Communication Quarterly

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rcqu20

Effects of Visual Spatial Structure on Textual Conversational Multitasking

Eli Dresner a & Segev Barak b

a Departments of Communication and Philosophy, Tel Aviv University,
b Department of Philosophy, Tel Aviv University,

Available online: 14 Feb 2009

To cite this article: Eli Dresner & Segev Barak (2009): Effects of Visual Spatial Structure on Textual Conversational Multitasking, Communication Quarterly, 57:1, 104-115

To link to this article: http://dx.doi.org/10.1080/01463370802664776

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Effects of Visual Spatial Structure on Textual Conversational Multitasking

Eli Dresner & Segev Barak

Textual computer-mediated communication gives rise to conversational multitasking—participation in several concomitant synchronous textual conversations. This study examined how this communication competence is affected by several visual parameters—the distance between conversation windows, number of windows, and window size. Results show that comprehension of concomitant textual conversations is not affected by the distance between conversation windows, and that the advantage of separating conversation threads into distinct windows is greater when overall window size is larger. This study considers the implications and applications of these results to communicative multitasking phenomena, in general, which become prevalent in technologically advanced societies.

Keywords: Computer-Mediated Communication; Interactive Written Discourse; Internet; Multitasking

The theory of communication competence has been used for several decades as a broad conceptual framework for variables that affect success in interpersonal communication (Parks, 1994). Communication competence was originally studied in the context of non-mediated, face-to-face (f2f) communication, where its main factors (or ingredients) included such capacities as body-language expression and comprehension, conversation management, and affective skills (Duran & Spitzberg, 1995). However, it is clear that the notion of communication competence can also apply to mediated interpersonal communication. In particular,
computer-mediated communication (CMC) gives rise to a large variety of new communication channels to which the analytic framework of communication competence is applicable.

In tune with earlier work on communication competence, Spitzberg (2006) provided a theoretical model for textual CMC competence. He broke the notion into three categories—motivation, knowledge, and skill—and under each category presented relevant empirical and theoretical work. In particular, under the title of “CMC skills” appear traits such as attentiveness, expressiveness, and coordination that characterize competent communicators both online and offline. The latter skill, coordination, is understood as including such factors as “navigation control, pace control, rapidity and responsiveness” (Spitzberg, 2006).

This research is concerned with a communication skill that is distinctive to textual CMC and falls under the aforementioned characterization of coordination skills—although it is not mentioned explicitly by Spitzberg (2006), and is often not acknowledged as a novel type of communication competence. The skill in question is conversational multitasking: the ability to take part (either passively or actively) in several synchronous textual conversations at the same time. This ability is demonstrated by users of a variety of textual Internet-based communication channels. Among them are chatrooms (such as Internet Relay Chat (IRC) or its Web-based look-alikes), textual virtual worlds, and instant-messaging applications. In each of these technological contexts a stream of text lines goes through a single window or several windows on the user’s computer screen, and this stream of text often includes more than one conversation thread at any given time. The competent user can keep track of these simultaneous, distinct threads.

Computer-mediated conversational multitasking, as part of the phenomenon of multitasking, in general, characterizes the generation of youngsters and young adults who have grown up in an environment rich with digital communication technologies. As elaborated by Rainie (2006), such so-called digital natives are now joining the work force, and they often manifest cognitive and behavioral characteristics alien to their employers (who quite often are digital immigrants). One such characteristic is multitasking, and Rainie urged employers to treat it as a new and valuable skill rather than negatively, as a breach of manners or inability to separate personal social activity from work. The interplay between cognitive and normative aspects of multitasking is discussed later.

Conversational multitasking is much more prevalent in the previously described textual, computer-mediated situations than in f2f auditory conversation (Herring, 1999). In everyday life there are, indeed, situations where several conversations take place at the same time—for example, a dinner table or a party. However, it seldom happens (if at all) that we follow more than one of these conversation threads for a long period of time. On the other hand, in interactive written discourse (IWD), this is often not the exception but rather the rule (Ferrara, Brunner, & Whittemore, 1991).

