Research Report

Psychological Science
21(10) 1383–1388

DOI: 10.1177/0956797610382126
http://pss.sagepub.com

Overheard Cell-Phone Conversations: When Less Speech Is More Distracting

Lauren L. Emberson[1,2], Gary Lupyan[3], Michael H. Goldstein[1], and Michael J. Spivey[4]

[1]Psychology Department, Cornell University; [2]Sackler Institute for Developmental Psychobiology, Weill-Cornell Medical College; [3]Psychology Department, University of Wisconsin–Madison; and [4]Cognitive Science Program, University of California at Merced

Abstract

Why are people more irritated by nearby cell-phone conversations than by conversations between two people who are physically present? Overhearing someone on a cell phone means hearing only half of a conversation—a "halfalogue." We show that merely overhearing a halfalogue results in decreased performance on cognitive tasks designed to reflect the attentional demands of daily activities. By contrast, overhearing both sides of a cell-phone conversation or a monologue does not result in decreased performance. This may be because the content of a halfalogue is less predictable than both sides of a conversation. In a second experiment, we controlled for differences in acoustic factors between these types of overheard speech, establishing that it is the unpredictable informational content of halfalogues that results in distraction. Thus, we provide a cognitive explanation for why overheard cell-phone conversations are especially irritating: Less-predictable speech results in more distraction for a listener engaged in other tasks.

Keywords

attention, speech, cell phones, unpredictability, dialogue, halfalogue, dual-task, distraction

Received 1/27/10; Revision accepted 3/30/10

Cell phones are ever present, contributing to ambient speech in public spaces. Although cell phones increase the overall amount of ambient speech, overhearing a cell-phone conversation is judged to be more irritating and intrusive than overhearing a dialogue in which both parties are heard (Monk, Carrol, Parker, & Blythe, 2004). Thus, there appears to be something "special" about overheard cell-phone conversations. In this report, we present evidence that the cognitive demands of overhearing a cell-phone conversation differ from those of overhearing other types of speech. Specifically, we show that overhearing different types of speech results in varying degrees of distraction in attentionally demanding tasks.

Conversations are acts of coordination between two individuals: Conversational partners share syntactic structures, lexical items, and even body posture (Shockley, Richardson, & Dale, 2009). It has been argued that the interaction and alignment of these cognitive processes increase the ease of speech production and communication (Garrod & Pickering, 2004). The same processes that aid dyadic conversation may also increase the predictability of successive utterances for a third-party listener.

Overhearing someone on a cell phone means hearing only half of an ongoing conversation—a "halfalogue." Compared with overhearing dialogue or monologue, overhearing half of a conversation means hearing less ambient speech overall, but the speech that is heard is less predictable. Although it has been well established that organisms process unpredicted or surprising events differently than predicted ones (Rescorla & Wagner, 1972), the effects of predictability on perceptual and cognitive processing are receiving renewed interest (Friston & Kiebel, 2009). Unpredicted experiences generate internal error or novelty signals (Amso, Davidson, Johnson, Glover, & Casey, 2005; Schultz, 2002), quickly and automatically drawing attention. For example, you are likely to turn in the direction of a loud, unexpected sound or to be distracted by a new billboard on your usual route to work. Recent work has established the influence of relative predictability on behaviors as diverse as decision making (Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006), language learning (Saffran, Aslin, & Newport, 1996), and visual processing (Enns & Lleras, 2008).

Corresponding Author:
Lauren L. Emberson, Department of Psychology, 211 Uris Hall, Cornell University, Ithaca, NY 14853
E-mail: lle7@cornell.edu

On the basis of these claims that predictability of information widely influences human behavior, we hypothesized that the relative unpredictability of a halfalogue will draw on limited attentional resources, resulting in poorer performance in concurrent tasks. We tested this hypothesis by having participants perform attentionally demanding tasks while overhearing different types of speech: a *dialogue* comprising both sides of a cell-phone conversation, a *halfalogue* comprising the speech of one speaker engaged in conversation (i.e., one half of a cell-phone conversation), or a *monologue* comprising one person's recap of a conversation. These types of speech differ in their relative predictability: The linguistic elements of a dialogue and a monologue are incrementally more predictable than those of a halfalogue. Comparison of performance across these three types of speech allowed us to evaluate the effects of two additional factors that could lead people to be more distracted when they overhear halfalogues: a speaker addressing a silent or absent listener (a factor shared by halfalogues and monologues) and overhearing conversation in general (a factor shared by halfalogues and dialogues).

