12-2007

Reasons why Mobile Telephone Conversations may be Annoying: Considerations and Pilot Studies

Nigel Ward
The University of Texas at El Paso, nigel@utep.edu

Anais G. Rivera

Alejandro Vega
The University of Texas at El Paso, avega5@miners.utep.edu

Comments:
Technical Report: UTEP-CS-07-60
Supplement to a paper to appear in *21st International Symposium on Human Factors in Telecommunication*, Kuala Lumpur, Malaysia, March 17-20, 2008.

Technical Report UTEP-CS-07-60

Reasons why Mobile Telephone Conversations may be Annoying: Considerations and Pilot Studies

Nigel G. Ward, Anais G. Rivera, Alejandro Vega

University of Texas at El Paso

December 27, 2007; updated January 3, 2008 (changes to Figure 1)

Abstract
Mobile telephone conversations in public places are often annoying to bystanders. Previous work has focused on the psychological and social causes for this, but has not examined the possible role of properties of the communication channel. In our paper "Do Bystanders and Dialog Participants Differ in Preferences for Telecommunications Channels?" (21st International Symposium on Human Factors in Telecommunication, 2008) we consider the possibility that a reason for the annoyance could be that bystander preferences differ from talker preferences, but conclude that this is in fact unlikely to be a major factor. This technical report provides supplemental information, specifically a broader view of the likely causes of annoyance and more details on the ten pilot studies and the data collected.

1. Possible Reasons

There are many factors that seem likely to be contributors to the annoyance felt. In addition to those discussed in the paper (Monk et al. 2004a, 2004b; Ward et al 2007), several other likely contributing factors have been identified by Ling (2004) and others. Figure 1 suggests some hypothesized causal paths. For example: There may be feelings of jealousy of the evident social involvement of the talker. There may be a sense of embarrassment at eavesdropping, or resentment at being involuntarily forced into that role. There may be annoyance at hearing content incongruous for the situation, due to incompatible expectations arising from the two contexts which every mobile conversation inhabits, or due to annoyance at the inability or unwillingness of the talker to communicate in a way that is considerate of bystanders. There may be differences in opinion regarding the proper etiquette for the use of mobile phones. There may be increased talker loudness due to the lack of sidetone (the system-generated audio feedback from the talker's mouth to his ear in traditional telephones).

