Technical Specification Group Services and System Aspects **TSGS#23(04)0069**
Meeting #23, Phoenix, AZ, USA, 15-18 March 2004
TSG-SA4#30 meeting *Tdoc S4 (04)0150*
February 23-27, 2004, Malaga, Spain

Source: **TSG SA WG4 (Dynastat)**[1]

Title: **Global Analysis Report for the 3GPP Conversation Tests for Packet-Switched Networks**

Document for: **Approval**

Agenda Item: **7.4.3**

1. Introduction

This contribution presents the work performed by Dynastat in its function as the Global Analysis Laboratory (GAL) for the 3GPP Conversation Tests for Packet Switched (PS) networks. Phase I of these tests are described in two test plans -- S4-030564 for conversation tests using the Adaptive Multi-Rate Narrow-Band (AMR-NB) codec and S4-030565 for conversation tests using the Adaptive Multi-Rate Wide-Band (AMR-WB) codec. The test plan for the Phase II tests are described in S4-030747 for conversation tests comparing various ITU-T standardized speech codecs. Document S4-030818 presents the GAL Test Plan proposed by Dynastat for characterizing the results of the conversation tests. It should be noted that this project is the first instance in 3GPP of conversation tests being used to characterize the performance of standardized speech codecs and the first instance of codecs being characterized for packet-switched networks. Moreover, the analyses reported in this document represent a new approach to evaluating the results of conversation tests.

2. Conversation Tests

The Phase 1 test plan described the methodology for conducting the conversation tests. In general, the procedure involved a pair of subjects located in different rooms and communicating over a simulated packet-switched network. The subjects were involved in a task, which required them to communicate in order to solve a specific problem. At the end of their task, each subject was required to rate various aspects of the quality of their "conversation." Each of these ratings involved a five-point scale with descriptors appropriate to the aspect of the conversation being rated. Table 1 shows a summary of the five rating scales. (The first row in each column shows the scale abbreviation that will be used throughout this report.)

Table 1. Summary of Rating Scales used in the Conversation Tests

VQ		US		IA		PC		GQ	
Voice Quality of your partner		Difficulty Understanding your partner		Interaction with your partner		Perception of impairments		Global Quality of the conversation	
5	Excellent	5	Never	5	Excellent	5	None	5	Excellent
4	Good	4	Rarely	4	Good	4	Not disturbing	4	Good
3	Fair	3	Sometimes	3	Fair	3	Slightly disturbing	3	Fair
2	Poor	2	Often	2	Poor	2	Disturbing	2	Poor
1	Bad	1	All the time	1	Bad	1	Very Disturbing	1	Bad

[1] Alan Sharpley
Dynastat, Inc
2704 Rio Grande
Austin, Texas, USA 78705
Email: sharpley@dynastat.com
Phone: +1-512-476-4797
FAX: +1-472-2883

Since each subject makes five ratings for each condition, there are five dependent variables involved in analyses of the response data. We would expect the ratings on the scales in Table 1 to show some degree of inter-correlation across test conditions. If, in fact, all five were perfectly correlated then we would conclude that they were each measuring the same underlying variable. In this scenario, we could combine them into a single measure (e.g., by averaging them) for purposes of statistical analyses and hypothesis testing. If, on the other hand, the ratings were uncorrelated, we would conclude that each scale is measuring a different underlying variable and should be treated separately in subsequent analyses. In practice, the degree of intercorrelation among such dependent variables usually falls somewhere between these two extremes. Multivariate Analysis of Variance (MANOVA) is a statistical technique designed to evaluate the results of experiments with multiple dependent variables and determine the nature and number of underlying variables. MANOVA was proposed in the GAL test plan for the conversation tests and was used extensively in the analyses presented in this report.

3. Experimental Design and Statistical Procedures

The two Phase I test plans, AMR-Narrowband (NB) and AMR-Wideband (WB), described similar experimental designs, each experiment involving 24 test conditions (*COND*) and 16 pairs of subjects. The test plans also specified that the experiments would be conducted by three Listening Laboratories (*LAB*), each in a different language: Arcon for North American English, NTT-AT for Japanese, and France Telecom for French.

Of the 24 conditions in both the NB and WB experiments, 18 were described as Symmetrical conditions (SYM), six as Asymmetrical (ASY). In the SYM conditions all subjects were located in a Quiet room, i.e., with no introduced background noise. The six ASY conditions were actually three pairs of conditions where one subject in each conversation-pair was located in a noisy background and the other subject was in the quiet. The data from these sets of paired conditions were sorted to effect a comparison of *sender in noise/receiver in quiet* and s*ender in quiet/receiver in noise* for the three conditions involving noise in the rooms.

The Phase II test plan described a single experiment involving 16 conditions conducted by one listening lab (France Telecom) but in two languages, French and Arabic.

For purposes of the GAL, the data from the three experiments, Phase I-NB, Phase I-WB, and Phase II were separated into five *Sets* of conditions for statistical analyses:

- *Set 1.* Phase I - NB/SYM conditions (1-18)
- *Set 2.* Phase I – NB/ASY conditions (19-24)
- *Set 3.* Phase I-WB/SYM conditions (1-18)
- *Set 4.* Phase I-WB/ASY conditions (19-24)
- *Set 5.* Phase II – Ph2 conditions (1-16)

For each of these five set of conditions, a three-step statistical process was undertaken to attempt to simplify the final analyses and arrive at the most parsimonious and unambiguous statistical method for characterizing the results of the conversation tests. These procedures involved the following steps:

Step 1) Compute an intercorrelation matrix among the dependent variables for the *Set* of conditions. Substantial inter-correlation among the dependent variables (i.e., correlation coefficients > .50 or < -.50) indicates that the number of dependent variables can be reduced -- that there is a reduced set of underlying variables accounting for the variance in the dependent variables.

Step 2) Conduct a MANOVA on the *Set* of scores for the effects of conditions (*COND*) in the *Set*, (18 *COND* for *Set 1*, *6 COND* for *Set 2*, etc.) ignoring other factors. The MANOVA procedure determines the linear combination of the dependent variables that best separates the linear combination of the independent variable, i.e., *COND*. The initial linear combination of dependent variables is the *root* that accounts for maximum variance in the independent variables -- it also represents the first underlying variable. A Chi-square test is conducted to determine the significance of the root. Subsequent roots are also extracted from the residual variance and tested with Chi-square for significance with each subsequent root being orthogonal to the preceding root. The number of significant roots indicates the number of significant underlying variables that account for the variance in the dependent variables.

Step 3) If there is only one significant root for the *COND* effect, the *Cannonical coefficients* for that root are used to compute a weighted average of the dependent variables to estimate the underlying variable. This composite dependent variable is then used in a univariate ANOVA to test the factors involved in the experiment. Such ANOVA's will produce results that are more parsimonious and less complicated than presenting the results in the multi-dimensional space which would be necessary with multiple dependent variables.

4. Phase I - Narrowband Test

Table 3 on the following page shows the 24 test conditions involved in the NB conversation tests. Also shown in the table are the Mean scores for each rating scale by condition and by listening lab. Each score shown in the table is the average of ratings from 32 subjects.

4.1 Narrowband Test - Symmetric conditions (*Set 1*)

The first step in the process described in the previous section is to examine the inter-correlations among the dependent variables for indications of underlying variables. Table 2 shows the inter-correlation matrix of the five dependent variables for the NB/SYM conditions. Absolute values of correlation above .50 have been bolded in the table. The table shows a high degree of inter-correlation among the dependent variables indicating the presence of a reduced set of underlying variables.

Table 2. Intercorrelations Among the Dependent Variables for the NB/SYM Conditions.

NB/S	VQ	US	IA	PC	GQ
VQ	1				
US	**0.6538**	1			
IA	0.3967	**0.5805**	1		
PC	**0.6097**	**0.7142**	**0.5616**	1	
GQ	**0.8102**	**0.6641**	0.4735	**0.6859**	1

The second step in the analysis is designed to determine how many underlying variables account for the variance in the five dependent variables. MANOVA for the effects of *COND* was conducted on the NB/SYM data – conditions 1-18. Table 4 summarizes the results of the MANOVA analysis. The table contains two sections. The top section shows the analysis for the main effect of *COND*. It includes the results of univariate ANOVA's for each of the five dependent variables followed by results for the Multivariate-ANOVA (i.e., the MANOVA) for the combination of dependent variables. In Table 4 we can see that the *COND* main effect is highly significant for each of the five individual dependent variables in the univariate ANOVA's as well as for the combination of dependent variables (MANOVA), i.e., the Pillai Trace and the associated F-statistic is highly significant in the MANOVA[1].

[1] For MANOVA, there is no single universally accepted procedure for hypothesis testing but rather a number of different methods. For the analyses that follow, we have chosen Pillai Trace and the associated F-statistic as the criterion for significance, primarily because of its robustness to violations of MANOVA assumptions.

