	6	
Codec	1	AMR
Codec Modes (HR/FR)	HR FR	5.9 kbit/s rate for Experiment 4 12.2 kbit/s rate for Experiment 5
BERs	0	Clear channel, no transmission errors
Input level	1	Nominal (-26dBov)
Acoustic Background Noise	3	Car, street, and babble noise
Background noise SNR	2	Low in 4A and 5A and high in 4B and 5B (see Table 2)
Input Characteristic	1	GSM transmit filtered
Codec references	**#**	**Notes**
All Experiments	1	the same AMR rate without NS
Other references	**#**	**Notes**
Direct		nominal level, GSM transmit filtered
MNRU		nominal level, GSM transmit filtered, Q= 12, ΔQ= 4
Ideal noise suppression simulation		4, 7, and 10 dB
Common Conditions	**#**	**Notes**
GSM Channel	0	No channel model
Number of talkers	4	2 male + 2 female primary talkers
Number of speech samples	28	7 Sentence-pairs/primary talker (6 for Test, 1 for Practice)
Listening instrument		Modified IRS handset
Listening Level	1	-15dBPa (79dB SPL) at Ear Reference Point
Listeners	24	Naive Listeners
Randomizations	6	6 groups of 4 listeners
Rating Scale	1	Comparison Category Rating
Replications	1	Original Presentation Only

5.3 Audio Presentation

The processed speech material was presented to groups of listeners, seated at separate, visually screened listening stations contained within an acoustically conditioned sound room meeting the requirements for an NC 20 acoustic facility. Presentation was made monaurally using a telephone handset. The handset was driven by a distribution amplifier set to deliver monophonic speech to the listener's preferred listening ear at an active level of -15 dBPa (79 dB SPL). The calibration was made using a B&K 4153 Artificial Ear with circumaural headphone adapter, 4134 Microphone element and 2610 Measurement Amplifier.

The processed speech files were stored within the main facility computer and presented to the listeners under program control as 16 kHz samples through a 16-bit D/A converter. The D/A converter was coupled to the input of the distribution amplifier through a Frequency Devices 9002 Eight-pole Elliptic Filter, which was set for a bandpass of 200-Hz to 3.4-kHz. Auxiliary filtering was performed to achieve an overall modified-IRS receive characteristic.

The listener responses were registered on auxiliary computers. One of these voting terminals is contained within each voting station. Voting was permitted following the completed presentation of each voting stimulus (in this experiment, two sentence-pairs). All seated listeners were required to register responses prior to the subsequent presentation of a new stimulus. Once a group of listeners was conditioned to the dynamics of the voting procedure, the voting response time for each presented stimulus was nominally four seconds for each presented stimulus.

Table 2:
Allocation of conditions for Experiments 4A, 4B, 5A, and 5B

Cond.	Noise	4A&5A SNR (dB)	4B&5B SNR (dB)	Exp.4 Bit rate (kbit/s)	Exp.5 Bit rate (kbit/s)	Reference	Processed	
							Codec	Ideal NS
1	Car	6	12	5.9	12.2	AMR	AMR	-
2	Street	9	15	5.9	12.2	AMR	AMR	-
3	Babble	9	15	5.9	12.2	AMR	AMR	-
4	Car	6	12	-	-	MNRU-16	MNRU-12	-
5	Car	6	12	-	-	Direct	MNRU-12	-
4'	Street	9	15	-	-	MNRU-16	MNRU-12	-
5'	Street	9	15	-	-	Direct	MNRU-12	-
4"	Babble	9	15	-	-	MNRU-16	MNRU-12	-
5"	Babble	9	15	-	-	Direct	MNRU-12	-
6	Car	6	12	5.9	12.2	AMR	AMR	4
7	Car	6	12	5.9	12.2	AMR	AMR	7
8	Car	6	12	5.9	12.2	AMR	AMR	10
9	Street	9	15	5.9	12.2	AMR	AMR	4
10	Street	9	15	5.9	12.2	AMR	AMR	7
11	Street	9	15	5.9	12.2	AMR	AMR	10
12	Babble	9	15	5.9	12.2	AMR	AMR	4
13	Babble	9	15	5.9	12.2	AMR	AMR	7
14	Babble	9	15	5.9	12.2	AMR	AMR	10
15	Car	6	12	5.9	12.2	AMR	AMR/NS1	-
16	Car	6	12	5.9	12.2	AMR	AMR/NS2	-
17	Car	6	12	5.9	12.2	AMR	AMR/NS3	-
18	Car	6	12	5.9	12.2	AMR	AMR/NS4	-
19	Car	6	12	5.9	12.2	AMR	AMR/NS5	-
20	Car	6	12	5.9	12.2	AMR	AMR/NS6	-
21	Street	9	15	5.9	12.2	AMR	AMR/NS1	-
22	Street	9	15	5.9	12.2	AMR	AMR/NS2	-
23	Street	9	15	5.9	12.2	AMR	AMR/NS3	-
24	Street	9	15	5.9	12.2	AMR	AMR/NS4	-
25	Street	9	15	5.9	12.2	AMR	AMR/NS5	-
26	Street	9	15	5.9	12.2	AMR	AMR/NS6	-
27	Babble	9	15	5.9	12.2	AMR	AMR/NS1	-
28	Babble	9	15	5.9	12.2	AMR	AMR/NS2	-
29	Babble	9	15	5.9	12.2	AMR	AMR/NS3	-
30	Babble	9	15	5.9	12.2	AMR	AMR/NS4	-
31	Babble	9	15	5.9	12.2	AMR	AMR/NS5	-
32	Babble	9	15	5.9	12.2	AMR	AMR/NS6	-

5.4 Scoring

Within experiments using the Comparison Mean-Opinion-Score (CMOS) method of assessment, the presented (referenced) sentence-pairs were scored by the listeners using a seven-point perceived quality comparison scale as: “Much better, Better, Slightly Better, About the same, Slightly Worse, Worse, and Much Worse”. The quality comparison designations were presented on the screen of the voting terminals and selected through the use of a pointing device. The voting screen was rendered neutral during the presentation of each new stimulus.

As all seated listeners completed their voting, the votes for all stations were transferred to the main facility computer prior to the presentation of subsequent new material. The votes of each group of listeners for each presentation set of speech material were stored as ASCII files within the main facility computer for subsequent analysis and presentation.

Upon completion of the listening sessions, all raw data were de-scrambled and consolidated into a single ASCII file, which was used for the statistical analysis.

6. Statistical Analysis

In Experiments 4&5, CMOS expresses the quality degradation of stimulus "B" relative to stimulus "A". In this test, stimulus A was in general the codec-processed speech without noise suppression, and stimulus B was in general the codec-processed speech with noise suppression (see Table 2 for precise descriptions). Positive CMOS values indicate a preference of stimulus B over stimulus A; negative values indicate a preference of stimulus A over stimulus B. Hence, positive CMOS values indicates a preference towards the NS-enabled vocoders. It should be noted that conditions 1, 2 and 3 were null conditions in the four experiments, i.e. stimuli A and B were identical. A total of 192 votes were cast per test condition for male and female talkers combined, or 96 for the gender-wise statistics.

As mentioned before, after the realization of the subjective sessions, a flaw was discovered in the presentation sequences given in the selection test plan. As a result, 96 votes had to be discarded for condition 7, which ended up only having valid data in the A/B presentation direction. Paired presentation experiment designs such as CCR experiments have a presentation direction effect, which is normally compensated by having data presented in both the A/B and B/A direction. Since this was not available for condition 7, an estimate of the direction bias was derived from the subject scores involving the AMR codec (conditions 1..3 and 6..32), and this bias was subtracted from the CMOS value for condition 7 in each experiment. Table 3 presents the CMOS values for the experimental data including this correction for condition 7.

In an effort to provide SMG 11 and SA4 with more reliable data, although limited by the Kyoto meeting date, COMSAT reran three of the four experiments, Experiments 4A, 5A and 5B. The choice of experiments to rerun was based on the overall trends observed for the ideal noise condition scores as well as the magnitude of bias correction necessary. The experiments with larger bias were rerun first. The supplemental data for Experiments 4A, 5A and 5B are presented in Table 4.

Tables 3 and 4 contain the basic statistical analysis data produced by COMSAT for AMR NS Selection Experiment 7. This is the data as provided to the Global Analysis Laboratory, except for the confidence interval columns. Each test condition received a total of 192 votes for combined talkers, or 96 votes per talker for gender-wise statistics. In the tables, *Cnd* represents the test condition number, *Y* is the Comparison Mean Opinion Score, and *SD* is the standard deviation. The upper and lower 95% confidence intervals are denoted as *+95%* and *-95%*, respectively, in the table.

