Using Language

Herbert H. Clark
(Stanford University)

Cambridge University Press, 1996,
xi+432 pp; hardbound, ISBN
0-521-56158-2, $54.95; paperbound,
ISBN 0-521-56745-9, $19.95

Reviewed by
Marilyn A. Walker
AT&T Labs Research

In *Using Language*, Herbert Clark proposes a broadly integrative theory of language as action. The book examines both social and cognitive aspects of language use, drawing from speech act theory (Austin 1965; Searle 1965; Allen and Perrault 1980), theories of discourse and dialogue (Reichman 1985; Grosz and Sidner 1986), and theories of social interaction (Goffman 1970; Brown and Levinson 1987; Sacks, Schegloff, and Jefferson 1974). In order to integrate these different perspectives, the book relies on empirical research by Clark and his students over the last fifteen years.

Clark begins the book with an overview of its central thesis that "language use is really a form of joint action" (p. 3), i.e., action carried out by an ensemble of people acting in coordination with one another. As a joint activity, conversation consists of a joint action and the individual actions by the conversational participants that constitute the joint action (Bruce 1975; Power 1974; Clark and Carlson 1982; Cohen and Levesque 1991; Grosz and Sidner 1990). Joint activities require coordination of both the **content** of the activity and the **process** by which the activity moves forward. The source of conversants' ability to coordinate is their **common ground**, the set of knowledge, beliefs and suppositions that they believe they share (Stalnaker 1978; Clark and Marshall 1981; Prince 1981). Common ground makes it possible for a speaker and a hearer to coordinate on what the speaker means and what the hearer understands the speaker to mean. In the remainder of the book, these core ideas are expanded and elaborated.

One core claim is that communicative acts are the primitive-level actions that all joint activities consist of. Speakers and hearers coordinate the production and interpretation of communicative acts through the mediation of a **signaling system** (Schelling 1960; Lewis 1969). The signs of the signaling system and their interpretation are often defined by convention, but the signal and its interpretation can also rely on the speaker's and hearer's common ground. Conversants accumulate common ground with joint signaling events, which move the conversants from one state of the conversation to the next. At the utterance level, each joint signaling event includes a **presentation phase**, in which the initiator presents a contribution, and an **acceptance phase**, in which the other conversants indicate the degree to which they understand and accept or reject that contribution. Conversants try to manage the production and interpretation of communicative acts with the **least collaborative effort**, i.e., with the smallest cumulative effort of the speaker and hearer combined. Since signaling systems encompass all types of signs, not merely those signaled by purely linguistic means, conversation can use any of a complex set of signals, such as facial expressions, gesture, speech, and shared awareness of actions and objects in a shared environment (Brennan 1990).

A second core claim is that dialogue has a layer of structure above the level of communicative acts. The conversants base their interpretation of each utterance-level act on the assumption that each utterance-level signaling event contributes to another joint action, namely some structured purpose, which defines a larger joint activity (goal or plan) at the discourse level (Bruce 1975; Power 1974; Allen and Perrault 1980; Litman 1985). At the discourse level, each joint signaling event consists of individual segments, or **sections** in Clark's terminology. A transition between two sections *s* and *t* depends on a set of relations that can hold between sections, such as *t* being subsequent to *s*, *t* being a part of *s*, or *t* being a digression from *s* (Reichman 1985; Litman 1985; Grosz and Sidner 1986). Each section requires the conversants to coordinate on the entry into the section and the exit from the section (Whittaker and Stenton 1988). One basis for coordination is the marking of transitions between sections in different ways, e.g., by discourse markers (Hirschberg and Litman 1993).

Computational linguists who read this book will notice that these core ideas are consistent in many ways with commonly assumed planning models of dialogue in computational linguistics. While Clark does not always make clear the relationship between his proposals and work in computational linguistics, many researchers in computational linguistics have used these ideas within computational frameworks that are more precise and testable. Clark argues that his perspective is inconsistent with planning models of dialogue, but Clark's view of planning mechanisms appears to reflect the state of the art circa 1971, when STRIPS was first proposed (Fikes and Nilsson 1971). He rejects all models based on planning because "people ... don't know in advance what they will actually do [because] they cannot get anything done without the others joining them, and they cannot know in advance what the others will do" (p. 319). However, it is not beyond the ability of current planning and control mechanisms to respond dynamically to a change in the environment or an unpredictable response by a conversational partner (Barto, Bradtke, and Singh 1995, inter alia). Furthermore, plan-revision mechanisms have been successfully used to model these phenomena (Cawsey 1993; Carletta 1992).

An additional thread that Clark integrates into his account is the effect of social relationships on language production and interpretation. Following Goffman (1970), and Brown and Levinson (1987), Clark claims that the production of each utterance-level signaling event is governed by a set of social constraints that derive from the social situation in which the conversation is carried out and the social relationship that holds between the conversants. These theories claim that it is primarily the orientation to social constraints that leads to many indirect forms of communicative acts. The use of planning representations in the interpretation of these **indirect speech acts** has been the focus of much work in computational linguistics (Perrault and Allen 1980; Litman 1985; McRoy and Hirst 1995), but these theories have had little impact on models of language production used in computational linguistics (with the exception of models reported by Hovy [1990] and Walker, Cahn, and Whittaker [1997]).