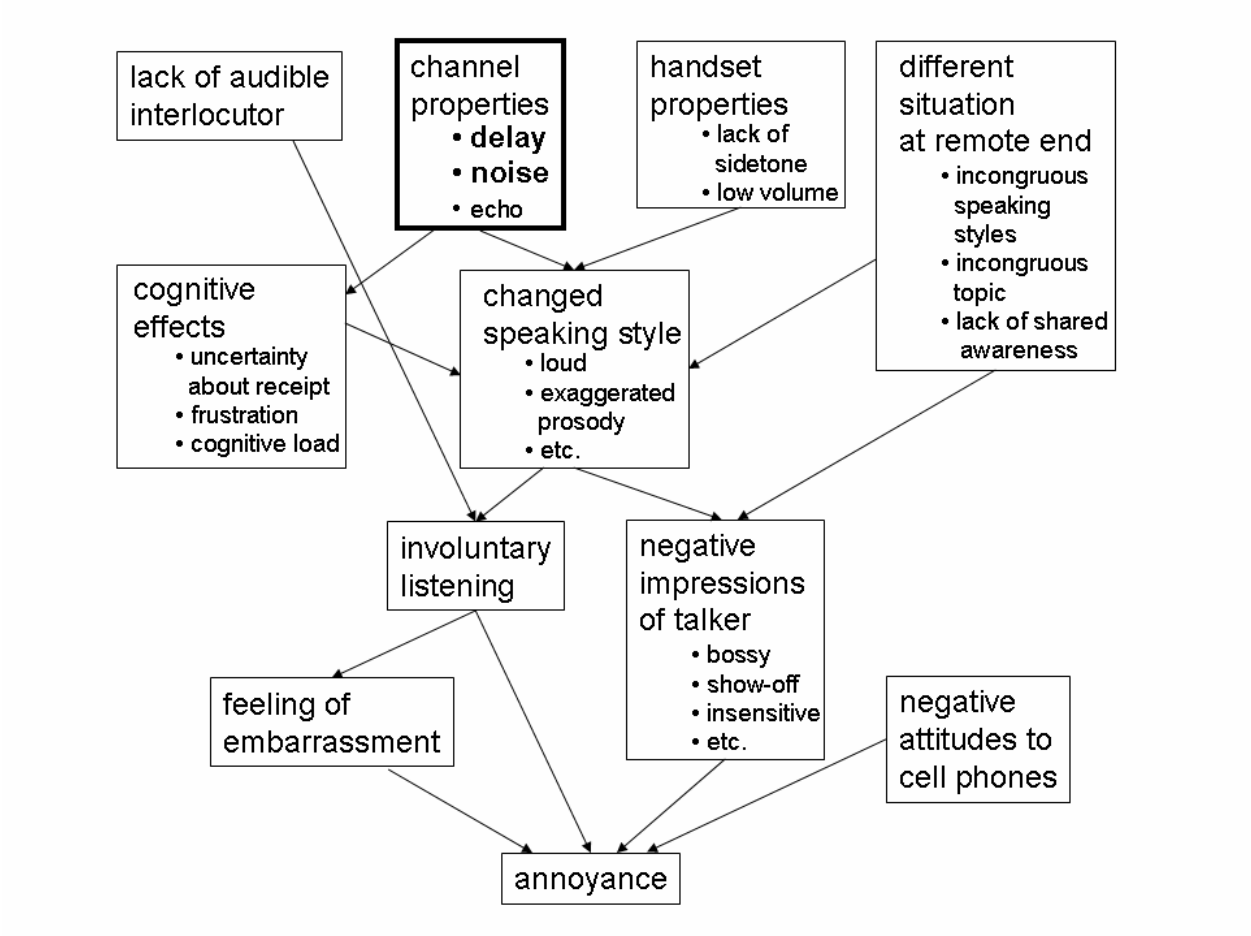


Figure 1: Some Possible Causal Chains Underlying Perceptions of Annoyance

In the paper we examine the contributions of factors of a different type: the quality of the telecommunications channel. The causal path is indirect: channel quality may affect the way that the talkers interact, and that the resulting changes in speaking style may be annoying to bystanders. This may be true not only for the downlink channel, directly affecting the local talker, but also for the uplink channel, where quality differences may affect the behavior of the remote talker, causing compensations by the near talker, which in turn cause annoyance.

2. Experiments and Results

We tried a succession of experimental methods, refining the setup and running more subjects every time it appeared that one method was unlikely to give the results we expected. This section tells the story.

7.1 First Attempt

In this attempt we gave each talker four cards, each containing seven numbers which had between one and seven digits. Each "dialog" consisted of the subjects reading numbers to each other switching off at each number. This was a variant of Task 1 of Kitawaki (1991), with the length of the numbers being

varied because we thought this would make it more necessary for the speakers to use turn-taking cues, specifically to signal prosodically when each number was complete.

We first let them practice with both Cn (GSM with 150ms delay) and Cd (no GSM and 300ms delay) encoding conditions. We then we recorded the third and fourth exchanges. After this we gave them a survey about these two and asked them to compare them. In these surveys we only had two options for each question, forcing them to choose either the first dialog or the second one. Then we had them listen to a pair of recordings of exchanges in both conditions and take another survey to rate which was more annoying. The recordings we asked them to judge were from a previous pilot session between two lab members.

The talkers questionnaires asked:

Q2T: Which line seemed to be worse quality?
Q3T: Which dialog did you feel was (would have been) more obnoxious to other people (if they had been) in the same room?
Q4T: Which time did you feel your dialog partner was less attentive and cooperative?

The "bystanders" questionnaires asked:

Q1B: Which dialog was more annoying?
Q2B: Which dialog was more noticeable?
Q3B: I found myself listening more to the:
Q4B: Which dialog was more intrusive?
Q6B: In which dialog did you think the line was worse quality?
Q7B: In which dialog do you think the talker(s) were less attentive, less cooperative and less polite?

Subject	**QT2**	*QT3*	QT4	*QB1*	QB2	QB3	QB4	**QB6**	QB7
1	**1**	*1*	1	*0*	0	0	0	**0**	0
2	**1**	*1*	0	*0*	0	1	0	**0**	0
3	**1**	*0*	1	*1*	0	0	1	**0**	1
4	**1**	*1*	1	*0*	1	1	0	**0**	0
5	**1**	*1*	1	*1*	1	1	1	**1**	1
6	**0**	*0*	0	*1*	0	0	1	**1**	1
7	**1**	*1*	1	*1*	0	0	1	**1**	1
8	**1**	*0*	1	*1*	0	0	1	**1**	1
9	**0**	*1*	0	*0*	1	1	0	**0**	0
10	**0**	*0*	0	*1*	1	1	0	**0**	0
mean	**0.7**	*0.6*	0.6	*0.6*	0.4	0.5	0.5	**0.4**	0.5

0 indicates Cn (GSM) and 1 indicates Cd (noGSM)

There seemed to be some tendency for the talkers to find Cd worse than Cn, whereas the bystanders had no clear preference (comparing QT2-3 with QB1-6). However the difference was obviously not significant.