Table 3. Test Conditions and Mean Scores for each Condition and for each Lab for the Narrowband Experiment

Narrowband - Experimental Parameters							Voice Quality			Understanding			Interaction			Perception			Global Quality		
Cond	**Rm-A**	**Rm-B**	**RC**	**PL**	**Mode**	**Del**	**Arcon**	**FT**	**NTT**	**Arcon**	**FT**	**NTT**	**Arcon**	**FT**	**NTT**	**Arcon**	**FT**	**NTT**	**Arcon**	**FT**	**NTT**
1	Quiet	Quiet	10^{-2}	0	6.7	300	3.47	3.81	3.28	3.94	4.06	4.34	3.78	3.69	4.63	4.00	3.84	4.13	3.56	3.53	3.34
2	Quiet	Quiet	10^{-2}	0	12.2	500	3.50	3.81	3.06	4.16	4.16	4.09	3.59	3.66	4.09	4.06	4.00	3.81	3.66	3.63	3.13
3	Quiet	Quiet	10^{-2}	0	12.2	300	3.81	3.63	3.47	4.16	3.94	4.34	3.88	3.72	4.56	4.19	3.84	4.19	3.88	3.56	3.53
4	Quiet	Quiet	10^{-2}	3	6.7	300	3.25	3.22	2.75	3.66	3.31	3.78	3.66	3.13	4.25	3.66	2.94	3.59	3.28	2.81	2.72
5	Quiet	Quiet	10^{-2}	3	12.2	500	3.44	3.38	2.84	3.69	3.66	3.63	3.72	3.38	4.00	3.84	2.94	3.72	3.50	2.94	2.72
6	Quiet	Quiet	10^{-2}	3	12.2	300	3.41	3.63	3.16	3.88	3.78	4.03	3.88	3.56	4.41	3.88	3.44	4.00	3.41	3.22	3.13
7	Quiet	Quiet	10^{-3}	0	6.7	300	3.91	4.16	3.41	4.19	4.47	4.44	3.94	4.00	4.84	4.34	4.38	4.31	3.78	4.00	3.50
8	Quiet	Quiet	10^{-3}	0	12.2	500	3.72	4.22	3.59	4.22	4.41	4.50	3.72	4.03	4.72	4.09	4.44	4.53	3.97	4.06	3.72
9	Quiet	Quiet	10^{-3}	0	12.2	300	4.00	4.56	3.47	4.38	4.69	4.44	4.03	4.38	4.72	4.44	4.78	4.31	4.16	4.50	3.44
10	Quiet	Quiet	10^{-3}	3	6.7	300	3.28	3.66	3.16	3.72	3.94	4.16	3.78	3.88	4.44	3.91	3.72	4.00	3.31	3.41	3.16
11	Quiet	Quiet	10^{-3}	3	12.2	500	3.75	3.84	3.19	4.13	3.97	4.31	3.81	3.56	4.38	3.94	3.91	4.13	3.66	3.69	3.25
12	Quiet	Quiet	10^{-3}	3	12.2	300	3.50	3.91	3.41	4.00	4.22	4.44	3.97	4.09	4.66	3.88	4.13	4.25	3.53	3.97	3.53
13	Quiet	Quiet	5×10^{-4}	0	6.7	300	3.91	4.25	3.59	4.19	4.63	4.47	4.06	4.16	4.72	4.38	4.59	4.44	4.00	4.25	3.59
14	Quiet	Quiet	5×10^{-4}	0	12.2	500	3.97	4.34	3.50	4.22	4.47	4.56	3.75	3.97	4.44	4.31	4.53	4.44	3.94	3.97	3.44
15	Quiet	Quiet	5×10^{-4}	0	12.2	300	4.03	4.44	4.03	4.53	4.50	4.75	4.09	4.19	4.88	4.47	4.50	4.69	3.97	4.19	3.97
16	Quiet	Quiet	5×10^{-4}	3	6.7	300	3.63	3.84	3.19	3.91	3.97	4.25	4.03	3.72	4.63	3.91	3.75	4.06	3.50	3.56	3.34
17	Quiet	Quiet	5×10^{-4}	3	12.2	500	3.66	3.88	3.22	4.03	4.22	4.25	3.78	3.78	4.34	4.13	4.13	4.09	3.69	3.78	3.19
18	Quiet	Quiet	5×10^{-4}	3	12.2	300	3.56	3.75	3.25	4.03	3.88	4.22	3.69	3.63	4.59	4.09	3.78	4.19	3.72	3.44	3.19
19	Car	Quiet	5×10^{-4}	3	12.2	300	3.16	3.63	2.88	3.13	2.97	3.34	3.84	3.06	3.88	3.66	2.72	3.66	3.41	2.53	2.81
20	Quiet	Car	5×10^{-4}	3	12.2	300	3.81	3.88	3.50	4.13	3.91	4.44	3.94	3.63	4.44	4.31	3.78	4.25	3.78	3.28	3.53
21	Cafeteria	Quiet	5×10^{-4}	0	6.7	300	3.69	4.06	3.13	3.59	3.69	3.88	3.97	3.53	4.38	4.13	3.44	4.00	3.78	3.28	3.16
22	Quiet	Cafeteria	5×10^{-4}	0	6.7	300	3.97	4.31	3.53	4.41	4.50	4.50	4.06	4.06	4.66	4.34	4.50	4.38	3.69	4.09	3.56
23	Street	Quiet	5×10^{-4}	0	12.2	500	3.66	4.03	3.25	3.53	3.72	4.16	4.00	3.47	4.28	3.94	3.44	4.22	3.81	3.31	3.22
24	Quiet	Street	5×10^{-4}	0	12.2	500	3.84	4.19	3.53	4.22	4.38	4.28	4.00	3.91	4.47	4.44	4.22	4.19	3.91	3.91	3.53

Rm-A/Rm-B (Noise environment)   RC (Radio Conditions)   PL (% Packet Loss)   Mode (Bit rate in kbps)   Del (Delay in msec)

The bottom section of Table 4 shows the Chi-square tests of the MANOVA roots. It shows only a single significant root (1 through 5), indicating that a single underlying variable accounts for the significant variation in the dependent variables for these conditions. The canonical coefficients for this root are also shown in the table and are used to compute the composite dependent variable that represents the underlying variable for the NB/SYM conditions. The composite dependent variable (**NB/S-CTQ** for **N**arrow**B**and/**S**ymmetric-**C**onversation **T**est **Q**uality) is used to characterize the ratings in the NB/SYM conditions. NB/S-CTQ scores for all conditions and all LAB's in *Set 1* are listed in the Appendix. Equation 1 shows the formula used to compute the composite score for the NB/SYM conditions.

Table 4. Results of MANOVA for *COND* for NB/SYM Conditions.

Univariate ANOVA's for Effect *COND* (df = 17, 1710)					
Dep.Var.	VQ	US	IA	PC	GQ
F-Rato	8.253	8.071	5.511	11.805	10.987
Prob.	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
MANOVA for Effect: *COND*					
Statistic	Value	F-Statistic	df	Prob	
Pillai Trace	0.162	3.376	85, 8550	**0.0000**	
Test of Residual Roots				Dep.Var.	Canon.Coeff. for Root 1-5
Roots	Chi-Square	df	Prob		
1 through 5	**292.5599**	**85**	**0.0000**	VQ	0.0426
2 through 5	73.4427	64	0.1963	US	0.0620
3 through 5	34.1441	45	0.8810	IA	-0.0015
4 through 5	11.2742	28	0.9979	PC	0.5664
5 through 5	4.2342	13	0.9884	GQ	0.4470

$$NB/S\text{-}CTQ = .0426*VQ + .0620*US - .0015 * IA + .5664 * PC + .4470 * GQ$$

Eq.1. Formula used to compute the Conversation Test Quality Score (NB/S-CTQ) for the conditions in Set 1.

The SYM conditions in the NB experiment are categorized by four experimental factors:

1. Radio conditions – 10^{-2}, 10^{-3}, and $5x10^{-4}$
2. Packet Loss – 0% and 3%
3. AMR-NB mode or bit rate – 6.7 kbps and 12.2 kbps
4. Delay – 300 msec and 500 msec

These conditions are assigned to two factorial experimental designs for analysing the effects of three of these factors. Table 5a shows the allocation of the 12 conditions used to evaluate the effects of Radio Conditions, Packet Loss, and Mode – with Delay held constant at 300 msec. Table 5b shows the allocation of the 12 conditions used to evaluate the effects of Radio Conditions, Packet Loss, and Delay – with Mode held constant at 12.2kbps.

Table 5a NB/SYM: Factorial Design for the Effects of Radio Cond., Packet Loss, and Mode.

No Noise - 300 msec delay				
6.7kbps / 0% PL			**6.7kbps / 3% PL**	
RC	Cond.#		RC	Cond.#
10^{-2}	1		10^{-2}	4
10^{-3}	7		10^{-3}	10
$5x10^{-4}$	13		$5x10^{-4}$	16
12.2kbps / 0% PL			**12.2kbps / 3% PL**	
RC	Cond.#		RC	Cond.#
10^{-2}	3		10^{-2}	6
10^{-3}	9		10^{-3}	12
$5x10^{-4}$	15		$5x10^{-4}$	18

Table 5b – NB/SYM: Factorial Design for the Effects of Radio Cond., Packet Loss, and Delay

No Noise - 12.2 kbps				
300 msec / 0% PL			**300 msec / 3% PL**	
RC	Cond.#		RC	Cond.#
10^{-2}	3		10^{-2}	6
10^{-3}	9		10^{-3}	12
$5x10^{-4}$	15		$5x10^{-4}$	18
500 msec / 0% PL			**500 msec / 3% PL**	
RC	Cond.#		RC	Cond.#
10^{-2}	2		10^{-2}	5
10^{-3}	8		10^{-3}	11
$5x10^{-4}$	14		$5x10^{-4}$	17

The composite dependent variable, NB/S-CTQ, was computed for the NB/SYM conditions using the equation shown in Eq.1. These composite scores were subjected to factorial ANOVA for the two experimental designs shown in Tables 5a and 5b. The results of those ANOVA's are shown in Tables 6 and 7, respectively.