Complementarily, Table 5 contains a rank-ordered presentation of the combined talker data in Table 4(a), 3(b), 4(b), and 4(c) for Experiments 4A, 4B, 5A, and 5B, respectively, grouped by impairment type. Upper and lower 95% confidence intervals are also reported. Statistically equivalent test conditions are indicated using Student's t-test Least Significant Difference (LSD) criterion, Tukey-Kramer's Honestly Significant Difference (HSD) Criterion, and Dunnet multiple pair comparison against a control (D). In the table, *Cnd* represents the test condition number, *Y* is the Comparison Mean Opinion Score, and the *+95%* and *–95%* columns represent the upper and lower 95% confidence interval, respectively. The *t* column shows which test conditions can be considered equivalent under the LSD criterion (indicated by contiguous vertical lines within each test factor) for a given impairment. The *HSD* column indicates which test conditions can be considered equivalent by the HSD criterion (indicated by contiguous vertical lines within each test factor). The LSD criterion is used to compare a pair of conditions, while the HSD criterion is used to determine groups of equivalency within a set of samples. The last column, *D*, indicates whether the CMOS for a test condition is significantly higher than (">"), equivalent to ("="), or significantly lower than ("<") the CMOS values for the control condition (indicated by "-").

Table 3(a)
CMOS and standard deviation for COMSAT's Experiment 4A (first realization)

Cnd.	Noise/SNR/Rate	Reference	Processed Codec/Ideal NS	Votes	Y(all)	SD(all)	+95%	-95%
1	Car/6dB/5.9	AMR	AMR	192	0.016	0.727	0.118	-0.087
2	Street/9dB/5.9	AMR	AMR	192	0.057	0.657	0.150	-0.036
3	Babble/9dB/5.9	AMR	AMR	192	-0.016	0.802	0.098	-0.129
4	Car/6dB	MNRU-16	MNRU-12	64	-0.453	0.795	-0.258	-0.648
4'	Street/9dB	MNRU-16	MNRU-12	64	-0.500	1.321	-0.176	-0.824
4"	Babble/9dB	MNRU-16	MNRU-12	64	-0.766	1.192	-0.474	-1.058
5	Car/6dB	Direct	MNRU-12	64	-1.469	1.181	-1.179	-1.758
5'	Street/9dB	Direct	MNRU-12	64	-1.297	1.191	-1.005	-1.589
5"	Babble/9dB	Direct	MNRU-12	64	-1.438	1.651	-1.033	-1.842
6	Car/6dB/5.9	AMR	AMR/4 dB	192	0.240	0.755	0.346	0.133
7	Car/6dB/5.9	AMR	AMR/7 dB	**96**	**0.389**	0.829	0.554	0.223
8	Car/6dB/5.9	AMR	AMR/10 dB	192	0.870	1.033	1.016	0.724
9	Street/9dB/5.9	AMR	AMR/4 dB	192	0.203	0.841	0.322	0.084
10	Street/9dB/5.9	AMR	AMR/7 dB	192	0.422	1.056	0.571	0.273
11	Street/9dB/5.9	AMR	AMR/10 dB	192	0.776	1.022	0.921	0.632
12	Babble/9dB/5.9	AMR	AMR/4 dB	192	0.224	1.027	0.369	0.079
13	Babble/9dB/5.9	AMR	AMR/7 dB	192	0.344	1.119	0.502	0.185
14	Babble/9dB/5.9	AMR	AMR/10 dB	192	0.557	1.143	0.719	0.396
15	Car/6dB/5.9	AMR	AMR/NS1	192	0.500	1.068	0.651	0.349
16	Car/6dB/5.9	AMR	AMR/NS2	192	0.682	1.157	0.846	0.519
17	Car/6dB/5.9	AMR	AMR/NS3	192	0.401	0.944	0.535	0.268
18	Car/6dB/5.9	AMR	AMR/NS4	192	0.620	1.037	0.766	0.473
19	Car/6dB/5.9	AMR	AMR/NS5	192	0.552	0.953	0.687	0.417
20	Car/6dB/5.9	AMR	AMR/NS6	192	0.734	1.057	0.884	0.585
21	Street/9dB/5.9	AMR	AMR/NS1	192	0.583	1.085	0.737	0.430
22	Street/9dB/5.9	AMR	AMR/NS2	192	0.542	1.028	0.687	0.396
23	Street/9dB/5.9	AMR	AMR/NS3	192	0.542	1.053	0.691	0.393
24	Street/9dB/5.9	AMR	AMR/NS4	192	0.599	1.107	0.756	0.442
25	Street/9dB/5.9	AMR	AMR/NS5	192	0.354	1.033	0.500	0.208
26	Street/9dB/5.9	AMR	AMR/NS6	192	0.807	1.240	0.983	0.632
27	Babble/9dB/5.9	AMR	AMR/NS1	192	0.302	1.045	0.450	0.154
28	Babble/9dB/5.9	AMR	AMR/NS2	192	0.318	1.087	0.471	0.164
29	Babble/9dB/5.9	AMR	AMR/NS3	192	0.208	1.101	0.364	0.053
30	Babble/9dB/5.9	AMR	AMR/NS4	192	0.344	1.042	0.491	0.196
31	Babble/9dB/5.9	AMR	AMR/NS5	192	0.266	0.942	0.399	0.132
32	Babble/9dB/5.9	AMR	AMR/NS6	192	0.500	1.176	0.666	0.334

Notes: - Cnd: Condition number; Noise: Background noise type; SNR is signal-to-noise ratio in dB; Rate is bit rate in kbit/s; NSx is the noise suppressor candidate x. Default parameters are nominal level, 1 transcoding.
- Y is Comparison Mean Opinion Score, SD is standard deviation, Se is Standard Error, +95% and -95% are respectively the upper and lower 95% confidence interval.
- The bias compensation for Condition 7 was 0.278 in Experiment 4A.

Table 3(b)
CMOS and standard deviation for COMSAT's Experiment 4B (first realization)

Cnd.	Noise/SNR/Rate	Reference	Processed Codec/Ideal NS	Votes	Y(all)	SD(all)	+95%	-95%
1	Car/12dB/5.9	AMR	AMR	192	0.005	0.506	0.077	-0.066
2	Street/15dB/5.9	AMR	AMR	192	0.068	0.623	0.156	-0.020
3	Babble/15dB/5.9	AMR	AMR	192	0.021	0.639	0.111	-0.070
4	Car/12dB	MNRU-16	MNRU-12	64	-0.359	1.029	-0.107	-0.612
4'	Street/15dB	MNRU-16	MNRU-12	64	-0.328	0.909	-0.105	-0.551
4"	Babble/15dB	MNRU-16	MNRU-12	64	-0.922	1.251	-0.615	-1.228
5	Car/12dB	Direct	MNRU-12	64	-1.703	1.353	-1.372	-2.035
5'	Street/15dB	Direct	MNRU-12	64	-1.266	1.043	-1.010	-1.521
5"	Babble/15dB	Direct	MNRU-12	64	-2.359	1.045	-2.103	-2.615
6	Car/12dB/5.9	AMR	AMR/4 dB	192	0.214	0.753	0.320	0.107
7	Car/12dB/5.9	AMR	AMR/7 dB	**96**	**0.517**	0.790	0.676	0.359
8	Car/12dB/5.9	AMR	AMR/10 dB	192	0.760	0.968	0.897	0.623
9	Street/15dB/5.9	AMR	AMR/4 dB	192	0.245	0.692	0.343	0.147
10	Street/15dB/5.9	AMR	AMR/7 dB	192	0.432	0.728	0.535	0.329
11	Street/15dB/5.9	AMR	AMR/10 dB	192	0.599	0.875	0.723	0.475
12	Babble/15dB/5.9	AMR	AMR/4 dB	192	0.146	0.799	0.259	0.033
13	Babble/15dB/5.9	AMR	AMR/7 dB	192	0.359	0.773	0.469	0.250
14	Babble/15dB/5.9	AMR	AMR/10 dB	192	0.406	0.820	0.522	0.290
15	Car/12dB/5.9	AMR	AMR/NS1	192	0.609	0.965	0.746	0.473
16	Car/12dB/5.9	AMR	AMR/NS2	192	0.667	0.962	0.803	0.531
17	Car/12dB/5.9	AMR	AMR/NS3	192	0.495	0.949	0.629	0.361
18	Car/12dB/5.9	AMR	AMR/NS4	192	0.635	1.094	0.790	0.481
19	Car/12dB/5.9	AMR	AMR/NS5	192	0.583	0.999	0.725	0.442
20	Car/12dB/5.9	AMR	AMR/NS6	192	0.813	1.062	0.963	0.662
21	Street/15dB/5.9	AMR	AMR/NS1	192	0.448	0.861	0.570	0.326
22	Street/15dB/5.9	AMR	AMR/NS2	192	0.448	0.791	0.560	0.336
23	Street/15dB/5.9	AMR	AMR/NS3	192	0.370	0.871	0.493	0.247
24	Street/15dB/5.9	AMR	AMR/NS4	192	0.646	0.904	0.774	0.518
25	Street/15dB/5.9	AMR	AMR/NS5	192	0.438	0.860	0.559	0.316
26	Street/15dB/5.9	AMR	AMR/NS6	192	0.677	1.049	0.825	0.529
27	Babble/15dB/5.9	AMR	AMR/NS1	192	0.318	0.758	0.425	0.211
28	Babble/15dB/5.9	AMR	AMR/NS2	192	0.193	0.837	0.311	0.074
29	Babble/15dB/5.9	AMR	AMR/NS3	192	0.099	0.783	0.210	-0.012
30	Babble/15dB/5.9	AMR	AMR/NS4	192	0.260	0.834	0.378	0.142
31	Babble/15dB/5.9	AMR	AMR/NS5	192	0.115	0.791	0.226	0.003
32	Babble/15dB/5.9	AMR	AMR/NS6	192	0.417	0.870	0.540	0.294