Thus, Clark provides a view of language use that integrates a number of perspectives, many of which have individually already been influential in computational linguistics. The integrative model that Clark presents has many complexities, but the book is accessible to readers with little or no background. The claims are nicely illustrated with excerpts from naturally occurring dialogues and backed up by empirical research by Clark and his students. What I find most remarkable about this book is the degree to which it reflects the convergence of various branches of discourse and dialogue theory on a set of common models based on theoretical perspectives in linguistics, psychology, sociolinguistics, philosophy, and computational linguistics. As a result, although the book is not written for a computational audience, it should be

of interest to computational linguists studying language as a means of acting in the world.

References

Allen, James F. and C. Raymond Perrault. 1980. Analyzing intention in utterances. *Artificial Intelligence*, 15:143–178.

Austin, J. L. 1965. *How to Do Things with Words*. Oxford University Press, New York.

Barto, Andrew G., Steven J. Bradtke, and Satinder P. Singh. 1995. Learning to act using real-time dynamic programming. *Artificial Intelligence*, 72(1–2):81–138.

Brennan, Susan E. 1990. *Seeking and Providing Evidence for Mutual Understanding*. Ph.D. thesis, Department of Psychology, Stanford University.

Brown, Penelope and Stephen C. Levinson. 1987. *Politeness: Some Universals in Language Usage*. Cambridge University Press.

Bruce, Bertram. 1975. Belief systems and language understanding. Technical Report AI-21, Bolt, Beranek and Newman.

Carletta, Jean C. 1992. *Risk Taking and Recovery in Task-Oriented Dialogue*. Ph.D. thesis, University of Edinburgh.

Cawsey, Alison. 1993. Planning interactive explanations. *International Journal of Man-Machine Studies*. 38:169–199.

Clark, Herbert H. and T. B. Carlson. 1982. Speech acts and hearer's beliefs. In Neil V. Smith, editor, *Mutual Knowledge*. Academic Press, New York, pages 1–37.

Clark, Herbert H. and Catherine R. Marshall. 1981. Definite reference and mutual knowledge. In Aravind K. Joshi, Bonnie L. Webber, and Ivan A. Sag, editors, *Elements of Discourse Understanding*. Cambridge University Press, pages 10–63.

Cohen, Philip and Hector Levesque. 1991. Teamwork. *Noûs*, 25:11–24.

Fikes, Richard E. and Nils J. Nilsson. 1971. STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, 2:189–208.

Goffman, Erving. 1970. *Strategic Interaction*. Blackwell, Oxford.

Grosz, Barbara J. and Candace L. Sidner. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12:175–204.

Grosz, Barbara J. and Candace L. Sidner. 1990. Plans for discourse. In Philip Cohen, Jerry Morgan, and Martha Pollack, editors, *Intentions in Communication*. The MIT Press, Cambridge, MA, pages 417–444.

Hirschberg, Julia and Diane Litman. 1993. Empirical studies on the disambiguation of cue phrases. *Computational Linguistics*, 19(3):501–530.

Hovy, Eduard H. 1990. Pragmatics and natural language generation. *Artificial Intelligence*, 43(2):153–197.

Lewis, David. 1969. *Convention.* Harvard University Press, Cambridge, MA.

Litman, Diane. 1985. Plan recognition and discourse analysis: An integrated approach for understanding dialogues. Technical Report 170, University of Rochester.

McRoy, Susan and Graeme Hirst. 1995. The repair of speech act misunderstandings by abductive inference. *Computational Linguistics*, 21(4):435–478.

Perrault, C. Raymond and James F. Allen. 1980. A plan-based analysis of indirect speech acts. *American Journal of Computational Linguistics*, 6:167–182.

Power, Richard. 1974. *A Computer Model of Conversation*. Ph.D. thesis, University of Edinburgh.

Prince, Ellen F. 1981. Toward a taxonomy of given-new information. In Peter Cole, editor, *Radical Pragmatics*. Academic Press, pages 223–255.

Reichman, Rachel. 1985. *Getting Computers to Talk Like You and Me*. The MIT Press, Cambridge, MA.

Sacks, Harvey, Emmanuel Schegloff, and Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking in conversation. *Language*, 50:325–345.

Schelling, Thomas C. 1960. *The Strategy of Conflict*. Harvard University Press, Cambridge, MA.

Searle, John R. 1965. What is a speech act? In M. Black, editor, *Philosophy in America*. Allen and Unwin, New York. Reprinted in John Searle, editor, *The Philosophy of Language*. Oxford University Press, 1971, pages 39–53.

Stalnaker, Robert C. 1978. Assertion. In Peter Cole, editor, *Syntax and Semantics, Volume 9: Pragmatics*. Academic Press, pages 315–332.

Walker, Marilyn A., Janet E. Cahn, and Stephen J. Whittaker. 1997. Improvising linguistic style: Social and affective bases for agent personality. In *Proceedings of the 1st Conference on Autonomous Agents, AGENTS 97*, pages 96–105.

Whittaker, Steve and Phil Stenton. 1988. Cues and control in expert-client dialogues. In *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics*, pages 123–130.

Marilyn Walker is a Principal Research Scientist at AT&T Labs Research. Her current research interests include spoken dialogue, models of performance for dialogue systems, and multimodal interfaces. She recently coedited a book entitled *Centering Theory in Discourse* (Oxford University Press), and a special issue of *Computational Linguistics* on empirical methods in discourse. Walker's address is AT&T Labs Research, D188, 180 Park Avenue, Florham Park, NJ 07932-0971; e-mail: walker@research.att.com