Supplementary material to:

Christian P. Janssen, Shamsi T. Iqbal, & Yun-Cheng Ju.

Sharing a Driver's Context with a Caller via Continuous Audio Cues to Increase Awareness about Driver State

Contents:

1. Detailed report of Experiment 2

In experiment 1 we found that when sounds were shared with the caller, the callers rated the driver as being more busy and the caller's talk segments were shorter. An open question is why the talk segments were shorter. Did callers reduce talk segment length because of their assessment of the driver's busyness (and the content of the sounds), or was this an artificial effect because the shared sounds interfered with the conversation? If the change was a response to the content of the shared sounds (i.e., to an assessment of busyness), we should only find such adaptation when content is shared that suggests busyness (e.g., driving sounds), but we should not find such adaptation when other, irrelevant sounds (e.g., nature sounds) are shared at the same pace. In contrast, if the change was due to interference of the sounds, we should find such adaptation regardless of what type of sound is being shared.

Method

Participants

Fifteen pairs took part. The drivers (12 male, 3 female, M_{age} = 31 year, SD_{age} = 7 year) and callers (7 male, 7 female, M_{age} = 32 year, SD_{age} = 10 year) were recruited following the same procedure as in experiment 1. Three pairs had also taken part in experiment 1.

All participants had strong English skills and all but one were native English speakers. All pairs knew each other: 3 were in a relationship, 5 were friends, 5 were colleagues, and 2 were colleagues and friends. Each participant received a gratuity.

To encourage appropriate attention, an additional gratuity was granted to the best performing pair similar to experiment 1.

Tasks & Materials

The materials were similar to those used in experiment 1. Drivers only drove on the difficult road. They communicated with the remote caller as before. We did not record physiological data in this experiment. Driving sounds were as before. Nature sounds were played at the same frequency as the driving sounds, with a new sound starting every three seconds. For these sounds we used fourteen hand-picked sounds from the national park service website (http://www.nature.nps.gov/sound/gallery.cfm). Sounds were clipped such that they were at most 2 seconds.

Design

The study had a 3 level within-subjects design. The factor was type of sound shared: no sounds, nature sounds, or driving sounds.

Measures

All measures were as before. In addition, on trials where sounds were shared we asked the caller to rate how much the sounds interfered with the conversation on a scale of 1 (no interference) to 5 (a lot of interference).

Procedure

Practice trials were as in experiment 1, with the exception that the driver only drove on the difficult road instead of both on the easy and the difficult road.

This was followed by six experimental dual-task blocks, with two trials each. The blocks varied in what sounds were shared: none, nature, or driving sounds. All three conditions were experienced during the first half of the experiment, with their order randomized. In the second half of the experiment, participants experienced all three conditions again, but in a different order than during the first half. This procedure provided data from four trials per condition.

After each trial, participants rated their experience on a five-point scale. The questionnaire was as before. We also asked the caller to rate whether the shared sounds interfered with their conversation on a five-point scale. After the experiment, participants filled out a general questionnaire. The experiment lasted about 2 hours.

Results

We present results by research question. For RQ 2 and 3 we compare performance between single- and dual-task using a One-way ANOVA (single- or dual-task).

When analyzing metrics for the dual-task conditions only (RQ 1, 2, and 3), we used a One-way ANOVA (sounds shared: none, nature, or driving), unless noted otherwise. A significance level of p = .05 is used. For each metric we report means and 95% within-subjects confidence intervals, following Morey (2008).