Legend: - Cnd: Condition number; Noise: Background noise type; SNR is signal-to-noise ratio in dB; Rate is bit rate in kbit/s; NSx is the noise suppressor candidate x. Default parameters are nominal level, 1 transcoding.
- Y is Comparison Mean Opinion Score, SD is standard deviation, Se is Standard Error, +95% and -95% are respectively the upper and lower 95% confidence interval.
- The bias compensation for Condition 7 was 0.149 in Experiment 4B.

Table 3(c)
CMOS and standard deviation for COMSAT's Experiment 5A (first realization)

Cnd.	Noise/SNR/Rate	Reference	Processed Codec/Ideal NS	Votes	Y(all)	SD(all)	+95%	-95%
1	Car/6dB/12.2	AMR	AMR	192	0.031	0.909	0.160	-0.097
2	Street/9dB/12.2	AMR	AMR	192	0.031	0.943	0.165	-0.102
3	Babble/9dB/12.2	AMR	AMR	192	-0.010	1.033	0.136	-0.157
4	Car/6dB	MNRU-16	MNRU-12	64	-0.688	1.562	-0.305	-1.070
4'	Street/9dB	MNRU-16	MNRU-12	64	-0.672	1.169	-0.385	-0.958
4"	Babble/9dB	MNRU-16	MNRU-12	64	-0.656	1.116	-0.383	-0.930
5	Car/6dB	Direct	MNRU-12	64	-2.109	1.010	-1.862	-2.357
5'	Street/9dB	Direct	MNRU-12	64	-1.453	1.332	-1.127	-1.780
5"	Babble/9dB	Direct	MNRU-12	64	-1.688	1.531	-1.312	-2.063
6	Car/6dB/12.2	AMR	AMR/4 dB	192	0.313	1.124	0.471	0.154
7	Car/6dB/12.2	AMR	AMR/7 dB	**96**	**0.612**	1.088	0.829	0.394
8	Car/6dB/12.2	AMR	AMR/10 dB	192	1.031	1.058	1.181	0.882
9	Street/9dB/12.2	AMR	AMR/4 dB	192	0.333	1.035	0.480	0.187
10	Street/9dB/12.2	AMR	AMR/7 dB	192	0.438	1.138	0.598	0.277
11	Street/9dB/12.2	AMR	AMR/10 dB	192	0.885	1.166	1.050	0.721
12	Babble/9dB/12.2	AMR	AMR/4 dB	192	0.120	1.140	0.281	-0.041
13	Babble/9dB/12.2	AMR	AMR/7 dB	192	0.438	1.110	0.594	0.281
14	Babble/9dB/12.2	AMR	AMR/10 dB	192	0.510	1.193	0.679	0.342
15	Car/6dB/12.2	AMR	AMR/NS1	192	0.651	1.223	0.824	0.478
16	Car/6dB/12.2	AMR	AMR/NS2	192	0.828	1.110	0.985	0.671
17	Car/6dB/12.2	AMR	AMR/NS3	192	0.646	1.171	0.811	0.480
18	Car/6dB/12.2	AMR	AMR/NS4	192	0.854	1.167	1.019	0.689
19	Car/6dB/12.2	AMR	AMR/NS5	192	0.510	1.219	0.683	0.338
20	Car/6dB/12.2	AMR	AMR/NS6	192	0.859	1.252	1.036	0.682
21	Street/9dB/12.2	AMR	AMR/NS1	192	0.734	1.201	0.904	0.565
22	Street/9dB/12.2	AMR	AMR/NS2	192	0.745	1.208	0.916	0.574
23	Street/9dB/12.2	AMR	AMR/NS3	192	0.521	1.198	0.690	0.351
24	Street/9dB/12.2	AMR	AMR/NS4	192	0.495	1.290	0.677	0.312
25	Street/9dB/12.2	AMR	AMR/NS5	192	0.458	1.161	0.623	0.294
26	Street/9dB/12.2	AMR	AMR/NS6	192	0.870	1.236	1.045	0.695
27	Babble/9dB/12.2	AMR	AMR/NS1	192	0.396	1.002	0.538	0.254
28	Babble/9dB/12.2	AMR	AMR/NS2	192	0.396	1.053	0.545	0.247
29	Babble/9dB/12.2	AMR	AMR/NS3	192	0.172	1.022	0.316	0.027
30	Babble/9dB/12.2	AMR	AMR/NS4	192	0.406	1.250	0.583	0.229
31	Babble/9dB/12.2	AMR	AMR/NS5	192	0.135	1.079	0.288	-0.017
32	Babble/9dB/12.2	AMR	AMR/NS6	192	0.536	1.166	0.701	0.372

Legend: - Cnd: Condition number; Noise: Background noise type; SNR is signal-to-noise ratio in dB; Rate is bit rate in kbit/s; NSx is the noise suppressor candidate x. Default parameters are nominal level, 1 transcoding.
- Y is Comparison Mean Opinion Score, SD is standard deviation, Se is Standard Error, +95% and -95% are respectively the upper and lower 95% confidence interval.
- The bias compensation for Condition 7 was 0.315 in Experiment 5A.

Table 3(d)
CMOS and standard deviation for COMSAT's Experiment 5B (first realization)