Table 6. Results of ANOVA of NB/S-CTQ for the Effects of Lab, Radio Conditions (RC), Packet Loss (PL), and Mode

ANOVA for Composite Variable NB/S-CTQ					
Source	Sum-of-Squares	df	Mean-Square	F-ratio	Prob
LAB	1.124	2	0.562	0.786	0.4559
RC	39.485	2	19.743	27.610	**0.0000**
PL	64.204	1	64.204	89.789	**0.0000**
MODE	9.736	1	9.736	13.616	**0.0002**
LAB*RC	10.367	4	2.592	3.625	**0.0061**
LAB*PL	4.424	2	2.212	3.093	**0.0457**
LAB*MODE	0.085	2	0.042	0.059	0.9424
RC*PL	0.634	2	0.317	0.444	0.6419
RC*MODE	1.759	2	0.879	1.230	0.2928
PL*MODE	0.511	1	0.511	0.714	0.3982
LAB*RC*PL	2.170	4	0.543	0.759	0.5522
LAB*RC*MODE	2.691	4	0.673	0.941	0.4395
LAB*PL*MODE	0.435	2	0.217	0.304	0.7379
RC*PL*MODE	0.910	2	0.455	0.636	0.5294
LAB*RC*PL*MODE	2.359	4	0.590	0.825	0.5095
Error	797.992	1116	0.715		
Total	938.884	1151			

Table 6 shows that the main effects for *Radio Conditions*, *Packet Loss*, and *Mode* are significant ($p<.05$) for the NB/S-CTQ composite variable as are the interactions of *LAB x RC* and *LAB x PL*. Figure 1 shows the NB/S-CTQ scores with 95% confidence-interval bars for the factors tested in Table 6. The significant interactions of *RC x LAB* and *PL x LAB* indicate that the pattern of scores for the levels of RC and PL were significantly different across the three LAB's.

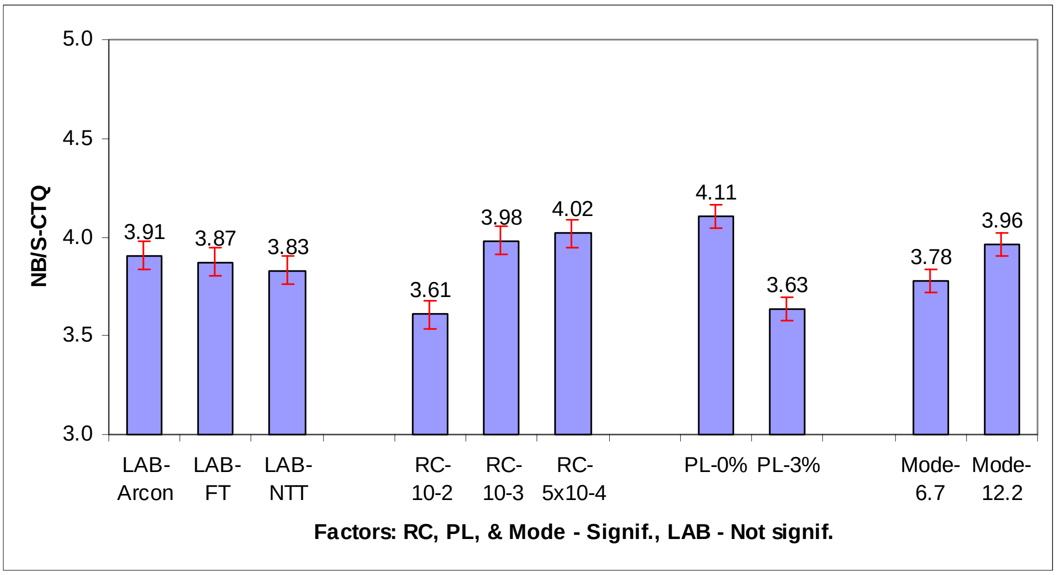

Fig.1 NB/S-CTQ Scores for the Effects of *LAB*, *Radio Conditions*, *Packet Loss*, and *Mode*

Table 7. Results of ANOVA of NB/S-CTQ for the Effects of *LAB*, *Radio Conditions* (RC), *Packet Loss* (PL), and *Delay*

ANOVA for Composite Variable NB/S-CTQ					
Source	Sum-of-Squares	df	Mean-Square	F-ratio	Prob
LAB	3.100	2	1.550	2.412	0.0901
RC	42.537	2	21.269	33.103	**0.0000**
PL	44.724	1	44.724	69.610	**0.0000**
DELAY	4.060	1	4.060	6.320	**0.0121**
LAB*RC	10.471	4	2.618	4.074	**0.0028**
LAB*PL	3.520	2	1.760	2.739	0.0651
LAB*DELAY	0.639	2	0.320	0.497	0.6083
RC*PL	0.101	2	0.051	0.079	0.9243
RC*DELAY	1.009	2	0.505	0.785	0.4563
PL*DELAY	0.373	1	0.373	0.580	0.4465
LAB*RC*PL	1.454	4	0.364	0.566	0.6875
LAB*RC*DELAY	4.464	4	1.116	1.737	0.1395
LAB*PL*DELAY	0.803	2	0.402	0.625	0.5355
RC*PL*DELAY	1.809	2	0.904	1.408	0.2452
LAB*RC*PL*DELAY	4.291	4	1.073	1.670	0.1547
Error	717.030	1116	0.643		
Total	840.386	1151			

The results in Table 7 show that the main effects for *Radio Conditions*, *Packet Loss*, and *Delay* are significant while only one interaction, *LAB x RC*, is significant. Figure 2 shows the NB/S-CTQ scores with 95% confidence-interval bars for the factors tested in Table 7.

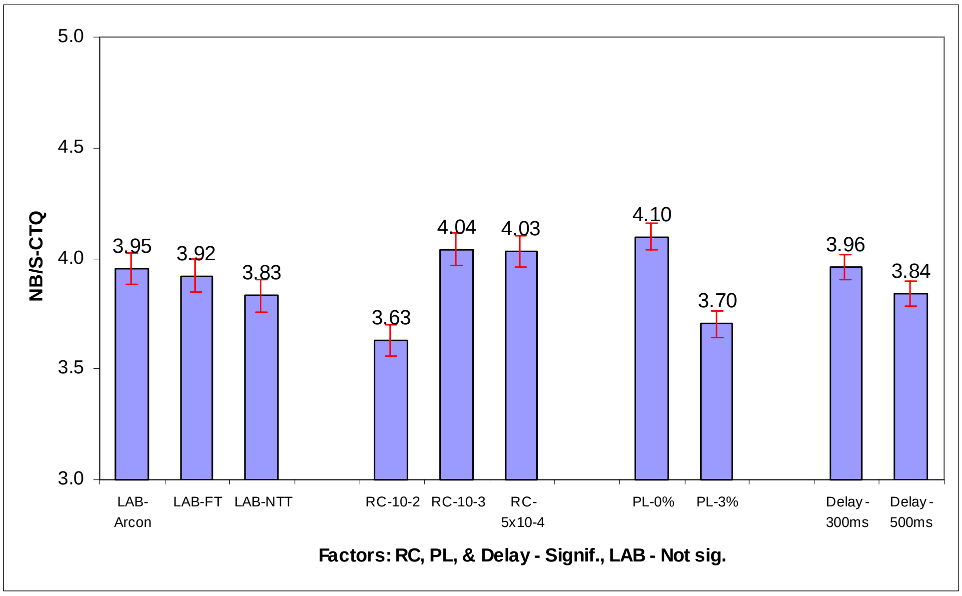

Fig.2 Mean NB/S-CTQ Scores for the Effects of *LAB*, *Radio Conditions*, *Packet Loss*, and *Delay*

4.2 Narrowband Test – Asymmetric Conditions (*Set 2*)

Table 8 shows the inter-correlation matrix for the dependent variables in the NB/ASY conditions. The degree of inter-correlation among the dependent variables suggests that a reduced set of underlying variables accounts for their variation.

Table 8. Inter-correlations Among the Dependent Variables for the NB/ASY Conditions.

WB/A	VQ	US	IA	PC	GQ
VQ	1				
US	**0.6006**	1			
IA	0.3483	**0.5576**	1		
PC	0.4402	**0.6499**	**0.5878**	1	
GQ	**0.6521**	**0.6449**	**0.5644**	**0.6769**	1

Table 9 shows the results of MANOVA for the effects of *COND* for the NB/ASY conditions. The analysis shows significant *COND* effects for all the univariate ANOVA's as well as for the MANOVA. The Chi-square tests of the MANOVA roots shows only a single significant root (1 through 5), indicating that a single underlying variable accounts for the significant variation in the dependent variables for these conditions. The canonical coefficients for this root are used to estimate the composite dependent variable that represents the underlying variable for the NB/ASY conditions. The composite dependent variable (**NB/A-CTQ** for **N**arrow**B**and/**A**symmetric-**C**onversation **T**est **Q**uality) is used to characterize the ratings in the NB/ASY conditions. NB/A-CTQ scores for all conditions and all LAB's in *Set 2* are listed in the Appendix. Equation 2 shows the formula that was used to compute the values of the composite variable, NB/A-CTQ, for characterizing the NB/ASY conditions.

Table 9. Results of MANOVA for *COND* for NB/ASY Conditions.

Univariate ANOVA's for Effect: *COND* (df = 5, 570)					
	VQ	US	IA	PC	GQ
F-Ratio	7.0483	22.3981	5.9907	13.3221	10.2000
Prob	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
MANOVA for effect: *COND*					
Statistic	Value	F-Ratio	df	Prob	
Pillai Trace	0.1849	4.3777	25, 2850	**0.0000**	
Test of Residual Roots				ependent	Canonical
Roots	Chi-Square	df	Prob	Variable	Coefficient
1 through 5	114.889	25	**0.0000**	VQ	0.0894
2 through 5	7.2281	16	0.9686	US	0.3420
3 through 5	2.6968	9	0.9751	IA	0.1851
4 through 5	0.3079	4	0.9893	PC	0.2761
5 through 5	0.039	1	0.8434	GQ	0.1074

$$NB/A\text{-}CTQ = .0894*VQ + .3420*US + .1851 * IA + .2761 * PC + .1074 * GQ$$

Eq.2. Formula used to compute the Conversation Test Quality Score (NB/A-CTQ) for the NB/ASY conditions.