What is the source of this discrepancy between online textual conversation and offline auditory talk? One possibility has to do with the limitations of
textual Internet synchronous communication channels; they receive input from many sources (i.e., speakers) and place this input in the exact chronological order in which it was received (Herring, 1999). This leads to intermingling of conversation threads within an IRC chatroom window, for example, and requires participants to maneuver within a chaotic conversational environment that is simply not allowed by the interactional coherence standards of f2f conversation. Another reason for the relative prevalence of conversational multitasking online may be that it is quite often done without awareness of conversation partners, and thus without upfront transgression of social norms. In f2f situations, other people can see that we divide our attention between several loci, and take offence; when two independent windows are open on a user’s computer screen, two-timing is done in private. This aspect of multitasking is returned to in the Discussion section.

However, underneath these technical and social rationales are cognitive, perceptual grounds for the fact that textual conversational multitasking seems to be more prevalent than auditory multitasking. It is arguably easier to multitask in a textual environment that we perceive with our eyes than in auditory conversation that we attend to with our ears; visual modality of perception helps chat users sort out distinct conversation threads (Dresner, 2005; Herring, 1999). The raw visual sensory data yielded by the eye is spatial in character, providing a two-dimensional representation of objective space (in front of the seeing eye). The human ear, on the other hand, collapses auditory input from all directions together and, therefore, does not represent the spatial location of auditory inputs. (This is notwithstanding the fact that further processing in our brain allows us to hear sounds as coming from [coarsely distinguished] distinct directions.) These different perceptual features affect communication, in general, and interpersonal synchronous communication, in particular. Thus, the relative unity of auditory input requires the strict structure of turn-taking mechanisms in auditory linguistic interaction, and renders multitasking difficult to accomplish. The spatial, metric structure of visually perceived text, on the other hand, can support various kinds of syntactic and semantic complexity, among which is the multiplicity of conversation threads in IWD. The fact that text lines accumulate on our screen for a (albeit short) period of time and that they are spatially related to each other seems to help us tie together lines that belong to the same conversation thread and to distinguish among different threads. In traditional, printed text, the same spatial structure is claimed to support other kinds of semantic complexity, such as more elaborate logical structure (McLuhan, 1962; Ong, 1982).

These observations were given initial support in a previous empirical study (Dresner & Barak, 2006), where the effects of two visual parameters on multitasking capabilities were tested. One result of this study was that two conversation threads that are intermingled within the same chat window are easier to follow if each is allocated a different color (as opposed to them both having the same color). Another result was that a significantly greater positive effect on understanding (as measured by recognition tests) was made by separating the
two textual conversation threads spatially and putting them within distinct chat windows. The spatial separation of the threads helped participants keep better track of each thread.

These initial, positive results give rise to further questions concerning the cognitive and perceptual underpinnings of the way vision helps support multitasking. How, exactly, is spatial separation used for this purpose? In this study, two hypotheses that consist in answers (albeit only partial answers) to this question were raised and tested. According to the first hypothesis, the advantage of placing distinct conversation threads in distinct windows on the screen arises from the fact that in this case each conversation “comes” from a distinct place in our visual field and, therefore, can be perceived and processed as a distinct, independent happening. If this is the case, then it may be expected that the greater the visual distance between the two loci in which the conversations take place, the easier it will be for users to make a cognitive distinction between the contents of the two conversations. Therefore, the following hypothesis is proposed:

H1: Two conversation threads that are presented in distinct windows on the user’s screen are easier to follow when the windows are farther apart from each other than when they are close.

The second hypothesis considers whether the observed advantage of two windows over a single window (in which 2 conversation threads are intermingled) is due to the following, different grounds. The significant result of the separation of the threads into distinct windows is that it recreates within each window the continuity and coherence that marks ordinary auditory conversation and uni-threaded textual conversation. Within each such conversation window all visually accessible text lines belong to the same thread; therefore, going back and forth among them helps the user keep track of the conversation. This is as opposed to the single-window case, where the threads are intermingled and much more effort is required to pick out previous lines that are relevant to the understanding of the current one. If this reasoning is correct, then we would expect that the beneficial effect on understanding of separation of conversation threads into distinct windows would be stronger when more text lines are concomitantly present on the screen at any given time—this way, the user can take better advantage (even if doing so unconsciously) of the “back-tracking” possibility hypothesized earlier. Therefore, the following hypothesis is considered:

H2: The advantage of separating conversation threads into distinct windows over intermingling them within the same window will be greater when the overall amount of text presented on the screen at any given moment is larger.