Cnd.	Noise/SNR/Rate	Reference	Processed Codec/Ideal NS	Votes	Y(all)	SD(all)	+95%	-95%
1	Car/12dB/12.2	AMR	AMR	192	0.047	0.820	0.163	-0.069
2	Street/15dB/12.2	AMR	AMR	192	-0.031	0.744	0.074	-0.137
3	Babble/15dB/12.2	AMR	AMR	192	-0.042	0.792	0.070	-0.154
4	Car/12dB	MNRU-16	MNRU-12	64	-0.422	1.020	-0.172	-0.672
4'	Street/15dB	MNRU-16	MNRU-12	64	-0.422	1.467	-0.062	-0.781
4"	Babble/15dB	MNRU-16	MNRU-12	64	-0.469	1.333	-0.142	-0.795
5	Car/12dB	Direct	MNRU-12	64	-1.703	1.305	-1.383	-2.023
5'	Street/15dB	Direct	MNRU-12	64	-1.594	1.561	-1.211	-1.976
5"	Babble/15dB	Direct	MNRU-12	64	-2.063	0.974	-1.824	-2.301
6	Car/12dB/12.2	AMR	AMR/4 dB	192	0.307	0.834	0.425	0.189
7	Car/12dB/12.2	AMR	AMR/7 dB	**96**	**0.354**	0.826	0.519	0.189
8	Car/12dB/12.2	AMR	AMR/10 dB	192	0.771	1.002	0.913	0.629
9	Street/15dB/12.2	AMR	AMR/4 dB	192	0.193	0.745	0.298	0.087
10	Street/15dB/12.2	AMR	AMR/7 dB	192	0.385	0.879	0.510	0.261
11	Street/15dB/12.2	AMR	AMR/10 dB	192	0.776	0.964	0.912	0.640
12	Babble/15dB/12.2	AMR	AMR/4 dB	192	0.104	0.850	0.224	-0.016
13	Babble/15dB/12.2	AMR	AMR/7 dB	192	0.313	1.001	0.454	0.171
14	Babble/15dB/12.2	AMR	AMR/10 dB	192	0.391	1.236	0.565	0.216
15	Car/12dB/12.2	AMR	AMR/NS1	192	0.802	1.079	0.955	0.649
16	Car/12dB/12.2	AMR	AMR/NS2	192	0.708	1.053	0.857	0.559
17	Car/12dB/12.2	AMR	AMR/NS3	192	0.625	1.005	0.767	0.483
18	Car/12dB/12.2	AMR	AMR/NS4	192	0.807	1.038	0.954	0.660
19	Car/12dB/12.2	AMR	AMR/NS5	192	0.641	1.093	0.795	0.486
20	Car/12dB/12.2	AMR	AMR/NS6	192	1.021	1.162	1.185	0.856
21	Street/15dB/12.2	AMR	AMR/NS1	192	0.635	0.905	0.763	0.507
22	Street/15dB/12.2	AMR	AMR/NS2	192	0.693	1.020	0.837	0.548
23	Street/15dB/12.2	AMR	AMR/NS3	192	0.479	0.965	0.616	0.343
24	Street/15dB/12.2	AMR	AMR/NS4	192	0.708	1.022	0.853	0.564
25	Street/15dB/12.2	AMR	AMR/NS5	192	0.464	0.891	0.590	0.337
26	Street/15dB/12.2	AMR	AMR/NS6	192	0.943	1.159	1.107	0.779
27	Babble/15dB/12.2	AMR	AMR/NS1	192	0.297	1.107	0.453	0.140
28	Babble/15dB/12.2	AMR	AMR/NS2	192	0.375	1.142	0.537	0.213
29	Babble/15dB/12.2	AMR	AMR/NS3	192	0.177	1.018	0.321	0.033
30	Babble/15dB/12.2	AMR	AMR/NS4	192	0.396	1.153	0.559	0.233
31	Babble/15dB/12.2	AMR	AMR/NS5	192	0.089	0.953	0.223	-0.046
32	Babble/15dB/12.2	AMR	AMR/NS6	192	0.557	1.143	0.719	0.396

Legend: - Cnd: Condition number; Noise: Background noise type; SNR is signal-to-noise ratio in dB; Rate is bit rate in kbit/s; NSx is the noise suppressor candidate x. Default parameters are nominal level, 1 transcoding.
- Y is Comparison Mean Opinion Score, SD is standard deviation, Se is Standard Error, +95% and -95% are respectively the upper and lower 95% confidence interval.
- The bias compensation for Condition 7 was 0.261 in Experiment 5B.

Table 4(a)
CMOS and standard deviation for COMSAT's Experiment 4A (second realization)

Cnd.	Noise/SNR/Rate	Reference	Processed Codec/Ideal NS	Votes	Y(all)	SD(all)	+95%	-95%
1	Car/6dB/12.2	AMR	AMR	192	-0.052	0.848	0.068	-0.172
2	Street/9dB/12.2	AMR	AMR	192	0.026	0.912	0.155	-0.103
3	Babble/9dB/12.2	AMR	AMR	192	0.010	0.927	0.141	-0.121
4	Car/6dB	MNRU-16	MNRU-12	64	-0.500	1.054	-0.242	-0.758
4'	Street/9dB	MNRU-16	MNRU-12	64	-0.734	1.073	-0.472	-0.997
4"	Babble/9dB	MNRU-16	MNRU-12	64	-0.359	0.932	-0.131	-0.588
5	Car/6dB	Direct	MNRU-12	64	-1.547	1.458	-1.190	-1.904
5'	Street/9dB	Direct	MNRU-12	64	-1.484	1.260	-1.176	-1.793
5"	Babble/9dB	Direct	MNRU-12	64	-0.953	1.385	-0.614	-1.292
6	Car/6dB/12.2	AMR	AMR/4 dB	192	0.266	0.919	0.396	0.136
7	Car/6dB/12.2	AMR	AMR/7 dB	192	0.599	1.003	0.741	0.457
8	Car/6dB/12.2	AMR	AMR/10 dB	192	0.672	1.131	0.832	0.512
9	Street/9dB/12.2	AMR	AMR/4 dB	192	0.198	0.956	0.333	0.063
10	Street/9dB/12.2	AMR	AMR/7 dB	192	0.359	1.039	0.506	0.212
11	Street/9dB/12.2	AMR	AMR/10 dB	192	0.620	1.119	0.778	0.461
12	Babble/9dB/12.2	AMR	AMR/4 dB	192	0.182	1.050	0.331	0.034
13	Babble/9dB/12.2	AMR	AMR/7 dB	192	0.276	1.108	0.433	0.119
14	Babble/9dB/12.2	AMR	AMR/10 dB	192	0.490	1.219	0.662	0.317
15	Car/6dB/12.2	AMR	AMR/NS1	192	0.479	1.189	0.647	0.311
16	Car/6dB/12.2	AMR	AMR/NS2	192	0.641	1.135	0.801	0.480
17	Car/6dB/12.2	AMR	AMR/NS3	192	0.531	1.058	0.681	0.382
18	Car/6dB/12.2	AMR	AMR/NS4	192	0.615	1.147	0.777	0.452
19	Car/6dB/12.2	AMR	AMR/NS5	192	0.411	1.159	0.575	0.248
20	Car/6dB/12.2	AMR	AMR/NS6	192	0.672	1.054	0.821	0.523
21	Street/9dB/12.2	AMR	AMR/NS1	192	0.495	1.093	0.649	0.340
22	Street/9dB/12.2	AMR	AMR/NS2	192	0.505	1.171	0.671	0.340
23	Street/9dB/12.2	AMR	AMR/NS3	192	0.349	0.964	0.485	0.213
24	Street/9dB/12.2	AMR	AMR/NS4	192	0.495	1.224	0.668	0.322
25	Street/9dB/12.2	AMR	AMR/NS5	192	0.177	0.987	0.317	0.038
26	Street/9dB/12.2	AMR	AMR/NS6	192	0.604	1.310	0.790	0.419
27	Babble/9dB/12.2	AMR	AMR/NS1	192	0.406	1.098	0.562	0.251
28	Babble/9dB/12.2	AMR	AMR/NS2	192	0.219	1.075	0.371	0.067
29	Babble/9dB/12.2	AMR	AMR/NS3	192	0.271	1.092	0.425	0.116
30	Babble/9dB/12.2	AMR	AMR/NS4	192	0.240	1.056	0.389	0.090
31	Babble/9dB/12.2	AMR	AMR/NS5	192	0.141	1.031	0.287	-0.005
32	Babble/9dB/12.2	AMR	AMR/NS6	192	0.510	1.126	0.670	0.351

Legend: - Cnd: Condition number; Noise: Background noise type; SNR is signal-to-noise ratio in dB; Rate is bit rate in kbit/s; NSx is the noise suppressor candidate x. Default parameters are nominal level, 1 transcoding.
- Y is Comparison Mean Opinion Score, SD is standard deviation, Se is Standard Error, +95% and -95% are respectively the upper and lower 95% confidence interval.

