BRAIN AND LANGUAGE **59,** 7–49 (1997)
ARTICLE NO. BL971812

# Neuropragmatics: Neuropsychological Constraints on Formal Theories of Dialogue

Bruno G. Bara,* Maurizio Tirassa,* and Marina Zettin*,†

** Centro di Scienza Cognitiva, Università di Torino, Italy; and † Presidio Ospedaliero Ausiliatrice, Torino, Italy*

We are interested in the validation of a cognitive theory of human communication, grounded in a speech acts perspective. The theory we refer to is outlined, and a number of predictions are drawn from it. We report a series of protocols administered to 13 brain-injured subjects and to a comparable control group. The tasks included direct and indirect speech acts, irony, deceits, failures of communication, and *theory of mind* inferences. All the predicted trends of difficulty are consistently verified; in particular, difficulty increases from direct/indirect speech acts to irony, from irony to deceits, and from deceits to failure recovery. This trend symmetrically shows both in the successful situation and in the failure situation. Further, failure situations prove more difficult to handle than the relevant successful situation. In sharp contrast with previous literature, there is no difference between the subjects' comprehension of direct and indirect speech acts. The results are discussed in the light of our theoretical approach. 

## 1. INTRODUCTION

Pragmatics of communication is the study of the competence involved in intentional communication, independently of the means used to communicate. Theories of communicative competence are usually expressed in terms of the mental states involved in communication. The problem with theories

Maurizio Tirassa conducted part of the research at the Institute of Psychology of the Medical Faculty of Milano. This research has been funded by the National Research Council (CNR) of Italy, Coordinate Project on *Planning and plan recognition in communication,* 95.04019.CT11. We express our gratitude to the following persons: Monica Bucciarelli for help in the development of the experimental protocols; Mauro Adenzato and Rita Ardito for administering them and, together with Davide Cipollone and Renato Miceli, for statistical advice; Steven Small for encouraging us to apply neuropsychology to pragmatics; Adam and Daniel Small for competent translation of children's jargon; Helen Gigley for her generous attempt to improve the editing of the paper; and Roberto Rago, head physician of Presidio Ospedaliero Ausiliatrice in Torino, the speech therapists, and all the personnel of the hospital, for their kind cooperation.

Address reprint requests to Maurizio Tirassa, Centro di Scienza Cognitiva, Università di Torino, via Lagrange 3, Torino 10123, Italy. E-mail: tirassa@psych.unito.it.



0093-934X/97 $25.00

of competence is that they are often underdetermined: it is not sufficient to develop a sound logical analysis, since the relevant behavior might, at least in principle, be produced by a number of different systems. In order to constrain theories of competence, the actual human performance must be taken into account.

In a Popperian fashion, we shall try to falsify a competence theory of human intentional communication, developed by Airenti, Bara, and Colombetti (1993a, 1993b), or, at least, to match it against competing ones. An interesting way to do this is to explore neuropsychological consequences of the theory. We have examined individuals suffering from the consequences of a closed head injury (henceforth, CHI), trying to understand whether their communicative skills had been impaired by the trauma. We have chosen such pathology because it is commonly observed that CHI patients have a peculiar metalinguistic impairment, which seems to endure permanently even after a satisfactory recovery of other cognitive skills. These symptoms, however, are not easily defined unless a sound theoretical framework is used. Our subjects were young people (as is common in CHI) with no focal lesions and no aphasic symptoms and well recovering from the other impairments commonly associated with CHI.

Our aim here is not to develop a theory of a normal cognitive process starting from neuropsychological data. Rather, we want to test a number of predictions that distinguish our theoretical approach. A welcome side effect of our research might be to help lay a framework in which to systematize some neuropsychological findings about communicative impairment following CHI which, to the present date, lack strong theoretical foundations.

Given the very initial status of neuropsychological studies in the area of interest, we needed to explore very basic tasks, the fundamentals of what constitutes full communication. In particular, we explored the following tasks: direct vs. indirect speech acts, irony, deceits, failures of each of the previous categories, and *theory of mind* inferences.

The first task, direct and indirect speech acts, is a classic topic of pragmatics (and one of the few which have been explored in a neuropsychological perspective); however, our view differs significantly from the one prevailing in the literature. The second and third tasks, irony and deceit, are part of what we call nonstandard situations, i.e., situations where there is something more than the obvious meaning of an utterance, and it is the task of the hearer to realize what it is. Failures are interesting in that they are quite common phenomena in communication, and it is necessary for an agent to handle them satisfactorily; further, failure recovery involves planning, which is known to be impaired in many CHI subjects. Finally, as an additional topic, we also explored so-called *theory of mind* inferences (i.e., the ability of an individual to realize that another individual's mental states have their own independence and to reason upon them); although not communicative per se, these inferences play a major ancillary role in communication.

The paper is articulated as follows: in the next section, we discuss the epistemological position underlying our view of communication, to lay the foundations for Section 3, where the theoretical approach we have adopted is presented. In Section 4, we describe the characteristics of the subjects and the methods we have used. The experiments are presented in Sections 5 (direct and indirect speech acts), 6 (irony), 7 (deceits), and 8 (*theory of mind* inferences). Section 9 discusses the general findings.

## 2. EPISTEMOLOGY

Ideally, a cognitive theory should describe both the *competence* and the *performance* of the system under analysis. The terms were originally used by Noam Chomsky (1957, 1965; see also Bara & Guida, 1984); the distinction they are meant to emphasize is nowadays a fundamental of cognitive science. We will follow Marr's (1977, 1982) thorough discussion of the topic, since it is best suited for our current goals.

According to Marr, a description of a system's competence corresponds to an abstract description of *what* such system (or part of it) does and *why*, independently of any processing detail. Marr gives the name of *computational theories* to descriptions of this sort; this is due to his underlying postulate that the mind is a computational device. We maintain, on the contrary, that biological minds are no such device (as demonstrated, e.g., by Searle, 1980, 1992). This position, however, does not affect descriptions of the system's competence, since, as Marr himself argues, real computational concerns typically regard the lower levels of analysis, where the actual processing machinery is taken into account. Thus, nothing of what we will say in this work depends on the acceptance or rejection of the computational postulate.

The theory of human communication we adopt here lies on the competence level. Current theories of human action and communication are often specified in terms of the mental states involved in the task (belief, intention, and so forth) and of the cognitive dynamics leading from a mental state to another. A state-of-the-art sample of this perspective may be found in the volume edited by Cohen, Morgan and Pollack (1990); our own approach will be outlined in the next section.

Mental states are analyzed in terms of the functional role they have in cognition. Belief, for example, is used to encode an agent's current knowledge of the world; intention is used to encode the agent's behavioral decisions, and so on. To take a naive example, one might theorize that Ann's belief that Bob has a candy, together with her desire to eat it, causes (by default) her intention to ask Bob for the candy. As a first approximation, let a more abstract version of the same idea be written like this:

$$\mathrm{BEL}_x\ \text{has}\ (y, p) \wedge \mathrm{WANT}_x\ \text{has}\ (x, p) \Rightarrow \mathrm{INT}_x\ \mathrm{DO}_x\ \text{ask}\ (y, p)$$

This formula comprises three terms and two relations. The first term, $\text{BEL}_x$ has ($y$, $p$), expresses agent $x$'s belief that agent $y$ has possession over object $p$. The second term, $\text{WANT}_x$ has ($x$, $p$), expresses $x$'s desire to have $p$. The third term, $\text{INT}_x$ $\text{DO}_x$ ask ($y$, $p$), expresses $x$'s intention to ask agent $y$ for $p$.

As regards to two relations, $\wedge$ corresponds to the logical connective AND and needs no further explanation in this context. The arrow $\Rightarrow$ indicates a default derivation, i.e., one that can be blocked whenever inappropriate to the context (Reiter, 1980). The classic example of default reasoning involves birds: we know that most birds can fly, so when we are told of a certain animal $x$ that it is a bird, we can reasonably assume that it can fly. But, if we then come to know that $x$ is a penguin (or that $x$ is an ostrich, or has a broken wing, and so on), we will have to retreat (block) our previous conclusion. Thus, the assertion that

$$\text{bird}\ (x) \rightarrow \text{flies}\ (x)$$

(i.e., if $x$ is a bird, then $x$ flies) is better written as the default assertion that

$$\text{bird}\ (x) \Rightarrow \text{flies}\ (x)$$

(i.e., if $x$ is a bird, then $x$ flies *unless there is further information that x cannot fly).*

We use a default formalization because a complete formal description of all of an individual's mental states is, in principle, impossible. In the candy example, Ann, although having both the belief that Bob has a sweet and the desire to eat it, might decide to say nothing, e.g., because she thinks that Bob is too selfish to give her the candy. Thus, knowing that an agent has a certain mental state can never guarantee the agent's action. It is impossible, in other words, to give a complete account of all the conditions necessary and sufficient for action (this problem is known in artificial intelligence under the name of *qualification problem:* McCarthy, 1980; see also Airenti & Colombetti, 1991).

On the competence level, the system is described as an abstract machine capable of producing a certain behavior. However, there is something more than this to the study of cognition. Although, in principle, any behavior may be produced by an indefinite number of different systems, only one of these is of interest to the human sciences, namely, the one corresponding to the human mind/brain. Theories of human competence, therefore, need be further constrained.

As a first thing, the validation of such theories depends on the adherence of the competence they describe to the characteristics exhibited by the real organism; i.e., on its capability to generate the relevant behavior. This corresponds to what Chomsky (1965) calls the *performance* of the system; what counts as a system's performance depends on the task considered and on the theory being developed. To study performance impairments after brain damage may help constrain the underlying competence theory.

Theories of communication, as we have discussed above, are usually cast in terms of the relevant mental states and of their dynamics, supposed to realize the process under analysis. Mental states are not localized, i.e., there seems to be no specific cerebral area where they reside. They can reasonably be supposed to result from the activity of large areas of the brain, each making its own contribution to the overall state of an individual's mind. When such areas are devoted to specific functions, local neural damages may produce specific impairments in the mental states an individual may entertain, or in their content, or in their dynamics.

A damage to the frontal cortex, for example, may imply problems with the individual's planning capabilities; i.e., in terms of a mental-state theory of action (e.g., Tirassa, 1997), with the complex dynamics of beliefs, long-term intentions, short-term intentions etc., involved in behavior planning. A different example comes from autism. This disorder is associated with the lack of a specific, dissociable ability to understand and reason upon the other individuals' mental states as different from those of one's own (Frith, Morton & Leslie, 1991; Leslie & Thaiss, 1992). Such ability requires specific types of mental states, other than ''simple'' belief (Leslie, 1994; see also next section). Autism then consists (at least) in the impossibility to form certain types of mental states, namely, those involved in the attribution of mental states to other individuals (Baron-Cohen, 1995).

The general approach we have adopted is to take a competence theory of human intentional communication, cast in terms of the mental states involved, and to understand how impairments in cognitive skills might possibly yield an impairment in this task. In other words: given a theory of competence, we are trying to constrain it on the basis of data from impaired performance of neuropsychological interest.

We have taken as a starting point the computational theory of communication developed by Airenti, Bara, and Colombetti (1993a, 1993b). We have derived from it a number of predictions of the performance of normal and brain-damaged individuals. Finally, we have tested some of these predictions through a series of experiments.

## 3. A COMPUTATIONAL THEORY OF SPEECH ACTS

Studies of the pragmatics of communication are far from a stable theoretical systemization. Stephen Levinson's *Pragmatics* (1983), a standard text of reference for the area, lists as many as 53 pages of different definitions of the term itself. Thus, neither our theoretically driven strategy, nor the detail of the theory we adopt, are unanimously agreed upon.

### *3.1. Cognitive Pragmatics*

Speech act theory (Bara, 1995; Leech, 1983; Levinson, 1983) is the most influential school in pragmatics, and appears particularly well-suited for neuropsychology. Further, it has now reached a satisfactory degree of formalization after Austin's (1962) and Searle's (1969) pioneering works, thanks to many works grounded in analytic philosophy and logic (e.g., Grice, 1989; Searle & Vanderveken, 1985; Sperber & Wilson, 1986; Strawson, 1964), computational linguistics (e.g., Appelt, 1985; Cohen, Morgan & Pollack, 1980), and artificial intelligence (e.g., Airenti, Bara & Colombetti, 1983; Allen & Perrault, 1980; Cohen & Levesque, 1991; Cohen & Perrault, 1979).

Within this framework, Airenti, Bara, and Colombetti (1993a, 1993b) have presented the bases of a computational theory of the cognitive processes underlying human communication. We give here a sketch of their theory, with no discussion of the formal apparatus or of the psychological evidence. We refer the interested reader to the articles cited in the bibliography; to discuss these aspects of our model here would make this paper longer than it already is, and would be likely to diminish its appeal to a readership which is mostly composed of neuroscientists.