RQ1: Caller’s perception of driver’s busyness & interference

Figure A.1 plots the effect of sound sharing on the rating of busyness. There was a significant effect of sound type on the caller's rating of the driver's busyness, F(2,

28) = 20.21, p < .001, η_p^2 = .59. A holm-corrected post-hoc test found that callers rated the driver to be more busy in the condition with driving sound sharing (M = 3.47, 95% CI = [3.18, 3.75]) than in the condition with nature sound sharing (M = 2.55, 95% CI = [2.31, 2.79]), p = .043. There was also a significant difference between the driving sound condition and the condition without sounds (M = 2.50, 95% CI = 2.25, 2.75]), p = .043. There was no significant difference between the nature sound condition and the condition without sound, p = .896. That is, the caller increased the busyness rating on average by 1 point when driving sounds were shared compared to the other two conditions. Sharing nature sounds did not influence the caller's perception of the driver's busyness.

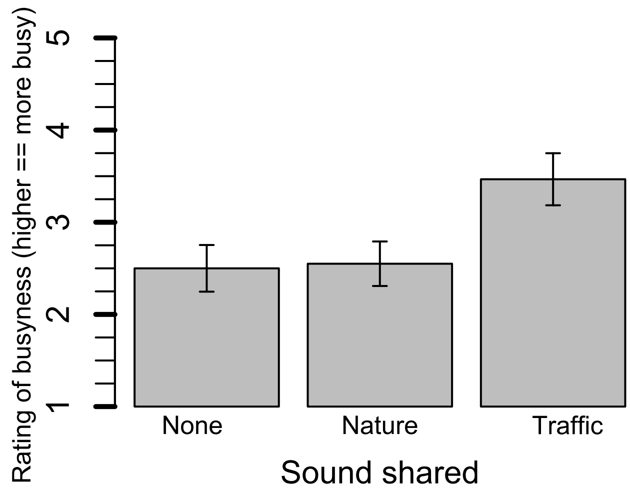

Figure A.1: Bar plot of the perceived busyness of the driver by the caller, given the sounds that were shared: no sounds, nature sounds, or traffic sounds. The caller rated the driver as being more busy when traffic sounds were shared. Error bars indicate 95% within-subjects confidence intervals.

We also asked participants to rate how interfering the nature and driving sounds were on the conversation for each trial. Figure A.2 plots these data. A one-way ANOVA found that driving sounds (M = 3.68, 95% CI = [3.46, 3.90]) were rated as more interfering by half a point on a five point scale than the nature sounds (M = 3.10, 95% CI = [2.88, 3.32]), $F(1, 14) = 16.03$, $p = .0013$, $\eta_p^2 = 0.53$.

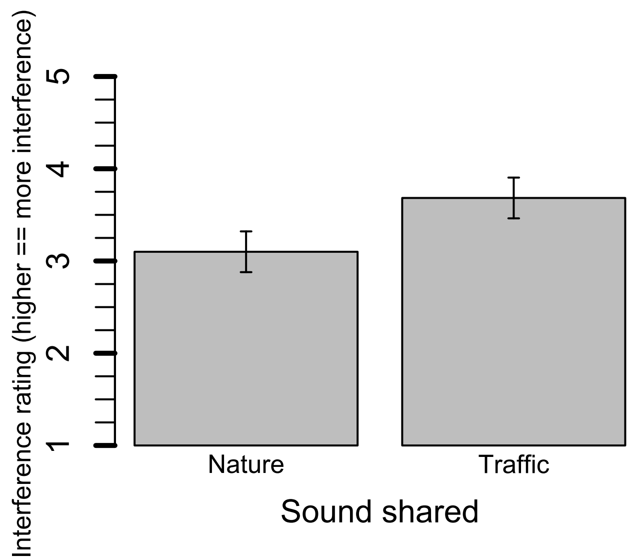

Figure A.2: Bar plot of the rating of interference of the sounds, as provided by the caller. The caller found the traffic sounds to be slightly more interfering with the conversation than the nature sounds. Error bars indicate 95% within-subjects confidence intervals.

These results suggest that there was indeed modest interference from sharing the driving sounds with the caller, but the sounds were useful.