The six NB/ASY conditions are distinguished by two factors. One factor has three levels with each level differing along a number of dimensions – Noise, Packet Loss, Mode, and Delay. These differences are listed in Table 3, but the factor will be referred to in the following analyses by the factor-name, *Noise*, noting that the conditions differ in more dimensions than noise alone. The second factor relates to the source of the noise. The noise is either in the room of the transmitting subject or in the room of the receiving subject. This factor will be referred to as *Room*. Table 10 shows the results of ANOVA for NB/A for the factors of *LAB*, *Noise*, and *Room*.

Table 10. Results of ANOVA of NB/A-CTQ for the Effects of *LAB*, *Noise*, and *Room*

ANOVA for Composite Variable - NB/A-CTQ					
Source	Sum-of-Squares	df	Mean-Square	F-ratio	Prob
LAB	7.091	2	3.546	5.656	**0.0037**
Noise	17.073	2	8.537	13.618	**0.0000**
Room	43.758	1	43.758	69.803	**0.0000**
LAB x Noise	3.280	4	0.820	1.308	0.2657
LAB x Room	2.388	2	1.194	1.905	0.1499
NOISE x Room	3.305	2	1.653	2.636	0.0725
LAB x Noise x Room	1.192	4	0.298	0.476	0.7538
Error	349.802	558	0.627		
Total	427.890	575			

The results of the ANOVA for NB/A-CTQ show that all three factors, *LAB*, *Noise*, and *Room*, are significant, but that none of the interactions are significant. Figure 3 shows the NB/A-CTQ scores with 95% confidence-interval bars for the three factors tested in Table 10.

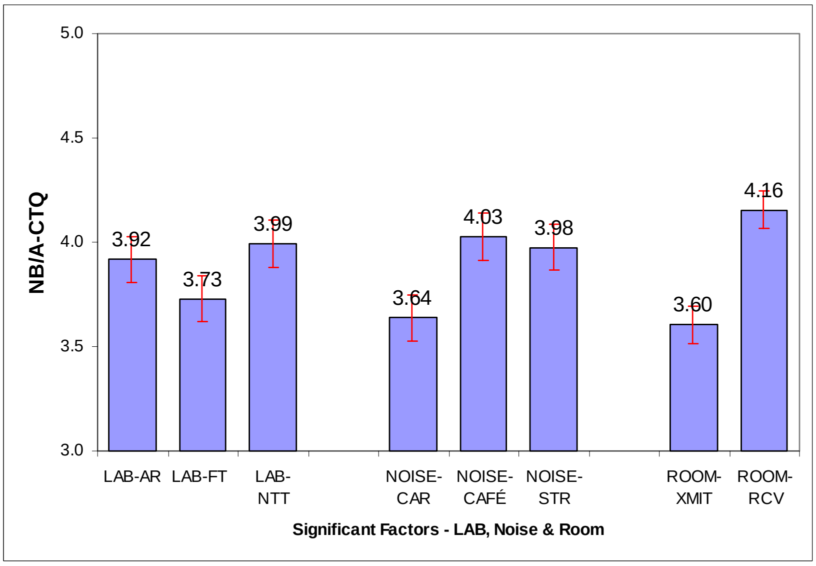

Fig.3 NB/A-CTQ Scores for the Effects of *LAB*, *Noise*, and *Room*

5. Phase I - Wideband Test

Table 12 on the next page shows the 24 test conditions involved in the AMR-WB conversation tests. Also shown in the table are the Mean scores for each rating scale by condition and by listening lab. Each score shown in the table is the average of ratings from 32 subjects.

5.1 Wideband Test – Symmetric Conditions (*Set 3*)

The initial step in the analyses is to examine the inter-correlation among the dependent variables for indications of underlying variables. Table 11 shows the inter-correlation matrix of the dependent variables for the WB/SYM conditions. Absolute values of correlation above .50 have been bolded in the table. The table shows a high degree of inter-correlation among the dependent variables indicating the presence of a reduced set of significant underlying variables.

Table 11. Intercorrelations Among the Dependent Variables for the WB/SYM Conditions.

WB/S	VQ	US	IA	PC	GQ
VQ	1				
US	**0.6559**	1			
IA	0.4902	**0.5121**	1		
PC	**0.5946**	**0.5854**	**0.5075**	1	
GQ	**0.7888**	**0.6832**	**0.5497**	**0.6612**	1

The second step in the analysis is designed to determine how many underlying variables account for the variance in the five dependent variables. MANOVA for the effects of *COND* was conducted on the WB/SYM data – conditions 1-18. Table 13 summarizes the results of the analysis. The top section shows the analysis for the main effect of *COND.* This section includes the results of the univariate ANOVA's for each of the five dependent variables followed by the results of the MANOVA. In the table we can see that the *COND* main effect is highly significant for each of the five individual dependent variables in the univariate ANOVA's as well as for the combination of dependent variables in the MANOVA.

The bottom section of the table shows the Chi-square test of the MANOVA roots or underlying variables extracted from the five dependent variables. In Table 13, only the first root (1 through 5) is significant, indicating that a single underlying variable accounts for the significant variation in the dependent variables for these conditions. The canonical coefficients shown in the table are used to estimate the composite dependent variable that represents this root or underlying variable. The composite dependent variable (**WB/S-CTQ** for **W**ide**B**and/**S**ymmetric**-C**onversation **T**est **Q**uality) is computed and used in the third step – ANOVA's to test and characterize the factors of interest in the Wideband/SYM conditions. WB/S-CTQ scores for all conditions and all LAB's for *Set 3* are listed in the Appendix. Equation 3 shows the formula that was used to compute the values of the composite variable, WB/S-CTQ, for characterizing the WB/SYM conditions.

Table 12. Test Conditions and Mean Scores for each LAB for the Wideband Experiment

Wideband - Experimental Parameters							Voice Quality			Understanding			Interaction			Perception			Global Quality		
Cond	**Rm-A**	**Rm-B**	**RC**	**PL**	**Mode**	**Del**	**Arcon**	**FT**	**NTT**	**Arcon**	**FT**	**NTT**	**Arcon**	**FT**	**NTT**	**Arcon**	**FT**	**NTT**	**Arcon**	**FT**	**NTT**
1	Quiet	Quiet	10^{-2}	0	12.65	RoHC	4.094	4.219	3.844	4.375	4.406	4.344	4.250	4.125	4.531	4.469	4.250	4.313	4.094	4.063	3.750
2	Quiet	Quiet	10^{-2}	0	12.65	-	4.000	4.438	3.969	4.219	4.844	4.531	4.063	4.375	4.719	4.281	4.406	4.313	3.781	4.313	4.000
3	Quiet	Quiet	10^{-2}	0	15.85	RoHC	4.125	4.281	4.125	4.375	4.500	4.688	4.313	4.188	4.656	4.500	4.281	4.594	4.281	4.094	4.219
4	Quiet	Quiet	10^{-2}	3	12.65	RoHC	3.875	3.719	3.719	4.188	4.094	4.031	3.906	4.094	4.281	4.344	3.844	4.063	3.875	3.531	3.594
5	Quiet	Quiet	10^{-2}	3	12.65	-	3.625	3.750	3.719	4.063	3.875	4.063	3.906	3.813	4.375	4.219	3.875	4.156	3.719	3.625	3.688
6	Quiet	Quiet	10^{-2}	3	15.85	RoHC	3.906	3.969	3.844	4.188	4.438	4.281	4.063	4.125	4.531	4.219	4.031	4.281	3.844	3.844	3.813
7	Quiet	Quiet	10^{-3}	0	12.65	RoHC	4.219	4.375	4.000	4.500	4.563	4.688	4.250	4.219	4.750	4.688	4.563	4.625	4.281	4.188	4.000
8	Quiet	Quiet	10^{-3}	0	12.65	-	4.063	4.469	4.063	4.281	4.688	4.719	4.219	4.250	4.688	4.313	4.469	4.688	4.156	4.250	4.219
9	Quiet	Quiet	10^{-3}	0	15.85	RoHC	3.875	4.625	3.938	4.344	4.750	4.531	4.156	4.375	4.750	4.438	4.500	4.531	3.938	4.375	4.063
10	Quiet	Quiet	10^{-3}	3	12.65	RoHC	3.969	4.313	3.969	4.188	4.500	4.406	4.125	4.125	4.656	4.469	4.188	4.531	4.031	3.938	3.969
11	Quiet	Quiet	10^{-3}	3	12.65	-	4.031	4.250	3.750	4.406	4.563	4.344	4.094	4.156	4.500	4.688	4.156	4.281	3.938	3.969	3.813
12	Quiet	Quiet	10^{-3}	3	15.85	RoHC	4.031	4.031	3.906	4.344	4.375	4.469	4.156	4.094	4.656	4.281	4.219	4.375	4.000	3.813	3.906
13	Quiet	Quiet	5×10^{-4}	0	12.65	RoHC	4.094	4.344	4.188	4.344	4.625	4.656	4.156	4.219	4.813	4.594	4.531	4.625	4.000	4.125	4.219
14	Quiet	Quiet	5×10^{-4}	0	12.65	-	4.094	4.594	4.063	4.469	4.813	4.594	4.156	4.438	4.750	4.500	4.563	4.563	4.156	4.375	4.094
15	Quiet	Quiet	5×10^{-4}	0	15.85	RoHC	4.188	4.469	4.031	4.469	4.688	4.656	4.438	4.313	4.781	4.594	4.469	4.594	4.375	4.156	4.063
16	Quiet	Quiet	5×10^{-4}	3	12.65	RoHC	3.938	3.969	3.906	4.250	4.531	4.406	4.000	3.969	4.625	4.250	4.156	4.375	3.844	3.875	4.000
17	Quiet	Quiet	5×10^{-4}	3	12.65	-	4.063	4.188	3.875	4.250	4.469	4.406	4.188	4.125	4.469	4.594	4.281	4.281	4.094	3.938	3.844
18	Quiet	Quiet	5×10^{-4}	3	15.85	RoHC	4.125	4.344	3.813	4.375	4.531	4.563	4.313	4.063	4.594	4.594	4.188	4.438	4.094	3.906	3.813
19	Car	Quiet	5×10^{-4}	3	12.65	RoHC	3.500	4.094	2.969	3.594	3.625	3.000	3.969	3.656	3.469	4.031	3.375	3.188	3.813	3.344	2.781
20	Quiet	Car	5×10^{-4}	3	12.65	RoHC	3.969	4.031	3.781	4.094	4.344	4.375	4.188	3.969	4.500	4.344	3.875	4.313	4.031	3.750	3.844
21	Cafeteria	Quiet	5×10^{-4}	0	12.65	-	3.750	4.375	3.656	3.781	4.375	3.875	3.938	4.094	4.063	4.313	3.969	3.844	3.813	3.813	3.344
22	Quiet	Cafeteria	5×10^{-4}	0	12.65	-	4.156	4.563	4.125	4.469	4.719	4.688	4.250	4.250	4.719	4.594	4.438	4.594	4.125	4.156	4.219
23	Street	Quiet	5×10^{-4}	0	15.85	RoHC	3.813	4.313	3.719	3.625	3.906	4.219	4.125	3.750	4.188	4.406	3.344	4.188	4.125	3.406	3.594
24	Quiet	Street	5×10^{-4}	0	15.85	RoHC	3.938	4.438	4.156	4.313	4.594	4.688	4.188	4.031	4.656	4.563	4.250	4.688	4.031	4.094	4.156