Given the assumption that the same analysis holds for both verbal and nonverbal communication, we adopt the terms *actor* and *partner* instead of the more traditional *speaker* and *hearer*. Further, we follow the convention that actor *x* is always a female (hence an actress) and partner *y* is always a male.

A major assumption of Airenti, Bara, and Colombetti's (1984) work is that a cognitive analysis of dialogue should distinguish between its *conversation* and *behavior* aspects. When one communicates, her aim is to achieve a certain effect on a partner, namely, to change his mental states and possibly to induce him to perform some action. But there is something more. The choice of a communicative way to attain this goal provides her with the subsidiary goal to follow the rules of conversation. Behavior and conversation goals have completely independent origins; thus, an agent's behavior may be noncooperative, in spite of his will to maintain a correct conversation:

[1] Ann: Could you please lend me your car for the weekend?
Bob: Sorry, I promised the kids to go camping with them in the lakes region.
Ann: Well, I'll ask Charles. Have a nice trip.

Each level requires a form of cooperation. Behavior cooperation requires that the agents act on the ground of a plan at least partially shared, which we call *behavior game* of *x* and *y*. For our purposes, action plans can be represented as trees of intentions, whose leaves are specified either as terminal, precisely defined actions, or as higher-level intentions to be worked out according to the context (Pollack, 1990). Besides, a behavior game specifies the typical situation in which it can be played.

The actions prescribed by a behavior game need not be logically necessary. Some of them may constitute a conventional, habitual part of the interaction, as is the case of the actions of greeting typically prescribed by the games governing meetings between persons. Greetings play no necessary role with respect to the rest of the interaction; nonetheless, it is considered impolite to neglect them.

The actual actions performed by the agents realize the moves of the behavior game they are playing. The meaning of a communicative action (either verbal or nonverbal or, more often, a mix of the two) is fully understood only when it is clear what move of what behavior game it realizes. Thus, we will consider speech acts as moves of behavior games; conversely, each move of a behavior game has a meaning value, and can therefore be considered as a communicative act.

Conversation cooperation is governed by a set of specific rules that, by analogy with behavior games, we call the *conversation game.* However, the conversation game need not be a shared representation of a multiagent plan, like behavior games; rather, it can be modeled as a set of tasks that each agent has to fulfill in a given sequence. Each task defines a specific phase of the comprehension/generation process. For example, the phase of understanding the speaker's meaning will be characterized by the task of recognizing the behavior game proposed by the actor (see below). Further, the conversation game specifies the chaining of the phases in both standard and nonstandard cases.

Communicative competence can thus be seen as an inferential metalevel, controlling base-level inferences which are carried out on shared representations of stereotyped patterns of interactions. To clarify this point, let us consider again example [1]. In any standard context, Bob's response would be taken as a justified rejection of Ann's request. The whole exchange can be explained by saying that:

> By her request, Ann proposes to Bob to play the behavior game:
> [2] [LEND-SOMETHING]:
> - *x* gives object1 to *y;*
> - *y* returns object1 to *x*.
>
> By his response, Bob rejects Ann's proposal, on the justification that he himself needs the object requested.

Conversation cooperation requires that Ann and Bob share the knowledge of the behavior game in [2]. The example is deliberately oversimplified: the actual game shared by Ann and Bob must describe the applicability conditions of the request (e.g., you will not ask someone to lend you his car unless he is a close friend of yours) as well as its possible outcomes (e.g., it has to be understood that *y* may legitimately refuse to comply with *x*'s request). It may also dictate actions necessary on a social, rather than logical, basis (e.g., the use of politeness formulas such as *please, thank you,* and so on). It is the sharedness of these knowledge structures that allows Ann and Bob to maintain conversation cooperation in spite of Bob's refusal to cooperate on the behavior level.

The reason for introducing games is that literal meaning is but the starting point for the full comprehension of an utterance. *Why is she saying this to me?* and *What does she want from me?* are the real questions to be answered. Suppose a colleague of yours says:

[3] I'd appreciate a cup of coffee

while you are working together in her office. Her utterance may be interpreted as a request of a pause, with reference to a sort of [WORK-TOGETHER] behavior game. The same statement, uttered in your house after dinner, might be recognized as a typical move of a [DINNER-TOGETHER] behavior game: a guest announces a desire that the host is bound to accomplish. Finally, the same utterance would be puzzling if someone you've never seen before suddenly pops in your office to produce it: either you are able to find a behavior game related to it (i.e., to understand what the actor wants you to think or do) or it will remain unexplained. The point is that in the latter case the context allows you to identify no game, mutually known to you and to the actor, which the utterance may be connected to.

Behavior games thus have a fundamental role in communication: an utterance may be assigned a full meaning only in the context provided for by a behavior game. The recognition of the behavior game bid by the actor does not bind the partner to play his role in the game. On the contrary, the partner can decide to accept or reject the proposed game, or to try to negotiate a different one, or even to let the conversation game interrupt.

As we have said, a behavior game can only be played when the agents share knowledge of it. The concept of sharedness requires some explanation. Airenti, Bara, and Colombetti (1993a) argue that communication requires, besides the standard apparatus of private beliefs and intentions, two specific mental states, namely, *shared beliefs* and *communicative intentions.* Although different versions of these two concepts can be found in the literature, we will define them here, since they are introduced in an original form.

The concept of *shared belief,* or *mutual knowledge,* was introduced by Lewis (1969) and Schiffer (1972); many versions thereof have been subsequently developed. We adopt a subjective view of shared belief; i.e., we assume that for an actor to have a shared belief means to have a certain belief and the belief that such belief is shared with one or more specific partners, or with a group of people, or with all human beings. In our model, all the inferences drawn during the understanding of an utterance are drawn in the space of shared beliefs of $x$ and $y$. Thus, sharedness is crucial in communication; indeed, it is a necessary condition of communication that each agent maintains a shared belief space (Clark, 1992).

We take shared belief to be a primitive, i.e., a specific mental state not reducible to a conjunction of standard private beliefs. This implies rejecting the prevailing view that shared belief is the result of an infinite nesting of beliefs (what $x$ believes, what $y$ believes that $x$ believes, what $x$ believes that $y$ believes that $x$ believes, and so on): human easiness in dealing with shared information rules out any such formula as cognitively implausible (Clark & Marshall, 1981). Rather, the connection between belief and shared belief is captured by a circular axiom, as informally suggested by Harman (1977):

$$\mathrm{SH}_{xy}\, p \equiv \mathrm{BEL}_{x}\, (p \wedge \mathrm{SH}_{yx}\, p),$$

where $\mathrm{SH}_{xy}\, p$ means that agents *x* and *y* share the belief that *p*. The formula reads as follows: from agent *x*'s viewpoint, agents *x* and *y* share the belief that *p* iff: (i) it is the case that *p*, and (ii) *y* and *x* share the belief that *p*. A formal model of shared belief has been given by Colombetti (1993). It is worth emphasizing that shared belief is a mental state held by an individual. It may happen, in other words, that *x* believes *p* to be shared by *x* and *y*, whereas *y* does not believe *p* to be shared by *y* and *x*. This, as will be discussed later, may give rise to failures (e.g., misunderstandings) and other nonstandard interpretation paths.

The above axiom is circular in the sense that, by distributing $\mathrm{BEL}_x$ on the conjunction, it is logically possible to derive an infinite number of nested implications of the type:

$$\begin{aligned}
&\mathrm{SH}_{xy}\, p \supset \mathrm{BEL}_{x}\, p\\
&\mathrm{SH}_{xy}\, p \supset \mathrm{BEL}_{x}\, \mathrm{BEL}_{y}\, p\\
&\mathrm{SH}_{xy}\, p \supset \mathrm{BEL}_{x}\, \mathrm{BEL}_{y}\, \mathrm{BEL}_{x}\, p\\
&\mathrm{SH}_{xy}\, p \supset \cdots
\end{aligned}$$

i.e., if *x* and *y* share that *p*, then *x* believes *p*, *x* believes that *y* believes *p*, *x* believes that *y* believes that *x* believes that *p*, and so on. However, in no case can a resource-bounded cognitive system derive more than a limited number of these implications. Finite nestings of beliefs play an important role in nonstandard communicative situations, particularly in the case of deceits, where some of these formulae can be negated, without affecting the derivation of the others.

The second mental state we postulate as a primitive is *communicative intention.* Successful communication has been described by Grice (1957) as the recognition of a particular set of mental states, among which are the intention to achieve a certain effect on the partner, and the intention that such intention be recognized. Such conditions have been strengthened by Strawson (1964) and Schiffer (1972). Airenti, Bara, and Colombetti (1993a) show that any definition including a finite number of nested intentions is too weak, because it will leave unrecognized part of the interactive situation, and therefore will not capture appropriately the notion of communication. On the other hand, for the reasons discussed above, no living system can be supposed to draw infinite chains of inferences.

This problem can be solved by defining the intention to communicate as comprising two components: the intention to share some fact, and the intention to share the whole intention to communicate. More precisely, *x* has the communicative intention that $p$ with respect to $y$, i.e., $x$ intends to communicate $p$ to $y$ (in symbols $\mathrm{CINT}_{xy}\ p$), when $x$ intends that the two following facts be shared by $y$ and $x$: that $p$, and that $x$ intends to communicate $p$ to $y$:

$$\mathrm{CINT}_{xy}\ p \equiv \mathrm{INT}_{x}\ \mathrm{SH}_{yx}\ (p \wedge \mathrm{CINT}_{xy}\ p)$$

From this formula an infinite number of logical implications may be derived of the type:

$$\mathrm{CINT}_{xy}\ p \supset \mathrm{INT}_{x}\ \mathrm{SH}_{yx}\, p$$
$$\mathrm{CINT}_{xy}\ p \supset \mathrm{INT}_{x}\ \mathrm{SH}_{yx}\ \mathrm{INT}_{x}\ \mathrm{SH}_{yx}\, p$$
$$\mathrm{CINT}_{xy}\ p \supset \cdots$$

(i.e., $x$ intends to share that $p$, $x$ intends to share that $x$ intends to share that $p$, and so on).

### *3.2. The Model of Dialogue*

We report here the basic model of dialogue developed by Airenti, Bara, and Colombetti (1993a). The general scheme of the comprehension/generation cycle goes as follows: actor $x$'s utterance is received by partner $y$, who represents its meaning; $y$'s mental states about the domain of discourse may be affected by the comprehension; then $y$ plans his next dialogue move, which is eventually generated. We distinguish five logically chained phases, described below, in $y$'s mental processes. The chaining of these five processes is controlled by the conversation game (see Fig. 1).

*Understanding literal meaning.* The starting point of this phase is the result of $y$'s analysis of $x$'s utterance as a literal illocutionary act, with addressee $y$, propositional content $p$, and literal illocutionary force $f$. The task set up by the conversation game for this phase is to recognize the actor's expression act. Once the expression act is recognized, the conversation game activates the process of understanding the speaker's meaning.

*Understanding speaker's meaning.* Here $y$ comprehends $x$'s communicative intentions. All the relevant inferences are drawn in the space of shared beliefs. The starting point is the literal meaning recognized, and the result is the recognition of the actor's communicative intentions. The task of this phase is to reach a state in which it is shared by $x$ and $y$ that $x$ has communicated her intention to play a certain move of a certain behavior game with $y$.