We employed two attentionally demanding tasks designed to assess different types of attention involved in real-world tasks, such as driving. The tasks had both verbal and nonverbal components, so that we could determine whether potential distraction effects were specific to constraints on verbal capacity. The *visual monitoring task* required participants to track a pseudorandomly moving dot using a circular mouse cursor (Fig. 1a). Participants were instructed to keep the dot as close to the center of the cursor as possible. This task required constant visual monitoring, similar to the vigilance required to stay within a traffic lane while driving. The *choice reaction time (choice RT) task* (Fig. 1b) required participants to maintain four target letters in short-term memory and to ignore distractors. This task required reorienting of attention similar to that required for responding to traffic signals.

Experiment 1: Overhearing Spontaneous Speech

Method

Twenty-four Cornell University undergraduates (6 male, 18 female; 3 left-handed; mean age = 19.45 years, *SD* = 1.16) participated in Experiment 1 for extra credit. This study complied with all regulations of Cornell University's institutional review board.

Materials. Two pairs of college roommates (all female) were given conversation starters (e.g., a news article, comic) to discuss over cell phones. Members of each pair were seated in separate, soundproof rooms during their discussions. For each conversation starter, the pair had a conversation for 2 min, and then each member of the pair recapped the conversation in a monologue. Thus, the sampling of conversational content was the same across the three speech conditions.

Each speaker was recorded using a wireless microphone. Each member of a pair was recorded on a separate channel of a wav file in order to facilitate a clean separation of the dialogue. Eight dialogues (one speaker per channel), eight halfalogues (derived from the nonselected dialogues; one speaker in stereo), and eight monologues (some recaps had to be concatenated) were used. All speech files were 60 s long and were volume-normalized.

In contrast to dialogue—a stream of continuous speech switching between two speakers—halfalogue consists of unpredictable onsets and offsets of speech from a single speaker. Our speech samples were consistent with this description: Speech was present 92% of the time (*SD* = 6.57%) in the dialogues and only 54.7% of the time (*SD* = 11.1%) in the halfalogues (Fig. 2a). Thus, participants heard substantially less speech when overhearing the halfalogues than when overhearing

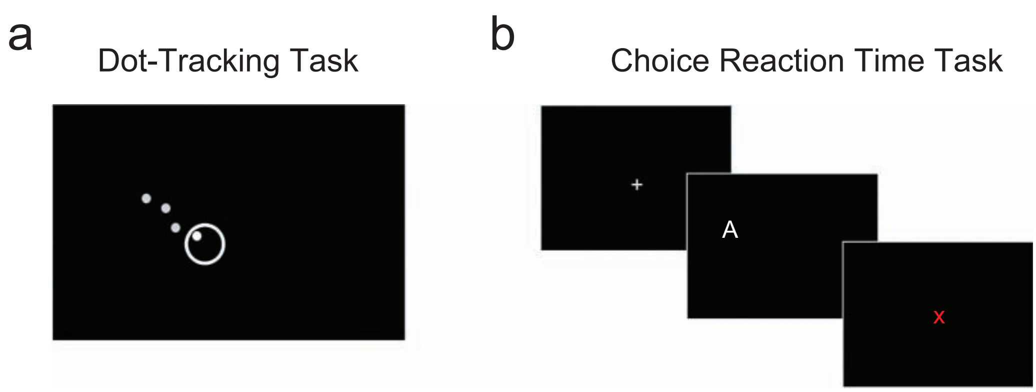

Fig. 1. Illustration of the attentionally demanding tasks employed in this study. In the visual monitoring task (a), participants were required to track a pseudorandomly moving dot using a circular mouse cursor. This task is nonverbal and requires continuous monitoring similar to the constant vigilance required to stay within a traffic lane while driving. The choice reaction time task (b) required participants to maintain four target letters (presented prior to the trial block) in short-term memory and respond only to these letters, ignoring distractors. Letters were presented for 400 ms each after the presentation of a fixation cross. The task was to press a key if a letter matched one of the target letters, and to refrain from pressing the key otherwise. Error feedback (a red "X") was presented if the subject missed a target letter or incorrectly responded to a distractor. The figure illustrates a trial on which the participant responded incorrectly. This verbal task requires the reorienting of attention similar to that required for responding to traffic lights.