7.2 Second Attempt

In the second attempt we began giving the bystanders different samples to judge. Specifically, we gave each set of judges a dialog from the immediately previous run, to enable us to compare talkers' and bystanders' opinions of the same dialog. We also decided to modify the surveys to give four choices instead of two: for each questions the options were "clearly the first", "probably the first", "clearly the second" and "probably the second". This was done for two reasons: first it was clear that many subjects did not like the forced choice and were choosing randomly, and we thought that this would reduce their frustration. Second, we thought this would enable us to gain more useful information, and in particular would give us pseudo-continuous values, making it possible to use a t-test to determine if a difference existed between talkers' and bystanders' judgments. The bystanders' version of the questionnaire appears as Appendix B.

We also dropped Q6B from the questionnaire.

Subject	QT2	**QT3**	QT4	QB1	QB2	QB3	**QB4**	QB7
1	0	**0**	0	2	0	3	**3**	2
2	0	**0**	0	0	2	0	**2**	2
mean	0	**0**	0	1	1	1.5	**2.5**	2

0 represents clearly Cn,
1 probably Cn,
2 probably Cd,
3 clearly Cd.

These two subjects seemed to defy the tendency seen in the first attempt.

We also examined the dialogs for quantitative differences in the speech signals, specifically looking for differences in loudness or in pitch range due to the channel differences, without success.

7.3 Third Attempt

By this point the basic experimental setup was reliably working, so we decided to try a more realistic situation. In this attempt we had two participants having a normal conversation while several bystanders in one of the rooms ate pizza. This setup was inspired by that developed by David Ponevac. Talkers were asked ``talk about anything, but try to have a good conversation, not too one-sided." After they talked for about 2 minutes using one channel, we had them wrap up and then talk for about 2 more minutes using the other channel.

After the conversations in both conditions we asked the participants, both the talkers and bystanders, to rate which conversation was more annoying. Then we had them listen to recordings of the number giving exchange taken from the previous attempt and rate those. The questions and rating options remained the same as on the previous attempt.
Some subjects only participated as bystanders; they did not complete a talker questionnaire.

There seemed to be a tendency opposite to that seen in the first attempt: the talkers seemed to find Cn worse than Cd (QT2 and QT3) and the bystanders to have a weaker dislike for Cn or no preference (QB1, QB2, QB3 and QB4). If we take reactions to noise as a baseline, it seems that the ill effects of delay were felt more strongly by bystanders than by talkers, matching our original hypothesis. However we were disappointed that the difference was not stronger.

Subjects**	QT2*	**QT3**	QT4	QB1	QB2	QB3	**QB4**	QB7
1	2	**2**	2	0	2	0	**3**	3
2	2	**1**	3	2	2	2	**1**	2
3				0	0	1	**0**	0
4				1	2	1	**2**	1
5				1	1	0	**2**	3
6				3		3	**1**	1
7	0	**0**	0	3	3	2	**2**	1
8	1	**2**	2	1	2	0	**0**	1
9				2	3	2	**3**	2
10				3	1	3	**1**	1
11	0	**0**	0	2	1	1	**1**	0
12	0	**0**	1	1	1	1	**1**	1
13				1	1	1	**1**	2
14				0	2	1	**0**	1
15				1	2	1	**3**	2
16				2	1	0	**0**	0
17				1	1	1	**1**	0
Average	0.83	**0.83**	1.3	1.4	1.6	1.2	**1.3**	1.2

We thought there might be several possible reasons for this. One reason might be that the bystanders seemed generally able to ignore the conversation, due to the fact that many of the talkers spoke rather softly, the fact that the bystanders were several meters away, and the fact that the bystanders had other things to occupy them, namely talking to fellow bystanders and eating pizza. Another possible reason might be that the content of the conversation seemed to have a large effect, for example it seemed that bystanders were more affected when the topic of the talkers' conversation was issues in grading in the class which all of the talkers and bystanders were taking.

