



# Telling it like it isn't

## – a comprehensive approach to analyzing verbal deception<sup>1</sup>

Swati Gupta<sup>1</sup>, Kayo Sakamoto<sup>1</sup> and Andrew Ortony<sup>1,2</sup>

<sup>1</sup>Program in Computational Social Cognition, Institute of High Performance Computing, Singapore

<sup>2</sup>Northwestern University, Evanston, IL 60208 USA

**Abstract** Verbal deception is everywhere—in interpersonal relationships, in politics, in advertising, and in courts of law. In all of these, and in many other domains, we often find people “telling it like it isn't.” This chapter is concerned with understanding what this seemingly simple idea really means. Much work has been done in several fields (e.g., communication studies, philosophy, linguistics) and on various aspects of verbal deception, but nowhere is there a comprehensive integration of the different views and observations that are to be found in the literature. In this chapter, we attempt to pull together the multifarious strands by proposing two taxonomies of verbal deception—one pertaining to verbal deception *types* and one to verbal deception *strategies*.

## 1 Introduction

In the satirical movie *The Invention of Lying*, Mark, the protagonist, lives in a full-disclosure world in which the concept of lying does not exist and in which everybody always expresses their every thought to their interlocutors. Quite by chance, Mark discovers a radically new concept—one totally alien to the culture in which he lives; he discovers the concept of lying. In his excitement he tries to explain the idea to his friends. He tells them that he said “something that isn’t,” but they have no idea what he means. He tries to explain by using blatant falsehoods such as “I’m a black Eskimo astronaut who invented the bicycle.” But his attempts are all a dismal failure. His friends simply accept what he says without a moment’s hesitation, being unable to imagine that what Mark or anyone else in their universe says can be anything other than the truth. We hope in this chapter to be more successful than Mark was in explaining the concept of verbal deception, but we have some sympathy for Mark, because a proper understanding of verbal deception, including lying, is more complicated than it might at first appear.

Many of the complexities associated with the nature of verbal deception, and in particular the nature of lying (even while telling the truth) have been detailed in the rigorous and influential work of Cristiano Castelfranchi and his colleagues (Castelfranchi and Poggi 1993; Vincent and Castelfranchi 1979). Indeed, his approach and his meticulous attention to detail have been an inspiration to all of us interested in a cognitive science approach to understanding linguistic behavior, and for many of us, it served as one of the key starting points for thinking about such issues. It is thus fitting that we dedicate this chapter to Cristiano’s many contributions to the field. Whereas much of his work on verbal deception has focused on lying (especially by telling the truth), lying is only a subset of several types of verbal deception, so with this in mind, we shall attempt in this chapter to situate his deception types in the broader spectrum of verbal deception.

At first blush, one might wonder why anybody bothers to study verbal deception. Surely, it’s quite simple—as the Oxford English Dictionary (OED) puts it, deception is “To cause to believe what is false; to mislead as to a matter of fact, lead into error, impose upon, delude, ‘take in’.” Authoritative a source as the OED is, however, things aren’t so simple. The OED’s definition is a fine definition of straightforward lying (although the definition that

the OED gives for lying is even less complex, namely, “to utter falsehood; to speak falsely”), but there is much more to verbal deception, and indeed to lying, than the OED’s characterizations. For example, the person(s) whom the speaker is intending to deceive (whom we shall refer to as the hearer(s)) might already believe the thing that is false, as when the child says “Look, mommy, the tooth fairy left this quarter under my pillow while I was asleep last night.” If the mother responds with something like “Yes, I told you that would happen,” she is being implicitly deceptive by allowing the child to continue to hold a false belief. Clearly, in this transaction, the parent does not “cause [the child] to believe what is false” because the child already believes it. Does the parent “mislead [the child] as to a matter of fact”? Well not really, since the child is already misled. Furthermore, the issue does not simply concern “a matter of fact”, but rather, it concerns what the mother *believes* to be a matter of fact (namely that there really is no tooth fairy, etc.). So, we would have to augment the OED definition so as to read something like “To cause [the hearer] to believe something that the speaker believes is false, or to allow [the hearer] to continue to believe something that the speaker believes is false.” This is already more complicated than the definition with which we started, but it is still not adequate. The second clause, “to allow [the hearer] to continue to believe...” lacks any mention of the need for the speaker to be in a position to correct a hearer’s pre-existing misapprehension.

Examples of the kind just discussed only scratch the surface of the difficulties we face when we try to provide a comprehensive analysis of verbal deception and lying. Scholars from several different disciplines who have studied verbal deception are well aware of such problems. Perhaps this is why they rarely start their analyses from dictionary definitions, except perhaps, as we do, to highlight the inadequacy of such definitions. However, many of them, (e.g., Turner et al. 1975; Vincent and Castelfranchi 1979) start with a characterization along the following lines:

Deception is an intentional act (of controlling information so as either (a) to make a hearer believe something that the speaker believes is false, or (b) to prevent the hearer from believing something that the speaker believes is true.

Provided that we augment this account to accommodate cases like the tooth fairy example, we have no problem with taking this as a starting point, but, only as a starting point, because there is more to be said and more to understand than is captured by this kind of abbreviated account. If we are to give a more comprehensive account, however, we are going

to have to rely upon a few key concepts—concepts that we shall now introduce and briefly discuss.

### ***1.1 “The truth”***

In the context of verbal deception, what is sometimes referred to as “the truth” and “a falsehood” (or “something that is true/false” etc.) takes on a rather different meaning than that which it has in philosophical discussions. The easiest way to understand it is to think of it as the truth or falsehood *as the speaker sees it*. More specifically, the truth is some state of affairs which, correctly or not, the speaker believes to be the case and wishes to conceal from the hearer. And a falsehood is some state of affairs which, correctly or not, the speaker believes not to be the case but which he or she wishes the hearer to come to, or continue to believe. In our characterizations, we will let P be the proposition representing the speaker’s conception of the “true” state of affairs in cases where the speaker seeks to conceal “the truth” from hearer. In cases where the speaker wants the hearer to come to or continue to believe a falsehood, we will let P be the proposition representing the speaker’s conception of the “false” state of affairs. We let Q represent the proposition that the speaker actually asserts in trying to deceive. As we shall see, it is not always the case that the speaker must believe that Q is false. All that is required is that in asserting Q, the speaker is expecting, or at least hoping, that the hearer will not or will conclude P, depending upon whether P is true or false.