*Communicative effect.* We define as the communicative effect on $y$ the set of all of $y$'s mental states, acquired or modified in accordance with $x$'s communicative intentions. A further condition is that such mental states are actually caused by the corresponding communicative intention; for instance, the fact that someone is trying to make us believe that it is raining must be the reason (or part of the reason) for believing that it is raining. Differently

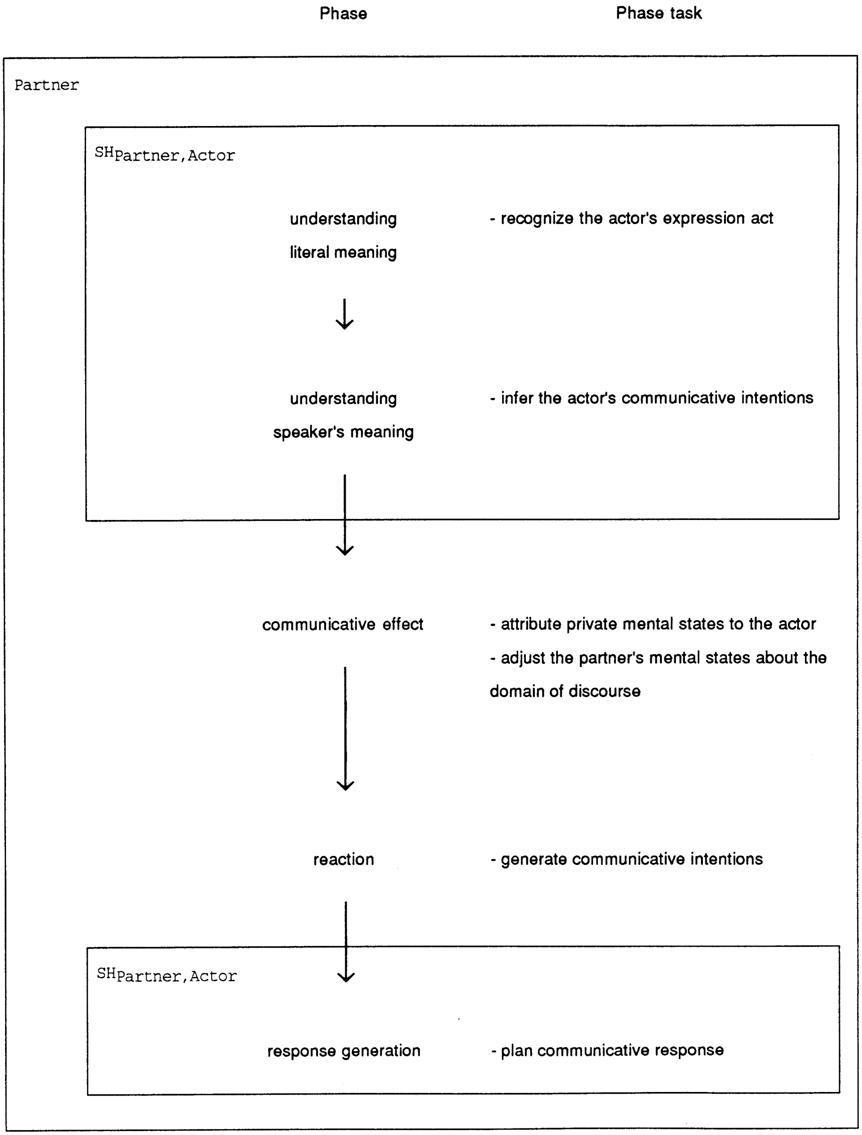


FIG. 1. The five phases of comprehension and production of a communicative act.

from the previous ones, this phase does not consist in a recognition; further, it involves private beliefs and the motivations (Airenti, Bara, & Colombetti, 1989), e.g., to understand that somebody is requesting us to lend her the car is a matter of shared knowledge on the use of language, but to decide whether to comply is definitely different, and involves private motivations. Thus, the conversation game leads *y* to question whether he should adhere to *x*'s communicative intentions; in particular, *y* has to decide whether to play the game proposed by *x* (*adjustment*); this may also require that private mental states of *x*'s be hypothesized by *y* (*attribution*: e.g., Bob might wonder what use Ann intends to make of his car).

*Reaction.* Here the intentions for the generation of a response are produced, which will be the input to the subsequent, final phase. From a conversation point of view, the response has to include information to *x* about the effects achieved by her speech act. Thus, the communicative intentions produced in the reaction phase result from the integration of the communicative effect (i.e., the output of the adjustment process) with the behavior game that *y* desires to play with *x*. The conversation metalevel dictates that the reaction be pertinent to the analysis performed in understanding the speaker's meaning. Thus, *y* has to take a stance about *x*'s communicative intentions, whether or not they have succeeded. The task of this phase is to plan the achievement of a communicative effect on *x*.

*Response generation.* Here an overt response is generated, which will be used by *x* as a feedback on the results of her communicative attempt.

## 4. EXPERIMENTAL SETTING

### *4.1. Rationale for the Experiments*

We are interested in neuropsychological correlates of pragmatic theories of communication. Our aim is not to build a theory of normal communication, starting from a collection of data from impaired performance. We already have such theory, developed on independent grounds, and are trying to falsify it, or at least to test it against competing ones, by an exploration of possible impairments in the communicative skills of brain-damaged individuals. A welcome side effect of this research might be to help build a framework for the comprehension of some consequences of brain damage.

Among the many potential neuropsychological implications of pragmatics, we have decided to start from the investigation of some relatively simple inferences involved in communication, with a certain emphasis on comprehension rather than generation. Our subjects were individuals who suffered from the consequences of a closed head injury (CHI). This pathology is known to imply a constellation of symptoms. Primary linguistic processes are usually preserved (although with possible exceptions, such as specific anomia: Semenza & Zettin, 1988, 1989). Communicative skills, however, are often impaired (e.g., Joanette, Goulet, Ska, & Nespoulous, 1986; McDonald, 1992; McDonald & van Sommers, 1993; Sherratt & Penn, 1990). These individuals' discourse is often confused and disorganized, with frequent intrusion of inappropriate associations, stereotypical perseverations, insinuations, or indirect contextual references. However, such subtle symptoms are not easily defined in the absence of a sound theory of normal communication (Joanette & Brownell, eds., 1990).

CHI symptoms are typically related to diffused tissular sufferance, or axonal damage. Thus, they can be hypothesized to result from the impairment of many different subsystems. This fits our current interests quite well: we do not aim here at associating specific deficits with specific foci, but at exploring how communication, a highly sophisticated process guaranteed to involve a number of different skills and subsystems, may be affected by brain injury.

As regards the methods suitable for investigations in this area, we wanted our protocols to leave the subjects' performance as unconstrained as possible. Communication is a peculiar activity, which seldom, if ever, prescribes anything like a ''correct answer'': most of the time, the ultimate meaning of an utterance is a matter of negotiation between the actor and the partner (Recanati, 1995; Tirassa, 1997). Also, open judgment tasks are strongly suggested in the study of CHI for neuropsychological reasons: as Shallice and Burgess (1991) point out, standard neuropsychological tests are unlikely to be useful in this syndrome, because they tend to provide for exactly the kind of structure the subject lacks.

We chose relatively simple and well-defined tasks because we wanted the results to be as clear as possible. Because the coordinates of the ground we are moving on are largely unknown, it is necessary to start with an assessment of some very basic capabilities before considering more complex aspects of the theory. Our first concern was therefore to understand to what extent the fundamentals of communicative ability were impaired or disrupted.

Thus, the tasks had to be both simple and unconstrained. Most of our protocols consisted in some variation on a standard design, which we will briefly describe in the next subsection. The variations, as well as the few protocols which were truly different, will be described in the relevant sections below. Our experiments explored different aspects of the theory, namely:

- the standard comprehension path and the problem of direct vs. indirect speech acts;
- two nonstandard paths, namely, irony and deceits;
- for each of the categories above, failures and their recovery;
- the ability to make inferences about another individual's mental states;

these so-called *theory of mind* inferences, although not communicative per se, play nonetheless a major role in the generation and comprehension of communications.

Most of our tasks were designed to explore inferences in comprehension, because most theories of pragmatics have been developed basically from an understander's point of view, and ours is no exception; most relevant work in neuropsychology has focused on understanding as well. A subset of protocols, however, dealt with some aspects of generation, in particular deceits and failure recovery.

## *4.2. Method*

The standard protocol is as follows. The subject is shown short (10–25 sec) scenes on a videotape recorder. Each scene involves two characters in an everyday context, such as a typical living room; a brief and simple dialogue occurs between them about some simple topic or event. A sample dialogue is the following (taken from the protocols about indirect speech acts, see Section 5 below):

[4] Child: Can I ride my bike, Mom?
Mother: It has a broken wheel.[1]

Needless to say, the actual sessions took place in Italian; here, we are translating everything into English. For the sake of convenience, we will further use the feminine form for the experimenters and the masculine form for the experimental subjects. Actually, sexes were equally distributed in each category.

Immediately after each scene, the subject is asked to report his understanding of it, or what he thinks might happen immediately after the episode portrayed. In [4], the experimenter asks: ''What did Mom mean by that?'' The goal is to understand whether the subject has correctly analyzed the two utterances; in particular, whether he has taken the mother's answer as implying that the bicycle cannot be used. In order to pass the task, thus, the subject can respond anything like ''That the child cannot go out,'' ''That she wants him to stay at home,'' or ''That he'll have to repair it,'' and so on.

In order to avoid the superimposition of possible short-term memory impairments, the subject can review the scene whenever he wants, with no limitation on the number of repetitions. The subject is left free to interpret the episode in whatever way he likes, without any intervention on the part of the experimenter and without any time constraint, until he says he has no more remarks to make. Then, he is shown the next scene.

The experimenter takes note of all the answers given by the subject, in their respective order. Further, the whole session is taperecorded. The tape is later reheard by two independent judges, in order to be certain of the subject's interpretation and to avoid possible biases on the part of the experimenters. The judges were two speech therapists; they were well acquainted

[1] In most protocols, the characters are mothers with children, or couples of children, and sometimes we also used marionettes. To use such ''childish'' protocols could imply some risk of emotional discomfort, since CHI subjects are typically very young (the mean age of our subjects was less than 23) and sometimes mishandled by physicians, therapists, and caretakers. This was not the case with our subjects: all of them were glad to participate in the experiment and enjoyed the novelty it represented. This might have been due to their previous confidence with the carefully trained personnel of the hospital. We thank Claudio Luzzatti (personal communication) for pointing at this potential problem.

with CHI in general and with the specific subjects in particular, but blind with respect to both the theoretical approach and the goal of the research.

In the case of an unclear answer (i.e., one from which the subject's comprehension cannot be classified as correct or incorrect) the experimenter poses a direct, closed question, asking the subject to choose among three different proposed interpretations of the scene: one is the ''correct'' interpretation, i.e., the one most people would reasonably give; one is just literal, but plausible anyway, and one makes no sense in the current context. The order of presentation of these options is randomized, e.g., in [4]:

''inferential'': Mom means that the child cannot go out;
''literal'': Mom means that the wheel is broken;
''out of context'': Mom means that Dad is coming home.

However, we needed to resort to these subprotocols only in four cases (two complex indirect speech acts and two ironies, see below), which makes this point substantially irrelevant.

There were 21 tasks. Their order of presentation was randomized for each subject. The average time for the whole protocol was approximately 30 min.

### 4.3. Subjects

The protocols were administered individually to 13 patients, 6 females and 7 males, with severe CHI of traumatic origin, who had been admitted to the hospital for post-acute neuropsychological, motor, and speech rehabilitation. None of them had a previous history of neurological or psychiatric disturbance. The demographic and clinical characteristics of the subjects are reported in Table 1. The control group consisted of the same number of individuals, of comparable sex, age and education, and having no history of neurological or psychiatric disease, or of alcohol or drug abuse.

The subjects were tested on a number of language and cognitive skills; the results are presented in Table 2 and briefly discussed in Section 9. We will only remark here that, although some subjects presented an impairment in metalinguistic skills, and sometimes anomia as well, none of them had any symptom of aphasia.

Aphasia impairs communication at a relatively low level, through a damage to the linguistic execution of decisions taken at the communication (i.e., pragmatic) level. High-level communicative competence, on the contrary, is often reasonably spared (Feyereisen & Seron, 1982; Foldi, Cicone & Gardner, 1983; Guilford & O'Connor, 1982; Holland, 1982; Joanette & Brownell, eds., 1990; Penn, 1987; Penn & Cleary, 1988). These results have led to therapeutic approaches aimed at the enhancement of alternative communicative channels (e.g., Davis & Wilcox, 1981, 1985; Howard & Hatfield, 1987), in addition to the more classical attempts at rehabilitation of strictly linguistic functions.

These results are in perfect agreement with our theoretical approach. We consider communication as a high-level competence, whose behavioral expression is realized through an understandable combination of any possible means of expression. Language, gestures, drawings, and so on, all may be used for communicative purposes. How a person communicates will be affected by an impairment of one of these *media* and particularly of such a preeminent one as language, but only insofar as the damage impedes the externalization of the individual's mental states. Communication and language can and should be kept independent of each other, as is also maintained by Chomskyan linguistics (Chomsky, 1980) and by many studies in animal cognition (e.g., Burling, 1993; Premack, 1986).