Rm-A/Rm-B (Noise environment) RC (Radio Conditions) PL (% Packet Loss) Mode (Bit rate in kbps) RoHC

Table 13. Results of MANOVA for *COND* for WB/SYM Conditions.

Univariate ANOVA's for Effect *COND* (df = 17, 1710)					
Dep.Var.	VQ	US	IA	PC	GQ
F-Rato	3.352	4.358	2.836	3.977	4.141
Prob.	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
MANOVA for Effect: *COND*					
Statistic	Value	F-Statistic	df	Prob	
Pillai Trace	0.076	1.549	85, 8550	**0.0010**	
Test of Residual Roots				Dep.Var.	Canon.Coeff. for Root 1-5
Roots	Chi-Square	df	Prob		
1 through 5	**132.5613**	**85**	**0.0007**	VQ	0.0685
2 through 5	43.3215	64	0.9779	US	0.3519
3 through 5	25.1721	45	0.9926	IA	0.1612
4 through 5	8.5498	28	0.9998	PC	0.2619
5 through 5	2.3528	13	0.9994	GQ	0.1565

$$WB/S\text{-}CTQ = .0685*VQ + .3519*US + .1612 * IA + .2619 * PC + .1565 * GQ$$

Eq.3. Formula used to compute the Conversation Test Quality Score (WB/S-CTQ) for the WB/SYM conditions

The SYM conditions in the WB experiment are categorized by four experimental factors:

1. Radio conditions – 10^{-2}, 10^{-3}, and $5x10^{-4}$
2. Packet Loss – 0% and 3%
3. AMR-NB mode or bit rate – 12.65 kbps and 15.85 kbps
4. RoHC

These conditions are assigned to two factorial experimental designs for analysing the effects through ANOVA of three of these factors. Table 14a shows the allocation of the 12 conditions used to evaluate the effects of Radio Conditions, Packet Loss, and Mode – with RoHC held constant. Table 14b shows the allocation of the 12 conditions used to evaluate the effects of Radio Conditions, Packet Loss, and RoHC – Mode held constant at 12.65kbps.

Table 14a WB/SYM: Factorial Design for the Effects of Radio Cond., Packet Loss, and Mode.

No Noise - RoHC				
12.65kbps / 0% PL			**12.65 kbps / 3% PL**	
RC	Cond.#		RC	Cond.#
10^{-2}	1		10^{-2}	4
10^{-3}	7		10^{-3}	10
$5x10^{-4}$	13		$5x10^{-4}$	16
15.85 kbps / 0% PL			**15.85 kbps / 3% PL**	
RC	Cond.#		RC	Cond.#
10^{-2}	3		10^{-2}	6
10^{-3}	9		10^{-3}	12
$5x10^{-4}$	15		$5x10^{-4}$	18

Table 14b – WB/SYM: Factorial Design for the Effects of Radio Cond., Packet Loss, and RoHC

No Noise - 12.65 kbps				
RoHC / 0% PL			**RoHc / 3% PL**	
RC	Cond.#		RC	Cond.#
10^{-2}	1		10^{-2}	4
10^{-3}	7		10^{-3}	10
$5x10^{-4}$	13		$5x10^{-4}$	16
No RoHC / 0% PL			**No RoHC / 3% PL**	
RC	Cond.#		RC	Cond.#
10^{-2}	2		10^{-2}	5
10^{-3}	8		10^{-3}	11
$5x10^{-4}$	14		$5x10^{-4}$	17

The composite dependent variable, WB/S-CTQ, was computed for the WB/SYM conditions and subjected to factorial ANOVA for the two experimental designs shown in Tables 14a and 14b. The results of the ANOVA's are shown in Tables 15 and 16, respectively.

Table 15. Results of ANOVA of WB/S-CTQ for the Effects of *Lab*, *Radio Conditions* (RC), *Packet Loss* (PL), and *Mode*

ANOVA for Composite Variable WB/S-CTQ					
Source	Sum-of-Squares	df	Mean-Square	F-ratio	Prob
LAB	6.525	2	3.263	6.521	**0.0015**
RC	6.899	2	3.450	6.895	**0.0011**
PL	14.332	1	14.332	28.646	**0.0000**
MODE	1.408	1	1.408	2.814	0.0938
LAB*RC	0.975	4	0.244	0.487	0.7450
LAB*PL	0.234	2	0.117	0.234	0.7914
LAB*MODE	0.044	2	0.022	0.044	0.9571
RC*PL	0.355	2	0.177	0.355	0.7015
RC*MODE	1.959	2	0.979	1.958	0.1417
PL*MODE	0.087	1	0.087	0.173	0.6776
LAB*RC*PL	0.452	4	0.113	0.226	0.9242
LAB*RC*MODE	2.245	4	0.561	1.122	0.3446
LAB*PL*MODE	0.109	2	0.054	0.109	0.8972
RC*PL*MODE	0.014	2	0.007	0.014	0.9863
LAB*RC*PL*MODE	0.997	4	0.249	0.498	0.7372
Error	558.336	1116	0.500		
Total	594.969	1151			

Table 15 shows that the main effects for *LAB*, *Radio Conditions*, and *Packet Loss* are significant for the WB/S-CTQ composite variable. The factor *Mode* is not significant nor are any of the interactions. Figure 4 shows the WB/S-CTQ scores with 95% confidence-interval bars for the factors tested in Table 15.

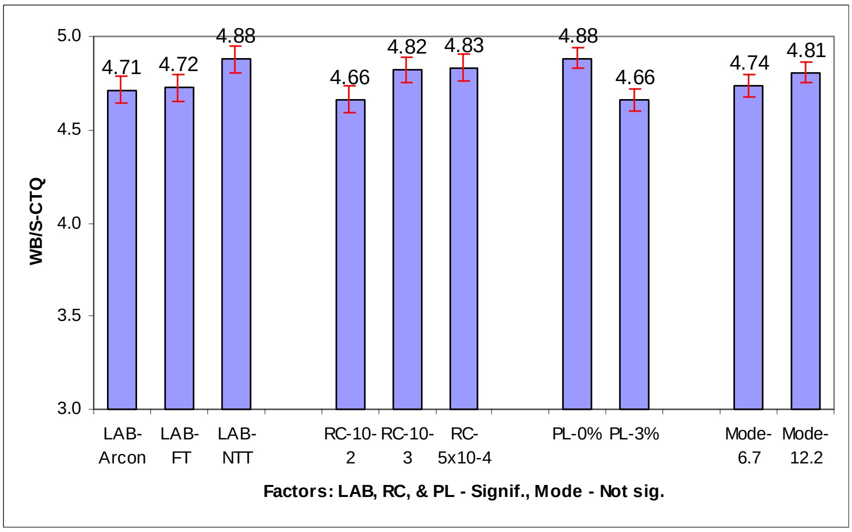

Fig.4. WB/S-CTQ Scores for the Effects of *LAB*, *Radio Conditions*, *Packet Loss*, and *Mode*

Table 16. Results of ANOVA of WB/S-CTQ for the Effects of *LAB*, *Radio Conditions* (RC), *Packet Loss* (PL), and *RoHC*

ANOVA for Composite Variable WB/S-CTQ					
Source	Sum-of-Squares	df	Mean-Square	F-ratio	Prob
LAB	5.243	2	2.621	5.102	**0.0062**
RC	13.590	2	6.795	13.225	**0.0000**
PL	19.414	1	19.414	37.785	**0.0000**
ROHC	0.073	1	0.073	0.142	0.7061
LAB*RC	0.801	4	0.200	0.390	0.8161
LAB*PL	2.456	2	1.228	2.390	0.0921
LAB*ROHC	0.698	2	0.349	0.680	0.5071
RC*PL	1.566	2	0.783	1.524	0.2183
RC*ROHC	0.244	2	0.122	0.237	0.7889
PL*ROHC	0.107	1	0.107	0.207	0.6490
LAB*RC*PL	0.981	4	0.245	0.477	0.7524
LAB*RC*ROHC	1.900	4	0.475	0.924	0.4490
LAB*PL*ROHC	2.022	2	1.011	1.968	0.1402
RC*PL*ROHC	0.498	2	0.249	0.485	0.6160
LAB*RC*PL*ROHC	0.847	4	0.212	0.412	0.8000
Error	573.401	1116	0.514		
Total	623.840	1151			

The results in Table 16 show that the main effects for *LAB*, *Radio Conditions*, and *Packet Loss* are significant. The factor *RoHC* is not significant nor are any of the interactions. Figure 5 shows the WB/S-CTQ scores with 95% confidence-interval bars for the factors tested in Table 16.