7.4 Fourth Attempt

In this attempt we had the bystanders sit next to the talkers. In the first experiment of this attempt the talkers had a regular conversation (subjects 1-4). In the second experiment we decided to control the content by going back to the multi-digit number exchange task from the first attempt. After the conversations we asked both the talkers and the bystanders to answer a survey. Then the talkers listened to recordings of themselves and answered the bystander survey. The questions remained the same as on attempt two.

Subjects	QT2	**QT3**	QT4	QB1	QB2	QB3	**QB4**	QB7
1	3	**1**	2	2	2	1	**3**	3
2	2	**2**	3	0	2	3	**2**	3
3				3	1	3	**1**	2
4				0	1	1	**3**	2
5	1	**3**	1	0	1	1	**2**	1
6	1	**1**	2	1	2	1	**2**	2
7				2	2	2	**2**	1
8				0	2	1	**2**	1
9				1	2	2	**1**	2

10				2	2	3	**1**	2		
Average	1.75	**1.75**	2	1.1	1.7	1.8	**1.9**	1.9		

Here it seemed clear that for most questions the answers of talkers and bystanders did not differ, however QB1 suggested that Cn was felt to be more annoying by bystanders, contrary to what was seen on the third attempt but matching the result of the first attempt.

7.5 Fifth Attempt

We then added a question asking the subjects to simply rate which conversations they preferred (QT5 and QB8), in order to allow direct comparison between talker and bystander judgments. We also reordered the questions.

Thus the questions were:
QT2: Which line seemed to be worse quality?
QT3: Which dialog did you feel was (would have been) more obnoxious to other people (if they had been on the room)?
QT4: Which time did you feel your dialog partner was less attentive and cooperative?
QT5: Which did you prefer?

QB1: Which dialog was more noticeable?
QB2: Which dialog was more intrusive?
QB3: I found myself listening more to the:
QB4: Which dialog was more annoying?
QB7: In which dialog do you think the talker(s) were less attentive, less cooperative, and less polite?
QB8: Which did you prefer?

The first experiment used multi-digit number exchange, and the other three free conversations.

Subject	QT2	*QT3*	QT4	**QT5**	QB1	QB2	QB3	*QB4*	QB7	**QB8**
1	2	*1*	2	**1**	1	2	2	*1*	2	**1**
2	1		1	**2**						
3					0	2	2	*1*	1	**1**
4					3	2	1	*3*	1	**3**
5	1	*3*	0	**3**	0	3	3	*1*	0	**3**
6	2	*0*	1	**3**	0	1	0	*1*	0	**3**
7					2	1	3	*0*	3	**1**
8					3	1	3	*1*	1	**3**
9	1	*3*	1	**1**	1	1	0	*3*	3	**0**
10	1	*2*	1	**1**	2	2	3	*1*	2	**2**
11					1	1	2	*1*	3	**1**
12					1	1	0	*2*	1	**1**
13	3	*2*	2	**0**	1	2	1	*2*	2	**0**
14					0	1	0	*1*	1	**2**
15					1	1	2	*2*	2	**1**
Average	1.6	*1.8*	1.1	**1.6**	1.1	1.5	1.6	*1.4*	1.6	**1.6**

Here there was a slight tendency for bystanders to find Cn more noticable (QB1), and perhaps faintly more annoying (QB4) although the talkers thought the opposite (QT3). On most other measures both talkers and bystanders did not seem to distinguish the two channels.

7.6 Sixth Attempt

At this point we began to wonder whether our original hypothesis (2) was correct, as the evidence seemed to leaning in the other direction, suggesting that Cn was relatively more annoying to bystanders. We even began to question hypothesis (1), that is, we began to wonder whether there was an effect to find, and if there was one, whether our methods were sufficient to reveal it. We decided to try a more extreme manipulation. Recalling the observation that the effects of line delay are less noticeable in dialog with long turns (Kitawaki 1991), we decided to accentuate the effects of delay by adopting Kitawaki's Task 1, in which the subjects read from two lists of single digit numbers, switching off at each digit, and attempting to do so as fast as possible. We used different lists in the two conditions, but subsets of the lists were the same, in order to collect data for direct comparison (discussed below). We conducted two experiments of this type; one with the usual 350 ms delay (subjects 1 and 2) and one with 500ms delay (subjects 3 and 4). The questions were the same as in attempt five.

