Article Number: LALI: 00861

a0005

Computer-Mediated Communication: A Cognitive Science Approach

S E Brennan and C B Lockridge, Stony Brook University (SUNY), Stony Brook, NY, USA

p0005 Human languages and the conventions for using them evolved with people interacting face-to-face. Likewise, face-to-face interaction is the key setting in which children acquire language. Despite these origins, more and more communication now takes place between people who are not copresent in the same space at the same time, via technologies such as e-mail, instant messaging, cell phones, voice mail, and videoconferencing. How do people adjust when communication is mediated? How is language processing affected? And how is conversation shaped by the medium in which it is conducted?

p0010 Consider this example: Early one morning, Calion is typing an e-mail message to his wife Aisha, who will soon be in her office in the English Department across campus. If Calion wants Aisha to meet him later for a bite to eat, he cannot simply say, "Meet me for Indian after class." Many things can go wrong. For instance, Calion needs to be confident that Aisha can receive the message (will she remember to plug her laptop into the campus network?), will be attentive enough to notice that a message has arrived (will she be too busy meeting with undergraduates to check e-mail?), will figure out what Calion intends (their common ground will likely enable her to figure out what he intends by "Indian" and "after class"), and is willing and able to commit herself to the action he proposes (or will she have a meeting or other commitment at the time he's proposing?). So after hitting the send key, Calion must await evidence that Aisha has received, understood, and committed to his invitation. For her part, Aisha doesn't simply read Calion's message and resolve to head out to the food court at the appropriate time; she sends an e-mail reply that gives evidence that she has received, understood, and accepted the invitation. Or if she needs to negotiate or clarify the plan, she may switch media and try to instant-message him; this will work only if they can both attend to their screens at the same time. If the expected e-mail response is not forthcoming soon enough, Calion may take the initiative to actively seek out evidence by calling Aisha on her cell phone.

p0015 The point is that communication does not succeed automatically, just because two people speak the same language, possess the same cognitive architecture, and know the same things. Regardless of the mode of communication, people jointly construct meanings by engaging in an active process of 'grounding', or seeking and providing evidence that they understand one another (Brennan, 1990, 2004; Clark and Brennan, 1991; Clark and Schaefer, 1989; Clark and Wilkes-Gibbs, 1986; Schober and Clark, 1989). Contributions to conversations are coordinated in two phases: a presentation phase and an acceptance phase (Clark and Schaefer, 1989). As Calion's invitation illustrates, an utterance does not count as an actual contribution to a conversation (nor as part of the interlocutors' common ground) until its acceptance phase is complete. After (or depending on the modality of communication, even while) one person presents an utterance, the addressee provides evidence of attention, understanding, and uptake. This evidence may be implicit, in the form of continued eye contact or a relevant next turn (as when an answer follows a question), or explicit, in the form of a rephrasing, a request for clarification, or a modification of what came before (Clark and Schaefer, 1989). Both speaker and addressee take responsibility for seeking and providing evidence; often who takes the initiative at any given moment depends on who can do so more easily (Brennan, 1990). In this way, interlocutors in a collaborative task adjust their individual effort in order to minimize the effort they expend jointly, in order to reach a grounding criterion, or degree of certainty that they understand one another sufficiently well for current purposes (Clark and Wilkes-Gibbs, 1986). AU:1

p0020 In the rest of this article, we will briefly present some robust findings about mediated communication and discuss them in the context of the grounding framework. The grounding framework conceptualizes mediated communication as a coordinated activity constrained by costs and affordances (Clark and Brennan, 1991). This framework is compatible with both experimental and descriptive findings about communication (whether electronic or face-to-face) and can be used to predict and explain how communication media shape language use.

Basic Findings About Mediated Communication: Speech and Visual Evidence

s0005

p0025 The richness associated with face-to-face conversation diminishes when communication goes electronic: for instance, prosody is absent when text is the currency of exchange rather than speech; spontaneous

facial expressions and gestures are lost when an interlocutor can't be seen; and conversational turns grow longer with voice mail or e-mail messages than with media that support more fine-grained interaction, such as electronic chat or telephone conversations.

p0030 Yet perhaps surprisingly, people are able to communicate quite clearly and easily over a wide variety of media, including those with relatively low bandwidth (e.g., text-based media); in fact, cognitive tasks tend to be accomplished just as well over lower-bandwidth media as face-to-face (for a comprehensive review, see Whittaker, 2002). Despite the common expectation that the more similar a medium is to face-to-face communication, the better communication should be, study after study has failed to confirm AU:2 this 'bandwidth hypothesis' (Brennan, 1990, 1994; Brennan and Lockridge, 2004; Karsenty, 1999; Ohaeri, 1998; Whittaker, 2002). Clearly, more bandwidth is not necessarily better. In fact, mediated communication sometimes offers tangible advantages over face-to-face conversation, especially when it is of value to be able to edit utterances, review them, or save them as a paper trail; when it is useful to broadcast them to many addressees at once; or when interlocutors' schedules prevent them from attending to a message at the same time.