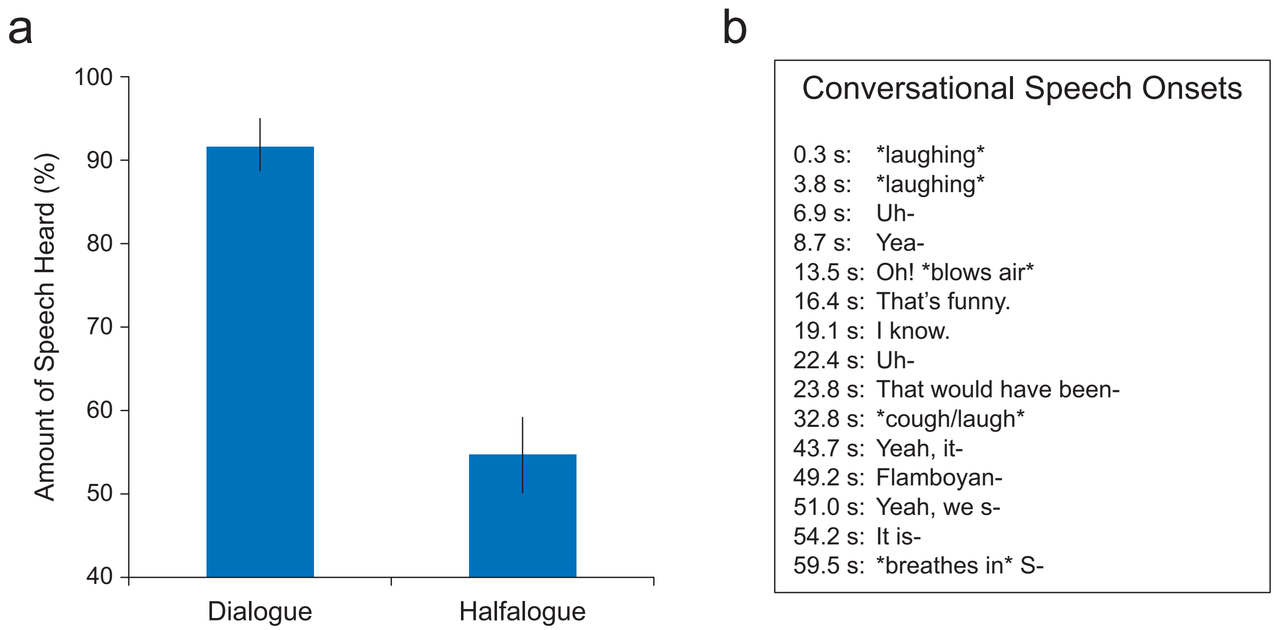

Fig. 2. Examination of the recorded cell-phone conversations. The graph (a) shows the mean percentage of time that speech was heard during the recorded speech samples used in both experiments, separately for the dialogues and the halfalogues. Error bars depict standard errors of the mean. The transcript (b) shows what was heard in the first 400 ms of each conversational onset for 1 min of a representative halfalogue; the transcript reveals that a notable and comprehensible amount of speech can be uttered in less than half a second.

the dialogues. Transcripts revealed that the conversational partners spoke simultaneously for short durations, which explains why speech was present more than 50% of the time in our halfalogues.

Procedure. Participants were seated at a computer and instructed that they would be completing two different tasks. One involved tracking a moving dot with the computer mouse, and the other involved responding to letters presented on the computer screen. They were given 1 min of practice with each of these tasks in silence. They were then instructed that they would be completing these tasks a number of times and would sometimes hear speech from the two computer speakers situated on either side of the monitor. Participants were asked to focus their attention on the attentionally demanding tasks.

During the experiment, participants were presented with the two tasks in alternating order for 32 blocks; each block was 60 s long—the duration of each speech file. The tasks were performed in four conditions: the three speech conditions and a silent condition that served as a baseline against which to compare performance in the speech conditions. In each set of 8 blocks, these four conditions were presented in shuffled order and paired with each task. At the end of the session, participants filled out a posttest questionnaire and were debriefed.

The visual monitoring task (see Fig. 1a) required continuous attention. Participants were asked to track a moving dot using a computer mouse. The dot followed a pseudorandom trajectory, shifting its direction by 1° to 20° every 40 ms. The dependent measure was the euclidean distance between the dot and the center of a ring-shaped mouse cursor, recorded every 40 ms. Greater distances corresponded to poorer ability to follow the trajectory of the dot.

The trials of the choice RT task (see Fig. 1b) were presented in minute-long blocks. Each block began with a display of four randomly selected English letters that served as the targets for that trial sequence. Each trial began with a fixation cross presented for 600 to 1,000 ms; one of the four target letters or a nontarget letter was then presented for 400 ms. Target letters were presented on 35% of the trials. Participants were allowed 1 s to respond. False alarms or misses were followed by a centrally presented red "X" indicating that the response was incorrect.