Subject	QT2	QT3	QT4	QT5	BQ1	BQ2	BQ2	BQ4	BQ7	BQ8
1	1	0	3	1	1	3	2	0	3	N/A*
2	2	0	1	3	0	1	0	2	2	N/A
3	0	1	2	1	2	1	2	1	1	2
4	1	1	3	1	1	2	2	1	2	3
Averages	1	0.5	2.25	1.5	1	1.75	1.5	1	2	2.5

*These subjects were mistakenly given the questionnaire without this question.

Even though we had attempted to accentuate the effects of delay, overall both talkers and bystanders still seemed to disprefer Cn more.

7.7 Seventh Attempt

In the previous experiments we had sought to measure overall impressions of substantial chunks of dialog. In this experiment we wanted to see if focused listening to specific short number sequences across the two conditions would reveal a clear preference. Using the recordings from the previous attempt we played matching sequences of numbers excised from both conditions to subjects and asked them to rate their preference after each pair.

Here there was a tendency for the subjects, here all acting as bystanders, to disprefer Cn, with the tendency being pronounced for Subject 1. The experimenters also listened to the samples, to see whether they could perceive any consistent difference in quality or annoyingness between samples from the two channels; here there was not much agreement on which of each pair was most annoying, and there was no overall tendency to prefer one channel over the other. Perhaps giving too much credence to our own impressions, we chose not to pursue this further.

1st Recording (350ms)	Numbers	Sub1	Sub2	Sub3
1	“024”	**2**	1	1
2	“102864973”	**2**	1	1
3	“47”	**1**	2	1
4	“79856247”	**1**	1	3
2nd Recording (350ms)				
1	“25”	**0**	1	1
2	“364”	**0**	1	1
3	“578954102360”	**2**	1	0
4	“679845”	**0**	2	3
5	“879641”	**0**	2	0
3rd Recording (600ms)				
1	“25”	**0**	0	2
2	“364”	**0**	1	0
3	“578954102360”	**0**	1	2
4	“679845”	**0**	2	1
5	“879641”	**0**	2	1
4th Recording (600ms)				
1	“024”	**3**	2	1
2	“102864973”	**3**	1	3
3	“47”	**0**	2	3
4	“79856247”	**0**	1	1
	Subject Average	**0.78**	**1.33**	**1.39**

Average (350ms)	Average (600ms)	Average of both
1.15	**1.19**	**1.17**

7.8 Eighth Attempt

Again wanting to obtain a clear effect, we increased the delay in Cd from 350ms to 600ms. In this attempt the subjects were asked to have regular conversations with each other in both conditions. They then answered the talker survey. Next then they listened to the recording of their own voice and answered the bystander survey. There were nine experiments of this type, with the final one including two bystanders who never took the talker role.

Subject	QT2	**QT3**	QT4	*QT5*	QB1	QB2	QB3	**QB4**	QB7	*QB8*
1	2	**1**	3	*1*	1	2	0	**2**	3	*1*
2	2	**1**	2	*1*	3	2	1	**1**	1	*3*
3	0	**2**	1	*2*	2	2	2	**1**	1	*2*
4	1	**1**	2	*2*	0	1	1	**2**	2	*1*
5	2	**2**	2	*0*	0	2	0	**2**	2	*0*
6	3	**3**	3	*0*	0	0	0	**2**	3	*0*
7	1	**1**	1	*2*	3	2	1	**1**	1	*3*
8	1	**1**	3	*3*	3	2	2	**1**	1	*2*
9	3	**2**	0	*2*	0	1	0	**2**	3	*0*
10	0	**1**	0	*3*	3	3	3	**1**	0	*3*
11	1	**1**	2	*3*	1	2	1	**1**	2	*2*
12	1	**3**	3	*0*	3	3	3	**3**	3	*0*
13	3	**0**	1	*3*	0	3	3	**1**	1	*0*
14	2	**1**	1	*2*	3	1	3	**3**	0	*3*
17	1	**2**	1	*2*	3	3	0	**1**	0	*3*
18	1	**1**	2	*2*	1	0	0	**0**	1	*2*
19	3	**3**	3	*0*	3	3	3	**0**	0	*3*
20	2	**2**	1	*0*	3	0	3	**0**	3	*2*
21					3	0	2	**1**	3	*0*
22					0	2	0	**3**	3	*3*
Average	1.6	**1.6**	1.7	*1.6*	1.75	1.7	1.4	**1.4**	1.65	*1.65*

As we had worsened Cd, we expected the numbers in the table to increase, relative to what was seen in the fourth attempt, however there was a slight opposite tendency.