Note that H2 goes beyond the weaker hypothesis that understanding of synchronously accumulating text is positively correlated with the amount of text that can be seen at each moment. This latter hypothesis has been empirically supported in several
studies (Duchnicky, 1984; Xu & Fu, 2002), and does not concern multitasking in particular. The hypothesis made here (H2) is that the positive effect of enlarged text windows on multitasking will be greater in the two-window case than in the single-window case, due to the considerations made earlier. Also, note that H1 and H2 are distinct, but not contradictory: It could be the case that both prove to be correct, representing two trajectories of the effect of spatial visual structure on the processing of IWD.

Method

A computerized chat-simulation program was used for testing the hypotheses experimentally. Two independent conversation threads were concomitantly presented to each participant. In Experiment 1, the two dialogues were presented in distinct windows, whereby the distance between the two windows varied between the two experimental groups. In Experiment 2, the size (length) of the windows was manipulated; and, in addition, the conversations were either intertwined within a single window or presented in two distinct windows. After being exposed to the dialogues, participants completed a multiple-choice test, assessing their recognition of factual details that were mentioned during the conversations. Our hypotheses predicted that in Experiment 1, participants presented with spatially distant windows will score more than participants presented with spatially closer windows (H1); and that in Experiment 2, the advantage of two separate windows over an intermingled window presentation will be enhanced by the enlargement of the windows (H2).

Participants

Thirty-two (22 women, 10 men) and 87 (62 women, 25 men; Experiments 1 and 2, respectively) undergraduate and graduate students participated in the experiments, as part of the requirements of a communication technologies course. The participants were 18 to 36 years old (mean age = 23.24), and were all native Hebrew speakers. The study was approved by a human subjects review board at the University.

Apparatus

Computerized chat simulation. The hypotheses were tested through the use of a computerized chat-simulation program (Bright-Aqua Technologies LTD., Israel), which can be found online at http://spirit.tau.ac.il/com. Two 20-lines long Hebrew conversation threads were presented to all participants. In one of these, two persons (a woman and a man) discuss football teams and food preferences, and in the other two persons (again a woman and a man) discuss Israeli politics. Both dialogues were fabricated, and are not claimed to be representative of typical chat conversations; the cognitive phenomena that are being examined in this study are of a general nature.
and, therefore, it was not required that actual chat conversation be exactly mimicked or reproduced.

The lines of text accumulate in typical chat mode—that is, each new line begins with the name of its author (followed by a semicolon), and appears below the previous line; when the window becomes full, new lines push older ones out of sight.

In Experiment 1, the dialogues were presented in two different conditions:

1. **Approximate two windows:** The two conversations were spatially separated, each presented in its own window (each window was 360 pixels high and 250 pixels wide—16 text lines long). The distance between the windows was 15 pixels (i.e., they were almost touching each other). A new line of text was added every 6.1 sec on the left window and 5.9 sec on the right window (the intervals were unequal to prevent simultaneous accumulation of lines in the 2 windows). The location of each of the two conversations on the screen (i.e., left or right) was counterbalanced; there was no significant difference in the participants’ scores between left and right location. (This remark applies also to all subsequent conditions, in both experiments, where participants were presented with 2 windows.) The windows closed 7 sec after the conversations ended.

2. **Distant two windows:** The two conversations were spatially separated and presented exactly as in the approximate two windows condition, except for the distance between the windows, which was extended in this condition to 495 pixels (i.e., as far from each other as the 17-in. screen allowed).

In Experiment 2, a 2 × 2 experimental design was used, with main factors of number of windows (1 or 2) and windows length (short or long). Thus, in this experiment, the dialogues were presented in four different conditions:

1. **One short window:** The conversation threads were intertwined within a single window (225 pixels high and 400 pixels wide—10 test lines long). A new line of text was added every 2.75 sec. The window closed 2.75 sec after the conversations ended.