### ***1.2 Common Ground***

Verbal deception often, but by no means always, occurs when a speaker is a respondent in a verbal exchange. Sometimes, however, the turns in the exchange are separated by so much time that an observer might not realize that the speaker’s utterance is actually a response to an earlier, temporally remote communication. In other cases, the speaker’s utterance is a response to an imagined, assumed, presupposed, or anticipated communication—an implicit communication. Examples of this abound. For instance, in the political arena, negative campaign advertisements frequently contain intentionally misleading information, for example by repudiating facts about an opponent. In such cases, where the originator of the

misleading claim is broadcasting to some large audience, there often is no identifiable explicit assertion to which the misleading claim is a response, although the facts being misrepresented are likely to be widely believed and possibly even documented in the public record<sup>2</sup>.

The implications of these complications for an analysis of deception can be finessed by simply assuming that a deceptive claim (spoken or written, although we shall focus on spoken) can usually be analyzed as a response to a prior utterance that might have been made, even if no such utterance was actually made. So when the mother comforts the child at bedtime, raising for the first time the idea that the tooth fairy will come during the night and exchange the tooth for money, we can assume some hypothetical prior utterance from the child such as “Why are we putting my tooth under my pillow?” or “What will happen to my tooth now that it’s fallen out?” Our analysis is in no way compromised by treating the speaker’s deceptive assertion as a response to only a possible prior utterance (a possible claim, question, accusation, etc.).

The specifics of an assumed prior utterance are often immaterial. To see this, we need the concept of *common ground* (Clark and Brennan, 1991)—the mutually shared knowledge to which each participant in a conversation assumes the other has access. This knowledge can be established by the *linguistic* context as when, in a conversation about national politics, a reference to “the president” is likely to be understood to have a different referent than in a conversation about the central administration of a university. In such cases, the common ground is established by the linguistic context together with assumptions about the beliefs of the interlocutors about one another. Common ground can also be established by the *non-linguistic* context—by the physical situation, the objects and events that each participant knows the other can (or in some cases, as in a telephone conversation, cannot) see. So, for example, when two people are watching a tennis match the mutual knowledge about what they are both watching serves the same function as a linguistically established context. The common ground establishes a common reference point in terms of which it is easier for each to understand what the other is talking about.

### ***1.3 Relevance and Satisfaction***

The reason that common ground is important is that it provides the background in terms of which an utterance can be judged as being relevant in a context. In the absence of some

very specific and rather improbable assumptions, a speaker who asks a stranger in the street the directions to the nearest bank could be forgiven for wondering what his interlocutor was talking about if the response to his query was “Hector Berlioz was addicted to opium.” Similarly, in the context of verbal deception, in most cases the speaker cannot deceive by uttering some completely random false remark. If, upon coming home very late from his office, a man’s wife explicitly (or implicitly) seeks to know why he is so late, and if the man has reasons for wishing to deceive her as to why by asserting something he knows to be false, we must assume that one of his goals is to respond with something that he believes will satisfy her as an answer. Blurting out “The square root of eight is seven,” while false, would usually not suffice in this regard<sup>3</sup>. So in most cases of deception, speakers intend their remarks to be relevant relative to whatever they take to be the common ground at the time, and insofar as what they say is in response to some actual or assumed question or challenge, they intend their remarks to satisfy their target hearers.

#### ***1.4 Grice’s Cooperative Principle***

Much of what we have discussed so far is covered in Paul Grice’s classic 1972 paper, *Logic and Conversation*. Grice proposed that when interlocutors engage in conversation they are implicitly committing themselves to abide by a contract (the Cooperative Principle) wherein their contributions to the conversation will be consistent with what Grice referred to as “the accepted purpose or direction of the talk exchange.” This principle he exemplified through four “maxims”. The first of these is the maxim of *Quantity*, which requires that the amount of information a speaker offers should be neither too much nor too little. The other three are the maxims of *Quality*, meaning “be truthful,” of *Relation*, meaning “be relevant,” and the maxim of *Manner*, meaning “be clear” (which, perhaps ironically, Grice wrote of as the need to “be perspicuous”). Grice viewed these maxims as comprising the background against which speakers expect normal conversations to take place, with the application of the maxims of course, always being relative to the context. Naturally, speakers frequently violate one or more these maxims, sometimes intentionally and blatantly, which Grice referred to as “flouting” and which constitutes an opting out of the Cooperative Principle. As will be discussed below, most (but not all) cases of verbal deception involve either secretly violating or explicitly flouting one or more of Grice’s maxims.

### ***1.5 Pragmatic implication***

There is a considerable literature, especially in philosophy of language, about the topic of pragmatic implication (e.g., Grant, 1958). The subtleties of the issue are not important for our purposes. For us it is sufficient to think of pragmatic implication as being a relation of implication between two propositions that depends on knowledge of the world rather than on semantics or pure logic. Suppose someone announces that they had a great view of the Alps as they flew to Venice. Many things follow from this, including the fact that they were probably in an airplane, and that the weather was probably clear. These inferences are perfectly reasonable, but they are not entailed by the assertion. They are licensed by the fact that they are the kinds of characteristics of the world that one would expect given the truth of the assertion. The relevance of this is that this relation between propositions enables us to talk about what might be reasonably inferred or concluded by the hearer from the speaker's assertion. In particular, we have to assume that whatever strategy of deception the speaker employs, a major aspiration is that the hearer does not somehow come to the conclusion that the to-be-concealed truth is in fact the truth. Thus, what the speaker asserts should neither logically entail nor pragmatically imply it—it should not suggest it.