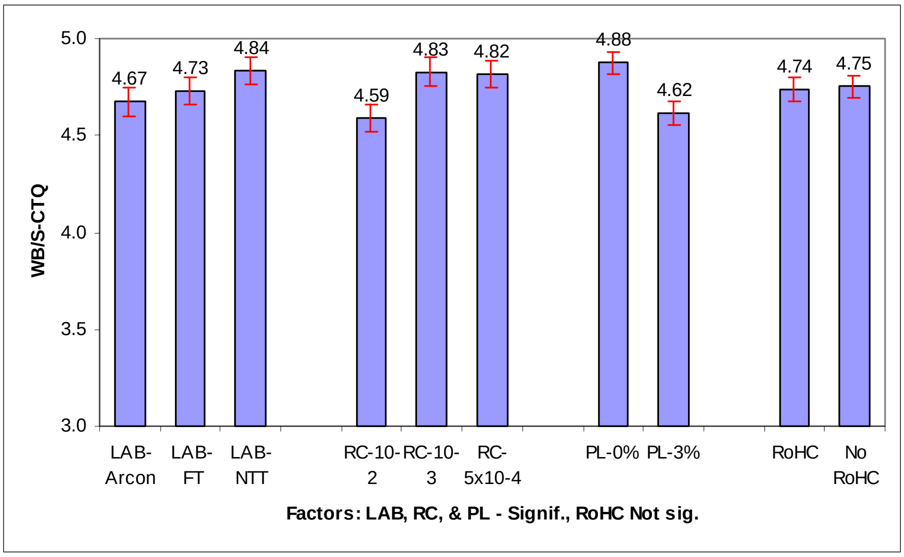

Fig.5 WB/S-CTQ Scores for the Effects of *LAB*, *Radio Conditions*, *Packet Loss*, and *RoHC*

5.2 Wideband Test – Asymmetric Conditions (*Set 4*)

Table 17 shows the inter-correlation matrix for the dependent variables in the WB/ASY conditions. The high degree of inter-correlation shown in the table suggests that a reduced set of underlying variables accounts for the variation in the five dependent variables.

Table 17. Inter-correlations Among the Dependent Variables for the WB/ASY Conditions.

WB/S	VQ	US	IA	PC	GQ
VQ	1				
US	**0.674**	1			
IA	**0.5568**	**0.6369**	1		
PC	**0.5466**	**0.6506**	**0.6648**	1	
GQ	**0.7217**	**0.7256**	**0.6917**	**0.7255**	1

Table 18 shows the results of MANOVA for the effects of *COND* for the WB/ASY conditions. The analysis shows significant *COND* effects for all the univariate ANOVA's as well as for the MANOVA. The Chi-square tests of the MANOVA roots shows only a single significant root (1 through 5), indicating that a single underlying variable accounts for the significant variation in the dependent variables for these conditions. The canonical coefficients for this root were used to compute the composite dependent variable that represents the underlying variable for the WB/Asymmetric conditions. The composite dependent variable (**WB/A-CTQ** for **W**ide**B**and/**A**symmetric**-C**onversation **T**est **Q**uality) is used to characterize the ratings in the WB/ASY conditions. WB/A-CTQ scores for all conditions and all LAB's for *Set 4* are listed in the Appendix. Equation 4 shows the formula that was used to compute the values of the composite variable, WB/A-CTQ, for characterizing the WB/ASY conditions.

Table 18. Results of MANOVA for *COND* for WB/ASY Conditions

Univariate ANOVA's for Effect: *COND* (df = 5, 570)					
	VQ	US	IA	PC	GQ
F-Ratio	8.3809	21.6252	8.1610	14.0989	10.9652
Prob	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
MANOVA for effect: *COND*					
Statistic	Value	F-Ratio	df	Prob	
Pillai Trace	0.1912	4.5331	25, 2850	**0.0000**	
Test of Residual Roots				ependent	Canonical
Roots	Chi-Square	df	Prob	Variable	Coefficient
1 through 5	118.4457	25	**0.0000**	VQ	-0.0970
2 through 5	11.1921	16	0.7975	US	0.8979
3 through 5	3.7996	9	0.9241	IA	-0.1103
4 through 5	1.8535	4	0.7627	PC	0.4136
5 through 5	0.0001	1	0.9920	GQ	-0.1042

WB/A-CTQ = -.0970*VQ + .8979*US - .1103 * IA + .4136 * PC - .1042 * GQ

Eq. 4. Formula used to compute the Conversation Test Quality Score (WB/ACTQ) for the WB/ASY conditions.

The six WB/ASY conditions are distinguished by two factors. One factor has three levels with each level differing along a number of dimensions – Noise, Packet Loss, Mode, and RoHC. These differences are listed in Table 12 but the factor will be referred to in the following analyses by the factor-name, *Noise*, noting that the conditions differ in more dimensions than noise alone. The second factor relates to the source of the noise and has two levels. The noise is either in the room of the transmitting subject or in the room of the receiving subject. This factor is referred to as *Room* in the following analyses. Table 19 shows the results of ANOVA for WB/A-CTQ for the factors of *LAB*, *Noise*, and *Room*.

Table 19. Results of ANOVA of WB/A-CTQ for the Effects of *LAB*, *Noise*, and *Room*

ANOVA for Composite Variable - WB/A-CTQ					
Source	Sum-of-Squares	df	Mean-Square	F-ratio	Prob
LAB	6.058	2	3.029	3.804	**0.0229**
NOISE	20.415	2	10.207	12.818	**0.0000**
ROOM	63.102	1	63.102	79.238	**0.0000**
LAB*NOISE	8.150	4	2.038	2.559	**0.0378**
LAB*ROOM	3.159	2	1.580	1.984	0.1386
NOISE*ROOM	2.188	2	1.094	1.374	0.2539
LAB*NOISE*ROOM	6.201	4	1.550	1.947	0.1013
Error	444.367	558	0.796		
Total	553.640	575			

The results of the ANOVA for WB/A-CTQ show that all three factors, *LAB*, *Noise*, and *Room*, are significant but only one of the interactions, *LAB x Noise* is significant. Figure 6 shows the WB/A-CTQ scores with 95% confidence-interval bars for the three factors tested in Table 19.

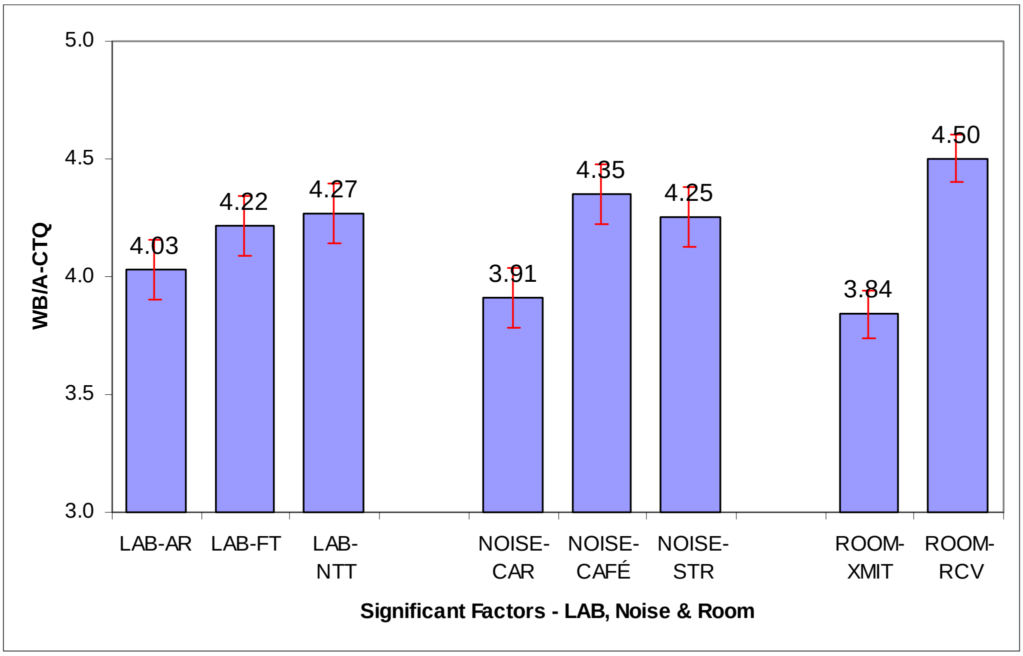

Fig.6 WB/A-CTQ Scores for the Effects of *LAB*, *Noise*, and *Room*

6. Phase II - ITU-T Codec Tests (*Set 5*)

Table 20 shows the test conditions involved in the conversation tests designed to compare the performance of standardized ITU-T codecs in packet switched networks. The test involves eight codecs and two levels of packet loss, 0% and 3%. Scores are shown for each of the five dependent variables by Condition and by Language (Language is referred to by factor-name *LAB* in the following analyses). Each score shown in the table is the average of ratings from 32 listeners.