## 5. THE STANDARD PATH: DIRECT VS. INDIRECT SPEECH ACTS

Let us now turn to a more detailed analysis of the experiments. To repeat, we explored the following issues:

TABLE 1
Demographic, Neurological, and Neuroimaging Data

| Subject | Age | Sex | Years of education | Preinjury vocational status | Months post onset | Days of coma | CT scan (frontal, temporal, parietal, occipital) |
|---|---|---|---|---|---|---|---|
| 1 | 19 | M | 10 | Mason | 26 | 150 | FT left |
| 2 | 23 | M | 13 | Student | 65 | 30 | FT left |
| 3 | 27 | F | 13 | Secretary | 17 | 15 | F bilateral, O right |
| 4 | 27 | M | 10 | Student | 145 | 30 | T right, F bilateral |
| 5 | 19 | F | 12 | Student | 27 | 100 | Diffused politrauma |
| 6 | 35 | M | 11 | Electronic technician | 63 | 22 | FT left |
| 7 | 18 | F | 8 | Serigrapher | 7 | 10 | T right |
| 8 | 24 | M | 18 | Student | 2 | 0 | O bilateral |
| 9 | 16 | M | 10 | Student | 4 | 8 | Diffused axonal trauma |
| 10 | 19 | F | 13 | Student | 5 | 40 | T right |
| 11 | 17 | F | 10 | Student | 16 | 20 | O bilateral |
| 12 | 23 | M | 10 | Workingman | 19 | 60 | FP right |
| 13 | 31 | F | 12 | Secretary | 38 | 20 | T right, FT left |

TABLE 2
Neuropsychological Data

a: Part 1

| Subject | Wechsler adult intelligence scale (WAIS) subtest: Picture completion | Block design | Picture arrangement | Picture assembling | IQ performance |
|---|---|---|---|---|---|
| 1[a] | — | — | — | — | — |
| 2 | 12 | 12 | 11 | 13 | 60 |
| 3 | 8 | 9 | 3 | 5 | 112 |
| 4 | 11 | 11 | 8 | 5 | 80 |
| 5 | 12 | 9 | 6 | 7 | 89 |
| 6 | 11 | 10 | 9 | 14 | 105 |
| 7 | 8 | 7 | 6 | 9 | 79 |
| 8 | 11 | 14 | 12 | 15 | 114 |
| 9 | 8 | 10 | 12 | 11 | 102 |
| 10 | 11 | 13 | 16 | 12 | 102 |
| 11 | 7 | 6 | 4 | 7 | 74 |
| 12 | 8 | 7 | 0 | 3 | 61 |
| 13 | 11 | 9 | 6 | 7 | 88 |

b: Part 2

| Subject | Frontal lobe tasks: Verbal fluency (colors, fruits, animals, cities) | Sentence contraction | Free association | Proverbs interpretation |
|---|---|---|---|---|
| 1 | 9.5 | 13 | 5.5 | 12 |
| 2 | 16.5 | 18 | 8 | 15 |
| 3 | 2.8 | 12 | 3.8 | 12 |
| 4 | 13.5 | 18 | 8 | 14 |
| 5 | 12 | 18 | 8 | 15 |
| 6 | 12.5 | 9 | 8.5 | 14 |
| 7 | 10.25 | 9 | 5.25 | 14 |
| 8 | 17 | 15 | 15.25 | 15 |
| 9 | 15.75 | 15 | 11.5 | 15 |
| 10 | 12.25 | 17 | 16.25 | 15 |
| 11 | 6.7 | 9 | 4 | 11 |
| 12 | 8.25 | 16 | 3.25 | 11 |
| 13 | 15.75 | 11 | 11.8 | 12 |

c: Part 3

| Subject | Comprehension: Token test | Verbal gramm. comprehension (%) | Memory: Verbal span | Spatial span | Story recall |
|---|---|---|---|---|---|
| 1[a] | 35 | 97 | 4 | — | 12.56 |
| 2 | 36 | 100 | 5 | 7 | 26.4 |
| 3 | 30 | 70 | 3 | 4 | 22.3 |
| 4 | 36 | 100 | 4 | 5 | 25.3 |
| 5 | 35 | 97 | 3 | 5 | 8.74 |
| 6 | 36 | 100 | 4 | 5 | 26.24 |

TABLE 2—*Continued*

| c: Part 3 | Comprehension | | Memory | | |
|---|---|---|---|---|---|
| Subject | Token test | Verbal gramm. comprehension (%) | Verbal span | Spatial span | Story recall |
| 7 | 35 | 100 | 3 | 5 | 16.89 |
| 8 | 36 | 100 | 4 | 5 | 24.3 |
| 9 | 32 | 100 | 4 | 6 | 23.66 |
| 10 | 36 | 100 | 5 | 7 | 16 |
| 11 | 30.5 | 98 | 4 | 4 | 12.07 |
| 12 | 36 | 100 | 5 | 5 | 12.03 |
| 13 | 35 | 98 | 6 | 5 | 18.7 |

d: Part 4

| Subject | Raven P.48 (%) | Attentional visual search |
|---|---|---|
| 1 | 85.4 | 15 |
| 2 | 88 | 58 |
| 3 | 41.7 | 13 |
| 4 | 87.5 | 55 |
| 5 | 75 | 53 |
| 6 | 95.83 | 40 |
| 7 | 81.25 | 46 |
| 8 | 83.32 | 56 |
| 9 | 93.75 | 32 |
| 10 | 75 | 51 |
| 11 | 58.3 | 34 |
| 12 | 47 | 13 |
| 13 | 83.3 | 57 |

[a] Subject 1 could not perform some tasks because of a Wolkmann syndrome with right arm paralysis.

- direct vs. indirect speech acts, and relevant failures (this section);
- two different nonstandard paths, namely, irony and deceits, and relevant failures (Sections 6 and 7);
- *theory of mind* inferences (Section 8).

We will present each topic separately, illustrating its relevance, its role within our theoretical approach, the experiments with the relevant predictions, and the results. A general discussion will follow in Section 9.

This section deals with the standard comprehension path: given an actor's utterance, the partner's task is to understand its relation to the behavior game in play, or to a newly proposed one. Irony and deceits are two different nonstandard paths; in both cases, there is something more to the utterance than would seem at a superficial glance. We also explored the subjects' performance on the failures of each of these three comprehension paths, in order to assess possible differences with the performance on the relevant cases of success. *Theory of mind* inferences are somehow different: the ability to reason upon other individuals' mental states is ancillary to communication rather than communicative per se. We decided to explore it because of its heavy involvement in dialogue; further, as far as we know, it had never been tested in these neuropsychological pathologies.

Our adopted theory allowed us to make qualitative predictions. We expected, e.g., that some types of utterances would turn out to be easier to understand than other ones, or harder, and so on. A competence theory clearly does not allow to formulate predictions of a quantitative nature.

To understand communication is to infer an actor's meaning, starting from the observation and analysis of whatever behavior of hers can be considered communicative (what we call *literal meaning*). Most theories of pragmatics, and all the few relevant neuropsychological investigations, adopt a Searlean view of speech acts understanding. Searle (1975) draws a distinction between direct and indirect speech acts. A *direct speech act* is performed when the literal meaning of an utterance corresponds straightforwardly to its intended one (in our terms, to the speaker's meaning), e.g., given the intention to induce the partner to close the window, the corresponding direct speech act would be the uttering of an imperative sentence:

[5] Close the window.

A speech act is *indirect* when the literal meaning and the speaker's meaning of an utterance do not correspond. Thus, any of an indefinite number of sentences may be used as an indirect request to close the window:

[6] a. Can you close the window?
b. Would you please close the window?
c. I wonder whether you would be so kind as to close the window.
d. The window is open.
e. What about that window?
f. How cold in here!
g. Have I ever told you that my grandfather died from pneumonia?
h. Come on, Johnny, be a polite kid.

According to Searle, the interpretation of [5] is straightforward in any standard context. The correct comprehension of any utterance in [6], on the contrary, depends on the previous recognition of the contextual inadequacy of a literal interpretation, e.g., in [6a], the understander has to realize that the actress's doubt about his physical and mental capabilities is rather unlikely in the current context, and wonder what else the actress might have meant. Searle (1975) also outlines the inferential path supposed to allow for the recovery of the indirect meaning; Perrault and Allen (1980) give a computational reformulation thereof. The machinery needed to recognize the contextual inadequacy of the literal interpretation can be provided for by, e.g., the maxims of conversation that Grice (1975) derived from his cooperative principle.

A similar explanation supposedly holds for other kinds of nonliteral interpretation as well (e.g., conversational implicatures, irony, etc.). The basic algorithm remains the same: recover the literal illocution and, if this is contextually inappropriate, look for an acceptable alternative.

This model, however, poses a number of problems. First, it depends completely on a straightforward identification of communication with language: how to distinguish direct and indirect ways to gesticulate? For example, is a ''meaningful'' gaze, accompanied by a slight movement of the head and a raising of the eyebrow, more direct or indirect than a window-oriented waving of the hand with a grumble? Even within the boundaries of a linguistic analysis, it is not at all clear how to arrange all literal illocutions with the three sentential moods allowed by many languages, Italian and English included (affirmative, interrogative, and imperative), or how to handle elliptical sentences (''The *window*, my God!''). The basic algorithm has also been criticized for its cognitive implausibility (Clark, 1979). The underlying assumption that the illocutionary level has an ontological status of its own may be questioned as well (e.g., Cohen & Levesque, 1985, 1990; Sperber & Wilson, 1986).

Our position, as should be apparent after Section 3, is that there exists no illocutionary level as such; and, consequently, that there can exist nothing like an indirect speech act—or a direct one. Every instance of dialogue understanding involves inferential activity: in no way can an utterance be understood with no inferences taking place at all (see Gazdar, 1981, and Levinson, 1983, for further arguments in favor of inferential theories of pragmatics). According to our own theory, an agent's observed behavior is understood as communicative when its role in a behavior game has been inferred, and that's all. Whether such behavior is a linguistic utterance (and, if so, whether it is affirmative, interrogative, imperative, elliptical, etc.), or a gesture, or a drawing, is irrelevant; and whether it is more or less straightforward is not a matter of directness or indirectness of the illocution, but of the difficulty of the inferential chain needed to recognize the game proposed.

To reject the difference between direct and indirect speech acts, indeed, does not imply that all utterances are equally easy to understand. The choice among the various utterances in [5]–[6a–h] may make a big difference: some of them are quite easy to understand in everyday contexts; others less; and some may be really hard. Analogously, we all have a rich experience of conversations where every single word has an unclear meaning; political as well as academic debates may be a very slippery ground, but not for reasons of ''indirectness'' as defined by Searle.

There are a few neuropsychological works on indirect speech acts (in particular, on pseudoidiomatic indirect requests, such as those in [6a–b]), e.g., Foldi (1987); Heeschen and Reischies (1979); Hirst, LeDoux, and Stein (1984); McDonald (1992); McDonald and van Sommers (1993); Weylman, Brownell, Roman, and Gardner (1989); and Wilcox, Albyn, and Leonard (1978). In her review of this literature, however, Stemmer (1994) remarks that the phenomena under investigation are never defined clearly or formally and that, for this reason, no finding, albeit suggestive, has a clear meaning. She also points at some difficulties with Searle's model, such as the notion of ''levels of indirectness'' (i.e., the fact, discussed above, that not all indirect speech acts are equally difficult to understand).

All these empirical findings seem to prove that indirect speech acts are more difficult to understand than direct ones. These results are in agreement with Searle's approach. Our data, however, are quite different: we found no evidence for a difference between direct and indirect utterances. Our protocols comprised two different series of data related to indirect speech acts.

### 5.1. Observations in the Wild

First, before starting the formal experimental session, there was a short, informal phase during which the experimenter tried to put the subject at ease, ''prepared the material'' for the experiments, explained what was going to happen, etc. During this informal phase, she nonchalantly addressed the subject with both direct and indirect utterances, some of which were indeed part of the experiment itself:

[8] *Direct speech acts*
a. What are your name and your age?
b. Are the instructions clear?
c. Tell me whether anything is unclear.

[8] *Indirect speech acts*
a. Will you see some scene with me?
b. Do you know the time?
c. Can I have that video?

The other experimenter, sitting by the first one and ''filling in modules,'' was actually evaluating the subject's reactions to the different classes of utterances, marking them as adequate or inadequate.

It has to be noticed that, although the ''indirect'' forms in [8] are seldom, if ever, used literally, they are definitely not idioms (as would be, for example, the expression ''to kick the buck'' used instead of ''to die''). Since their literal interpretation cannot be ruled out in principle, forms like [8] are indeed considered prototypical indirect speech acts, to the point that most of the examples discussed in philosophy of language (Searle, 1975), artificial intelligence (Perrault & Allen, 1980), and neuropsychology (Foldi, 1987; Hirst, LeDoux, & Stein, 1984; Weylman, Brownell, Roman, & Gardner, 1989; Wilcox, Albyn, & Leonard, 1978) belong to this class. It also has to be noticed that forms like those in [8] are, as far as it is known, universally diffused in human languages (Brown & Levinson, 1987; see also Levinson, 1983, for further discussion and a rejection of the idiomatic interpretation of indirect speech acts).

Thus, theories accepting the direct/indirect distinction would reasonably predict that the utterances in [8] impose a greater inferential load on the understander than those in [7]. Our prediction, on the contrary, was that no difficulty should be encountered with either class, because of the obviousness of the behavior game proposed.