RQ2: Adaptation of the conversation

Conversational turn length. When comparing performance between single- and dual-task, it was found that the caller's median turn length (M = 12.07 sec, 95% CI = [10.46, 13.68]) was longer than the driver's median turn length (M = 8.66 sec, 95% CI = [7.05, 10.27]), $F(1, 14) = 6.454$, $p = .0235$, $\eta_p^2 = .32$. That is, the caller's conversational turn was approximately 30% longer than the driver's turn. There was no significant effect between turn length in single- and dual-task and no significant interaction, $p > .1$. Figure A.3 plots these data, to allow comparison with the results from Experiment 1 (see Figure 4 in main paper).

Within dual-task trials, the caller's conversational turn ($M = 12.54$ sec, 95% CI = [10.74, 14.33]) was again longer than the driver's turn ($M = 8.79$ sec, 95% CI =

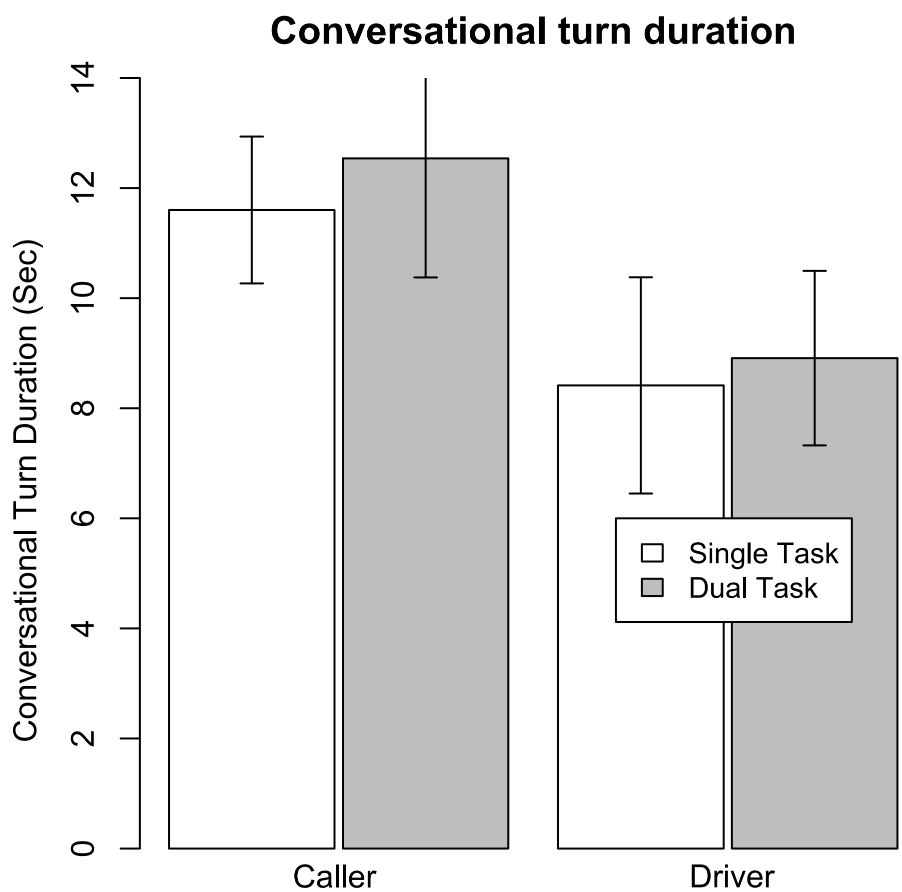

Figure A.3: Bar plot of conversational turn lengths in single-task (white bars) and dual-task (grey bars). The caller's turns (left two bars) are longer compared to the driver's turns. Error bars indicate 95% within-subjects confidence intervals.

[7.11, 10.71]), $F(1, 14) = 5.851$, $p = .0298$, $\eta_p^2 = 0.29$. That is, the caller's turn length was approximately 40% longer than the driver's turn.

There was no effect of sound sharing and no significant interaction between speaker and condition, $p > .1$.