Table 4(b)
CMOS and standard deviation for COMSAT's Experiment 5A (second realization)

Cnd.	Noise/SNR/Rate	Reference	Processed Codec/Ideal NS	Votes	Y(all)	SD(all)	+95%	-95%
1	Car/6dB/12.2	AMR	AMR	192	0.021	0.898	0.148	-0.106
2	Street/9dB/12.2	AMR	AMR	192	0.031	0.805	0.145	-0.083
3	Babble/9dB/12.2	AMR	AMR	192	0.109	0.870	0.232	-0.014
4	Car/6dB	MNRU-16	MNRU-12	64	-0.250	1.512	0.120	-0.620
4'	Street/9dB	MNRU-16	MNRU-12	64	-0.484	1.357	-0.152	-0.817
4"	Babble/9dB	MNRU-16	MNRU-12	64	-0.453	1.068	-0.191	-0.715
5	Car/6dB	Direct	MNRU-12	64	-1.313	1.332	-0.986	-1.639
5'	Street/9dB	Direct	MNRU-12	64	-0.891	1.323	-0.567	-1.215
5"	Babble/9dB	Direct	MNRU-12	64	-1.453	1.321	-1.130	-1.777
6	Car/6dB/12.2	AMR	AMR/4 dB	192	0.328	0.916	0.458	0.199
7	Car/6dB/12.2	AMR	AMR/7 dB	192	0.422	1.010	0.565	0.279
8	Car/6dB/12.2	AMR	AMR/10 dB	192	0.807	1.063	0.958	0.657
9	Street/9dB/12.2	AMR	AMR/4 dB	192	0.203	0.878	0.327	0.079
10	Street/9dB/12.2	AMR	AMR/7 dB	192	0.427	0.968	0.564	0.290
11	Street/9dB/12.2	AMR	AMR/10 dB	192	0.703	1.098	0.858	0.548
12	Babble/9dB/12.2	AMR	AMR/4 dB	192	0.260	0.877	0.385	0.136
13	Babble/9dB/12.2	AMR	AMR/7 dB	192	0.380	0.952	0.515	0.245
14	Babble/9dB/12.2	AMR	AMR/10 dB	192	0.458	1.092	0.613	0.304
15	Car/6dB/12.2	AMR	AMR/NS1	192	0.594	1.019	0.738	0.450
16	Car/6dB/12.2	AMR	AMR/NS2	192	0.729	1.180	0.896	0.562
17	Car/6dB/12.2	AMR	AMR/NS3	192	0.438	1.042	0.585	0.290
18	Car/6dB/12.2	AMR	AMR/NS4	192	0.776	1.096	0.931	0.621
19	Car/6dB/12.2	AMR	AMR/NS5	192	0.604	1.102	0.760	0.448
20	Car/6dB/12.2	AMR	AMR/NS6	192	0.729	1.144	0.891	0.567
21	Street/9dB/12.2	AMR	AMR/NS1	192	0.625	1.056	0.774	0.476
22	Street/9dB/12.2	AMR	AMR/NS2	192	0.536	1.058	0.686	0.387
23	Street/9dB/12.2	AMR	AMR/NS3	192	0.411	1.025	0.556	0.267
24	Street/9dB/12.2	AMR	AMR/NS4	192	0.583	1.080	0.736	0.431
25	Street/9dB/12.2	AMR	AMR/NS5	192	0.281	0.984	0.420	0.142
26	Street/9dB/12.2	AMR	AMR/NS6	192	0.719	1.213	0.890	0.547
27	Babble/9dB/12.2	AMR	AMR/NS1	192	0.313	1.071	0.464	0.161
28	Babble/9dB/12.2	AMR	AMR/NS2	192	0.240	0.995	0.380	0.099
29	Babble/9dB/12.2	AMR	AMR/NS3	192	0.255	0.894	0.382	0.129
30	Babble/9dB/12.2	AMR	AMR/NS4	192	0.401	0.987	0.541	0.261
31	Babble/9dB/12.2	AMR	AMR/NS5	192	0.161	0.868	0.284	0.039
32	Babble/9dB/12.2	AMR	AMR/NS6	192	0.500	1.102	0.656	0.344

Legend: - Cnd: Condition number; Noise: Background noise type; SNR is signal-to-noise ratio in dB; Rate is bit rate in kbit/s; NSx is the noise suppressor candidate x. Default parameters are nominal level, 1 transcoding.
- Y is Comparison Mean Opinion Score, SD is standard deviation, Se is Standard Error, +95% and -95% are respectively the upper and lower 95% confidence interval.

Table 4(c)
CMOS and standard deviation for COMSAT's Experiment 5B (second realization)

Cnd.	Noise/SNR/Rate	Reference	Processed Codec/Ideal NS	Votes	Y(all)	SD(all)	+95%	-95%
1	Car/12dB/12.2	AMR	AMR	192	-0.026	0.618	0.061	-0.113
2	Street/15dB/12.2	AMR	AMR	192	0.073	0.720	0.175	-0.029
3	Babble/15dB/12.2	AMR	AMR	192	-0.026	0.802	0.087	-0.139
4	Car/12dB	MNRU-16	MNRU-12	64	-0.344	1.130	-0.067	-0.621
4'	Street/15dB	MNRU-16	MNRU-12	64	-0.656	1.011	-0.408	-0.904
4"	Babble/15dB	MNRU-16	MNRU-12	64	-0.891	1.236	-0.588	-1.193
5	Car/12dB	Direct	MNRU-12	64	-1.844	0.877	-1.629	-2.059
5'	Street/15dB	Direct	MNRU-12	64	-2.016	1.134	-1.738	-2.293
5"	Babble/15dB	Direct	MNRU-12	64	-2.203	0.962	-1.967	-2.439
6	Car/12dB/12.2	AMR	AMR/4 dB	192	0.359	0.800	0.473	0.246
7	Car/12dB/12.2	AMR	AMR/7 dB	192	0.682	0.843	0.801	0.563
8	Car/12dB/12.2	AMR	AMR/10 dB	192	0.776	0.964	0.912	0.640
9	Street/15dB/12.2	AMR	AMR/4 dB	192	0.224	0.784	0.335	0.113
10	Street/15dB/12.2	AMR	AMR/7 dB	192	0.510	0.880	0.635	0.386
11	Street/15dB/12.2	AMR	AMR/10 dB	192	0.646	0.965	0.782	0.509
12	Babble/15dB/12.2	AMR	AMR/4 dB	192	0.281	0.894	0.408	0.155
13	Babble/15dB/12.2	AMR	AMR/7 dB	192	0.443	0.925	0.574	0.312
14	Babble/15dB/12.2	AMR	AMR/10 dB	192	0.563	1.062	0.713	0.412
15	Car/12dB/12.2	AMR	AMR/NS1	192	0.854	0.943	0.988	0.721
16	Car/12dB/12.2	AMR	AMR/NS2	192	0.766	0.905	0.894	0.638
17	Car/12dB/12.2	AMR	AMR/NS3	192	0.656	0.835	0.774	0.538
18	Car/12dB/12.2	AMR	AMR/NS4	192	0.807	0.970	0.945	0.670
19	Car/12dB/12.2	AMR	AMR/NS5	192	0.703	0.921	0.833	0.573
20	Car/12dB/12.2	AMR	AMR/NS6	192	1.068	1.093	1.222	0.913
21	Street/15dB/12.2	AMR	AMR/NS1	192	0.682	0.948	0.816	0.548
22	Street/15dB/12.2	AMR	AMR/NS2	192	0.505	0.812	0.620	0.390
23	Street/15dB/12.2	AMR	AMR/NS3	192	0.417	0.852	0.537	0.296
24	Street/15dB/12.2	AMR	AMR/NS4	192	0.646	1.043	0.793	0.498
25	Street/15dB/12.2	AMR	AMR/NS5	192	0.422	0.782	0.533	0.311
26	Street/15dB/12.2	AMR	AMR/NS6	192	0.948	1.001	1.090	0.806
27	Babble/15dB/12.2	AMR	AMR/NS1	192	0.401	1.044	0.549	0.253
28	Babble/15dB/12.2	AMR	AMR/NS2	192	0.406	0.881	0.531	0.282
29	Babble/15dB/12.2	AMR	AMR/NS3	192	0.281	0.918	0.411	0.151
30	Babble/15dB/12.2	AMR	AMR/NS4	192	0.443	0.990	0.583	0.303
31	Babble/15dB/12.2	AMR	AMR/NS5	192	0.302	0.876	0.426	0.178
32	Babble/15dB/12.2	AMR	AMR/NS6	192	0.578	1.010	0.721	0.435

Legend: - Cnd: Condition number; Noise: Background noise type; SNR is signal-to-noise ratio in dB; Rate is bit rate in kbit/s; NSx is the noise suppressor candidate x. Default parameters are nominal level, 1 transcoding.
- Y is Comparison Mean Opinion Score, SD is standard deviation, Se is Standard Error, +95% and -95% are respectively the upper and lower 95% confidence interval.