We want to emphasize the importance of these observations *in the wild.* Communication is no formal task to human beings; rather, it is a means to engage in and maintain social interactions with other individuals who are active planners and understanders in their turn. Thus, an individual's task is definitely not to solve a technical problem about language, but to involve other individuals in the cooperative activity of conversation (e.g., aphasics' communicative performance is known to be reasonably satisfactory in everyday contexts, in spite of their linguistic impairment: de Bleser & Weisman, 1986). Protocols [7] and [8] aimed at exploring spontaneous (understanding of) communication in informal contexts. A similar procedure might have been to ask the subjects' relatives to provide for the same kind of evaluation; however, they could not be expected to adequately handle formal definitions of direct vs. indirect speech acts.

### *5.2. Standard Experiments*

We also made experiments on indirect speech acts during the formal session, with scenes like those described in Section 4. The contexts considered here were relatively more difficult than those in [7] and [8]. Difficulty, as discussed above, depends on the inferential path needed to get to the speaker's meaning from the literal meaning. As is common in linguistics, we rely on a shared background among the experimenter, the subject, and the reader, to guarantee our intuitions about what should be easier or harder for an individual in everyday contexts.

The protocols were as follows:

[9] Child: Mom, can I ride my bike?
Mother: It has a broken tire.
Comment: This is the example we discussed as [4] above. The task is passed if the subject realizes that Mom will not allow the child to use the bike, or that she is reminding him that he will have to repair it.

[10] Child: Mom, can I have some chocolate?
Mother: You have already eaten too much.
Comment: The task is passed if the subject understands that the child will be given no more chocolate.

[11] (The characters are looking outside through a window).
Child: Mom, let's go out and play.
Mother: It's raining.
Comment: The task is passed if the subject understands that the characters will not go out, or that they will have to take an umbrella with them, etc.

In each case, there are *two* indirect speech acts to be interpreted, since both the child's requests and the mother's ever-repressing answers are indirect speech acts according to Searle's definition. This should put a considerable inferential load on the interpretation, and our subjects were expected to encounter more problems with [9]–[11] than with [8].

### *5.3. Failures of Indirect Speech Acts*

The third series of protocols on indirect speech acts aimed at exploring the subjects' ability to handle failures. Communication fails whenever it does not achieve the goals it was planned for. Different classes of failures may be distinguished, according to what phase fails (Airenti, Bara & Colombetti, 1993b). The conversation goal to let the partner understand fails when he misunderstands the speaker's meaning, or when he understands no meaning at all. The behavior goals to play one's own role in a game, and to induce the partner to do the same in his turn, fail respectively when the partner refuses to acknowledge the agent's move as a valid one, or when he refuses to do his part.

We used two scenes here regarding, respectively, the failure of a direct and of an indirect speech act:

[12] *Direct speech act*
(Child B is taking candies out of a box and eating them).
Child A: Give me a candy.
Child B: No.

[13] *Indirect speech act*
(The window is closed).
Child A: Can you open the window?
Child B: No.

The subject is asked what Child A can do to recover from the failure. Strategies for recovery depend on the perceived cause of the failure. To just realize that something has gone wrong is not sufficient: it is also necessary to understand *what* has gone wrong, e.g., in the case of a request, it may have happened that the partner has not heard it, or has not understood it, or has misunderstood it, or is unwilling to comply with it, etc.

Since a failure may be due to any of an indefinite number of possible causes (which can be seen as a reformulation of the infamous *qualification* *problem* and *frame problem:* McCarthy, 1980; McCarthy & Hayes, 1968), we let the subjects free to make hypotheses about the situations proposed and the possible ways out. Thus, an appropriate answer could be something like ''She can promise something to the other child if he gives her a candy'' and so on. The answer ''She can ask again'' was considered ambiguous, because it does not discriminate between a simple repetition of the failed utterance (which would be appropriate only if the partner is supposed not to have heard the first attempt), and the planning of a different one (such as ''Come on, you have so many candies, and I just want to taste one''). In this case, the experimenter asked the subject exactly what sentence he thought Child A would utter.

Consistently with our hypothesis on direct and indirect speech acts, we expected no differences between [12] and [13]. In fact, both cases are simple speech acts. We predicted the relevant inferences to be saved at least in part, because a dramatic loss in the ability to handle failures (both of one's own and of the partner's) would severely hamper social interactions.

However, we expected the subjects to encounter more difficulties with these scenes than with the successful ones, because of the greater attentional and inferential load needed for an appropriate handling of failures. When a failure occurs, the actor has to realize that what the partner thinks is not what the actor wanted him to think. Further, the adoption of a strategy for recovery is a form of planning, and CHI patients often suffer from an impairment in planning skills, specially when the frontal lobes are involved.

A couple of answers sounded like ''Child A can beat Child B and take the candies.'' We decided to consider them inappropriate, because of their inappropriateness in the friendly context presented. The repair of a communication failure cannot involve *whatever* action may achieve the desired goal, but mandates for a more subtle form of planning.

### *5.4. Results*

All our predictions were confirmed, as shown in Tables 3 and 4. The control group solved the tasks without any problem. None of our subjects made a single error in understanding simple speech acts, either direct or indirect. As predicted by our theory, and in sharp contrast with the findings of previous research, all the CHI subjects comprehended all sentences *in the wild,* without regard to their directness or indirectness. As regards complex indirect speech acts, there were 37 correct answers out of 39 ($13 \times 3$) items. Two subjects made one mistake each (5% of the total), which is not statistically significant; further, both were able to pick the ''inferential'' answer out of the three possibilities presented in the closed judgment task.

These findings are fully consistent with our theoretical approach. They, on the contrary, contradict all previous neuropsychological research and undermine the traditional approach to speech acts in terms of illocutions. If

TABLE 3
Results for Class of Items

| | Test: 13 brain-injured subjects | | | | | | |
|---|---|---|---|---|---|---|---|
| | Direct speech acts ($n = 3$) | Simple indirect s.a. ($n = 3$) | Complex indirect s.a. ($n = 3$) | Irony ($n = 3$) | Deceits ($n = 3$) | Failures (as a whole) ($n = 4$) | Theory of mind ($n = 2$) |
| Correct | 39 (100%) | 39 (100%) | 37 (95%) | 36 (92%) | 31 (79%) | 29 (56%) | 25 (96%) |
| Wrong | 0 | 0 | 2 (5%) | 3 (8%) | 8 (21%) | 23 (44%) | 1 (4%) |
| Comparison with controls (Student's $t$) | | | | $t = 1.897$ $p < .07$ | $t = 2.889$ $p < .008$** | $t = 3.888$ $p < .0007$** | |

TABLE 4
Results for Failures Only

| | Test: 13 brain-injured subjects | | | |
|---|---|---|---|---|
| | Direct s.a. ($n = 1$) | Indirect s.a. ($n = 1$) | Irony ($n = 1$) | Deceits ($n = 1$) |
| Correct | 8 (62%) | 8 (62%) | 8 (62%) | 5 (38%) |
| Wrong | 5 (38%) | 5 (38%) | 5 (38%) | 8 (62%) |
| Comparison with controls (Student's $t$) | $t = 2.739$ $p = .01$** | $t = 2.739$ $p = .01$** | $t = 2.739$ $p = .01$** | $t = 4.302$ $p < .0002$** |

cognitively impaired subjects present no difference in the comprehension of direct and indirect speech acts, then there is no reason to think that the latter impose a greater inferential load than the former, as would be predicted in a Searlean perspective.

The difference between our results and those found in the literature needs be explained. We substantially agree with the points made by Stemmer (1994) in her review: previous research was inadequate from both a theoretical and an experimental viewpoint. The topic of indirect speech acts was not integrated in an overall approach to the pragmatics of communication; the kind of indirect speech acts considered had not been defined satisfactorily, and the strict adherence to Searle's analysis prevented any discussion about the ''level of indirectness'' of the various utterances (i.e., about their difficulty). This theoretical confusion mirrors in the frequent ambiguity of the protocols used.

Further, previous research invariably used closed judgment tasks, which have two major drawbacks. The first is that they are too rigid to allow a satisfactory exploration of the subject's interpretation of a situation (Stemmer, 1994). The second has been discussed above: closed judgment tasks are generally not recommended in CHI, because a subject's difficulties may be masked if he can exploit the structure of the task instead of imposing his own one on it (Shallice & Burgess, 1991). It must be noticed that, as a consequence, using open judgment tasks should yield a *worse*, rather than a better, performance in CHI subjects, i.e., the odds were against us. Thus, the contrast between our data and the previous ones is made even sharper.

Finally, the neuropsychological characteristics of the subjects studied in previous research were often ill defined.

We tried to avoid all these problems. Throughout this work, we make use of a well-defined theoretical approach. We aim at exploring a two-way relationship between pragmatics and neuropsychology. On the one hand, we are looking for a coherent theoretical framework in which to integrate clinical findings about human communication; on the other hand, we hope to find clinical evidence in support of the particular framework we have adopted.

Thus, we argue that Stemmer's concerns are undebatable, and that unambiguous clinical results can only emerge through the use of strong theoretical foundations. As regards the particular problem discussed in this section, indirect speech acts, we also argue that Searle's model cannot provide for such foundations.

Our results on the direct/indirect controversy may give rise to an objection, which we need to rule out. The objection is that our tasks might have been too easy for our subjects not to pass them, i.e., that we might have incurred in a ceiling effect. This has been hypothesized both by Marco Colombetti (personal communication) and by one of the referees.

We do not view this point as pertinent. The theoretical difference between direct and indirect speech acts is qualitative, not quantitative. In the direct case, no inference is supposed to take place beyond the recovery of the literal meaning: there simply is nothing more to the speaker's meaning than the literal meaning, so that understanding the latter corresponds straightforwardly to understanding the former. In the indirect case, on the contrary, the path to be followed is rather complex: once the literal meaning has been recovered, its contextual inappropriateness must be recognized, and a different, nonliteral meaning must be looked for (Searle, 1975). Within the range of indirect speech acts, quantitative differences in difficulty are supposed to obviously exist in a continuum, but this has seldom been perceived as particularly interesting (Stemmer, 1994).

In this framework, it seems quite reasonable to look for neuropsychological evidence on the border between the direct and the indirect cases, rather than in some random point along the quantitative continuum of indirect speech acts. Indeed, this is exactly what has been commonly done: the evidence found in previous research mostly concerns the difference between direct speech acts and *very simple* indirect speech acts, i.e., the predicted qualitative threshold. Of the five main references on this topic, thus, three (namely, Hirst, LeDoux, & Stein, 1984; Weylman, Brownell, Roman, & Gardner, 1989; Wilcox, Albyn, & Leonard, 1978) only tested utterances like ''Can you do X?'' vs. ''Do X''; the level of difficulty is the same of our ''simple'' protocols. The remaining two studies tested also (Foldi, 1987), or only (Heeschen & Reischies, 1979), more difficult utterances, but not harder than our ''complex'' protocols.

We explored both the qualitative breakpoint (i.e., the difference between direct and simple indirect speech acts: tasks [7] vs. [8]) and the quantitative continuum (with some more difficult tasks: [9]–[11]). As regards the former, our tasks were at the standard level of difficulty of previous research; as regards the latter, they were at the level of difficulty of the most difficult tasks used in previous research. We will also remind, once again, that we used open judgment tasks, whereas all the previous research used closed tasks; this means that our tasks were actually more difficult to CHI patients than those previously designed.

A Searlean approach in terms of illocutions would have been validated by a significant difference across the predicted threshold between direct and simple indirect speech acts, i.e., between tasks [7] and [8]. The difference between our results and the previous ones is that we found no difference here. If no experimental breakdown can be found across a theoretically motivated, qualitative threshold, then there is no reason to suppose that such threshold really exists. Thus, the point is not that our data here show no difference, but that we were *not* looking for statistical significance here. Once the standard prediction is disconfirmed, our theory appears more suitable than the competing ones, because it denies the very idea of a qualitative threshold.

As regards the quantitative continuum of indirect speech acts (tasks [8]–[11]), we found no statistically significant difference. Such difference is more likely to be found with more difficult utterances (such as [6g–h] above), but this is irrelevant to the present study. What we are saying is not that CHI subjects have no inferential problems in communication, but that our approach can describe and explain their problems better than the classical illocutionary framework.

Our theory allows for breakdowns in the individual performance at any degree of difficulty, because we deny the existence of qualitative differences: there are *always* inferences to be drawn in understanding communication. As the difficulty of these inferences increases, breakdowns in the individual performances will become correspondingly more likely, and that's all: there can be no specific breaking points. Where competing theories would predict a specific, *theoretically motivated* breaking point, we found nothing at all. While this crucial negative findings has no explanation in standard theories, it has a simple one in ours: the inferential load is too low to pose problems to CHI subjects.