Again no clear difference between talker and bystanders was seen, neither between QT2 and QB1 or between QT5 and QB8.

7.8.1 Eight and a Halfth Attempt

In the course of attempt eight, we did one run in which the characteristics of the Cd channel were exaggerated: the delay was increased to 800ms and the audio quality to CD-quality: 44.1kHz, 16 bit (Subjects 15 & 16).

Subject	QT2	**QT3**	QT4	*QT5*	QB1	QB2	QB3	**QB4**	QB7	*QB8*
15	2	**1**	1	*1*	1	2	0	**2**	1	*1*
16	1	**1**	1	*3*	2	1	1	**1**	0	*3*

Here we expected to see a truly obvious effect, but, as usual, the judgments were not extreme.

7.9 Ninth Attempt

For this attempt we went back to 350ms for Cd. We also changed the questionnaires to clarify that the annoyance rating should be given based on the speaking style rather than the content (QB6). Thus the questionnaires became as follows:

QT2: Which line seemed to be worse quality?
QT3: Which dialog did you feel was (would have been) more annoying to other people (if they had been) in the same room?
QT4: Which time did you feel your dialog partner was less attentive and cooperative?
QT5: Which did you prefer?
QB1: Which dialog was more annoying?
QB4: In which dialog do you think the talker(s) were less attentive, less cooperative, and less polite?
QB5: Which was more annoying in content?
QB6: Now, not thinking about the content, rather just the speaking style, if these were cell phone conversations you were overhearing which would you prefer?

This added question became moot, however, since we also decided to go back to Kitawaki's task 1, the exchange of single digit numbers. Another difference was that we asked subjects to first listen to another conversation from a previous session and then to their own. In the tables below the QB questions were asked first, with the subject acting as a bystander to someone else's conversation, and the QR questions were the same questions asked with the subject acting as a "relistener", that is, an after-the-fact bystander to his own voice. In the tables below the first was an experiment similar to attempt eight, with the difference being that the fast single digit number format was used but the old questionnaires were used. In the second table the new surveys were used.

We also semi-formalized the experiment protocol at this point, as seen in Appendix A.

Subject	QT2	**QT3**	QT4	*QT5*	QB1	QB2	QB3	**QB4**	QB7	*QB8*
1	2	**2**		0	1	3	0	**3**	2	0
2	2	**2**	2	1	1	2	0	**3**	3	0
Average	2	**2**	2	0.5	1	2.5	0	**3**	2.5	0

Subject	QT2	**QT3**	QT4	*QT5*	QB1	QB4	**QB5**	QB6	QR1	QR4	QR5	*QR6*
1	2	**2**	3	*0*	2	1	**2**	1	2	2	2	*1*
2	2	**2**	3	*0*	0	2	**0**	3	0	2	0	*3*
3	3	**2**	2	*0*	2	2	**1**	0	1	0	1	*2*
4	2	**1**	2	*0*	1	0	**2**	0	3	0	2	*0*
Average	2.25	**1.75**	2.5	*0*	1.25	1.25	**1.25**	1	1.5	1	1.25	*1.5*

Although few subjects were run, there appears to be an opposite effect: the bystanders were more annoyed by Cn than were the talkers (relative to the annoyance felt by Cd).

7.10 Tenth Attempt

The protocol was kept the same but the surveys were redesigned. Instead of a forced answer design, the survey now consisted of two scales, labeled from 0-10 (Moeller 2000), on which subjects were to place a mark at the place corresponding to their rating of the channel they were using or overhearing. The talkers version of this questionnaire is in Appendix C. In addition to providing this rating, subjects answered two other questions with comments on the reasons for these ratings, giving us more insight into what they thought of the two lines.

QT1: Your opinion of the connection you have just been using.
QT2: What differences did you notice between the two connections?

QT3: What do you think affected you ratings of the two connections?
QB1: Independent of content, the way people talk may result in annoyance, please give your opinion on the sample.
QB2: What differences did you notice between the two connections?
QB3: What do you think affected your ratings of the two connections?
QR1: Independent from content, the way people talk may result in annoyance, please give your opinion of the sample.
QR2: What differences did you notice between the two connections?
QR3: What do you think affected your ratings of the two connections?