2. **One long window:** This condition was identical to the one short window condition, except for the length (height) of the window, which was extended to 450 pixels (20 text lines long).

3. **Two short windows:** The two conversations were spatially separated, each presented in its own window (each window was 113 pixels high and 400 pixels wide—5 text lines long). A new line of text was added every 5.6 sec on the left window and 5.4 sec on the right window (the intervals were unequal to prevent simultaneous accumulation of lines in the 2 windows). The windows closed 5.5 sec after the conversations ended.

4. **Two long windows:** This condition was identical to the two short windows condition, except for the length (height) of the windows, which was extended to 250 pixels (10 text lines long).

Thus, in all four conditions presentation duration was the same (approximately 110 sec). The total amount of text presented at any given time was identical for short
windows and was doubled for long windows, but was equal for one or two windows within each window-length condition. Presentation time of each line was shortened in Experiment 2 compared to Experiment 1 to avoid a possible ceiling effect, reflected in very high scores among all groups when the windows are extended in height, as a pilot study indicated. Thus, a shorter presentation time increased the difficulty of the task and allowed differences between the groups to be seen.

Multiple-choice test. The participants’ ability to follow the conversations was measured using a hard-copy, multiple-choice test that was completed immediately after the conversations ended. The multiple-choice examination was the same for all conditions and consisted of 20 questions, 10 regarding each conversation thread (grouped together under distinct headers—first, 10 questions about first dialogue and then 10 questions about the other. The order of the dialogues was counterbalanced; there was no significant difference in the participants’ scores between the 2 dialogue presentation orders used). Each question had five answers (4 of which were distracters, and 1 was the correct answer). All questions were factual, concerned with specific details that were mentioned during the conversations. The internal reliability of this test Cronbach’s alpha was 0.96 and 0.99 in Experiments 1 and 2, respectively. Total test scores ranged from 45% to 95% correct answers with a mean of 67.19% (SD = 16.70) in Experiment 1 and from 15% to 70% correct answers with a mean of 43.91% (SD = 14.77) in Experiment 2.

Procedure

Participants were seated by the experimenter in front of a computer screen, where the instructions for the experiment were presented. The instructions preceding each condition described what the participant should expect. When the participants completed reading the instructions, he or she pressed a button, after which the windows with the textual conversations that they were instructed to follow appeared on the screen. Each participant was randomly assigned one of the experimental conditions (n = 16 and n = 21–23 per group in Experiments 1 and 2, respectively). The order of the conditions was random. The participants were passive observers of the textual conversations—they could not take part in them.

When the conversations ended, the chat windows closed, and then a message appeared prompting the participants to receive from the experimenter the (hard copy) multiple-choice test examining their understanding of the textual conversations they viewed. After completing the multiple-choice test, participants completed general information forms including age, gender, and experience with computers and Internet chat.

Results

Experiment 1

There was no difference between the mean proportion of correct answers in the approximate two windows group (M = 66.88, SD = 15.15) and the distant two windows group (M = 67.50, SD = 10.17), t(30) = 0.14, p = .89.
Experiment 2

Mean proportion of correct answers was analyzed in a $2 \times 2$ two-way analysis of variance (ANOVA), with main factors of number of windows (1 or 2) and windows length (short or long). Figure 1 presents the means and standard errors of proportion of correct answers in the four experimental groups. As can be seen, the two-window groups scored higher than the one-window group under both window-size conditions.

Moreover, the difference between the one-window and two-windows groups was considerably higher in the long windows, compared to the short windows. These findings were confirmed by a two-way ANOVA, which yielded a significant main effect of number of windows, $F(1, 83) = 48.61, p < .0001$, as well as a significant interaction of number of Windows $\times$ Windows Length, $F(1, 83) = 11.55, p < .001$. Post hoc comparisons (Fisher PLSD) showed significant differences between the one-window and two-windows groups under the long-windows condition ($p < .00001$; Cohen’s $d = 2.31$), as well as under the short-windows condition ($p < .025$; Cohen’s $d = 0.75$). The proportion of correct answers did not correlate with level of computer or Internet use, nor with chat or instant-messaging software experience, in either of the two experiments.