Results

We hypothesized that the relative predictability of speech affects how distracting it is for people to overhear it. Given that halfalogue speech is less predictable than either dialogue or monologue, we predicted that overhearing halfalogue would result in a significant decrease in performance on a concurrent task, whereas overhearing dialogue and monologue would not. To test this prediction, we compared performance in the speech and silent conditions, separately for each task.

Analyses of variance (ANOVAs) on performance in the choice RT task revealed a significant effect of condition (silent, dialogue, halfalogue, monologue) on overall performance (hits and correct rejections), $F(3, 69) = 2.87$, $p < .05$, and on hit rate alone, $F(3, 69) = 3.03$, $p < .05$, but not on false alarms, $F < 1$. We next conducted simultaneous pair-wise comparisons (Tukey's tests) of mean hits for the four conditions. The only significant difference was between the silent condition (M = 94.7%) and the halfalogue condition (M = 90.4%), $t(23) = 2.88$, $p < .05$. Planned Dunnett's t tests comparing baseline performance with performance in each speech condition revealed the same pattern of results for overall performance: Performance in the silent condition was significantly different from performance in the halfalogue condition, $t(23) = 2.74$, $p < .05$, but not different from performance in the dialogue and monologue conditions, $p > .05$. We also examined the effect of condition using d' as the dependent measure (see Fig. 3b). Planned Dunnett's t tests corroborated the previous findings that only overhearing halfalogue reliably affected performance compared with baseline (i.e., the silent condition), $t(23) = 2.21$, $p < .05$. However, the omnibus ANOVA for d' was not significant, $F(3, 69) = 1.67$, $p = .18$, likely because of the differential effects of condition on hits and false alarms. There were no RT differences across conditions. These results are consistent with our hypothesis that overhearing a halfalogue, but not overhearing a monologue or dialogue, leads to significant distraction.

In the continuous tracking task, participants' mean distance from the moving target was 56.9 pixels (SD = 15.5). There were no gross differences in tracking error across conditions. Because increases in error would result from exogenous attentional shifts to the overheard speech, we predicted that any effect of overheard speech on tracking performance would be specifically related to the time course of the speech stream. We compared changes in tracking error for analogous time windows in halfalogue and dialogue speech to examine specific effects of the predictability of overheard speech on visual monitoring performance. Although the time course of reflexive attentional shifts as a response to spontaneous speech is unknown, previous research has demonstrated that nonlinguistic auditory cues have an effect on visual attention by 400 ms after their onset (Spence, Nicholls, Gillespie, & Driver, 1998). Accordingly, we calculated the change in tracking error for halfalogues by subtracting error in the 400 ms before the onset of speech from error in the 400 ms after the onset of speech, and we calculated the change in tracking error for dialogues by subtracting error in the 400 ms before a change of speaker from error in the 400 ms after a change of speaker. (See Fig. 2b for the conversational onsets during these 400-ms time windows in a portion of a sample halfalogue.) Consistent with the hypothesis that overhearing a halfalogue is attentionally demanding, this analysis revealed a significant increase in tracking error after the onset of halfalogue speech (M = 2.80, SD = 5.71), $t(22) = 2.40$, $p = .012$, and no corresponding increase in error after a speaker change in the dialogue condition (M = 0.43, SD = 4.26), $t(22) = 0.494$, $p = .31$ (see Fig. 3a).

Thus, in both attentionally demanding tasks, there was a decrease in performance when participants overheard a halfalogue. There was no reliable change in performance when they overheard a dialogue or monologue.

Experiment 2: Overhearing Filtered Speech

Relative to dialogues, halfalogues are less predictable in both their informational content and their acoustic properties. Whereas dialogues involve the nearly continuous switching of speech between speakers, speech onsets and offsets are more unpredictable in halfalogues. If these unpredictable acoustic changes were solely responsible for the detrimental effects of overhearing halfalogues, then filtering out informational content while leaving the low-level acoustic content the same would still result in significant distraction. But if the increased distraction observed when participants overheard halfalogues