The Cd delay in this occasion was set to 400 ms. A total of 20 subjects participated and 10 experiments were run. The first two subjects were given a double dose of the experiment, since we thought that this would be more efficient, however they were clearly fatigued by the second set of comparisons and their opinions did not seem thoughtful. The following table shows the results:

Subject	Δ=gsm-nogsm (Talkers)	Δ=gsm-nogsm (Bystanders)	Δ=gsm-nogsm (Relisteners)
1*	0	-10	0
1(D)*	0	10	0
2*	0.9	0.2	0.6
2(D)*	-3.1	-0.6	-1
3	1.9	0.9	3.4
4	4	1.5	-1.6
5	0.6	4	-3
6	-5	-2	2
7	-4	2	3
8	2	2	-2
9	-4	5	-2
10	3	-3	-1
11	1	-1	-3
12	0.5	1.5	0.9
13	-1	-3	1
14	-2	3	-4
15	-5	5	-2
16	1.5	-2	3
17	0.1	-2	2
18	0.1	-1	-1
19	1	-2	3
20	-2	0	2
Averages (total)	-.043	0.39	0.01
Averages (without doubled runs)	-0.41	0.49	0.04

The averages here suggest that talkers preferred Cd, but the bystanders preferred Cn, Matched-pair two-tailed t-tests (subjects 3-20) indicated that this tendency was weak ($P \sim 0.4$); we estimate that we would have to run about 100 subjects at this rate to obtain a significant response. It was also clear that the variation was large, as illustrated by the scatterplot in the paper, Figure 2 of Ward et al. (2008).

We also noted that subjects were fairly enthusiastic about diagnosing the properties of the two channels, which most were able to do correctly. This raises the possibility that their judgments were affected by

this, for example, they may have noticed the delay, thought that delay was annoying, and rated Cd poorly for that reason, rather than merely using their impressions.

3. Summary

The results are inconclusive, but some methodological lessons can be drawn, as discussed in the paper (Ward et al. 2008).

Acknowledgments

We thank: David Novick and Sebastian Möller,for insight and discussion, Udit Sajjanhar for the initial implementation, David Ponevac for the first pilot studies, and Ana I. Schwartz for comments on the experiment design. This work was supported in part by the National Science Foundation under awards EIA-0080940 and IIS-0415150 and DARPA.

References

(Deneger 1992) GSM 06.10 lossy speech compression library. Jutta Degener. kbs.cs.tu-berlin.de/~jutta/toast.html, accessed October 2006.

(ITU 2005) The E-model, a computational model for use in transmission planning. ITU-T Recommendation G.107. International Telecommunication Union. 1998, revised 2005.

(Kitawaki 1991) Delay Effects on Speech Quality in Telecommunications. Nobuhiko Kitawaki and Kenzo Itoh. *IEEE Journal on Selected Areas in Communications*, 9, pp 586-593. 1991.

(Ling 2004) The Mobile Connection: *The Cell Phone's Impact on Society.* Rich Ling, Morgan Kaufmann, 2004.

(Moeller 2000) *Assessment and Prediction of Speech Quality in Telecommunications*. Sebastian Moeller, Springer, 2000.

(Monk 2004a) Why are Mobile Phones Annoying? Andrew Monk, Jenni Carroll, Sarah Parker and Mark Blythe. *Behavior and Information Technology*, 23, pp 33-41, 2004.

(Monk 2004b) Hearing only one side of normal and mobile phone conversations. Andrew Monk, Evi Fellas and Eleanor Ley.i, 23, pp 301-305, 2004.

(Ponevac 2006) Effects of Line Delay and Novice on Bystanders' Perceptions of Mobile Phone Conversations. David Ponevac. University of Texas at El Paso, Department of Computer Science, Masters Thesis, 2006.

(Ward et al. 2008) Do Bystanders and Dialog Participants Differ in Preferences for Telecommunications Channels? Nigel G. Ward, Anais G. Rivera, Alejandro Vega. 21st International Symposium on Human Factors in Telecommunication, to appear.

Appendix A

Detailed Protocol for "Dialog-Based Evaluation of Mobile Phone Infrastructure"

Nigel Ward and Alejandro Vega, September 14, 2007.

Preparation

1. Set up cables, connections, and recorders on both rooms and do a sound check.
2. Prepare list of numbers and determine starting condition (gsmmachine2 or test400 found on ~avega5/gsmEx/ on linux machines).
3. Print out surveys and label properly.
4. If session number is odd do gsm first, then nogsm(test400). If session number is even do nogsm(test400) first, then gsm.