## **2 What is Verbal Deception?**

We have already indicated that an informal account of verbal deception will not be adequate for our purposes. Accordingly, we propose the following more formal account:

verbal deception occurs when a speaker, S, expresses or intimates some proposition, Q, to a hearer, H, because there is some aspect of S's perceived world (representable by a proposition, P) with respect to which S has the intention of deceiving H. To realize this intention, S creates a belief-manipulation<sup>4</sup> goal, G, of establishing or maintaining in H a particular mental state relative to P. There are usually many strategies through which S can achieve G, but insofar as S is a boundedly-rational agent, S's choice of strategy will be determined by some sort of cost-benefit analysis strongly constrained by social-psychological considerations relating to S's relationship to H.

Of particular interest to us in this chapter are the strategies speakers use to deceive their hearers. Our ultimate goal is to identify and characterize these strategies, and to organize

them into a coherent and psychologically plausible classification scheme. But we will not only address the question of how speakers achieve their goals through verbal deception, we will also consider the different kinds, or types, of verbal deception available to them. In addressing these issues, we shall focus on two main bodies of literature, namely, work coming out of pragmatics and speech act theory on the one hand, and out of communication theory on the other. We also discuss some contributions from philosophy, but because the boundary between pragmatics conceived of as part of the philosophy of language and pragmatics conceived of as a subfield of linguistics is so indistinct, we make no sharp distinction between them.

It often happens that when different disciplines address the same general problem, both the focal questions and the methodologies they employ differ. They also often use different terminologies and different levels of analysis with the result that they produce interesting analyses that are somewhat independent of each other. In what follows, we hope to bring together these two perspectives on verbal deception. We believe that much is to be learned from both even though they sometimes have different emphases and different approaches, and we think that a more comprehensive account of verbal deception can be achieved by considering both than can be achieved by considering either alone.

The most important contribution to analyses of verbal deception from pragmatics concerns the role of the Grice's notion of Conversational Implicature (e.g., Adler, 1997; Meibauer, 2005). Implicature arises when a speaker means to imply or suggest one thing while saying something else (Grice 1975), that is, when the intended meaning differs from the literal meaning of the utterance. However, unlike the case of metaphoric language wherein the relation between the literal meaning and the intended meaning is usually at the semantic level (e.g., "the interviewer *crucified* the politician," "the president's speech went over *like a lead balloon*"), in the case of implicature the relation is usually one in which what is meant at the propositional level can somehow be surmised from the literal meaning at the pragmatic level. So the general idea is that in many cases, verbal deception is accomplished through false implicature, that is, through implicature that the speaker believes to be false, while believing the literal meaning is true. The field of pragmatics has also inspired speech act theory and artificial intelligence (e.g., BDI—Belief, Desires, Intentions) analyses of verbal deception of the kind advanced by Castelfranchi and colleagues (e.g., Castelfranchi and Poggi, 1993; Vincent and Castelfranchi, 1979). In addition, we believe

that the analysis of Chisholm and Feehan (1977) in terms of the omission-commission distinction is crucial to an understanding of the intentions underlying deception. The omission-commission distinction has to do with the fact that a person may seek to deceive either through what he or she says (commission), or through what he or she fails to say (omission).

The literature in communication theory, on the other hand, has focused on identifying dimensions on which the different types of deception strategies vary (e.g., Burgoon et al., 1996; Hopper and Bell, 1984; McCornack, 1992; Turner et al., 1975), generating a taxonomy of deception strategies (e.g., Hopper and Bell, 1984; Metts, 1989; Turner et al., 1975), and identifying the social goals for deception—goals such as not hurting one’s partner’s feelings, and avoiding conflict and unpleasant situations(e.g., Levine et al., 2002; Metts, 1989). Although basic concepts from pragmatics play a role in some of the research in communication theory, the goals and methods of the research efforts have generally been rather different.

### **3 Existing approaches to classifying verbal deception types**

More than 20 different types of verbal deception have been discussed in the different literatures, and there have been numerous proposals as to how to organize them into meaningful classification scheme. These proposals, can be roughly grouped into three classes: (1) those that have the Cooperative Principle at their core (Burgoon et al., 1996; McCornack, 1992; Metts, 1989; Turner et al., 1975), (2) those that follow a “speaker intention” approach (Bradac, 1983; Chisholm and Feehan, 1977; Vincent and Castelfranchi, 1979), and (3) those based on attempts to discover the underlying dimensions of deception (e.g., Hopper and Bell, 1984).

#### ***3.1 Cooperative Principle approaches***

As already mentioned, Grice’s Cooperative Principle and its four maxims of Quality, Quantity, Relevance, and Manner, specify how interlocutors in conversation normally subscribe to and comply with an implicit contract to be cooperative. The four maxims, which constitute the core of the Cooperative Principle, are, as expressed by Grice: (1) be as informative

as is required for the current purpose of the conversation, and do not be more informative than is required (the maxim of Quantity), (2) do not say what you believe to be false, do not say anything for which you lack adequate evidence (Quality), (3) make your contribution relevant to the conversation/context (Relevance), and (4) do not be ambiguous, be brief and orderly (Manner).

Three related strands of research in Communication Theory exemplify this approach, starting with Turner et al.'s (1975) tripartite taxonomy of deception strategies into Distortion, Concealment, and Diversionary responses. Subsequently, McCornack (1992) extended Turner et al.'s idea and, relating it to the Cooperative Principle, proposed the *Information Manipulation Theory* (IMT). The third strand is Burgoon et al.'s (1996) *Interpersonal Deception Theory* (IDT), which goes beyond IMT by incorporating a construct they called Personalization.