p0035 Some studies have documented media-based differences in efficiency among task-oriented conversations (higher efficiency is when the same task is accomplished just as well in less time or with fewer words). In comparisons of different configurations of speech, handwriting, teletyping, and video, Chapanis and colleagues found early on that remote communication is much less efficient without speech; the only way to substantially improve a medium's efficiency is to add a voice channel (Chapanis *et al*., 1972; Chapanis *et al*., 1977; Ochsman and Chapanis, 1974). The ability to coordinate using speech typically makes a task more efficient by a factor of two or more.

p0040 Yet adding a video channel to a medium that already includes speech may do nothing to improve either performance or efficiency (Chapanis *et al*., 1972; Chapanis *et al*., 1977; Ochsman and Chapanis, 1974; Whittaker, 2002; Williams, 1977). Of course, this depends on what visual information is transmitted: For cognitive or physical tasks where the focus is on the task activity, there are few if any benefits to seeing a partner's face (Fish *et al*., 1993; Gaver *et al*., 1993; Whittaker, 1995, 2002), despite repeated attempts by telephone companies and teleconferencing researchers to supply disembodied talking heads along with people's voices. (Seeing the face of a remote interlocutor can, however, have effects upon interpersonal social judgments, affiliation, or adversarial situations involving negotiation; see Whittaker, 2002 for a review). Visual information other than faces, such as views of the objects or task under discussion, can be very useful in task-directed communication (Anderson *et al*., 2000; Brennan and Lockridge, 2004; Clark and Krych, 2004; Kraut *et al*., 2002; Whittaker, 1995, 2002). The impact of a particular kind of visual information can be explained by the role it plays in grounding.

p0045 Consider the task of giving someone driving directions. This is easiest when both partners can see and point at the same map. In one study of remote communication (Brennan, 1990, 2004), two partners had the same map displayed on their screens and could speak freely to one another. One, the director, knew the target location, and directed the other, a matcher, to move his car icon to the target. Half the time the director could see on her map where the matcher's car was, and half the time she could not (the situation was asymmetrical; the matcher saw his own car icon in both conditions). When directors had visual evidence about matchers' understanding, matchers quite literally came to use icon motion to replace their turns in the conversation. And directors could quickly tell when matchers understood where the target location was, so directors took responsibility for deciding when it was time to move on to the next trial. In trials without such evidence, directors waited for matchers to tell them when they understood well enough to move on. Trials with visual evidence also took less than half as long as those without, because pairs could ground in parallel; that is, while the director presented a description, the partner conducted the acceptance phase simultaneously by silently moving his icon (see **Figure 1**). Without visual evidence (see **Figure 2**), he had to give verbal evidence, speaking after the director's description, which made the granularity of interaction much larger.

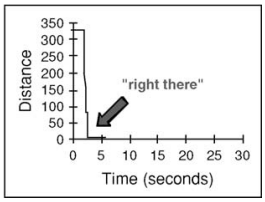
Grounding in Mediated Communication

s0010

p0050 In mediated communication, interlocutors typically inhabit different times and/or different places, so some aspects of coordination can be more difficult than in face-to-face conversation, particularly if people are limited to a medium that does not facilitate grounding or if techniques for grounding within the medium are unknown. **Table 1**, adapted from Clark and Brennan (1991), compares key affordances of face-to-face conversation with those of four other communication media.

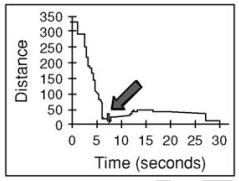
p0055 Grounding in communication can be decomposed into various sub-tasks bearing distinct costs (Clark and Brennan, 1991; Brennan, 1998), with the idea that people must adapt techniques for grounding to

D: ok
now we're gonna go over to
M-Memorial Church?
and park right in
right there
that's *good.*
M: *that's* rude
to park in the church.
D: hheh heh



f0005 **Figure 1** In this example, D can see M's icon, which provides immediate visual evidence about how M understands D's description. The exchange occurs as M moves his icon toward the target location described by D. D takes the initiative to propose that M has the right location using a deictic cue (''right there'') after only 6 seconds. The graph shows the convergence of icon to target over time, with the point at which the icon reaches the target marked on the graph by an arrow.