## 6. NONSTANDARD PATH I: IRONY

After recognizing the literal meaning, the partner's task is to understand the speaker's meaning, i.e., to comprehend the actor's communicative intentions. In this phase, the partner may also recognize the actor's attempt to pursue a nonstandard path.

Many nonstandard paths involve some form of exploitation. The term *exploitation* is borrowed from Grice's (1975) discussion of how his maxims of cooperation can be exploited to convey conversational implicatures. Analogously, the rules of the model proposed by Airenti, Bara, and Colombetti (1993b) can be exploited to lead the dialogue toward nonstandard paths. Irony is the most common example of exploitation; others are *as-if* situations, metaphors, hyperboles, etc.

Here, our discussion will be restrained to the simplest cases of irony, such as:

[14] You have been so kind!

uttered to a clerk insolently unwilling to comply with a legitimate request.

What makes the utterance ironic is that the agents share that precisely the opposite of its literal meaning holds. This blocks the application of the rule stating that:

$$SH_{yx}\ DO_{x}\ \text{express}\ (x, BEL_{x}\ p) \Rightarrow SH_{yx}\ CINT_{xy}\ p$$

(i.e., in the shared belief space, if actor *x* expresses the belief that *p*, then by default she intends to share *p* with *y*).

As shown by this simple example, exploitations are ways to play with sharedness: statement [14], obviously an ironic one in the context described, would be taken seriously if reported to a third person unaware of the specific circumstances. In other words, an ironic intent does not affect how rules are chained in a sequence, but how they are used.

Often (but not necessarily) the ironic intent can be recognized by reversing the meaning expressed. Grice's (1975) definition of irony is limited to such simple cases of reversal. His position, however, has been criticized for two reasons. First, some ironic statements are not understood by this simple rule. Sperber and Wilson (1986) report an excellent quotation from Voltaire's *Candide:* ''When all was over and the rival kings were celebrating their victory with Te Deums in the respective camps . . . .'' Further, Grice's account leaves unclear why *p* should be interpreted as an ironic non-*p* and not as a lie (Morgan, 1990).

Both Sperber and Wilson and Morgan offer alternative interpretations. According to Sperber and Wilson, irony is ''echoic,'' in that it echoes the attitudes and thoughts of somebody else than the speaker herself; according to Morgan, irony is ''transparent pretense'' on the part of the speaker to believe the bizarre thing she has just said.

Our thesis is that the only way to distinguish an ironic utterance from a serious one is to consider what knowledge the actor takes to be shared with the partner. Actually, if the actor wrongly believes that some crucial piece of knowledge is shared with the partner, irony cannot be recognized. Let us consider the examples we gave to our subjects:

[15] (A child is building a big but questionably balanced Lego structure. A girl enters who inadvertently destroys the construction).
Child: Thanks for the help!

[16] (Child A is taking chocolates out of a box and eating them).
Child A: This is the last one.
Child B: Can I have it?
(Child A unwraps the last chocolate, eats it, and hands the empty tinfoil to Child B).
Child B: How generous of you!

[17] (Children A and B are racing in a yard. Child A wins; Child B arrives with a remarkable delay and puffing).
Child A: You run real fast.
Child B: Give me a break!

In each case, the subject's task is to explain what, in his opinion, the child meant by her utterance. To understand the three scenes he had to realize that the actors share the belief that to destroy a toy building is no helpful behavior, that a kind person should be willing to share what she is eating with other people (particularly after a request thereof) and, respectively, that a faster runner should have won the race. Any answer from which it could be inferred that the subject had understood irony would be considered correct, e.g., ''He was joking,'' ''He was laughing at the other child,'' etc.

Our predictions were of a certain difficulty with these tasks, for the reasons we have mentioned above, in particular because of the difficulty of handling different, possibly contrasting, representations of the actors' mental states. This characteristic is common to all exploitations: the understander has to keep in mind both the standard path and the nonstandard path in order to compare them and decide what is going on inside the actor's mind. Consequently, we expected the subjects to encounter more difficulties here than with the standard path, at least as far as simple examples are considered in each category.

### *6.1. Failures of Irony*

Irony, like any other communicative act, may fail. The protocol we used to test our subjects' comprehension of such failures was the following:

[18] (Child A wears a very ugly and dirty T-shirt)
Child B: Lovely shirt you're wearing today!
Child A (rejoicing): You can't have it.

The question we asked was ''What can Child B do now?'' Here, the subject has to understand that the Child B's intention to be ironic failed because Child A did not recognize it and took Child B's statement seriously.

We expected the subjects to encounter more difficulties in handling irony when it fails than when it is successful, for the reasons discussed about the failures of the standard path. Further, we predicted failures of irony to be more difficult than failures of the standard path, for the same reason why successful irony was predicted to be more difficult than the successful standard path. In other words, we predicted the following trends: (i) that nonstandard paths are more difficult than the standard one; (ii) that, within each category, a failure is more difficult than the corresponding successful case; (iii) that, as a consequence of (i) and (ii), failures of nonstandard paths are more difficult than failures of the standard path.

### 6.2. Results

The control group solved all the tasks without any error. The results of the experimental group are reported in Tables 3 and 4. There were three errors (8% of the answers), two of which were corrected with the closed judgment task. This is only suggestive of a certain difficulty with irony: in fact, the comparison with the control group is near significant (Student's $t = 1.897$, $p < .07$). The result seems to confirm that, on the average, CHI patients encounter no real problems with simple everyday communication. Once again, it is an easy guess that they would be less successful in harder situations. It is particularly difficult, however, to make quantitative predictions within this topic, since, as we all know, normal persons themselves may experience some problems with difficult cases of irony.

## 7. NONSTANDARD PATH II: DECEITS

The other type of nonstandard communicative situation we explored concerns the relationship between the mental states communicated by the actor and the private ones she actually entertains, or is supposed to entertain.

Airenti, Bara, and Colombetti (1993b) define a deceit as a premeditated rupture of the rules governing sincerity in the behavior game in play. For example, the German spy Mata Hari may inform the French general Nivelle that there will be a German attack to Verdun, although she knows that there will be no such attack. Her utterance is a deceit in that the game she is playing with the partner commits her to sincerity about this kind of topic (although not about *any* kind of topic, of course).

Contrary to what happens with irony, a deceit must pass undetected to be successful. When the partner realizes that the actor is attempting a deceit, he may either denounce it or pretend credulity, while possibly planning a counterdeceit. These situations correspond to distinct mental states of the partner, which can be formally described by different nestings of the SH operator. Figure 2 presents some possibilities. While there is no principled limit to the complexity of the situations (i.e., to the nesting of the SH operator inside a sequence of BEL operators), humans are incapable of handling deceits beyond a small number of levels.

Actually, deceits are among the most complex social phenomena, as is also shown by the difficulty to prove their presence even in nonhuman Primates, however complex their social life may be (Byrne & Whiten, 1988; Whiten & Byrne, 1988). Deceits require planning and understanding intricate communicative or behavioral strategies, which always include explicit reasoning of the most complex kind about the partner's mental states. Accordingly, it is reasonable to expect some difficulty in planning and understanding even simple deceits.

We decided to partially change the protocol design here, because scenes involving deceits would be either too long or too complicated to understand.

x utters: "There will be an attack to Verdun tomorrow."

Let p = There will be an attack to Verdun tomorrow.

Assume that $BEL_x$ ~p, i.e., x is attempting to deceive y.

| | |
|---|---|
| $SH_{yx}$ p | Successful deceit: y trusts x and assumes p to be shared. |
| $BEL_y$ ~p<br>$BEL_y$ $SH_{xy}$ p | Although y does not recognize the deceit, he does not believe p anyway (e.g., he thinks that x is mistaken). |
| BELy ~p<br>BELy $BEL_x$ ~p<br>$BEL_y$ $BEL_x$ $SH_{yx}$ p | The deceit is detected, but y believes that x has not realized the failure of her attempt. |
| $BEL_y$ ~p<br>$BEL_y$ $BEL_x$ ~p<br>$BEL_y$ $BEL_x$ $BEL_y$ ~p<br>$BEL_y$ $BEL_x$ $BEL_y$ $SH_{yx}$ p | The deceit is detected, y realizes that x has recognized the failure, but thinks that x has not recognized that y has understood that x has recognized that the deceit was detected. |
| ... | |
| $SH_{yx}$ ~p | It is shared that the deceit has been detected, e.g., because y has denounced it. |

FIG. 2. Possible situations after an attempted deceit.

We used one comprehension task, [19], and asked the subjects to directly plan and perform two other deceits: [20] and [21].

[19] *Where is the ring?* The Princess and the Villain together hide a ring. Both exit. The Villain sneaks in again and moves the ring to another place, then exits.
The subject is asked:
a. According to the Princess, where will the Villain look for the ring when he comes back?
b. According to the Villain, where will the Princess look for the ring when she comes back?
Comment: the subject has to understand that both expect the other to look for the ring in the original hideout, although one of them, the Villain, holds a different belief about where the ring actually is.

[20] *The Princess and the Villain.* The subject is shown two puppets, the Princess and the Villain. The Villain is then put away. The subject is

told that the Villain is chasing the Princess, and is requested to suggest a hideout to the Princess. The Princess is hidden in the place chosen by the subject. Then the Villain enters, and asks the subject where the Princess is. After the subject has replied, the Villain exits and the subject is asked to explain why he said whatever he said.
Comment: in order to pass the task, the subject has to lie, pointing the Villain to a wrong place, while letting him believe it is the right one.

[21] *Chocolates or biscuits?* The subject is shown some chocolates and some biscuits, and asked which he likes better. Then, he is told that the Villain will also choose either, that he (the Villain) will choose first and that, being a Villain, he will always choose what the subject would choose. Then the Villain enters, and asks the subject ''Which sweet do you like better?'' Again, after the subject has replied, the Villain exits and the subject is asked to explain why he has said what he has.
Comment: the task here is to lie to the Villain about the preferred sweet, so to have him choose the less preferred one. In the evaluation of the performance, the possibility of a double deceit must also be controlled: the subject may suppose that the Villain would suspect a deceit, therefore choosing the sweet the subject had *not* indicated. In this case, paradoxically, the task is passed if the subject does *not* lie. No subject pushed the nesting beyond the second level.

### *7.1. Failures of Deceits*

We explored the failure of deceits with the following protocol:

[22] (Child A has a toy that Child B always asks for himself. There are two boxes in the living room, one red and one yellow. When Child A sees Child B coming, she hides the toy in the yellow box. Child B enters).
Child B: Could I have a box?
Child A: Yes.
Child B: Which one can I have?
Child A points at the red box.
Child B takes the *yellow* box and goes away.

We asked the subject what Child A might do after this. As in the other cases of failure, our prediction was of a poor performance: worse than that on the successful case, and worse than that on failures of the standard path and of irony. As we have discussed, deceits are among the most complex cognitive phenomena in communication.

### *7.2. Results*

The results of the experimental group are reported in Tables 3 and 4; the control group solved the tasks without a single error. Here we find the first statistically significant difference with the control group, which is an unsurprising confirmation of our predictions, given the heavy load imposed by deceits on an agent's inferential and planning capabilities. Of the 13 subjects, errors were made by two in [19], none in [20], six in [21], and eight in [22]. Considering the successful deceits taken as a whole (tasks [19], [20], and [21]), thus, there were 8 errors out of $13 \times 3 = 39$ items (21% of the total). The comparison with the control subjects is highly significant: Student's $t = 2.889$, $p < .008$**). As for the failure (task [22]), there were 5 errors out of $13 \times 1 = 13$ items (38%); the comparison with the controls is highly significant: Student's $t = 4.302$, $p < .0002$**.

Decomposition into subitems brings some interesting result as to the relative difficulty of the tasks. Given our theoretical analysis, task [22] was expected to be the hardest, because it requires the recognition of the failure and the planning of a suitable strategy for recovery; in this case, further, both involve a relatively complex nesting of mental states. Task [21] also involves a complex nesting, but should be easier than [22], because to plan a deceit is simpler than to recognize and repair a failed one. We had no way to predict the relative difficulty of the two tasks remaining. Task [20] requires planning, while [19] only requires understanding. In terms of the mental states involved, however, [20] is easier than [19], because it requires a single nesting, whereas the latter requires two independent nestings (namely, what each character will believe the other will believe). To sum up, we expected one of two trends here:

[19] < [20] << [21] < [22], or:
[20] < [19] << [21] < [22] ($p < .01$**).

We explored these possibilities with Page's test (Page, 1963). The latter is highly significant ($L = 246$, $p < .01$**), confirming our predictions and also showing that CHI subjects find it harder, ceteris paribus, to deal with the *nesting* of mental states than with the *planning* of simple lies. This trend also shows if only the three successful cases (i.e., [20] < [19] < [21]) are considered: Page's test gives $L = 81$, $p < .05$*.