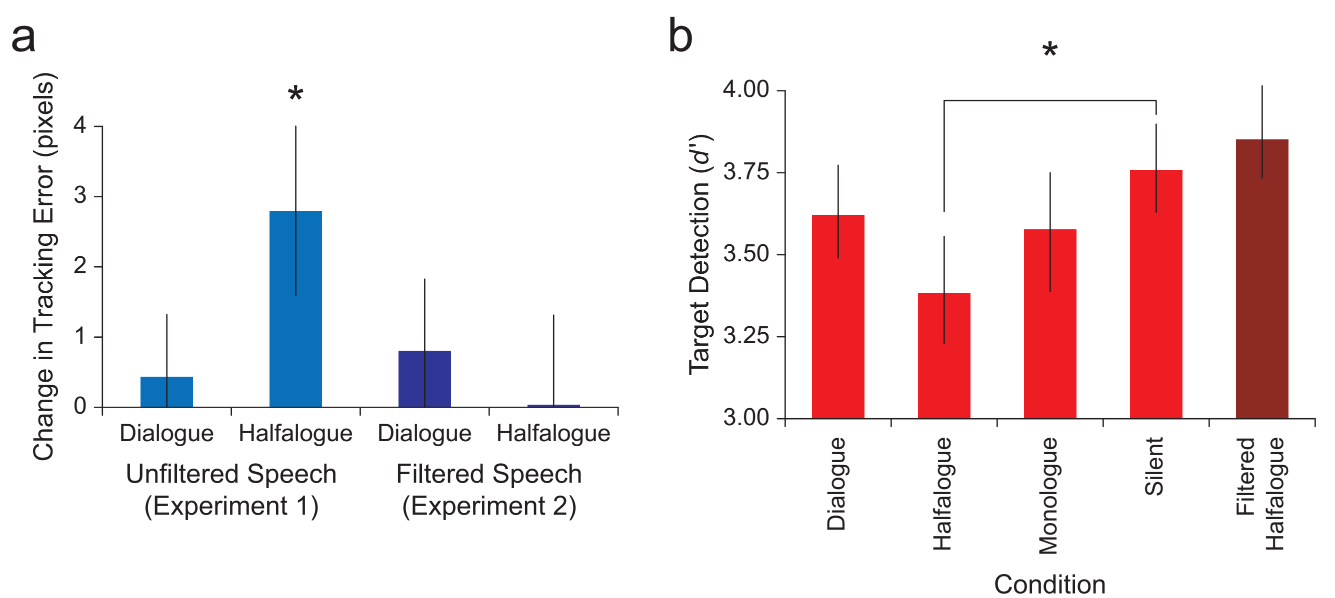

Fig. 3. Experimental results: (a) mean change in tracking error on the visual monitoring task and (b) d' on the choice reaction time task as a function of condition. Tracking error was measured as euclidean distance from the center of the mouse cursor to the dot in pixels. Change in tracking error was calculated by subtracting error in the 400 ms before speaker onset from error in the 400 ms after speaker onset (halfalogue condition) or by subtracting error in the 400 ms before speaker change from error in the 400 ms after speaker change (dialogue condition). Error bars depict standard errors of the mean, and asterisks indicate a significant difference from zero (a) or a significant difference between conditions (b), $p < .05$.

was due to the relative unpredictability of the content of half of a conversation, then filtering the speech should remove the distraction resulting from overhearing halfalogues. Thus, in Experiment 2, we isolated the effects of the unpredictable acoustics of the halfalogue speech by filtering out the information content (i.e., rendering the speech incomprehensible) while maintaining the essential acoustic characteristics of speech onsets and offsets.

Method

Seventeen additional participants (1 male, 16 female; 1 left-handed; mean age = 19.81 years, *SD* = 1.32) were recruited for Experiment 2. All procedures and materials were identical to those for Experiment 1 with one crucial exception: All sound files were low-pass filtered such that only the fundamental frequency of the speech was audible. This manipulation rendered the speech incomprehensible while maintaining the essential acoustic characteristics of speech; the effect was similar to hearing someone speak under water.

Results

Unlike in Experiment 1, no attentional deficits were found when participants overheard filtered halfalogues. For the visual monitoring task, we again examined the difference in tracking error between the 400-ms windows before and after speech onset (filtered halfalogue) or a change in speaker (filtered dialogue). As shown in Figure 3a, there were no significant changes in tracking error (all ts < 1) in either the filtered-dialogue condition (M = 0.821, SD = 4.68) or the filtered-halfalogue condition (M = 0.056, SD = 5.81). For the choice RT task, there were no significant differences in d' between conditions (all ts < 1) and no main effect of condition (F < 1; results for the filtered-halfalogue condition are shown in Fig. 3b). An ANOVA comparing results for the silent and halfalogue conditions in Experiments 1 and 2 (comprehensible vs. filtered speech) revealed a marginally significant interaction between experiment and condition, $F(1, 39) = 3.32$, p = .076. Together, these results support our hypothesis that the behavioral deficits resulting from overheard halfalogue in Experiment 1 were directly related to the relative unpredictability of information in halfalogue speech, and not the differences in acoustics across the speech conditions.