D: you're in the upper
far far upper right-hand corner of
the screen it says Sea Street?
M: *yah*
D: *way* at the top?
M: yeh *[icon in correct location]*
D: you're
you're just a little bit on the road
and the
corner of your car is touching A
of Sea
but you're mostly off the road
the road is to your right
just a- touching *the car*
M: *the road* is to the right of
the car?
D: put the road-put the car right on
the road and you'll overlap me.
M: ok



f0010 **Figure 2** In this example, D cannot see M's icon and so grounding depends on the verbal evidence of understanding sought by D and provided by M. After the icon reaches the correct location there follows a lengthy period of grounding before they reach their grounding criteria and can conclude that they understand one another.

the affordances and constraints of the current medium in order to meet these costs. Sub-tasks that incur grounding costs include getting a partner's attention in order to initiate communication (startup costs); producing a presentation by speaking or typing or in some other manner (production costs); timing the placement of feedback (asynchrony costs) or of a conversational turn (speaker-change costs); pointing, demonstrating, or gesturing in order to refer or clarify content (display costs); awaiting, reading, or listening to a partner's utterance (reception costs); monitoring the partner's focus of attention and, if the dialog is task-oriented, any relevant activities or tangible products that make up the task (monitoring costs), preventing misunderstandings or repairing errors caused by self or partner (repair costs), and maintaining politeness (face-management costs) [based on Clark and Brennan, 1991 and Brennan and Ohaeri, 1999]. Discussing a few of these costs will help show how grounding shapes behavior.

p0060 Startup and monitoring costs are low for people who are physically copresent because they can easily monitor what a partner is doing, assess when the time is right for an interruption, and initiate a conversation by speaking to get the partner's attention (for review of physical proximity effects, see Kraut *et al*., 2002). Startup is more costly for a video conference, since participants must arrange to be present in appropriately equipped facilities at the same time. Starting up a telephone call is unpredictable on a landline, as people are often away from such telephones; but with proliferating cell phones, calls find addressees regardless of their locations and so startup costs are somewhat lower.

p0065 Production costs are typically higher for text than for speech because most people find it harder to type than to speak, so typed utterances tend to be shorter than spoken utterances. In one study, people were more likely to sacrifice politeness when typing than when speaking when it took more words to frame a polite utterance (e.g., inviting a partner's input using hedges), but not when it took the same number of words to be polite (e.g., inviting the partner's input with questions); moreover, individuals with faster typing speeds used more politeness devices per 100 words than those who typed slowly (Brennan and Ohaeri, 1999). This finding demonstrates that people who communicate remotely do not actually become depersonalized or cease to care about politeness (as some social psychological theories have suggested), but that when they must struggle to meet production costs they do this at the expense of something else, such as face-management. It also illustrates that grounding costs are not independent of one another; often one cost must be traded off against another, and such trade-offs are made differently in different media.

p0070 As another example, consider repair costs: When communication is cotemporal, such as with voice, text-based chat, and instant messaging, the grain of interaction is small, and turns tend to be shorter, less formal, and more numerous than in larger-grained text-based media (such as letters or e-mail). So any errors or misunderstandings can be addressed quickly, and repair costs are relatively low (more so for speech than for text, since production costs are higher for text).

t0005 **Table 1** Affordances of Communication Media

	Media				
Affordances of Media	*Face-to-face*	*Video conference*	*Telephone*	*Instant messaging or chat*	*E-mail*
(1) *Physical co-presence*: Participants share a physical environment, including a view of what each is doing and looking at.	++	??	−−	−−	−−
(2) *Visibility*: One participant sees another, but not necessarily what the other is doing or looking at.	++	+	−−	−−	−−
(3) *Audibility*: One participant can hear another.	++	++	++	−−	−−
(4) *Cotemporality*: Messages are received without delay (close to the time that they are produced and directed at addressees), permitting fine-grained interactivity.	++	??	++	??	−−
(5) *Simultaneity*: Participants can send and receive messages at the same time, allowing communication in parallel.	++	??	++	??	−−
(6) *Sequentiality*: Participants take turns in an orderly fashion in a single conversation at a time; one turn's relevance to another is signaled by adjacency.	++	++	++	−−	−−
(7) *Reviewability*: Messages do not fade over time.	−−	−−	−−	++	++
(8) *Revisability*: Messages can be revised before being sent.	−−	−−	−−	++	++

[a]Present in a particular medium: ++; present to a limited extent: +; present in some systems: ??; absent: −−. Physical co-presence (1), the hallmark of face-to-face communication, nearly always includes affordances (2) through (5). Adapted from Clark & Brennan, 1991.