Table 5(a)

Sorted CMOS presentation organized by impairment for COMSAT Experiment 4A (**second** realization)

Cnd.	Noise/SNR/Rate	A x B	Votes	Y(all)	+95%	-95%	t	HSD	D
4"	Babble/9dB	MNRU16 x MNRU12	64	-0.359	-0.131	-0.588			
4	Car/6dB	MNRU16 x MNRU12	64	-0.500	-0.242	-0.758			
4'	Street/9dB	MNRU16 x MNRU12	64	-0.734	-0.472	-0.997			N/A
5	Car/6dB	Direct x MNRU12	64	-0.953	-0.614	-1.292			
5"	Babble/9dB	Direct x MNRU12	64	-1.484	-1.176	-1.793			
5'	Street/9dB	Direct x MNRU12	64	-1.547	-1.190	-1.904			
8	Car/6dB/12.2	*AMR x AMR/10 dB*	192	0.672	0.832	0.512			>
20	Car/6dB/12.2	AMR x AMR/NS6	192	0.672	0.821	0.523			>
16	Car/6dB/12.2	AMR x AMR/NS2	192	0.641	0.801	0.480			>
18	Car/6dB/12.2	AMR x AMR/NS4	192	0.615	0.777	0.452			>
7	Car/6dB/12.2	*AMR x AMR/7 dB*	192	0.599	0.741	0.457			>
17	Car/6dB/12.2	AMR x AMR/NS3	192	0.531	0.681	0.382			>
15	Car/6dB/12.2	AMR x AMR/NS1	192	0.479	0.647	0.311			>
19	Car/6dB/12.2	AMR x AMR/NS5	192	0.411	0.575	0.248			>
6	Car/6dB/12.2	*AMR x AMR/4 dB*	192	0.266	0.396	0.136			>
1	Car/6dB/12.2	AMR x AMR	192	-0.052	0.068	-0.172			-
11	Street/9dB/12.2	*AMR x AMR/10 dB*	192	0.620	0.778	0.461			>
26	Street/9dB/12.2	AMR x AMR/NS6	192	0.604	0.790	0.419			>
22	Street/9dB/12.2	AMR x AMR/NS2	192	0.505	0.671	0.340			>
21	Street/9dB/12.2	AMR x AMR/NS1	192	0.495	0.649	0.340			>
24	Street/9dB/12.2	AMR x AMR/NS4	192	0.495	0.668	0.322			>
10	Street/9dB/12.2	*AMR x AMR/7 dB*	192	0.359	0.506	0.212			>
23	Street/9dB/12.2	AMR x AMR/NS3	192	0.349	0.485	0.213			>
9	Street/9dB/12.2	*AMR x AMR/4 dB*	192	0.198	0.333	0.063			=
25	Street/9dB/12.2	AMR x AMR/NS5	192	0.177	0.317	0.038			=
2	Street/9dB/12.2	AMR x AMR	192	0.026	0.155	-0.103			-
32	Babble/9dB/12.2	AMR x AMR/NS6	192	0.510	0.670	0.351			>
14	Babble/9dB/12.2	*AMR x AMR/10 dB*	192	0.490	0.662	0.317			>
27	Babble/9dB/12.2	AMR x AMR/NS1	192	0.406	0.562	0.251			>
13	Babble/9dB/12.2	*AMR x AMR/7 dB*	192	0.276	0.433	0.119			=
29	Babble/9dB/12.2	AMR x AMR/NS3	192	0.271	0.425	0.116			=
30	Babble/9dB/12.2	AMR x AMR/NS4	192	0.240	0.389	0.090			=
28	Babble/9dB/12.2	AMR x AMR/NS2	192	0.219	0.371	0.067			=
12	Babble/9dB/12.2	*AMR x AMR/4 dB*	192	0.182	0.331	0.034			=
31	Babble/9dB/12.2	AMR x AMR/NS5	192	0.141	0.287	-0.005			=
3	Babble/9dB/12.2	AMR x AMR	192	0.010	0.141	-0.121			-

Legend: - Cnd: Condition number; Noise: Background noise type; SNR is signal-to-noise ratio in dB; Rate is bit rate in kbit/s; NSx is the noise suppressor candidate x. Default parameters are nominal level, 1 transcoding.

- Y is Comparison Mean Opinion Score, SD is standard deviation, Se is Standard Error, +95% and -95% are respectively the upper and lower 95% confidence interval.

Table 5(b)

Sorted CMOS presentation organized by impairment for COMSAT Experiment 4B (*first* realization)

Cnd.	Noise/SNR/Rate	A x B	Votes	Y(all)	+95%	-95%	t	HSD	D
4'	Street/15dB	MNRU16 x MNRU12	64	-0.328	-0.105	-0.551			
4	Car/12dB	MNRU16 x MNRU12	64	-0.359	-0.107	-0.612			
4"	Babble/15dB	MNRU16 x MNRU12	64	-0.922	-0.615	-1.228			N/A
5'	Street/15dB	Direct x MNRU12	64	-1.266	-1.010	-1.521			
5	Car/12dB	Direct x MNRU12	64	-1.703	-1.372	-2.035			
5"	Babble/15dB	Direct x MNRU12	64	-2.359	-2.103	-2.615			
20	Car/12dB/12.2	AMR x AMR/NS6	192	0.813	0.963	0.662			>
8	Car/12dB/12.2	*AMR x AMR/10 dB*	192	0.760	0.897	0.623			>
16	Car/12dB/12.2	AMR x AMR/NS2	192	0.667	0.803	0.531			>
18	Car/12dB/12.2	AMR x AMR/NS4	192	0.635	0.790	0.481			>
15	Car/12dB/12.2	AMR x AMR/NS1	192	0.609	0.746	0.473			>
19	Car/12dB/12.2	AMR x AMR/NS5	192	0.583	0.725	0.442			>
7	Car/12dB/12.2	*AMR x AMR/7 dB*	**96**	**0.517**	**0.676**	**0.359**			>
17	Car/12dB/12.2	AMR x AMR/NS3	192	0.495	0.629	0.361			>
6	Car/12dB/12.2	*AMR x AMR/4 dB*	192	0.214	0.320	0.107			=
1	Car/12dB/12.2	AMR x AMR	192	0.005	0.077	-0.066			-
26	Street/15dB/12.2	AMR x AMR/NS6	192	0.677	0.825	0.529			>
24	Street/15dB/12.2	AMR x AMR/NS4	192	0.646	0.774	0.518			>
11	Street/15dB/12.2	*AMR x AMR/10 dB*	192	0.599	0.723	0.475			>
21	Street/15dB/12.2	AMR x AMR/NS1	192	0.448	0.570	0.326			>
22	Street/15dB/12.2	AMR x AMR/NS2	192	0.448	0.560	0.336			>
25	Street/15dB/12.2	AMR x AMR/NS5	192	0.438	0.559	0.316			>
10	Street/15dB/12.2	*AMR x AMR/7 dB*	192	0.432	0.535	0.329			>
23	Street/15dB/12.2	AMR x AMR/NS3	192	0.370	0.493	0.247			>
9	Street/15dB/12.2	*AMR x AMR/4 dB*	192	0.245	0.343	0.147			=
2	Street/15dB/12.2	AMR x AMR	192	0.068	0.156	-0.020			-
32	Babble/15dB/12.2	AMR x AMR/NS6	192	0.417	0.540	0.294			>
14	Babble/15dB/12.2	*AMR x AMR/10 dB*	192	0.406	0.522	0.290			>
13	Babble/15dB/12.2	*AMR x AMR/7 dB*	192	0.359	0.469	0.250			>
27	Babble/15dB/12.2	AMR x AMR/NS1	192	0.318	0.425	0.211			>
30	Babble/15dB/12.2	AMR x AMR/NS4	192	0.260	0.378	0.142			>
28	Babble/15dB/12.2	AMR x AMR/NS2	192	0.193	0.311	0.074			=
12	Babble/15dB/12.2	*AMR x AMR/4 dB*	192	0.146	0.259	0.033			=
31	Babble/15dB/12.2	AMR x AMR/NS5	192	0.115	0.226	0.003			=
29	Babble/15dB/12.2	AMR x AMR/NS3	192	0.099	0.210	-0.012			=
3	Babble/15dB/12.2	AMR x AMR	192	0.021	0.111	-0.070			-

Legend: - Cnd: Condition number; Noise: Background noise type; SNR is signal-to-noise ratio in dB; Rate is bit rate in kbit/s; NSx is the noise suppressor candidate x. Default parameters are nominal level, 1 transcoding.

- Y is Comparison Mean Opinion Score, SD is standard deviation, Se is Standard Error, +95% and -95% are respectively the upper and lower 95% confidence interval.

- The bias compensation for Condition 7 was 0.149 in Experiment 4B.