Introduction

1. Welcome subjects and ask them to sit wherever they please (preferably the couch)
2. Ask them to fill out the consent forms.
3. When the consent forms are filled out, ask them to accompany you to the computer in the room and overview how the experiment will be done *"Ok, this is how the experiment will be run. you will having two conversations through two computers, one in this room and one in the other room. you will talk through the microphone on the table and hear the other person through the head set. The small microphone on the headset will record your voice onto a small recorder which will later be used in the experiment. Your conversations will be brief single digit exchanges from the number sheet on the table. As you exchange numbers, you will say one number, then wait for the other person to say their first number, then you say your second number and so on. These exchanges should be done as fast as possible."*
4. After the explanation, ask the subjects to choose which room they would like to be in. Once they sit, review the exchange process in order to make sure they do it properly. Adjust headset microphone and ask them to sit at same distance from the microphone on the desk. Have them start the first exchange and press record on the small recorders.
5. Once they finish exchanging the first line of numbers, stop the recorder, have them answer the first scale on the survey labeled "T", change the experiment condition (on both machines), start up the recorders, and have them exchange the next line of numbers.
6. After the exchanges are finished instruct them to go back to room B and have them answer the rest of the survey labeled "T".
7. After they are done, play the recordings of the immediately previous experiment(on Sony recorder from room A*, plugged into speakers) and have them answer the survey labeled "B".
8. When they finish with that play them back their own recordings, in separate rooms through speakers, of the conversations they just had and have them answer the survey labeled "R". Instruct them in each of these steps properly.
9. After they finish answering the surveys and making comments, review surveys and ask if anything is unclear, pay the subjects (or not depending if they are obtaining credit for

the experiment itself) and give a vague explanation to why their participation is important to this study and its purpose (vaguely!!).Finally, thank them for coming.

*10. *The locations (folder and file number) of the recordings from the immediately previous experiment will be recorded on the surveys at the end of the experiment for reference and use for future experiments.*

Appendix B

Dialog-Based Evaluation of Mobile Phone Infrastructure

Date _____

Bystanders

Session# _____

Dialog Pair # __

1. Which dialog was more noticeable?
clearly the first ☐, probably the first ☐, probably the second ☐, clearly the second ☐

2. Which dialog was more intrusive?
clearly the first ☐, probably the first ☐, probably the second ☐, clearly the second ☐

3. I found myself listening more to the:
clearly the first ☐, probably the first ☐, probably the second ☐, clearly the second ☐

4. Which dialog was more annoying?
clearly the first ☐, probably the first ☐, probably the second ☐, clearly the second ☐

5. Did you feel any other difference between the two dialogs? _____

If so, what?

6. Do you have any idea why your perceptions of the two dialogs may have differed?

If so, why?

7. In which dialog do you think the talker(s) were less attentive, less cooperative, and less polite?
clearly the first ☐, probably the first ☐, probably the second ☐, clearly the second ☐

Appendix C

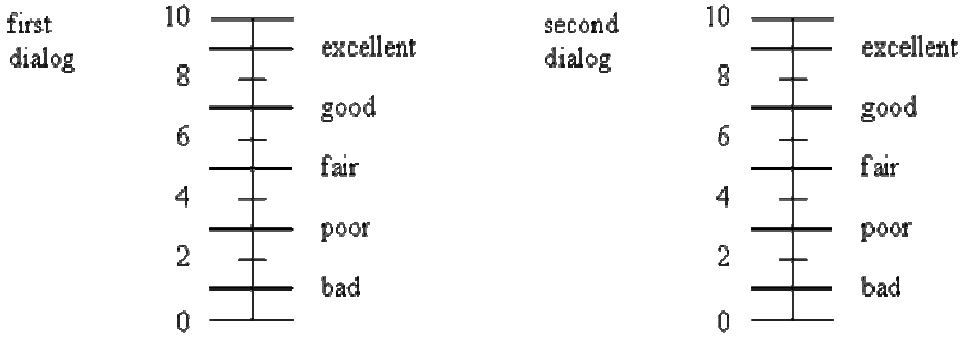
T

Dialog-Based Evaluation of Mobile Phone Infrastructure

date _____
session _____
subject A B
recording# 1 _____
recording# 2 _____

Phase 1

A. Your opinion of the connection you have just been using.
(Please place a line crossing the axis at the appropriate point.)



B. What differences did you notice between the two connections?

C. What do you think affected your ratings of the two connections?