Turner et al. (1975), took as their starting point the basic notions of distortion and concealment—two general forms of what they called “information control,” namely, different ways in which information can be manipulated. They had subjects recall and record a conversation with someone close to them in which “important matters were at issue or stake”. Subjects were then asked to mark their own utterances in the conversation as either completely honest or not completely honest. The experimenters were particularly interested in two types of responses namely those that their respondents actually reported saying during the conversation, and those that respondents claimed they would have said had they been totally honest. Results showed that subjects classified the majority (61.5 %) of their utterances as not completely honest, indicating some level of information control. A more fine-grained analysis enabled the researchers to further classify Distortion and Concealment into sub-types (e.g., lies, exaggerations, half-truths, and secrets). They also found a third form of information control that they referred to as Diversionary responses. From these results, Turner et al. (1975) identified the following seven types of deception:

1. Lies: what respondents said contradicted what they would have said had they been totally honest,
2. Exaggerations (type 1): respondents gave more information they would have given had they been totally honest,
3. Exaggerations (type 2): respondents overstated something by using superlative modifiers that they wouldn't have used had they been totally honest,

4. Half-truths (type 1): respondents disclosed only part of what they considered to be a complete disclosure,
5. Half-truths (type 2): respondents overstated something by using modifiers that gave less than an honest evaluation of the situation,
6. Secrets: respondents remained silent when in fact they had something to say,
7. Diversionary responses: what respondents said was different from or irrelevant to what they would have said had they been totally honest.

While Turner et al. (1975) suggested three forms of information control—Concealment, Distortion and Diversionary responses—all having to do with message *content*, Bavelas et al. (1990), appealing to the case of equivocation, argued that deceptive messages can also be generated by manipulating the *manner* in which the information is presented. McCornack (1992) followed Bowers et al.'s (1977) suggestion that “deviant” messages mislead by violating Gricean maxims, and concluded that Turner et al.'s (1975) categories of Distortion, Concealment, and Diversionary responses, as well as Bavelas et al.'s (1990) Equivocation are directly related to Grice's maxims of Quantity, Quality, Relevance, and Manner. For this reason, McCornack (1992) proposed his *Information Manipulation Theory* (IMT), according to which deceptive messages are constructed by violating one or more of the Grice's maxims. He tested this idea by asking subjects to generate messages in response to three specific situations, and then analyzing their responses in terms of IMT. These situations were generated in another experiment in which subjects were asked to write a detailed description of a situation in which they had verbally deceived a dating partner, including a description of the relationship and the details of the conversation. McCornack also reviewed taxonomies that had been proposed by other people and showed that they could easily be accommodated by IMT.

McCornack (1992) was highly critical of “taxonomy or strategy-based” approaches to studying deceptive messages. He argued that such approaches lead to inconsistent systems of categorization, and that examples of deceptive messages presented in such approaches are too often generated using recall methods as in, for instance, Metts (1989), with the result that the examples are not representative of deception in everyday communication. For these reasons, McCornack advocated a “dimensional” approach in terms of which deceptive messages can be conceptualized as “a potentially infinite class of specific message forms resulting from the manipulation of information in particular, characteristic ways” (p. 3). Nevertheless, McCornack realized that there remained a class of deceptive messages not

covered in his analysis, namely, messages involving deception by implication, and the fact that there exists this form of deception, he stated, “suggests a realm of deceptive messages beyond the immediate scope of IMT: messages that mislead not through the manipulation of information, but through the generation of deceptive implicatures” (p. 14). This notion of deceiving by falsely implicating is discussed at length by Adler (1997). Meibauer (2005), motivated by the work of Adler and others, presented a detailed speech act analysis of deceptive false implicatures within the framework of Grice’s Cooperative Principle.

A variant of McCornack’s (1992) IMT was proposed by Burgoon et al. (1996) who postulated five dimensions, namely, Veridicality, Completeness, Directness/Relevance, Clarity, and Personalization. Of these, four are essentially the same as McCornack’s, with Veridicality being related to the maxim of Quality, Completeness to the maxim of Quantity, Directness/Relevance to the maxim of Relevance, and Clarity being the analog of the maxim of Manner. The new, fifth, dimension of Personalization, Burgoon et al. took to represent the speaker’s own thoughts, opinions and feelings about the conveyed information. Personalization is related to the concept of “disassociation” or “verbal nonimmediacy” so that speakers can manipulate the “ownership” of their utterance by using strategies that disassociate them from the information presented. Nonimmediacy refers to the implied relationship between the speaker and the subject of the information being discussed. It shifts descriptions of events by, for example, using abstract language or generalizations instead of concrete details, and obscuring the agent of the action, opinion or belief (e.g., “It is said that”). Burgoon et al. argued that even though Personalization cannot be readily related to Grice’s Cooperative Principle, it is a background assumption of all discourse so that any violation of it could cause a hearer to be misled. According to this view, an utterance belongs to the speaker who utters it and is thus supposed to reflect the speaker’s opinion and feelings, unless explicitly stated otherwise. Burgoon et al. tested their overall theory by assessing the degree to which speakers can vary and hearers can recognize deceptive messages in terms of the five postulated dimensions. In one experiment, they engaged participants in two interviews each. In one of the interviews subjects had to be truthful, and in the other, they had to be deceptive (to a different interviewer) about the same questions. In another study, subjects were to be both truthful and deceptive during a single interview. Results, based on self-reports and observer ratings showed that all deceptive messages involved some level of information manipulation along one or more of the five dimensions of IDT. Because the models proposed by Burgoon et al. (1996) and by McCornack (1992) are di-

mensional models, any particular case of verbal deception can be thought of as one of an infinite number of possible points in a multi-dimensional space.

Finally, although not explicitly cast in terms of the Cooperative Principle, there is other work that can also be seen in this light. For example, in her analysis of deception in close relationships, Metts (1989) employed as her starting point three main types of deception: Falsification, Omission, and Distortion. She argued that all forms of deception proposed in the literature fall on a continuum with Falsification at one end, Omission at the other, and with the region around the midpoint representing various types of Distortion such as Exaggeration, Minimization, and Evasion. It is not difficult to conceptualize her notion of Falsification as a violation of the maxim of Quality, and her notion of Omission as a violation of Quantity. Finally, Metts's Distortion could be viewed as a combination of violations of Manner and of Relevance (since she also includes cases of evasion as an instance of Distortion).

### ***3.2 Speaker intentions approaches***

Cooperative Principle approaches to classifying different types of verbal deception focus on what a speaker *says*. An alternative approach is to focus on what the speaker *intends*. Viewed in this way, three apparently quite disparate lines of work (e.g., Bradac, 1983; Chisholm and Feehan, 1977; Vincent and Castelfranchi, 1979) can be grouped together.