Table 5(c)

Sorted CMOS presentation organized by impairment for COMSAT Experiment 5A (second realization)

Cnd.	Noise/SNR/Rate	A x B	Votes	Y(all)	+95%	-95%	t	HSD	D
4	Car/6dB	MNRU16 x MNRU12	64	-0.250	0.120	-0.620			
4"	Babble/9dB	MNRU16 x MNRU12	64	-0.453	-0.191	-0.715			
4'	Street/9dB	MNRU16 x MNRU12	64	-0.484	-0.152	-0.817			N/A
5'	Street/9dB	Direct x MNRU12	64	-0.891	-0.567	-1.215			
5	Car/6dB	Direct x MNRU12	64	-1.313	-0.986	-1.639			
5"	Babble/9dB	Direct x MNRU12	64	-1.453	-1.130	-1.777			
8	Car/6dB/12.2	*AMR x AMR/10 dB*	192	0.807	0.958	0.657			>
18	Car/6dB/12.2	AMR x AMR/NS4	192	0.776	0.931	0.621			>
16	Car/6dB/12.2	AMR x AMR/NS2	192	0.729	0.896	0.562			>
20	Car/6dB/12.2	AMR x AMR/NS6	192	0.729	0.891	0.567			>
19	Car/6dB/12.2	AMR x AMR/NS5	192	0.604	0.760	0.448			>
15	Car/6dB/12.2	AMR x AMR/NS1	192	0.594	0.738	0.450			>
17	Car/6dB/12.2	AMR x AMR/NS3	192	0.438	0.585	0.290			>
7	Car/6dB/12.2	*AMR x AMR/7 dB*	192	0.422	0.565	0.279			>
6	Car/6dB/12.2	*AMR x AMR/4 dB*	192	0.328	0.458	0.199			>
1	Car/6dB/12.2	AMR x AMR	192	0.021	0.148	-0.106			-
26	Street/9dB/12.2	AMR x AMR/NS6	192	0.719	0.890	0.547			>
11	Street/9dB/12.2	*AMR x AMR/10 dB*	192	0.703	0.858	0.548			>
21	Street/9dB/12.2	AMR x AMR/NS1	192	0.625	0.774	0.476			>
24	Street/9dB/12.2	AMR x AMR/NS4	192	0.583	0.736	0.431			>
22	Street/9dB/12.2	AMR x AMR/NS2	192	0.536	0.686	0.387			>
10	Street/9dB/12.2	*AMR x AMR/7 dB*	192	0.427	0.564	0.290			>
23	Street/9dB/12.2	AMR x AMR/NS3	192	0.411	0.556	0.267			>
25	Street/9dB/12.2	AMR x AMR/NS5	192	0.281	0.420	0.142			=
9	Street/9dB/12.2	*AMR x AMR/4 dB*	192	0.203	0.327	0.079			=
2	Street/9dB/12.2	AMR x AMR	192	0.031	0.145	-0.083			-
32	Babble/9dB/12.2	AMR x AMR/NS6	192	0.500	0.656	0.344			>
14	Babble/9dB/12.2	*AMR x AMR/10 dB*	192	0.458	0.613	0.304			>
30	Babble/9dB/12.2	AMR x AMR/NS4	192	0.401	0.541	0.261			>
13	Babble/9dB/12.2	*AMR x AMR/7 dB*	192	0.380	0.515	0.245			>
27	Babble/9dB/12.2	AMR x AMR/NS1	192	0.313	0.464	0.161			=
12	Babble/9dB/12.2	*AMR x AMR/4 dB*	192	0.260	0.385	0.136			=
29	Babble/9dB/12.2	AMR x AMR/NS3	192	0.255	0.382	0.129			=
28	Babble/9dB/12.2	AMR x AMR/NS2	192	0.240	0.380	0.099			=
31	Babble/9dB/12.2	AMR x AMR/NS5	192	0.161	0.284	0.039			=
3	Babble/9dB/12.2	AMR x AMR	192	0.109	0.232	-0.014			-

Legend: - Cnd: Condition number; Noise: Background noise type; SNR is signal-to-noise ratio in dB; Rate is bit rate in kbit/s; NSx is the noise suppressor candidate x. Default parameters are nominal level, 1 transcoding.
- Y is Comparison Mean Opinion Score, SD is standard deviation, Se is Standard Error, +95% and -95% are respectively the upper and lower 95% confidence interval.

Table 5(d)

Sorted CMOS presentation organized by impairment for COMSAT Experiment 5B (second realization)

Cnd.	Noise/SNR/Rate	A x B	Votes	Y(all)	+95%	-95%	t	HSD	D
4	Car/12dB	MNRU16 x MNRU12	64	-0.344	-0.067	-0.621			
4'	Street/15dB	MNRU16 x MNRU12	64	-0.656	-0.408	-0.904			
4"	Babble/15dB	MNRU16 x MNRU12	64	-0.891	-0.588	-1.193			N/A
5	Car/12dB	Direct x MNRU12	64	-1.844	-1.629	-2.059			
5'	Street/15dB	Direct x MNRU12	64	-2.016	-1.738	-2.293			
5"	Babble/15dB	Direct x MNRU12	64	-2.203	-1.967	-2.439			
20	Car/12dB/12.2	AMR x AMR/NS6	192	1.068	1.222	0.913			>
15	Car/12dB/12.2	AMR x AMR/NS1	192	0.854	0.988	0.721			>
18	Car/12dB/12.2	AMR x AMR/NS4	192	0.807	0.945	0.670			>
8	Car/12dB/12.2	*AMR x AMR/10 dB*	192	0.776	0.912	0.640			>
16	Car/12dB/12.2	AMR x AMR/NS2	192	0.766	0.894	0.638			>
19	Car/12dB/12.2	AMR x AMR/NS5	192	0.703	0.833	0.573			>
7	Car/12dB/12.2	*AMR x AMR/7 dB*	192	0.682	0.801	0.563			>
17	Car/12dB/12.2	AMR x AMR/NS3	192	0.656	0.774	0.538			>
6	Car/12dB/12.2	*AMR x AMR/4 dB*	192	0.359	0.473	0.246			>
1	Car/12dB/12.2	AMR x AMR	192	-0.026	0.061	-0.113			-
26	Street/15dB/12.2	AMR x AMR/NS6	192	0.948	1.090	0.806			>
21	Street/15dB/12.2	AMR x AMR/NS1	192	0.682	0.816	0.548			>
11	Street/15dB/12.2	*AMR x AMR/10 dB*	192	0.646	0.782	0.509			>
24	Street/15dB/12.2	AMR x AMR/NS4	192	0.646	0.793	0.498			>
10	Street/15dB/12.2	*AMR x AMR/7 dB*	192	0.510	0.635	0.386			>
22	Street/15dB/12.2	AMR x AMR/NS2	192	0.505	0.620	0.390			>
25	Street/15dB/12.2	AMR x AMR/NS5	192	0.422	0.533	0.311			>
23	Street/15dB/12.2	AMR x AMR/NS3	192	0.417	0.537	0.296			>
9	Street/15dB/12.2	*AMR x AMR/4 dB*	192	0.224	0.335	0.113			>
2	Street/15dB/12.2	AMR x AMR	192	0.073	0.175	-0.029			-
32	Babble/15dB/12.2	AMR x AMR/NS6	192	0.578	0.721	0.435			>
14	Babble/15dB/12.2	*AMR x AMR/10 dB*	192	0.563	0.713	0.412			>
13	Babble/15dB/12.2	*AMR x AMR/7 dB*	192	0.443	0.574	0.312			>
30	Babble/15dB/12.2	AMR x AMR/NS4	192	0.443	0.583	0.303			>
28	Babble/15dB/12.2	AMR x AMR/NS2	192	0.406	0.531	0.282			>
27	Babble/15dB/12.2	AMR x AMR/NS1	192	0.401	0.549	0.253			>
31	Babble/15dB/12.2	AMR x AMR/NS5	192	0.302	0.426	0.178			>
12	Babble/15dB/12.2	*AMR x AMR/4 dB*	192	0.281	0.408	0.155			>
29	Babble/15dB/12.2	AMR x AMR/NS3	192	0.281	0.411	0.151			>
3	Babble/15dB/12.2	AMR x AMR	192	-0.026	0.087	-0.139			-

Legend: - Cnd: Condition number; Noise: Background noise type; SNR is signal-to-noise ratio in dB; Rate is bit rate in kbit/s; NSx is the noise suppressor candidate x. Default parameters are nominal level, 1 transcoding.
- Y is Comparison Mean Opinion Score, SD is standard deviation, Se is Standard Error, +95% and -95% are respectively the upper and lower 95% confidence interval.