Table 20. Test Conditions and Scores for each Condition and Lab (Language) for the Codec (Phase II) Experiment

Phase II Experiment			VQ		US		IA		PC		GQ	
			Voice Quality		Understand		Interaction		Perception		Global Quality	
Cond	**PL**	**Codec, Mode**	**FR**	**AB**	**FR**	**AB**	**FR**	**AB**	**FR**	**AB**	**FR**	**AB**
1	0%	AMR-NB, 6.7kbit/s	4.250	3.938	4.438	4.281	4.125	4.063	4.125	4.281	4.094	3.781
2	0%	AMR-NB, 12.2kbit/s	4.406	4.125	4.563	4.375	4.281	3.906	4.156	4.094	4.063	3.781
3	0%	AMR-WB, 12.65kbit/s	4.344	4.406	4.594	4.500	4.344	4.125	4.313	4.375	4.219	4.000
4	0%	AMR-WB, 15.85kbit/s	4.500	4.438	4.719	4.469	4.344	4.125	4.219	4.250	4.313	4.000
5	0%	G. 723., 6.4 kbit/s	4.219	4.094	4.469	4.094	4.063	3.875	4.031	4.063	3.938	3.688
6	0%	G.729, 8kbit/s	4.156	4.344	4.469	4.281	3.969	3.938	4.188	4.063	3.875	3.813
7	0%	G.722, 64 kbit/s + plc	4.406	4.281	4.625	4.438	4.188	4.094	4.250	4.031	4.125	3.719
8	0%	G.711 + plc	4.406	4.438	4.563	4.438	4.125	4.031	4.125	4.313	4.094	3.875
9	3%	AMR-NB, 6.7kbit/s	3.781	3.656	4.000	4.031	3.813	3.563	3.781	3.656	3.750	3.250
10	3%	AMR-NB, 12.2 kbit/s	4.156	3.938	4.375	4.281	4.000	3.719	3.938	3.969	3.656	3.594
11	3%	AMR-WB, 12.65kbit/s	4.375	4.094	4.375	4.344	4.063	3.844	4.094	4.000	4.063	3.813
12	3%	AMR-WB, 15.85kbit/s	4.125	4.094	4.531	4.344	4.031	3.938	4.031	4.000	4.063	3.656
13	3%	G. 723.1, 6.4 kbit/s	3.906	3.531	4.438	3.906	3.906	3.750	3.656	3.625	3.656	3.344
14	3%	G.729, 8kbit/s	4.063	3.906	4.344	4.031	4.031	3.781	3.938	4.000	3.719	3.531
15	3%	G.722, 64 kbit/s + plc	4.438	4.438	4.531	4.500	4.125	4.094	4.156	4.125	4.063	3.969
16	3%	G.711 + plc	4.438	4.438	4.563	4.500	4.250	4.031	4.125	4.375	4.094	4.031

PL – Packet Loss LAB: FR–French, AB-Arabic

Table 21 shows the inter-correlation matrix for the dependent variables in the Phase II experiment. The moderate degree of inter-correlation shown in the table suggests that a reduced set of underlying variables may account for the variation in the five dependent variables.

Table 21. Inter-correlations Among the Dependent Variables for the Codec Conditions.

WB/S	VQ	US	IA	PC	GQ
VQ	1				
US	0.4684	1			
IA	**0.5025**	**0.5352**	1		
PC	0.4765	0.4160	**0.5060**	1	
GQ	**0.5987**	**0.5318**	**0.6204**	**0.6053**	1

Table 22 shows the results of MANOVA for the effects of *COND* for the Phase II experiment. The analysis shows significant *COND* effects for all the univariate ANOVA's as well as for the MANOVA. The Chi-square tests of the MANOVA roots shows only a single significant root (1 through 5), indicating that a single underlying variable accounts for the significant variation in the dependent variables for these conditions. The canonical coefficients for this root were used to compute the composite dependent variable that represents the underlying variable for the Phase II conditions. The composite dependent variable (**Ph2-CTQ** for **Ph**ase**2-C**onversation **T**est **Q**uality) is computed and used to characterize the ratings in the Phase II experiment. Ph2-CTQ scores for all conditions and all LAB's for *Set 5* are listed in the Appendix. Equation 5 shows the formula that was used to compute the values of the composite variable, Ph2-CTQ, for characterizing the Phase II conditions.

Table 22. Results of MANOVA for *COND* for the Phase II Conditions.

Univariate ANOVA's for Effect: *COND* (df = 15, 1008)					
	VQ	US	IA	PC	GQ
F-Ratio	5.6360	2.4296	2.6802	2.5353	4.2549
Prob	**0.0000**	**0.0017**	**0.0005**	**0.0010**	**0.0000**
MANOVA for effect: *COND*					
Statistic	Value	F-Ratio	df	Prob	
Pillai Trace	0.1168	1.607	75, 5040	**0.0008**	
Test of Residual Roots				Dependent	Canonical
Roots	Chi-Square	df	Prob	Variable	Coefficient
1 through 5	**122.2608**	**75**	**0.0005**	VQ	0.5995
2 through 5	32.4394	56	0.9951	US	0.0860
3 through 5	19.2889	39	0.9966	IA	-0.0092
4 through 5	10.4532	24	0.9924	PC	0.0459
5 through 5	2.5753	11	0.9952	GQ	0.2778

$$\text{Ph2-CTQ} = .5995*VQ + .0860*US - .0092 * IA + .0459 * PC + .2778 * GQ$$

Eq.5. Formula used to compute the Conversation Test Quality Score (Ph2-CTQ) for the Phase II conditions

The 16 Phase II conditions are distinguished by two factors, *Codec* and *Packet Loss*. Table 23 shows the results of ANOVA for Ph2-CTQ for these factors.

Table 23. Results of ANOVA of Ph2-CTQ for the Effects of *Codec* and *Packet Loss*

ANOVA for Composite Variable - Ph2-CTQ					
Source	Sum-of-Squares	df	Mean-Square	F-ratio	Prob
LAB	5.708	1	5.708	11.930	**0.0006**
CODEC	27.436	7	3.919	8.192	**0.0000**
PL	10.330	1	10.330	21.592	**0.0000**
LAB*CODEC	1.698	7	0.243	0.507	0.8297
LAB*PL	0.065	1	0.065	0.136	0.7123
CODEC*PL	7.088	7	1.013	2.116	**0.0395**
LAB*CODEC*PL	1.454	7	0.208	0.434	0.8811
Error	474.606	992	0.478		
Total	528.384	1023			

The results of the ANOVA for Ph2-CTQ show that all three factors, *LAB*, *Codec*, and *Packet Loss*, are significant as well as the interaction *Codec x Packet Loss*. Figure 7 shows the Ph2-CTQ scores with 95% confidence-interval bars for the factors tested in Table 23. Figure 8 illustrates the interaction of *Codec x Packet Loss*.

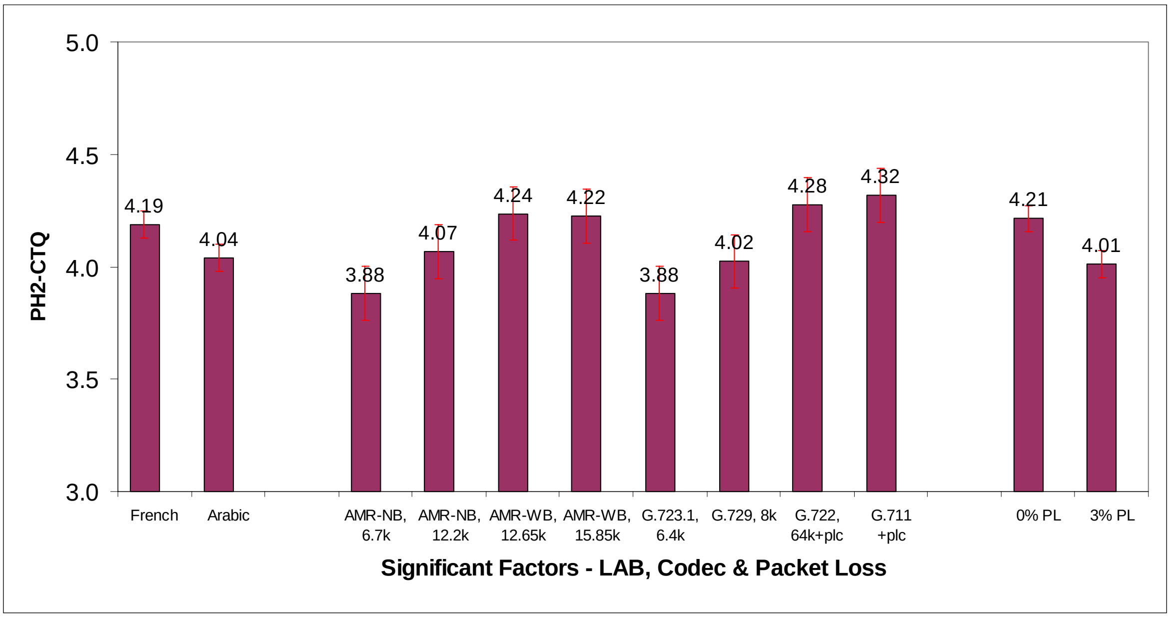

Fig.7 Ph2-CTQ Scores for the Effects of *LAB*, *Codec*, and *Packet Loss*

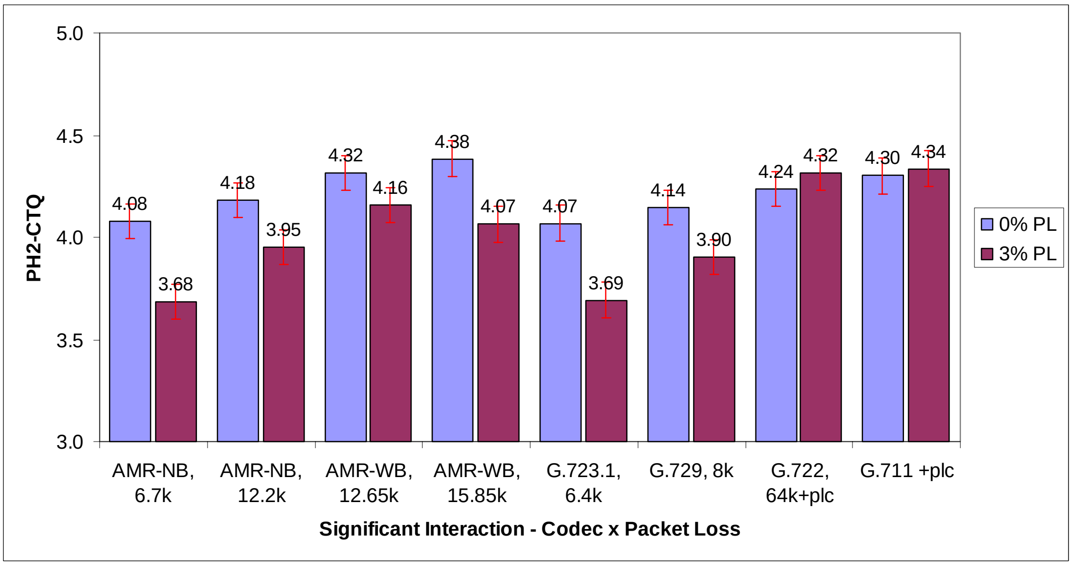

Fig.8 Ph2-CTQ Scores Showing the Interaction of Factors *Codec* and *Packet Loss*