Chisholm and Feehan's (1977) insightful analysis of deception revolved around the important distinction between deception by commission and deception by omission. These authors defined deception by commission as an act through which a speaker, *S*, *contributes causally* to a hearer, *H*, acquiring or continuing to believe a falsehood, or to *H*'s ceasing to believe, or being prevented from acquiring a particular truth. In a similar vein (but with the emphasis now on allowing rather than causing), they defined deception by omission as an act whereby *S allows* *H* to acquire or continue to believe a falsehood, or to cease to believe or to continue without the belief of a particular truth. Thus both deception by commission and by omission admit of four cases, namely, (1) *H*'s acquiring, or (2) continuing to believe, a falsehood, and preventing (3) *H* from acquiring, or (4) ceasing to believe, a truth. We worry a little about Chisholm and Feehan's account of deception by omission because their notion of allowing seems to require that *S* could have prevented *H* from embracing or holding onto the mistaken belief in question, but did not. However, this means that *S* must

have engaged in a mental act, for example, the act of deciding or choosing not to say something, in which case Chisholm and Feehan's notion of allowing means that S did in fact contribute causally to H's state of belief. And if this is so, then the difference between deception by commission and deception by omission lies not in the difference between contributing causally and allowing (because ultimately allowing results in contributing causally), but rather in the kind of act (communicative vs. mental) that contributed to H's belief state.

However, even though we have reservations about Chisholm and Feehan's criterion for distinguishing deception by omission from deception by commission, we believe that their identification of eight different forms of deception is an important and valuable contribution, and one that has significantly influenced our own approach to verbal deception. As we will discuss in more detail later, we take their two sets of four deception types and collapse them into one set, the elements of which we refer to as "belief manipulation goals" which, we argue, can be achieved by various deception "strategies." For example, one such goal is to get H to acquire a belief about something that S believes is false. Such a goal can be achieved by several different strategies, including fabricating a false story, or overstating, using excessively abstract language.

Bradac (1983) studied the relationship between language and social relationships, and their mutual entailments, maintaining that humans hold three kinds of beliefs: beliefs about the world, beliefs about beliefs, and beliefs about the consequences of their utterances. He further argued that the intersection of these beliefs with intentionality gives rise to several discrete types of communicative phenomena (Bowers and Bradac, 1982; Scott, 1977). Bradac proposed three strategies for verbal deception: Lies, Secrets, and Evasions. These, together with true revelations, he considered to be related but distinct communicative acts. This account, although explicitly considering the role of intention, is a very broad and general approach to communication, which, while accommodating deception, is not specific to it. It thus fails to help with respect to deception strategies, which is an issue of primary concern to us.

Some of the most important work on verbal deception that takes into account the role of speaker intentions is that of Vincent and Castelfranchi (1979) who analyzed verbal deception, and in particular "indirect lying or lying while telling the truth", in terms of agent assumptions and beliefs, intentions and goals. By indirect lies, they meant deceptive messages in which the utterance itself is believed to be true by S, but through that true utterance, S

has the goal of causing H to believe that something is true, when in fact S believes that it is false. Taking this goal-based analysis approach as a heuristic, Vincent and Castelfranchi generate a taxonomy of eight communicative strategies through which indirect deception can be achieved. However, they also maintain that these strategies can be used for communicative purposes other than indirect deception. The eight strategies<sup>5</sup> they identify are:

1. **Insinuation:** an indirect lie wherein S wants H to believe what is implied by S's utterance while S believes that what is implied is false. Consider an example given by the authors, where a child on his first day at a new school says to his classmate, "My dad works at the BBC," intending H to infer that his father has a prestigious job, when the truth is that his father works there as a janitor.
2. **Reticence:** S leaves something (that S believes is false) unsaid and wants H to understand (1) what that unsaid thing is from the context, (2) that it is intentionally being left unsaid, and (3) the goal behind leaving it unsaid. For example, consider a situation in which, when H says to S, speaking about some third person, John, "I hope John doesn't drink too much", S wants (out of spite, perhaps) to falsely suggest that John does indeed drink excessively without actually saying so. Under these conditions, S, knowing that John is Irish, might exploit what he takes to be a stereotype by saying "Well, he *is* Irish!"
3. **Half-truth:** S tells only part of the truth and conceals another part that S believes is relevant to H, thereby getting H to make some false assumptions. For example, when a child only admits that he punched his little sister, when the truth is that he beat her up<sup>6</sup>.
4. **Precondition or Presupposition faking:** S says something that S believes is true that presupposes something that S believes is false. Consider Vincent and Castelfranchi's (1979) example of a guest at a wedding saying to another, "It's a pity that Anne and Mark had to be at Burleigh this weekend." Here there could be several false presuppositions including that Anne and Mark had been invited to the wedding, that S knows Anne and Mark very well, and that they would have come had they not had to address some important issue at Burleigh.
5. **Deliberate ambiguity (or equivocation):** S says something that has two interpretations, where S believes that one of them is true and the other is false, and wants H to infer the false one. The authors give Bolinger's (1973) example of an advertisement which says, "No heat costs less than oil heat," where one (true) interpre-

tation is that not using any heating will cost less than heating using oil, and the other (questionable, if not outright false) interpretation is that oil-heating is cheaper than any other kind of heating.

6. Obfuscation: S says something that S believes is true but believes that H will not be able to fully understand it and will therefore be deceived, as might happen when S intentionally uses technical jargon or euphemisms to H that S believes H is unlikely to understand.
7. Pretending to lie: S says something that S believes is true but wants H to believe that it is false. Vincent and Castelfranchi (1979) give the example of a burglar, who, when caught by the police and interrogated about his partner's whereabouts, tells them the correct location, hoping that the police won't believe him and will therefore look in the wrong places, while at the same time insulating himself from charges of lying.
8. Pretending to act or joke: S says something that S believes is true as a joke, and wants H to believe that S believes that it is not true, thereby pretending solidarity and complicity.