Table 5 provides the most convenient presentation of the test data for statistical comparisons, in particular using the HSD and Dunnet criteria. Overall across the 4 experiments, it can be seen that the rank-order of the NS-enabled solutions was in all cases higher than AMR without NS. However, for car and street noise, most of the NS-enabled AMR solutions were significantly preferred than AMR without NS, while for babble noise a smaller number of NS solutions were significantly preferred to AMR without NS. The cases with no significant differences were as follows:

- Using the HSD criterion, NS 3 and 5 were equally preferred in Experiment 4A to AMR without NS for street noise. Under the Dunnet criterion, only NS 5 was equivalent to AMR without NS under street noise. For babble noise, both HSD and Dunnet indicate that NS 2, 3, 4, and 5 were equally preferred to AMR without NS.
- In Experiment 4B, NS 2, 3, 4 and 5 were equivalent under the HSD criterion to AMR without NS when babble noise is present. Under the Dunnet criterion, only NS 2, 3, and 5 were equivalent to AMR without NS when babble noise is present..
- In Experiment 5A, NS 5 under street noise and NS 1, 2, 3, 4, and 5 under babble noise were equivalent to AMR without NS using the HSD criterion. With the Dunnet criterion, NS 5 under street noise and NS 1, 2, 3, and 5 under babble noise were again equivalent to AMR without NS.
- In Experiment 5B, all NS solutions were significantly preferred to AMR without NS for any of the three noises (using either HSD or Dunnet criteria).

As regards the ideal noise suppression scores, Table 6 summarizes the HSD equivalencies of the ideal NS cases from Table 5. It can be seen that there is a trend that indicates that for the car noise cases, 4 and 7 dB ideal NS were equally preferred for low SNR cases and 10 and 7 dB ideal NS were equally preferred for high SNR cases. Additionally, 7 and 10 dB cases were also equally preferred in the low SNR, 5.9 kbit/s, car noise case. In the presence of street noise, 4 and 7 dB ideal NS were equally preferred for low SNR; however, in this case, 7 and 10 dB ideal NS were also equivalently preferred. No distinction in preference was observed for the ideal NS cases with AMR at 5.9 kbit/s in high SNR street noise, while for the higher bit rate (12.2 kbit/s), 7 and 10 dB ideal NS were also equally preferred, but significantly more preferred than 4 dB ideal NS. For all babble noise cases but the high SNR at 5.9 kbit/s case, all three ideal SNR cases were equally preferred.

Table 7 contains a summary of the rank-orders (irrespective of statistical significance) observed for the six NS solutions studied in Experiments 4 and 5. Overall, it can be seen that NS 6 held the highest-ranking position in 8 out of 9 cases. NS 1, 2, and 4 also showed up in the upper ranking positions. Alternatively, NS 3 and 5 most of the time occupied the lower ranking positions. This is likely to indicate that NS 6 was the most aggressive noise suppressor, while NS 3 and 5 had a more conservative behavior.

Table 6:
Summary of HSD criterion for Ideal Noise Suppression in Experiments 4&5

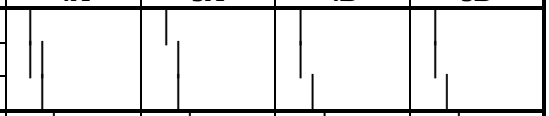
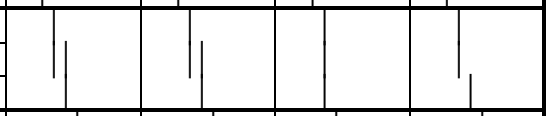
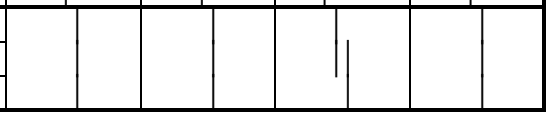
Noise Type	Ideal NS	Low SNR (6/9 dB)		High SNR (12/15 dB)	
		4A	5A	4B	5B
Car	10				
	7				
	4				
Street	10				
	7				
	4				
Babble	10				
	7				
	4				

Table 7:
Summary presentation of rank-order for the NS candidates in Experiments 4&5

4A			4B			5A			5B		
Car	**Street**	**Babble**	**Car**	**Street**	**Babble**	**Car**	**Street**	**Babble**	**Car**	**Street**	**Babble**
6	6	6	6	6	6	4	6	6	6	6	6
2	2	1	2	4	1	2	1	4	1	1	4
4	1	3	4	1	4	6	4	1	4	4	2
3	4	4	1	2	2	5	2	3	2	2	1
1	3	2	5	5	5	1	3	2	5	5	5
5	5	5	3	3	3	3	5	5	3	3	3

7. Conclusion

COMSAT performed Experiments 4&5 of the AMR NS Selection Phase in the Spanish language in compliance with the test plans [1,2]. These experiments were designed to evaluate the performance of the six candidate noise suppression algorithms in the presence of three different noise types within the same experiment (Car, street, and babble noise) for different signal-to-noise ratios (SNR) and two fixed bit rates (5.9 and 12.2 kbit/s), for a total of four subjective experiments. The six noise suppression candidate algorithms are presented in this contribution using blinded designations NS1 through NS6.

The assessment method specified in the Test Plan was a modified version of the Category Comparison Rating (CCR) method defined in [3]. Here, the reference signal, instead of being the unprocessed ("Direct") speech, was the speech processed through the AMR codec without noise suppression for all conditions involving the NS candidates. In the data reported, a positive CMOS value indicates preference towards the test sample, and a negative value indicates preference towards the reference sample.

After the performance of the subjective sessions, COMSAT discovered that up to version 2.0 of the Test Plan, there was a flaw in the presentation sequences for Experiments 4 & 5. This flaw caused condition 7 (7 dB Ideal noise suppression for car noise) to be systematically paired with the incorrect reference (condition 38, instead of 39) when the B/A direction was played. The problem was discovered in time for the English-language listening laboratory (Nortel) to perform the experiments using corrected presentation sequences. However, half of the votes (96) had to be discarded for condition 7 from COMSAT's analysis in each of the four experiments in their first realization. Due to time constraints, it was possible for COMSAT to repeat only Experiments 4A, 5A and 5B. The second realization of Experiments 4A, 5A and 5B used the corrected presentation sequences, and contain the full 192 votes for condition 7. Additional measures were taken to compensate the CCR presentation direction bias for condition 7 in Experiment 4B, which could not be repeated.

As regards the ideal noise suppression scores, a trend indicates that for the car noise cases, 4 and 7 dB ideal NS were equally preferred for low SNR cases and 10 and 7 dB ideal NS were equally preferred for high SNR cases. Additionally, 7 and 10 dB cases were also equally preferred in the low SNR, 5.9 kbit/s, car noise case. In the presence of street noise, 4 and 7 dB ideal NS were equally preferred for low SNR; however, in this case, 7 and 10 dB ideal NS were also equivalently preferred. No distinction in preference was observed for the ideal NS cases with AMR at 5.9 kbit/s in high SNR street noise, while for the higher bit rate (12.2 kbit/s), 7 and 10 dB ideal NS were also equally preferred, but significantly more preferred than 4 dB ideal NS. For all babble noise cases but the high SNR at 5.9 kbit/s case, all three ideal SNR cases were equally preferred.

Overall, NS 6 held the highest-ranking position in 8 out of 9 cases. NS 1, 2, and 4 also showed up in the upper ranking positions. Alternatively, NS 3 and 5 most of the time occupied the lower ranking positions. The modified CCR approach in Experiments 4&5 allowed the subjects to directly compare the AMR NS candidates with the AMR without NS under noisy background conditions. In this context, two main dimensions were of significance: impairments in speech quality and noise level (or amount of noise reduction). The test results, when compared to the ACR results presented by in the Global Analysis document [5], indicate that subjects keyed on the noise reduction parameter, rather than in the impairment dimension. Hence, the higher ranking of some candidates seem to indicate more aggressive NS solutions.

References

[1] SMG11SQ, *"Test Plan for the AMR Specification for the AMR-NS Selection Phase"*; Tdoc SMG11 288/99.

[2] SMG11SQ, *"Processing Functions for the GSM AMR NS Selection Tests"*; Tdoc SMG11 281R/99

[3] ITU Rec.P.800, *"Methods for Subjective Determination of Transmission Quality"*; Geneva, August 1996.

[4] COMSAT Laboratories, *"Host Laboratory Processing for ETSI/AMR Noise Suppression Selection Tests"*, Tdoc SMG11 417/99

[5] ARCON, *"Global analysis results"*, Revision 10, Tdoc SMG11 S4-99470 Rev.2; Kyoto, December 1999.