Discussion

Both results reported earlier enhance our understanding of the way IWD is perceived and processed, and of the way its processing helps support conversational multitasking. The first, negative result tells us that conversation windows are placed in a visual space that is conceived of as being discrete, in a sense. The distance between the windows does not affect the degree to which the conversations in them are distinguished—what matters is that the two conversations take place in distinct loci. This

Figure 1  Effects of Number of Windows and Window’s Length on Proportion of Correct Answers. Means and standard errors of proportion of correct answers in the four experimental groups. The advantage of two over one window was fortified when the length (height) of windows was extended. Asterisks indicate significant difference between one-window and two-windows groups.
is contrary to other contexts—both on and off screen—in which it seems that the greater the proximity between events, the more we are inclined to think of them as related to each other. The two conversation windows were placed very close to each other in one of the conditions, but this did not create any effect of “crowding.” This fact suggests that several (more than 2) windows could be open on a user’s computer screen without reduction in his or her ability to multitask (provided that the overall rate of text accumulation is kept constant), as the diminishing distances between windows will not reduce the user’s ability to distinguish between the different conversational loci. Also, the same fact may have implications to some social aspects of life online: When we multitask, we perceive ourselves as jumping between distinct, unrelated (social) places, not as being involved in a complex social situation in which several discussions occur at the same time.

The other, positive result confirming \( H2 \) opens the door for further research on the way synchronous text is perceived. The hypothesis that better understanding is achieved when conversations are separated into distinct windows because of visual backtracking can be examined more thoroughly by using eye-movement research methodology (Rayner & Pollatsek, 1994). Such research could provide us with a more detailed and accurate picture of the way visual attention is drawn to previous lines in each window and then returned to the current, most recently added text. These results, in turn, may have applications in the domain of human–computer interface (HCI): They could help us design online environments in which multitasking is more easily realized.

Of course, the degree to which multitasking is realizable and beneficial depends on various important factors other than the technological one. In particular, various aspects of the content of the conversational tasks that are to be juggled together may affect the possibility and utility of multitasking. For example, it can be asked whether social conversations are more easily multitasked than work-related discussions, or whether content that demands less concentration (be it either work-related or social) renders multitasking more easily achieved and useful. Are discussions that are unrelated to each other more easily multitasked, or are things the other way around? Also, it seems reasonable to hypothesize that the participants’ level of engagement and interest in the discussions should influence their ability to follow them. The legitimacy of (and interest in) all these questions and avenues of research are certainly not ignored or downplayed by this research; its framework is not naïve (and fallacious) technological determinism. Rather, it is our view that the new technological setting gives rise to new affordances that should be analyzed and understood, together with the study of the various ways in which people put them to use.

Another important and interesting cluster of considerations relevant to this study has to do with the fact that the phenomenon of conversational multitasking is not limited to the computer screen. New communication technologies allow us more than ever before to engage in several communication tasks at the same time. People—notoriously, students—send and receive text messages on their mobile phones while attending face-to-face communicative situations (Ling, 2004), business men...
and women read e-mails on their Blackberries during meetings (Hallowell, 2006), we are often active on the phone and online at the same time, and the list could go on and on. In resonance with some of the questions raised in the previous paragraph, some lament this widespread multitasking (which is certainly more prevalent among young people) as a dangerous process of a culturally induced acquisition of attention deficit disorder. Others take this phenomenon to be a natural and positive aspect of modern life, especially for digital natives.

Does the research reported here bear in any way on the discussion of this more general issue? We argue that it does, in at least two ways. First, note that the examples of multitasking considered in the previous paragraph are in accord with the theoretical outlook presented in the first section of this article. It is difficult to juggle together two or more auditory conversations—be they f2f or technologically mediated—for the perceptual reasons outlined earlier; therefore, in most cases of multitasking, we see a text channel multitasked with another (auditory or textual, f2f or mediated) conversational thread. Thus, some of the findings concerning on-screen multitasking—those reported in this article and others that future research may yield—could be applicable to the analysis of hybrid situations (involving on-screen and off-screen communication) of the kind we are considering now.