## 8. THEORY OF MIND INFERENCES

As we have repeatedly discussed, understanding communication often requires reasoning upon another agent's supposed mental states. The skill to simulate another individual's mental states is peculiar to the human species and, possibly, to a few other anthropomorphic Primates such as chimpanzee (Premack & Woodruff, 1978). At least in our species, it seems to require a specialized, modular machinery dubbed *Theory of Mind Mechanism,* or ToMM (Leslie, 1987, 1994). A disorder in the development of ToMM is associated with childhood autism (Baron-Cohen, 1995; Baron-Cohen, Leslie & Frith, 1985; Leslie & Thaiss, 1992), while its later disruption might be associated with schizophrenia (Frith, 1992).

An active ToMM is required whenever the complexity of the current interaction goes beyond what is in the individual's memory, i.e., always, the exceptions being highly conventional and script-like contexts (such as participating in a ceremony, or buying bread and milk), where nobody would seriously wonder about the speaker's mental states, unless something very peculiar happens.

There are currently a large number of experimental studies in this area. Most aim at understanding what species, beside humans, possess the ability to simulate conspecifics' mental states (e.g., Byrne & Whiten, 1988; Premack & Woodruff, 1978; Whiten, 1991), or at clarifying its development in the child (e.g., Astington, Harris, & Olson, 1988; Perner, 1991; Wellman, 1990; see also Karmiloff-Smith, 1992, for an overview).

Although we are more interested in communication here, rather than in ToMM per se, we wanted to assess our subjects' capability to perform basic ToMM tasks, because of their crucial role in conversation. Further, we are not aware of any data in the literature about ToMM performance in brain damaged persons. We expected no impairment of this ability, or at most a very mild one, because a dramatic damage to the ToMM would correspondingly hamper any social interaction but the simplest (Frith, 1992): simulating other individuals is necessary for any kind of high-level social behavior, whether communicative or not. The literature, on the other hand, is unanimous in evaluating the social life of these brain-damaged patients as substantially preserved.

We used two classical tasks of the area, known under the names of ''Maxi task'' (Wimmer & Perner, 1983) and ''Smarties task'' (Perner, Leekam, & Wimmer, 1987):

[23] *Maxi task.* Two persons (the subject, and a foil conventionally named Maxi) watch while the experimenter puts a chocolate bar in a drawer. Maxi then leaves the room. The experimenter opens the drawer, takes out the chocolate and hides it elsewhere. The subject is then asked where Maxi will look for the chocolate when he comes back.
Comment: the task is passed if the subject answers that Maxi will look into the drawer.

[24] *Smarties task.* The subject is shown a closed Smarties box and asked to guess what is inside. Obviously, he answers there are Smarties. Then the box is open and he can see there is a pencil instead. The box is closed again. Then the subject is asked to guess what Maxi will answer when asked the same question.
Comment: since Maxi has not seen the pencil, the correct answer is ''Smarties.''

To pass these tasks, the subject has to realize the relationships between perception, belief, and action, and the autonomy of each individual's mental states. It is easy to see how closely the structure of these tasks resembles that of the other ones we devised. This similarity is intrinsic to our approach to communication, that strongly emphasizes the reciprocal simulation of the agents involved in conversation.

### *8.1. Results*

The control group solved both tasks without any error. The results of the experimental group are reported in Table 3. There was only one error on the Maxi task and none on the Smarties task. This confirms our expectations: possible problems in communication after CHI are not due to a diminished ability in reasoning on other individuals' mental states.

## 9. GLOBAL RESULTS

All our major predictions were confirmed. As a first thing, we found no evidence for the greater difficulty of indirect with respect to direct speech acts, that would be predicted by most studies in pragmatics and in neuropsychology. We have already discussed the topic and have presented some hypotheses as to why our data differ from those available in the relevant literature (see Section 5).

An issue worth discussing here, before turning to a more detailed analysis of the data, is that of response time. We have not reported the time it took to the subjects to solve (or not to solve) the tasks. As Kosslyn and Intriligator (1992) point out, time is an important factor with tasks of this kind, in that a response time of, say, 5 min instead of 20 sec may indicate a severe impairment in the underlying cognitive processes, which would pass undetected if only the correctness of the answer is taken into account. Having taperecorded all the protocols, we found that the time taken by the control and the experimental groups was comparable; actually, the total time for completing the protocol never varied more than 10 min on an average of 30 min.

### *9.1. Trends in Difficulty*

The theory predicts various differences in processing difficulty. Nonstandard paths, taken as a whole, are expected to be more difficult than the standard path. Within the standard path, further, simple speech acts (whether direct or indirect) are expected to be easier than complex ones. Among the nonstandard paths, the increase should be from irony to deceits to failures.

In analyzing these trends, we collapsed the data of direct and simple indirect speech acts. We put them together for two reasons. The first is theoretical: we assume them to be processed in the same way, and to be absolutely equivalent. The second reason is empirical: none of our 13 subjects made even a single error in understanding them.

In short, our predicted trend of difficulty is the following:

simple (direct and indirect) speech acts $<$ complex speech acts $<<$ irony $<$ deceits $<$ failures

This hypothesis is confirmed by the data: Page's test gives $L = 462$ ($p < .01$**).

It is important to remember that, for each pragmatic category, we chose relatively simple tasks. There can be little doubt that a complex irony (such as Voltaire's, quoted in Section 6) is harder to comprehend than a very simple deceit (such as a simple yes/no lie: ‘‘Was it you who ate the jam?’’ ‘‘No’’). Our data prove that a *simple* irony is easier than a *simple* deceit, but harder than a *simple* standard path, and so on. We would never derive from this trend that *any* irony is easier than *any* deceit, and harder than *any* standard path.

Had we used more difficult tasks, we might have found more differences between the experimental and the control groups. As we have already discussed in Section 5, however, our aim was not to prove that communication is necessarily damaged after CHI. Instead, we wanted to start our investigation from the very fundamentals of communication, such as can be found in everyday, colloquial communication. We found that communication *is* actually impaired after CHI, although in quite a subtle way. We believe that this result was due to the adoption of a clear and detailed theory of communication, which provided us with a set of formal definitions and allowed us to clearly define the scope of our analyses and to formulate precise predictions, although of a qualitative nature.

There would have been little point in the exploration of more complex aspects of the theory, without a prior understanding of the extent to which our subjects could handle basic ones. Once departed from basic tasks, further, it becomes more difficult to decide what is more or less simple, and why. An ironic utterance, for example, may be difficult to understand because it requires a complex inferential chain, or because it requires a heavy use of some unusual world knowledge.

### *9.2. Failures*

Failures are a complex topic, from both a theoretical and a neuropsychological point of view. Let us review our predictions and results.

The impairment in the subjects' performance on failures taken as a whole, compared to the control group, is highly significant (Student's $t = 3.888$, $p < .0007$**; see Table 3). Decomposing this result into subitems (Table 4), all the performances prove significantly worse than those of the control group: on direct and indirect speech acts (in both cases $t = 2.739$, $p <$

TABLE 5
Comparison between the Performance on the Successful and the Failure Case

| | Test: 13 brain-injured subjects | | | |
|---|---|---|---|---|
| Failure | Direct s.a. | Indirect s.a. | Irony | Deceit |
| Comparison with the successful case (Wilcoxon, tied) | $Z = 2.236$ $p < .03$* | $Z = 2.070$ $p < .04$* | $Z = 1.897$ $p < .06$ | $Z = 2.239$ $p < .03$* |

.01**), on irony ($t = 2.739$, $p < .01$**) and, with a high significance, on deceits ($t = 4.382$, $p < .0002$**).

A second prediction was that, within each pragmatic category (direct speech acts, indirect speech acts, irony and deceits), the performance had to be worse on the examples of failure than on the successful ones: to repair a failure requires that the actor recognize it, identify its causes and plan a suitable alternative strategy. For each category, we used Wilcoxon's test to compare the performance on the failure to that on the average of the three examples of success. As can be seen from Table 5, the prediction is generally confirmed (except in the case of irony, which is, however, near significant).

Finally, we expected failures to show a trend similar to that of the successful cases, i.e., we expected failures of deceits to be harder than failures of irony, and failures of irony harder than failures of the standard path. Page's test disconfirms this hypothesis by a very slight interval ($L = 159$ against a minimum of 163 for a $p < .05$ significance). Remarkably, however, there is no difference between the failures of direct and of indirect speech acts, which adds further evidence to our hypothesis on the topic.

### *9.3. Other Neuropsychological Correlations*

We had no particular predictions about the possible correlations between the subjects' performance on the tasks about communication and on collateral neuropsychological tests. Although we did not perform a statistical evaluation of this topic, some observations may be of interest.

Four subjects made no errors at all: 9, 10, 11, and 13. Three of these performed relatively well on collateral tests. Subject 11 presents a more severe cognitive impairment; this apparent contradiction might be due to her rehabilitation program, part of which is currently concentrating on planning.

Subject 5 only made mistakes on deceits. Globally, her cognitive damage is mild, but her memory is more severely impaired, in both its short- and long-term components.

Two subjects, 1 and 8, only had trouble with failure recovery. Both present frontal symptoms and a damage to the working memory. In general, working memory seems to be the function most sensitively related to the communicative performance.

The remaining six subjects made errors in two or three categories each. Here we find those with the oldest onset of pathology: 2, 4, and 6 (respectively, 65, 145, and 63 months). These three subjects are logorrheic, possibly as an attempt to compensate for their communicative deficit, but often get tangled up in their own talk, and leave the partner the responsibility to find a way out. However, they would typically deny having any problem whatsoever and always try to impress favorably upon the partner. Subject 2 is, within this group, the less impaired from a general point of view; he was able to solve closed judgment task on complex indirect speech acts and on deceits, but not on failure recovery. The last three subjects (3, 7, and 12) are the most severely impaired from a cognitive viewpoint.

Clearly, these observations might be strengthened, and lend themselves to statistical analysis, with a greater number of subjects. As a first comment, we find them reasonable: working memory is likely to be heavily involved in the kind of reasoning required for the handling of deceit, where complex nestings of mental states have to be taken into account, i.e., it is likely to play a major role in complex ToMM reasoning. As regards planning, it is crucial to failure recovery: not surprisingly, it is impaired in the subjects who find these tasks particularly puzzling.

## 10. FUTURE WORK

As we have already stated, this work did not aim at building a cognitive theory from neuropsychological data. Our goal was rather to test some predictions drawn from an independently developed theory of communicative competence, to constrain it with data from impaired performance, and to match it against competing ones.

We are quite satisfied with the data collected. When a competing theory exists, as in the case of the direct/indirect controversy, our data validate our thesis. Further, our theory predicts the trend of difficulty exhibited by CHI subjects in the experiments. No existing theory of pragmatics had proposed such a precise and justified trend before.

The work presented here is meant as an initial step. Planned developments include a follow-up of these subjects after their recovery, and a broadening of the analysis to cover other pathologies (e.g., Alzheimer's disease and other forms of dementia, Parkinson's disease, etc.), as well as to subjects with no neuropsychological involvement (e.g., elderly persons). We also plan to study focal lesions; in particular, to investigate possible differences between pathologies of the right and of the left hemisphere. A good idea would be, of course, to use tests like those proposed here as a first screening, to identify the most interesting subjects and to focus on them with greater attention. This will require tests for the effective exploration of less fundamental aspects of communication. Finally, we are developing an approach to discourse planning; this is particularly interesting from a neuropsychological point of view, since many pathologies seem to affect the generation of behavior of one's own rather than the comprehension of the behavior of other individuals'.

## REFERENCES


Airenti, G., Bara, B. G., & Colombetti, M. 1983. Planning perlocutionary acts. *Proceedings of the 8th International Joint Conference on artificial intelligence.* Karlsruhe.

Airenti, G., Bara, B. G., & Colombetti, M. 1984. Planning and understanding speech acts by interpersonal games. In B. G. Bara & G. Guida (Eds.), *Computational models of natural language processing.* Amsterdam: North-Holland.

Airenti, G., Bara, B. G., & Colombetti, M. 1989. Knowledge for communication. In M. M. Taylor, F. Néel & D. G. Bouwhuis (Eds.), *The structure of multimodal dialogue.* Amsterdam: North-Holland.

Airenti, G., Bara, B. G., & Colombetti, M. 1983a. Conversation and behavior games in the pragmatics of dialogue. *Cognitive Science,* **17,** 197–256.

Airenti, G., Bara, B. G., & Colombetti, M. 1993b. Failures, exploitations and deceits in communication. *Journal of Pragmatics,* **20,** 303–326.