p0075 In closing, the grounding framework is a usefulvantage point from which to view, understand, and predict the effects of new media upon communication. The abundance and portability of new communication programs and devices (PDAs, added cell phone functionality such as digital photography, more extensive wireless networks, unobtrusive methods for eye-tracking, multimedia Internet content, etc.) will continue to make it even easier for mediaphiles to switch mid-conversation from one medium to another, as in our opening example of Calion's e-mail invitation to Aisha. Recently the New York Times chronicled a man and his BlackBerry (a portable wireless device for e-mail and instant messaging): "He once saw a romantic interest walk into a bar and immediately called her on her cell phone. 'I saw her look at the phone and put me right to voice mail,' he said, still indignant. But then he sent her a BlackBerry message, which made her laugh and prompted her to walk over and find him." The ability to spontaneously switch media within the same conversation enables increasingly flexible and innovative techniques for grounding.

See also: Computer-mediated communication (00703); Context and common ground (01088); Dialogue and Interaction (00792); E-mail, Internet, Chatroom talk, pragmatics of (00377); Multimodal Interaction with computers (04362); Pauses and Hesitations: A psycholinguistic approach (00796); Psycholinguistics: Overview (00788).

Bibliography

Anderson A H, Smallwood L & MacDonald R (2000). 'Video data and video links in mediated communication: What do users value?' *International Journal of Human-Computer Studies 52*, 165–187.

Brennan S E (1990). *Seeking and providing evidence for mutual understanding.*' Unpublished doctoral dissertation, Stanford University, Stanford, CA.

Brennan S E (1998). 'The grounding problem in conversation with and through computers.' In Fussell S R & Kreuz R J (eds.) *Social and cognitive psychological approaches to interpersonal communication*. Hillsdale, NJ: Erlbaum. 201–225.

Brennan S E (2004). 'How conversation is shaped by visual and spoken evidence.' In Trueswell J & Tanenhaus M (eds.) *World situated language use: Psycholinguistic, linguistic, and computational perspectives on bridging the product and action traditions*. Cambridge, MA: MIT Press.

Brennan S E & Lockridge C B (2004). 'Monitoring an addressee's visual attention: Effects of visual co-presence on referring in conversation.' Unpublished manuscript.

Brennan S E & Ohaeri J O (1999). 'Why do electronic conversations seem less polite? The costs and benefits of hedging.' Proceedings, International Joint Conference on Work Activities, Coordination, and Collaboration (WACC '99). San Francisco, CA: ACM. 227–235.

Chapanis A, Ochsman R & Parrish R (1972). 'Studies in interactive communication I: The effects of four communication modes on the behavior of teams during cooperative problem solving.' *Human Factors 14*, 487–509.

Chapanis A, Ochsman R & Parrish R (1977). 'Studies in interactive communication II: The effects of four communication modes on the linguistic performance of teams during cooperative problem solving.' *Human Factors 19,* 487–509.

Clark H H & Brennan S E (1991). 'Grounding in communication.' In Resnick L B, Levine J & Teasley S D (eds.) *Perspectives on socially shared cognition.* Washington, DC: APA. 127–149 [Reprinted in Baecker R M (ed.) *Groupware and computer-supported cooperative work: Assisting human-human collaboration*. San Mateo, CA: Morgan Kaufman. 222–233.].

AU:3

Clark & Krych (2004). 'Speaking while monitoring addressees for understanding.' *Journal of Memory and Language 50,* 62–81.

AU:4

Clark H H & Schaefer E F (1989). 'Contributing to discourse.' *Cognitive Science 13,* 259–294.

Clark H H & Wilkes-Gibbs D (1986). 'Referring as a collaborative process.' *Cognition 22,* 1–39.

Fish R, Kraut R & Root R (1993). 'Video as a technology for informal communication.' *Communications of the ACM 36,* 48–61.

Gaver W, Sellen A & Heath C (1993). 'One is not enough: Multiple views in a media space.' In *Proceedings of CHI '93: Human Factors in Computing Systems.* New York: ACM Press. 335–341.

Hanna J E & Brennan S E (2004). 'Using a speaker's eye-gaze during comprehension: A cue both rapid and flexible.' Abstract, *17th Annual CUNY Conference on Human Sentence Processing*. MD: College Park.

Karsenty L (1999). 'Cooperative work and shared context: An empirical study of comprehension problems in side-by-side and remote help dialogues. *Human-Computer Interaction 14(3),* 283–315.