Second, and more generally, the joint consideration of on-screen and off-screen multitasking sheds some light on the interesting interplay between cognitive capacities, technology, and social norms. As noted in the first section of this article, one possible reason due to which auditory, f2f multitasking is not prevalent is that it is censored by current social norms: It is just not polite to talk to one person and appear to be listening to another conversation that takes place near by. However, what is the source of this norm? A plausible conjecture is that it is grounded (at least in part) in cognitive considerations: It is implicitly assumed that people cannot pay attention to more than one auditory thread of conversation; therefore, an attempt to listen to another conversation is interpreted as disinterest in the conversation that one is currently engaged in.

If we turn our attention now back to the on-screen cases of multitasking, we see a picture that is different in two ways. First, multitasking of this sort is typically done in private, without conversation partners being aware of it: I can open three conversation widows on my screen at the same time and, if I am competent enough, none of my three interlocutors will know that he or she is sharing my time and attention with others. Second, at least according to some anecdotal information gathered by these researchers, on-screen multitasking is not viewed by Internet users as a breech of social norms at all: Several experienced “cybernauts” told us that among their friends it is quite legitimate to lead several on-line discussions at the same time. Why is such time sharing legitimate online? Again, it is plausible that the reason is the cognitive feasibility of doing so, demonstrated empirically (e.g., in this research). Thus, we see here the same pattern of interplay between cognition and social norm as before, but this time mediated by technology: The new technological setting alters the cognitive viability of a certain kind of behavior and, consequently, normative appraisal of such behavior changes.
The more concrete upshot of these considerations, coupled together with the ones preceding them in this discussion, is this: As exemplified by the phenomenon of communicative multitasking, new communication technologies make possible (e.g., cognitively possible) new kinds of behavior, and thereby challenge extant social norms. Our assessment of such novel kinds of behavior should not be strictly conservative, cherishing and upholding existing norms, nor should it embrace unthinkingly whatever technology offers us. Rather, concrete cases and contexts of use should be distinguished from each other and evaluated. In some cases, the result of such evaluation could be an assessment that older norms and preconceptions should be given up, and in other cases that measures should be taken against technologically induced processes that are detrimental to individuals and to society.

Finally, the limitations of this study, some of which were already pointed out earlier, should be explicitly acknowledged. First, we used a recognition, closed test as the single method for measuring the degree to which participants followed the conversation that rolled in front of their eyes. As noted in Dresner and Barak (2006), the reason for this choice is the relative difficulty of the cognitive task that the participants in the experiments were faced with: We wanted to avoid a flooring effect and, therefore, opted for a relatively easier test. However, it is certainly the case that other measures can and should be applied. For example, participants could be asked to take part in the conversations in some key points, and the adequacy of their contributions could be assessed and compared. Alternatively, the test could go beyond the basic level of recognition and examine participants’ ability to apply the contents they were exposed to and possibly to relate them to each other. A second point is that we limited ourselves to social conversations. As already noted earlier, variability in multitasking abilities among different kinds of content (along trajectories that need to be elaborated and specified) may be expected, and should be researched. Third, it should be noted that this study used a small, relatively homogenous sample, which obviously limits the external validity of our findings. Further studies of the phenomena discussed here should be conducted, involving larger samples and more variegated demographics. Moreover, variability in cultural and linguistic context was not considered at all in this research, and should be given attention in the future. Are there differences among cultures in their approach to multitasking, and do such differences affect on- and off-screen multitasking? Can bilinguals multitask more easily than monolinguals by employing distinct languages for distinct conversational tasks? These seem like questions worth looking into.

In sum, the foregoing discussion of the implications, possible applications, and limitations of this study all indicate that its subject matter is of interest and importance, raising questions in a variety of domains—communication, cognitive psychology, sociology, and HCI design. Further research into conversational multitasking may yield insights into communicative and cognitive linguistic processes, as well as useful theoretical inputs into the future development of better communication tools.
References