Median talk segment length for caller. When comparing single- with dual-task performance, there was no significant difference between talk segment length in single-task ($M = 1.88$, 95% CI = [1.77, 2.01]) and dual-task ($M = 1.87$, 95% CI = [1.76, 1.99]), $F(1, 14) = 0.03$, $p > .1$.

Within the dual-task trials, no significant differences were found between the three sound sharing conditions, $F(2, 28) = 0.07$, $p > .1$. Figure A.4 plots these data, to allow comparison with the results from Experiment 1 (see Figure 5 in main paper). The mean values and 95% Confidence Intervals were as follows: no sounds (M = 1.90 sec, 95% CI = [1.68, 2.12]), nature sounds ($M = 1.87$, 95% CI = [1.68, 2.06]), and traffic sounds ($M = 1.85$, 95% CI = [1.70, 2.01]).

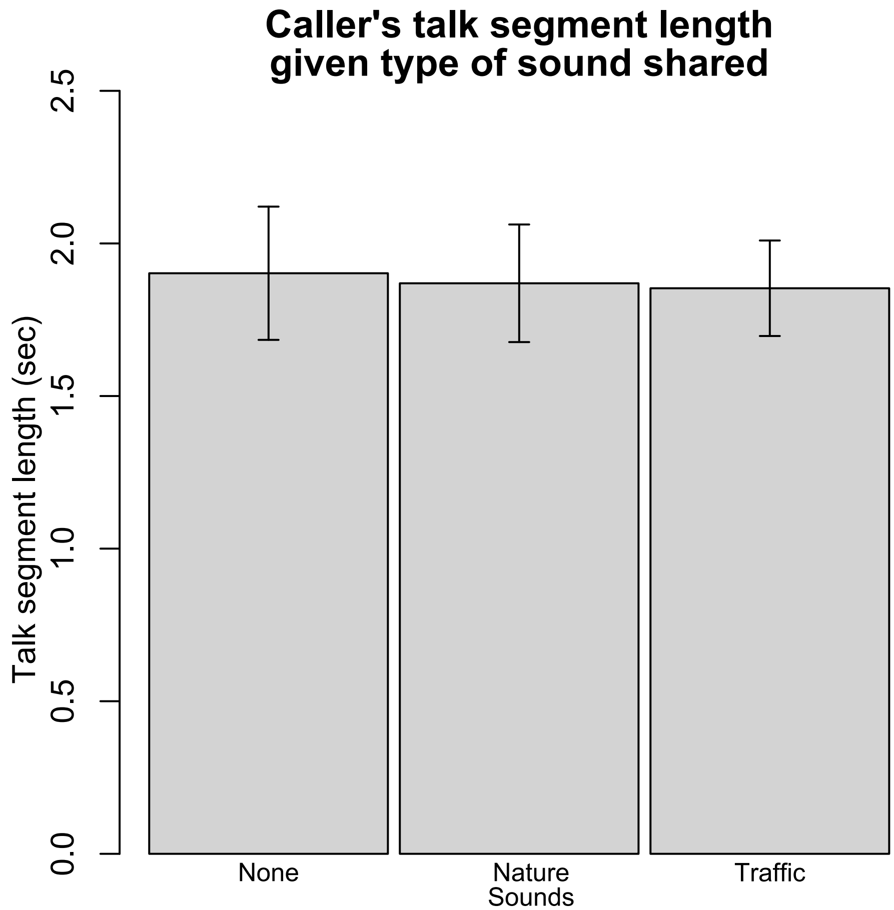

Figure A.4: Bar plot of the length of the caller's talk segments given the type of sounds that were shared. There was no significant effect of sharing no sounds, nature sounds, or traffic sounds. Error bars indicate 95% within-subjects confidence intervals.