Although their account is restricted to cases in which the surface form of the utterance is believed to be true by the speaker, Vincent and Castelfranchi's (1979) approach provides a deep account of how speaker's assumptions, intentions and goals interact with different types of linguistic actions to generate deception. We find this kind of detailed speech act analysis (see also, Castelfranchi and Poggi, 1993) most appealing, comprising as it does a BDI theory of action (i.e., a theory based on beliefs, desires, and intentions). The analysis that we shall shortly present is in the same spirit, although it addresses a broader range of deception strategies to include cases in which the surface form of the utterance is believed to be false by the speaker.

### ***3.3 A Multidimensional scaling approach***

Hopper and Bell (1984), concerned that too much of the literature on deception in communication focused on lies, undertook to present what they considered to be a broader and more empirically-based conceptualization of deceptive communication. Accordingly, they set out to identify the dimensions underlying deception in general, on the basis of which to propose a typology of deception strategies. To do this, they compiled a list of 120 English

words and idioms that people use to talk about various aspects of deception and then had judges select those deemed to be “most central to deception,” of which 46 were identified. These were then used as stimuli for similarity judgments, the data from which were subjected to multidimensional scaling. The results suggested three underlying dimensions of Evaluation (right or wrong), Detectability (easy or hard to detect), and Premeditated (planned or unplanned). Then, on the basis of a cluster analysis, the authors proposed a typology of six deception strategies:

1. Fictions (comprising lexical items such as “make-believe,” “exaggeration,” “myth,” and “white lie”): aspects of messages that make them imaginative or counter-factual,
2. Playing (e.g., “joke,” “tease,” and “kidding”): deceptions perpetrated for the purpose of amusement,
3. Lies (e.g., “dishonesty,” and “lie”): false verbal statements intended to deceive,
4. Crimes (e.g., “con,” “conspiracy,” and “entrapment”): acts that are prescribed as crimes in the criminal justice system,
5. Masks (e.g., “hypocrisy,” “evasion,” and “concealment”): deception by obscuring the truth, and
6. Unlies (e.g., “distortion,” “false implication,” and “misrepresentation”): deception through implication.

Hopper and Bell’s (1984) analysis is interesting and suggestive, but because it deals with all forms of deception (not just verbal deception), we think that it is perhaps too course grained for our purposes. Many of the words they used on the grounds that they were judged to be “most central to deception” (e.g., “conspiracy,” “forgery”) are not readily applicable to verbal deception, if applicable at all. We thus cannot assume that these same (and only) three dimensions are applicable to *verbal* deception. Furthermore, because some of their categories have to do with speakers’ intentions, some with properties of (speech) acts, and some with reactions of hearers, their classification scheme lacks the coherence that one might hope for of a comprehensive account of verbal deception. For example, categories such as “playings” and crimes pertain to the social (or antisocial) goals that are achieved by deception, with the result that different kinds of deceptive strategies end up in the same category. For example, exaggeration, irony and make-believe are grouped together into fictions, whereas from the perspective of, for instance, speakers’ communicative goals, these rhetorical devices are really quite dissimilar.

### 3.4 Summary

The three approaches discussed in this section are summarized in Table 1. Since verbal deception is heavily dependent upon context, a pragmatics perspective is obviously very relevant. Grice's (1975) Cooperative Principle lies at the heart of this perspective and it provides a good theoretical basis for systematically addressing the problem of verbal deception. However, its strength lies in its ability to furnish explanations of the properties of deceptive messages, rather than providing insight into the functional aspect of verbal deception. In other words, it focuses on the message rather than on the speaker. Moreover, the framework it provides is rather vague and by itself does not provide enough support for further exploration. On the other hand, the speaker intention approach is more functional in nature. It is also potentially useful for modeling purposes because it allows one to explore how agents might reason with their beliefs, intentions and their goals to generate deceptive communicative behavior. In what follows, we shall attempt to exploit the advantages of both approaches with a view to providing a comprehensive model of verbal deception.

**Table 1:** Criteria for classifying verbal deception types by author(s)

Underlying Criterion		Classification	
Cooperative Principle	<i>Turner et al., 1975</i> Three forms of Information Control: Distortion, Concealment, and Diversionary responses	(1) Lies (2) Exaggeration (Type 1) (3) Exaggeration (Type 2) (4) Half-truth (Type 1)	(5) Half-truth (Type 2) (6) Secrets (7) Diversionary responses
	<i>Metts, 1989</i> Continuum with Omission at one end and Falsification at the other, with the central region being Distortion	(1) Omission (2) Distortion (3) Falsification	
	<i>McCornack, 1992</i> Four maxims of Quality, Quantity, Relevance, and Manner	“a potentially infinite class of specific message forms resulting from the manipulation of information in particular, characteristic ways”	
	<i>Burgoon et al., 1996</i> Four maxims of Quality, Quantity, Relevance and Manner, plus a fifth dimension of Personalization	“many possibilities [with] contrasting combinations of the information dimension qualities”	

Speaker intentions	<i>Chisholm &amp; Feehan, 1977</i> Omission/Commission dichotomy	(1) Commission (S contributes causally to): H's acquiring or continuing to believe a falsehood; H's being prevented from acquiring, or ceasing to believe a truth.	(2) Omission (S allows H to): acquire or continue to believe a falsehood, cease to believe or continue without the belief of a truth.
	<i>Vincent &amp; Castelfranchi, 1981</i> goal analysis model in terms of agent assumptions, intentions and goals	(1) Insinuation (2) Reticence (3) Half-truths (4) Precondition or Presupposition faking (6) Obfuscation	(6) Deliberate ambiguity (7) Pretending to lie (8) Pretending to act or joke
	<i>Bradac, 1983</i> Intersection of beliefs and intentions suggests several types of related but distinct communicative acts	(1) Lying (2) Secrecy (3) Evasion	
Multi-dimensional scaling	<i>Hopper &amp; Bell, 1984</i> Multidimensional scaling of lexical items revealing dimensions of: Evaluation, Detectability, and Premeditation	(1) Fictions (2) Playing (3) Lies	(4) Crimes (5) Masks (6) Unlies

## 4 A Comprehensive Approach to Analyzing Verbal Deception

Having reviewed several proposals from the literature as to how to organize the different types of verbal deception, we now present our own proposals for a comprehensive, integrated account. Our account comprises twelve basic types.