Discussion

We examined the effect of overhearing different types of ambient speech on performance in tasks that tap into the diverse attentional demands in daily life. Using two markedly different attentional tasks, we found that overhearing a cell-phone conversation led to distraction. Because overhearing a cell-phone conversation entails access to only half of a dialogue, the speech content is less predictable than that of a full conversation. We found no evidence for decreased performance when participants overheard more predictable speech (i.e., dialogue and monologue), which indicates that the observed attention deficits were not related to overheard speech in general, but were specifically related to the unpredictability of overhearing one half of a conversation. Thus, overheard cell-phone speech may be especially distracting because it is less predictable than other forms of ambient speech.

In a second experiment, we demonstrated that this effect is not due to mere acoustic unpredictability: Overhearing the same speech when it was filtered and made incomprehensible did not produce the same pattern of results. Thus, speech comprehension is necessary to produce this behavioral effect of predictability. Although it is beyond the scope of the current work to determine which specific aspects of language processing are responsible for this effect, our results suggest that there is an increased cost of processing halfalogue speech when it is present, and that the effect is not due to increased processing during the silences of a halfalogue (e.g., a participant filling in the conversation). The tracking results are consistent with the time course of cross-modal attentional shifts in the case of nonlinguistic auditory stimuli (Spence et al., 1998). However, because this is the first investigation into attentional costs of hearing unpredictable speech, the time course of these attentional effects is largely unknown.

The finding that relative predictability of ambient speech has significant behavioral impact is consistent with the broader theoretical view that the goal of perceptual-cognitive systems is to reduce uncertainty (Friston & Kiebel, 2009). In this view, the relative predictability of information affects perceptual and cognitive processing generally and could drive behavioral changes in diverse cognitive domains.

Finally, it has been established that engaging in a cell-phone conversation impairs driving performance (e.g., producing increased braking latency) compared with engaging in conversation with a passenger in the vehicle (Haigney & Westerman, 2001) or listening to the radio or books on tape (Strayer & Johnston, 2001). The current findings demonstrate that simply overhearing a cell-phone conversation is sufficient to reduce performance in concurrent, attentionally demanding tasks that were designed to reflect two attentional components of driving. These results suggest that a driver's attention can be impaired by a passenger's cell-phone conversation. Future work in this area would benefit from extending this line of inquiry to tasks involving driving simulators and eye tracking.

Acknowledgments

The authors would like to thank Andrew Webb for help with data collection, transcription, and analysis of speech files; Dani Rubenstein, Noa Hertzman, and David Kalkstein for help with data collection; and Translink-Vancouver for inspiration.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

References

Amso, D., Davidson, M.C., Johnson, S.P., Glover, G., & Casey, B.J. (2005). Contributions of the hippocampus and striatum to simple association and frequency learning. *NeuroImage*, *27*, 291–298.

Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., & Dolan, R.J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*, 876–879.

Enns, J.T., & Lleras, A. (2008). What's next? New evidence for prediction in human vision. *Trends in Cognitive Sciences*, *12*, 327–333.

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principal. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*, 1211–1221.

Garrod, S., & Pickering, M.J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*, 8–11.

Haigney, D., & Westerman, S.J. (2001). Mobile (cellular) phone use and driving: A critical review of research methodology. *Ergonomics*, *44*, 132–143.

Monk, A., Carrol, J., Parker, P., & Blythe, M. (2004). Why are mobile phones annoying? *Behaviour & Information Technology*, *23*, 33–41.

Rescorla, R.A., & Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A.H. Black & W.F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York, NY: Appleton-Century-Crofts.

Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron*, *36*, 241–263.

Shockley, K., Richardson, D.C., & Dale, R. (2009). Conversation and coordinative structures. *Topics in Cognitive Science*, *1*, 305–319.

Spence, C., Nicholls, M., Gillespie, N., & Driver, J. (1998). Cross-modal links in exogenous covert spatial orienting between touch, audition, and vision. *Perception & Psychophysics*, *60*, 544–557.

Strayer, D.L., & Johnston, W.A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological Science*, *12*, 462–466.