7. Summary

For each of the five sets of conditions in the Packet-Switched Conversation Tests, analysis by MANOVA revealed a single underlying variable that accounts for the significant variation in the five opinion rating scales, VQ, US, IA, PC, and GQ. Conversation Test Quality (CTQ) scores were computed for each set of conditions. The CTQ scores were analysed through ANOVA to characterize the conditions involved in the Conversation Tests.

8. References

S4-030564 Test Plan for the AMR Narrow-Band Packet Switched Conversation Test

S4-030565 Test Plan for the AMR Wide-Band Packet Switched Conversation Test

S4-030747 Test plan for Packet Switched Conversation Test. Comparison of quality offered by different speech coders.

S4-030818 Proposed Test Plan for Global Analysis of PSS Conversation Tests

Appendix

Conversation Test Composite Dependent Variable Scores by Condition and Lab

Set 1 - Narrowband/SYM Experimental Parameters							NB/S-CTQ Scores			
Cond	**Rm-A**	**Rm-B**	**RC**	**PL**	**Mode**	**Del**	**Arcon**	**FT**	**NTT**	**Average**
1	Quiet	Quiet	10^{-2}	0	6.7	300	3.801	3.730	3.792	3.774
2	Quiet	Quiet	10^{-2}	0	12.2	500	3.884	3.852	3.524	3.753
3	Quiet	Quiet	10^{-2}	0	12.2	300	4.047	3.728	3.906	3.893
4	Quiet	Quiet	10^{-2}	3	6.7	300	3.491	2.919	3.221	3.210
5	Quiet	Quiet	10^{-2}	3	12.2	500	3.682	2.994	3.279	3.318
6	Quiet	Quiet	10^{-2}	3	12.2	300	3.669	3.376	3.619	3.555
7	Quiet	Quiet	10^{-3}	0	6.7	300	4.094	4.222	3.959	4.092
8	Quiet	Quiet	10^{-3}	0	12.2	500	4.037	4.278	4.168	4.161
9	Quiet	Quiet	10^{-3}	0	12.2	300	4.305	4.656	3.936	4.299
10	Quiet	Quiet	10^{-3}	3	6.7	300	3.634	3.603	3.638	3.625
11	Quiet	Quiet	10^{-3}	3	12.2	500	3.828	3.820	3.749	3.799
12	Quiet	Quiet	10^{-3}	3	12.2	300	3.730	4.059	3.940	3.910
13	Quiet	Quiet	5×10^{-4}	0	6.7	300	4.197	4.445	4.069	4.237
14	Quiet	Quiet	5×10^{-4}	0	12.2	500	4.145	4.296	4.008	4.150
15	Quiet	Quiet	5×10^{-4}	0	12.2	300	4.256	4.373	4.378	4.336
16	Quiet	Quiet	5×10^{-4}	3	6.7	300	3.733	3.691	3.751	3.725
17	Quiet	Quiet	5×10^{-4}	3	12.2	500	3.927	3.984	3.706	3.872
18	Quiet	Quiet	5×10^{-4}	3	12.2	300	3.920	3.648	3.753	3.774

Set 2 - Narrowband/ASY Experimental Parameters							NB/A-CTQ Scores			
Cond	**Rm-A**	**Rm-B**	**RC**	**PL**	**Mode**	**Del**	**Arcon**	**FT**	**NTT**	**Average**
19	Car	Quiet	5×10^{-4}	3	12.2	300	3.438	2.929	3.429	3.265
20	Quiet	Car	5×10^{-4}	3	12.2	300	4.077	3.750	4.205	4.011
21	Cafeteria	Quiet	5×10^{-4}	0	6.7	300	3.838	3.579	3.858	3.759
22	Quiet	Cafeteria	5×10^{-4}	0	6.7	300	4.209	4.359	4.307	4.292
23	Street	Quiet	5×10^{-4}	0	12.2	500	3.772	3.579	4.015	3.789
24	Quiet	Street	5×10^{-4}	0	12.2	500	4.172	4.178	4.142	4.164

Set 3 - Wideband/SYM - Experimental Parameters							WB/S-CTQ Scores			
Cond	**Rm-A**	**Rm-B**	**RC**	**PL**	**Mode**	**Del**	**Arcon**	**FT**	**NTT**	**Average**
1	Quiet	Quiet	10^{-2}	0	12.65	RoHC	4.764	4.682	4.729	4.725
2	Quiet	Quiet	10^{-2}	0	12.65	-	4.551	5.011	4.896	4.819
3	Quiet	Quiet	10^{-2}	0	15.85	RoHC	4.817	4.751	5.054	4.874
4	Quiet	Quiet	10^{-2}	3	12.65	RoHC	4.528	4.353	4.443	4.441
5	Quiet	Quiet	10^{-2}	3	12.65	-	4.418	4.210	4.519	4.382
6	Quiet	Quiet	10^{-2}	3	15.85	RoHC	4.531	4.603	4.705	4.613
7	Quiet	Quiet	10^{-3}	0	12.65	RoHC	4.903	4.873	5.052	4.943
8	Quiet	Quiet	10^{-3}	0	12.65	-	4.685	4.919	5.099	4.901
9	Quiet	Quiet	10^{-3}	0	15.85	RoHC	4.693	5.006	4.972	4.890
10	Quiet	Quiet	10^{-3}	3	12.65	RoHC	4.642	4.685	4.880	4.736
11	Quiet	Quiet	10^{-3}	3	12.65	-	4.769	4.716	4.723	4.736
12	Quiet	Quiet	10^{-3}	3	15.85	RoHC	4.662	4.613	4.855	4.710
13	Quiet	Quiet	5×10^{-4}	0	12.65	RoHC	4.744	4.882	5.090	4.905
14	Quiet	Quiet	5×10^{-4}	0	12.65	-	4.796	5.065	5.011	4.957
15	Quiet	Quiet	5×10^{-4}	0	15.85	RoHC	4.929	4.921	5.049	4.966
16	Quiet	Quiet	5×10^{-4}	3	12.65	RoHC	4.549	4.639	4.836	4.675
17	Quiet	Quiet	5×10^{-4}	3	12.65	-	4.727	4.697	4.746	4.723
18	Quiet	Quiet	5×10^{-4}	3	15.85	RoHC	4.813	4.677	4.880	4.790

Set 4 - Wideband/ASY - Experimental Parameters							WB/A-CTQ Scores			
Cond	**Rm-A**	**Rm-B**	**RC**	**PL**	**Mode**	**Del**	**Arcon**	**FT**	**NTT**	**Average**
19	Car	Quiet	5×10^{-4}	3	12.65	-	3.694	3.618	3.168	3.493
20	Quiet	Car	5×10^{-4}	3	12.65	RoHC	4.141	4.322	4.526	4.330
21	Cafeteria	Quiet	5×10^{-4}	0	12.65	RoHC	3.828	4.348	4.008	4.062
22	Quiet	Cafeteria	5×10^{-4}	0	12.65	22	4.469	4.650	4.800	4.640
23	Street	Quiet	5×10^{-4}	0	15.85	23	3.706	3.872	4.323	3.967
24	Quiet	Street	5×10^{-4}	0	15.85	24	4.340	4.490	4.785	4.539

Set 5 - Phase II Experimental Parameters			Ph2-CTQ Scores		
Cond	**PL**	**Codec, Mode**	**French**	**Arabic**	**Average**
1	0%	AMR-NB, 6.7kbit/s	4.218	3.938	4.078
2	0%	AMR-NB, 12.2kbit/s	4.314	4.052	4.183
3	0%	AMR-WB, 12.65kbit/s	4.329	4.303	4.316
4	0%	AMR-WB, 15.85kbit/s	4.455	4.313	4.384
5	0%	G. 723., 6.4 kbit/s	4.155	3.981	4.068
6	0%	G.729, 8kbit/s	4.108	4.182	4.145
7	0%	G.722, 64 kbit/s + plc	4.342	4.129	4.235
8	0%	G.711 + plc	4.323	4.279	4.301
9	3%	AMR-NB, 6.7kbit/s	3.791	3.577	3.684
10	3%	AMR-NB, 12.2 kbit/s	4.028	3.875	3.951
11	3%	AMR-WB, 12.65kbit/s	4.278	4.035	4.157
12	3%	AMR-WB, 15.85kbit/s	4.139	3.991	4.065
13	3%	G. 723.1, 6.4 kbit/s	3.871	3.514	3.692
14	3%	G.729, 8kbit/s	3.986	3.818	3.902
15	3%	G.722, 64 kbit/s + plc	4.331	4.301	4.316
16	3%	G.711 + plc	4.340	4.331	4.336