### 4.1 Specification of Verbal Deception Types

1. **Fabrication:** By fabrication, we mean an outright lie wherein S simply makes up a false story in order to deceive H. We distinguish two sub-types of fabrication: Simple Fabrication, and Fabrication for False Inference.

In *Simple Fabrication* S makes up a story that he or she wants H to believe. For example, S is a mean-spirited student and wants to deceive H during a test. S knows that

H needs to know the capital of Turkey, and knowing that it is not Istanbul, he tells H that the capital of Turkey is Istanbul.

In *Fabrication for False Inference*, S intends to deceive H primarily with respect to what the fabricated story implies. Unlike the case of Simple Fabrication, the fabricated story here is a means to an end rather than an end in itself. For example, Steve, S, is jealous of Harry, H, because he, Steve, is very attracted to Harry's girlfriend. Steve has no reason to believe anything bad about Harry, and has never seen Harry in the company of disreputable people. Nevertheless, Steve tells Harry's girlfriend that he often sees Harry hanging out with some really bad people, thereby intending to imply that Harry can't be trusted. It is this implication of the fabrication that is the nexus of the deceit. Fabrication for False Inference also covers cases of false presuppositions. For example, a child, H, who doesn't know about Santa Claus, asks his mother, S, on Christmas morning where the presents under the Christmas tree came from. The mother, wanting her child to believe there is a Santa Claus says, "Santa Claus left them last night while we were all sleeping," which, as well as being a fabrication, presupposes that there exists a Santa Claus. In this case, S deceives H with respect not only to what she says, but also with respect to the presupposition of what she says, which is the primary goal of her utterance.

2. **False Implicature:** As already discussed, the term "implicature" refers to the implied or suggested meaning of an utterance when that meaning is not (logically) entailed by the utterance. False Implicatures are conversational implicatures in which the literal meaning is true, but the implied meaning is false, a nice example of which is cited by Meibauer (2005) as a classic. It pertains to a ship's captain who every day enters into the ship's log the fact that the first mate was drunk. One day the captain is unable to complete the log, so the responsibility to do so falls to the first mate, who of course notices the daily references to his drunkenness. To retaliate against the captain, the first mate enters into the log the true statement that "Today, the captain was not drunk," intending thereby to falsely suggest that the captain's not being drunk was an unusual event. Although this example does not involve an actual speaker and an actual hearer, the logic of the example is the same as if the first mate were the speaker, S, and the captain (or indeed anybody else) was the hearer, H.

3. **Denial:** A denial is a rejection of the alleged truth of a proposition; the speaker indicates that he or she believes the proposition in question is false. We distinguish two forms of denial, namely, Simple Denial and Denial for False Inference.

In *Simple Denial* the proposition being contradicted is itself the nexus of deception. For example, a mother, H, knows full well that it was her child, S, who took all the cookies from the cookie jar. Knowing this, she says to her child “You took the cookies, and I told you not to”. The child, S, denies her assertion that it was he who took the cookies, saying, “I didn’t take them. It wasn’t me.”

In *Denial for False Inference*, S, by denying a true proposition, implies something else that is false. For example, consider a scenario in which Sara, S, is a prime witness in a murder trial. Convinced that the accused person is guilty, she testifies that she heard a gunshot. During cross-examination, the defense attorney, H, says “You said that at the time of the gun shot, you were in the bathroom. And did you not testify that you were drying your hair?” In fact, Sara was indeed blow drying her hair at that time, and was not really sure that she heard a gun shot at all. To hide her uncertainty, she contradicts the defense attorney and denies that she was drying her hair, saying, “No, I didn’t say that.” In this scenario, S wants to deceive H about her uncertainty concerning hearing the gunshot. The intended implication of her denial that she was using her hair dryer is the falsehood that she was able to clearly hear the gunshot. So the nexus of deception lies in the inference she hopes will be made from the denial rather than in the denial itself.

4. **Half-truth:** Half-truths involve revealing only part of the truth, while concealing another part that involves the nexus of deception. We have identified two sub-types of Half-truth: Simple Half-truth and Half-truth for False Inference.

In *Simple Half-truth*, the proposition being concealed is itself the nexus of deception. For example, a man, S, gets home late because he had a long meeting running into the early evening, and so left work much later than usual. Then, he went to see his girlfriend. When he gets home, his wife asks him why he is so late. The man doesn’t want his wife to learn the truth, so he tells her "We had a long meeting that went on into the evening." In this example, what S says is the truth, but it is not the whole truth; it is a Half-truth.

In *Half-truth for False Inference*, S, by revealing only part of the truth and concealing a true proposition, implies something else that is false. For example, a small girl, S,

instead of finishing her homework, went out to play with her friends. It was chilly and damp outside, and when she got back home, she began to feel feverish. The next day, her teacher, H, asks her why her homework wasn't finished. Wanting H to believe that it wasn't her fault, S says, "I got sick yesterday, that's why." What S says is the truth, but it is only a Half-truth, and S utters it with the intention of (falsely) implying that it wasn't her fault that she didn't finish her homework.

5. **Abstraction:** In abstraction, S frames her utterance in a way that is sufficiently general or broad that it hides the more specific proposition that S intends to conceal from H. Thus, the proposition with respect to which S wants to deceive H implies what S actually says, but is not implied by it, while what S actually says is true. There are two sub-types of Abstraction: Simple Abstraction and Abstraction for False Inference.

In *Simple Abstraction*, the proposition that is generalized is the nexus of deception. For example, Sally, S, is going to her boyfriend's place at the city center, but does not want her mother, H, to know. When Sally's mother asks her where she is going, Sally, tells says that she's going to a friend's place. Thus, by generalizing from the more specific "boyfriend's place" to the more general or abstract "friend's place," S seeks to conceal from H the fact that she is going to her boyfriend's place.