Airenti, G., & Colombetti, M. 1991. Artificial intelligence and the representation problem. In M. De Glas & D. Gabbay (Eds.), *Proceedings of the 1st World Conference on the Fundamentals of Artificial Intelligence, Paris.* Paris: Angkor.

Allen, J. F., & Perrault, C. R. 1980. Analyzing intention in utterances. *Artificial Intelligence,* **15,** 143–178.

Appelt, D. E. 1985. *Planning English sentences.* Cambridge: Cambridge Univ. Press.

Astington, J. W., Harris, P. L., & Olson, D. R. (Eds.) 1988. *Developing theories of mind.* Cambridge: Cambridge Univ. Press.

Austin, J. L. 1962. *How to do things with words.* London: Oxford University Press. [2nd ed. revised by J. O. Ormson & M. Sbisà. London: Oxford University Press, 1975].

Bara, B. G. 1995. *Cognitive science: A developmental approach to the simulation of the mind.* Hillsdale, NJ: Erlbaum.

Bara, B. G., & Guida, G. 1984. Competence and performance in the design of natural language systems. In B. G. Bara & G. Guida (Eds.), *Computational models of natural language processing.* Amsterdam: North-Holland.

Baron-Cohen, S. 1995. *Mindblindness. An essay on autism and theory of mind.* Cambridge, MA: MIT Press.

Baron-Cohen, S., Leslie, A. M., & Frith, U. 1985. Does the autistic child have a 'theory of mind'? *Cognition,* **21,** 37–46.

Brown, P., & Levinson, S. C. 1987. *Politeness: Some universals in language use.* Cambridge: Cambridge Univ. Press.

Burling, R. 1993. Primate calls, human language, and nonverbal communication. *Current Anthropology,* **343,** 25–53.

Byrne, R. W., & Whiten, A. (Eds.). 1988. *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans.* Oxford: Clarendon Press.

Chomsky, N. 1957. *Syntactic structures.* The Hague: Mouton.

Chomsky, N. 1965. *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Chomsky, N. 1980. *Rules and representation.* New York: Columbia Univ. Press.

Clark, H. H. 1979. Responding to indirect speech acts. *Cognitive Psychology,* **11,** 430–477.

Clark, H. H. 1992. *Arenas of language use.* Chicago, IL: University of Chicago Press.

Clark, H. H., & Marshall, C. R. 1981. Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding.* Cambridge: Cambridge Univ. Press.

Cohen, P. R., & Levesque, H. J. 1985. Speech acts and rationality. *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics,* Chicago, IL.

Cohen, P. R., & Levesque, H. J. 1990. Rational interaction as the basis for communication.

In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in communication.* Cambridge, MA: MIT Press.

Cohen, P. R., & Levesque, H. J. 1991. Confirmation and joint action. *Proceedings of the 12th International Joint Conference on Artificial Intelligence,* Sydney.

Cohen, P. R., Morgan, J., & Pollack, M. E. (Eds.). 1990. *Intentions in communication.* Cambridge, MA: MIT Press.

Cohen, P. R., & Perrault, C. R. 1979. Elements of a plan-based theory of speech acts. *Cognitive Science,* **3,** 177–212.

Colombetti, M. 1993. Formal semantics for mutual belief. *Artificial Intelligence,* **62,** 341–353.

Davis, A. G., & Wilcox, M. J. 1981. Incorporating parameters of natural conversation in aphasia treatment. In R. Chapey (Eds.), *Language intervention strategies in adult aphasia.* Baltimore, MD: Williams & Wilkins.

Davis, A. G., & Wilcox, M. J. 1985. *Adult aphasia rehabilitation: Applied pragmatics.* Windsor: NFER-Nelson.

de Bleser, R., & Weisman, H. 1986. The communicative impact of non-fluent aphasia on the dialog behavior of linguistical unimpaired partners. In F. Lowenthal & F. Vandamme (Eds.), *Pragmatics and education.* New York: Plenum.

Feyereisen, P., & Seron, X. 1982. Nonverbal communication and aphasia: A review. *Brain and Language,* **16,** 191–236.

Foldi, N. S. 1987. Appreciation of pragmatic interpretation of indirect commands: Comparison of right and left hemisphere brain-damaged patients. *Brain and Language,* **31,** 88–108.

Foldi, N. S., Cicone, M., & Gardner, H. 1983. Pragmatic aspects of communication in brain-damaged patients. In S. J. Segalowitz (Ed.), *Language functions and brain organization.* New York: Academic.

Frith, C. D. 1992. *The cognitive neuropsychology of schizophrenia.* Hove, UK/Hillsdale, NJ: Erlbaum.

Frith, U., Morton, J., & Leslie, A. M. 1991. The cognitive basis of a biological disorder: Autism. *Trends in Neurosciences,* **14,** 433–438.

Gazdar, G. 1981. Speech act assignment. In A. K. Joshi, B. L. Webber & I. A. Sag (Eds.), *Elements of discourse understanding.* Cambridge: Cambridge Univ. Press.

Grice, H. P. 1957. Meaning. *Philosophical Review,* **67,** 377–388.

Grice, H. P. 1975. Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics, vol. 3: Speech acts.* New York: Academic Press.

Grice, H. P. 1989. *Studies in the way of words.* Cambridge, MA, & London: Harvard University Press, 1989.

Guilford, A., & O'Connor, J. 1982. Pragmatic functions in aphasia. *Journal of Communication Disorders,* **15,** 337–346.

Harman, G. 1977. Review of ''Linguistic Behavior'' by Jonathan Bennett. *Language,* **53,** 417–424.

Heeschen, C., & Reischies, F. 1979. *On the ability of brain damaged patients to understand indirect speech acts.* Internal report. Berlin: Free University of Berlin, Institute of Physiology.

Hirst, W., LeDoux, J. E., & Stein, S. 1984. Constraints on the processing of indirect speech acts: Evidence from aphasiology. *Brain and Language,* **23,** 26–33.

Holland, A. 1982. Observing functional communication of aphasic adults. *Journal of Speech and Hearing Disorders,* **47,** 50–56.

Howard, D., & Hatfield, F. M. 1987. *Aphasia therapy: Historical and contemporary issues.* Hillsdale, NJ: Erlbaum.

Joanette, Y., & Brownell, H. H. (Eds.). 1990. *Discourse ability and brain damage: Theoretical and empirical perspectives.* New York: Springer.

Joanette, Y., Goulet, P., Ska, B., & Nespoulous, J.-L. 1986. Informative content of narrative discourse in right-brain-damaged right-handers. *Brain and Language,* **29,** 81–105.

Karmiloff-Smith, A. 1992. *Beyond modularity: A developmental perspective on cognitive science.* Cambridge, MA: MIT Press.

Kosslyn, S. M., & Intriligator, J. M. 1992. Is cognitive neuropsychology plausible? The perils of sitting on a one-legged stool. *Journal of Cognitive Neuroscience,* **4,** 96–106.

Leech, G. N. 1983. *Principles of pragmatics.* London: Longman.

Leslie, A. M. 1987. Pretense and representation: The origins of ''theory of mind.'' *Psychological Review,* **94,** 412–426.

Leslie, A. M. 1994. *Pretending* and *believing:* Issues in the theory of ToMM. *Cognition,* **50,** 211–238.

Leslie, A. M., & Thaiss, L. 1992. Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition,* **43,** 225–251.

Levinson, S. C. 1983. *Pragmatics.* Cambridge: Cambridge Univ. Press.

Lewis, D. 1969. *Convention. A philosophical study.* Cambridge, MA: Harvard Univ. Press.

Marr, D. 1977. Artificial intelligence: A personal view. *Artificial Intelligence,* **9,** 37–48.

Marr, D. 1982. *Vision: A computational investigation into the human representation and processing of visual information.* San Francisco: Freeman.

McCarthy, J. 1980. Circumscription—A form of non-monotonic reasoning. *Artificial Intelligence,* **13,** 27–39.

McCarthy, J., & Hayes, P. J. 1969. Some philosophical problems from the standpoint of artificial intelligence. In B. Meltzer & D. Michie (Eds.), *Machine intelligence 4.* Edinburgh: Edinburgh Univ. Press.

McDonald, S. 1992. Differential pragmatic language loss after closed head injury: Ability to comprehend conversational implicature. *Applied Psycholinguistics,* **13,** 295–312.

McDonald, S., & van Sommers, P. 1993. Pragmatic language skills after closed head injury: Ability to negotiate requests. *Cognitive Neuropsychology,* **10,** 297–315.

Morgan, J. 1990. Comments on Jones and on Perrault. In P. R. Cohen, J. Morgan & M. E. Pollack (Eds.), *Intentions in communication.* Cambridge, MA: MIT Press.

Page, E. B. 1963. Ordered hypotheses for multiple treatments: A significance test for linear ranks. *Journal of the American Statistical Association,* **58,** 216–230.

Penn, C. 1987. Compensation and language recovery in the chronic aphasic patient. *Aphasiology,* **1,** 235–245.

Penn, C., & Cleary, J. 1988. Compensatory strategies in the language of closed head injured patients. *Brain Injury,* **1,** 3–17.

Perner, J. 1991. *Understanding the representational mind.* Cambridge, MA: MIT Press.

Perner, J., Leekam, S. R., & Wimmer, H. 1987. Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology,* **5,** 125–137.

Perrault, C. R., & Allen, J. F. 1980. A plan-based analysis of indirect speech acts. *American Journal of Computational Linguistics,* **6,** 167–182.

Pollack, M. E. 1990. Plans as complex mental attitudes. In P. R. Cohen, J. Morgan & M. E. Pollack (Eds.), *Intentions in communication.* Cambridge, MA: MIT Press.

Premack, D. 1986. *Gavagai! Or the future history of animal language controversy.* Cambridge, MA: MIT Press.

Premack, D., & Woodruff, G. 1978. Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences,* **1,** 515–526.

Recanati, F. 1995. The alleged priority of literal interpretation. *Cognitive Science,* **19,** 207–232.

Reiter, R. 1980. A logic for default reasoning. *Artificial Intelligence,* **13,** 81–132.

Schiffer, S. R. 1972. *Meaning.* Oxford: Oxford Univ. Press.

Searle, J. R. 1969. *Speech acts: An essay in the philosophy of language.* London: Cambridge Univ. Press.

Searle, J. R. 1975. Indirect speech acts. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics, vol. 3: Speech acts.* New York: Academic Press.

Searle, J. R. 1980. Minds, brains, and programs. *Behavioral and Brain Sciences,* **3,** 417–424.

Searle, J. R. 1992. *The rediscovery of the mind.* Cambridge, MA: MIT Press.

Searle, J. R., & Vanderveken, D. 1985. *Foundations of illocutionary logic.* Cambridge: Cambridge Univ. Press.

Semenza, C., & Zettin, M. 1988. Generating proper names: A case of selective inability. *Cognitive Neuropsychology,* **5,** 711–721.

Semenza, C., & Zettin, M. 1989. Evidence from aphasia for the role of proper names as pure referring expressions. *Nature,* **342,** 678–679.

Shallice, T., & Burgess, P. W. 1991. Deficits in strategy application following frontal lobe damage in man. *Brain,* **114,** 727–741.

Sherratt, S. M., & Penn, C. 1990. Discourse in a right-hemisphere brain-damaged subject. *Aphasiology,* **4,** 539–560.

Sperber, D., & Wilson, D. 1986. *Relevance. Communication and cognition.* Oxford: Blackwell.

Stemmer, B. 1994. A pragmatic approach to neurolinguistics: Requests (re)considered. *Brain and Language,* **46,** 565–591.

Strawson, P. F. 1964. Intention and convention in speech acts. *Philosophical Review,* **73,** 439–460.

Tirassa, M. 1997. Mental states in communication. *Proceedings of the 2nd European Conference on Cognitive Science.* Manchester, UK.

Wellman, H. M. 1990. *The child's theory of mind.* Cambridge, MA: MIT Press.

Weylman, S. T., Brownell, H. H., Roman, M., & Gardner, H. 1989. Appreciation of indirect requests by left- and right-brain-damaged patients: The effects of verbal context and conventionality of wording. *Brain and Language,* **36,** 362–377.

Whitten, A., & Byrne, R. W. 1988. Tactical deception in primates. *Behavioral and Brain Sciences,* **11,** 233–244.

Whiten, A. (Ed.). 1991. *Natural theories of mind. Evolution, development and simulation of everyday mindreading.* Oxford: Blackwell.

Wilcox, M. J., Albyn, D. G., & Leonard, L. B. 1978. Aphasics' comprehension of contextually conveyed meaning. *Brain and Language,* **6,** 362–377.

Wimmer, H., & Perner, J. 1983. Belief about beliefs. Representation and constraining functions of wrong beliefs in young children's understanding of deception. *Cognition,* **13,** 103–128.