Kraut R E, Fussell S R & Brennan S E (2002). 'Understanding effects of proximity on collaboration: Implications for technologies to support remote collaborative work.' In Hinds P & Kiesler S (eds.) *Distributed work*. Cambridge, MA: MIT Press. 137–162.

Lee J (2004). 'A BlackBerry throbs, and a wonk has a date.' *New York Times Sunday Styles, Section 9,* May 30. 1–2.

Ochsman R B & Chapanis A (1974). 'The effects of 10 communication modes on the behavior of teams during cooperative problem-solving.' *International Journal of Man–Machine Studies 6,* 579–619.

Ohaeri J O (1998). 'Group processes and the collaborative remembering of stories.' Unpublished doctoral dissertation, State University of New York at Stony Brook.

Whittaker S (1995). 'Rethinking video as a technology for interpersonal communications: theory and design implications.' *International Journal of Man-Machine Studies 42,* 501–529.

Whittaker S (2002). 'Theories and methods in mediated communication.' In Graesser A, Gernsbacher M & Goldman S (eds.) *The Handbook of Discourse Processes.* Hillsdale, NJ: Erlbaum. 243–286.

Whittaker S J, Brennan S E & Clark H H (1991). 'Coordinating activity: An analysis of interaction in computer-supported cooperative work.' In *Proceedings of CHI '91: Human Factors in Computing Systems.* New Orleans, LA: Addison-Wesley. 361–367.

Williams E (1977). Experimental comparisons of face-to-face and mediated communication. *Psychological Bulletin 16,* 963–976.

Article Number: LALI: 00861

Author Query Form

Book: Encyclopedia of Language and Linguistics
Article No.: 00861

Dear Author,

During the preparation of your manuscript for typesetting some questions have arisen. These are listed below. Please check your typeset proof carefully and mark any corrections in the margin of the proof or compile them as a separate list. Your responses to these questions should be returned within seven days, by email, to Dr Jon Oberlander, email: J.Oberlander@ed.ac.uk, and copied to MRW Production, email: lali_proofs@elsevier.com

Query	Details Required	Author's response
AU1	Schober and Clark is not in ref list; please supply full reference.	
AU2	Brennan 1994 is not in ref list; please supply full reference.	
AU3	Please supply year of publication of reprint.	
AU4	Please supply author initial(s).	
AU5	Please provide descriptive text for the sound clips.	
AU6	Please provide photograph for author biography.	

Article Number: LALI: 00861

Non-Print Items

Abstract:

In successful communication, two people converge on the belief that they are talking about the same thing. Now that much communication takes place over electronic media that bridge time and space, many activities formerly conducted face-to-face are no longer so. We present a theoretical framework for conceptualizing conversation as a collaborative activity, that of 'grounding' (Clark and Brennan, 1991). This framework is consistent with experimental and descriptive findings about mediated communication, particularly for task-oriented dialogues, and is useful for explaining and predicting how a medium shapes communication.

AU:6

Biography:

Susan Brennan holds joint appointments at Stony Brook University in Psychology and Computer Science. Her Ph.D. is from Stanford University and her M.S. is from MIT's Architecture Machine Group (the Media Lab), where she worked on computer-generated caricatures and teleconferencing interfaces. She has conducted research on natural language processing and human-computer interaction at Atari, Apple, and HP Labs. She currently uses behavioral and eye-tracking techniques to study the interpretation, production, and adaptation of spontaneous speech in interactive settings, including how interlocutors converge in perspective, lexical choices, and dialect. She serves as consulting editor of *Psychological Science*; previously she served as associate editor of *Discourse Processes* and consulting editor of *Computational Linguistics*.

Biography:

Calion B. Lockridge received his B.A. from Langston University and his M.A. from Stony Brook University, where he is currently a Ph.D. student. He previously held a W. Burghardt Turner Fellowship and is currently an NRSA Fellow, supported by the National Institutes of Health's Institute for Deafness and Other Communication Disorders (NIDCD). His research interests include referential communication, the impact of individual differences in working memory span on language use in conversation, and the effects of visual copresence on communication.

Keywords: Multimedia, E-mail, Face-to-face conversation, Dialogue, Interaction, Coordination, Speaking, Feedback, Grounding, Common ground, Repair

Multi Media Component:

AU:5

Sound Clip 1 Given by Susan Brennan of audio data from Brennan, 1990.

Sound Clip 2 Given by Susan Brennan of audio data from Brennan, 1990.

Author Contact Information:

Susan Brennan
Department of Psychology
State University of New York at Stony Brook
Stony Brook, NY 11794-2500
New York
USA
susan.brennan@sunysb.edu