Median talk segment length for driver. When comparing single- with dual-task performance, there was no significant difference between talk segment length in single-task (M = 1.73, 95% CI = [1.56, 1.90]) and dual-task (M = 1.74, 95% CI = [1.58,1.91]), $F(1, 14) = 0.008$, $p > .1$.

Within the dual-task trials, no significant differences were found between the three sound sharing conditions, $F(2, 28) = 0.028$, $p > .1$. The means and 95% Confidence Intervals were as follows: no sounds ($M = 1.75$ sec, 95% CI = [1.63, 1.86]), nature sounds ($M = 1.75$ sec, 95% CI = [1.56, 1.95]), and traffic sounds ($M = 1.73$ sec, 95% CI = [1.59, 1.87]).

Taken together, in this experiment we found no evidence to support the hypothesis that the caller adapted their conversation style based on the sharing of sounds. The data also did not show that the conversation adaptation that was observed in Experiment 1 was due to sound interference. If that was the case, we should have observed a significant difference here between the condition without sounds and the two conditions with sounds (nature and traffic).

RQ3: Effects on the driver's performance

Driving performance: Coherence score. When comparing performance between single- and dual-task, it was found that the coherence score was higher in single-task ($M = 0.879$, 95% CI = [0.867, 0.891]) compared to dual-task ($M = 0.848$, 95% CI = [0.835, 0.861]), $F(1, 14) = 14.12$, $p = .002$, $\eta_p^2 = .50$.

Within the dual-task trials, there was no effect of sound sharing, $p > .1$. The mean values and 95% Confidence Intervals were as follows: no sounds ($M = 0.847$, 95% CI = [0.841, 0.853]), nature sounds ($M = 0.844$, 95% CI = [0.836, 0.851]), and traffic sounds ($M = 0.853$, 95% CI = [0.847, 0.859]).

Driving performance: Delay in reacting to lead car. When comparing performance between single- and dual-task, there was a significant difference between performance in single-task and dual-task trials, $F(1, 14) = 17.92$, $p < .001$, $\eta_p^2 = .56$. Delays were shorter in single-task ($M = 0.71$ sec, 95% CI = [0.60 ,0.82]) compared to dual-task ($M = 1.02$ sec, 95% CI = [0.91, 1.13]).

Within dual-task, there was no significant difference between the different sound conditions, $p > .1$. The mean values and 95% Confidence Intervals were as follows: no sounds ($M = 1.04$ sec, 95% CI = [0.97, 1.12]), nature sounds ($M = 0.99$ sec, 95% CI = [0.91, 1.07]), and traffic sounds ($M = 1.04$ sec, 95% CI = [0.97, 1.10]).

Driving performance: Modulus. There was no significant difference in modulus values between single- ($M = 1.25$, 95% CI = [1.23, 1.27]) and dual-task trials ($M = 1.25$, 95% CI = [1.22, 1.27]), $p > .1$.

Within the dual-task trials, there was no significant difference between the sound conditions, $p > .1$. The mean values and 95% Confidence Intervals were as follows: no sounds ($M = 1.25$, 95% CI = [1.24, 1.27]), nature sounds ($M = 1.24$, 95% CI = [1.22, 1.25]), and traffic sounds ($M = 1.25$, 95% CI = [1.23, 1.27]).

Driving performance: Fixed distance with the lead car. The difference between single- and dual-task was marginally significant, $F(1, 14) = 3.93$, p = .067. The trend in the data was that the difference from the ideal was larger in dual-task ($M = 10.72$ ft., 95% CI = [10.02, 11.42]) compared to single-task trials ($M = 9.81$ ft., 95% CI = [9.11, 10.51]).

Within dual-task trials, there was no difference between sound conditions, $p > .1$. The mean values and 95% Confidence Intervals were as follows: no sounds ($M = 10.88$, 95% CI = [10.42, 11.34]), nature sounds ($M = 10.60$, 95% CI = [10.06, 11.14]), and traffic sounds ($M = 10.68$, 95% CI = [10.27, 11.08]).