In *Abstraction for False Inference*, S, by generalizing a proposition, implies something else that is false. Suppose this same Sally, S, intends to go to the movies, but does not want her mother to know because she knows that her mother expects her to go to the library to study there for an impending exam. When Sally's mother asks, "Where are you going?" Sally says, "I'm going to the city center", hoping that her mother will (falsely) infer that she is going to the library.

There are similarities between Half-truth and Abstraction in that both conceal a proposition from H. In Half-truth, the proposition is concealed by being omitted, whereas in Abstraction it is concealed by being generalized. Since, in a sense, Abstraction also involves omission, it might be considered a special case of Half-truth. However, we keep them separate partly because they violate different Gricean maxims, as will be seen in the next section.

6. **Contrived Distraction:** In Contrived Distraction S is evasive by finding some pretext to urgently change the subject, and it is by virtue of this that S attempts to deceive H. For example, a child, H, is a firm believer in Santa Claus. His friends at school tell him that there is no Santa Claus. Troubled, he comes back home and says to his mother, S,

“My friends say that Santa Claus isn’t real.” His mother, S, wanting H to continue believing in the existence of Santa Claus, tries to distract him by saying "Never mind that for now. Show me what homework you have today."

7. **Overstatement:** Deception through overstatement arises when S exaggerates an aspect of something featured in the proposition about which she intends to deceive H. There are two sub-types of overstatement: Simple Overstatement and Overstatement for False Inference.

In *Simple overstatement* what is being overstated is itself the nexus of deception. For example, H, has invited her future mother-in-law, S, for dinner. On eating the food, the mother-in-law-to-be thinks that it is at best, average; not too good, but not too bad either. Out of politeness, she says to H, who has no idea what S thinks about the food, “This is a delicious meal.” In this example, S overstates by describing what she considers to be a so-so meal as “delicious.”

In *Overstatement for False Inference*, by overstating a proposition, S implies something else that is false. That is, the nexus of the deception is not the proposition itself (as in Simple Overstatement), but the inference that S hopes H will draw from the proposition. For example, in a radio commercial, the announcer, S, is explaining the benefits of a new cream that he claims removes pimples. In fact, he knows that the cream is not very effective and not worth buying, because it removes only a few pimples and it takes two weeks to do so. But to do his job, he says to the listeners, H, "This cream will remove most of your pimples in just two weeks. In this case, S knowingly overstates the effectiveness of the cream, and in doing so, implies (and intends to imply) that people should buy the cream if they have pimples.

8. **Understatement:** Understatement is the opposite of overstatement and involves stating something as less than it is by underspecifying an aspect of something featured in the proposition about which S wants to deceive H. Again, there are two sub-types of understatement, Simple Understatement and Understatement for False Inference.

In *Simple understatement* the proposition in which the understatement is embedded is itself the nexus of deception. For example, in order to cause listeners, H, to stop believing that the 2010 BP oil spill in the Gulf of Mexico was very serious, BP’s then CEO Tony Hayward, S, said, “The Gulf of Mexico is a very big ocean. The amount of volume [sic] of oil and dispersant we are putting into it is tiny in relation to the total water volume.” In saying this, S intended H to infer that the amount of oil and dispersant

was not significant. But in fact the amount of oil and dispersant was huge, so S was understating the quantity of oil and dispersant.

*Understatement for False Inference* occurs when, by understating a proposition, S implies something else that is false. For example, a child, S, is sick and has a high fever, but hates taking medicines. His mother, H, says "Take this medicine now and your fever will quickly go down." The child, wanting his mother to stop believing that he needs to take the medicine, says, "I don't have much fever now." By understating the degree of his fever, S implies that he now does not need to take any medicine, which is false.

9. **Augmentation:** In augmentation, S adds something gratuitous to the truth, thereby deceiving H. For example, Steven, S, is dating his secretary, Mary, but does not want another colleague, H, to know because Steven and Mary are not yet ready to make their relationship public. One day H, being suspicious, asks Steven, "Did I see you having dinner with Mary last night?" In order to prevent H from coming to believe the truth, S replies, "Yes you did; we used it as an opportunity to discuss some important project issues," where "...it was a good opportunity to discuss some important project issues" is the augmented part.
10. **Equivocation:** In cases of equivocation, S deliberately says something ambiguous so as to avoid being committed to the interpretation that is true. Consider a situation in which a professor, H, is considering hiring John, an ex-student of a colleague, Smith, from a different institution. The professor asks Smith whether he would recommend John. Smith wants the professor to hire John because, even though John is not very bright, Smith is very fond of him and knows that John desperately needs a job. In order to prevent the professor from coming to believe the truth that John is not very bright, while also wanting to free himself of a real commitment, Smith, S, says, "I can't recommend John highly enough." In this way, the burden of getting to the truth is placed on the H. Does Smith mean "John is so good that there is no limit to how highly I can recommend him" or something more like "John is not good enough for me to recommend him sufficiently highly to impress you." Another example is provided by Bolinger (1973) in reporting an advertisement that says: "No heat costs less than oil heat," a line which carries the same "bipolar" ambiguity. In a sense equivocation provides S with "plausible deniability" should it be needed. More often, however, H will recognize the ambiguity and realize that the negative interpretation is called for.

11. **Obfuscation:** Obfuscation occurs when S deceives H by deliberately using confusing or complicated language in the hope that H will either be unable to understand the exact meaning or will misunderstand. Vincent and Castelfranchi's (1979) examples clarify what this means. They say that euphemisms, technical jargon, and verbal cosmetics are examples of obfuscation, and advertisers often use this strategy to deceive buyers. For example, an advertiser might use the term "coney" to refer to the type of fur on a fur-coat, hoping that people don't know what it is and so might think that it is something exotic, when in fact it simply means "rabbit fur".
12. **Pretending to Lie:** This somewhat ironic form of deception occurs when S says what he believes to be true, but hopes that H will think that he is lying. We have already discussed Vincent and Castelfranchi's example of a burglar, S, who when asked about the whereabouts of his regular partner in crime, actually tells the truth, thinking that the police won't believe him, and will thus spend more time trying to ascertain the "true" location of his partner elsewhere.

## ***4.2 New Approaches to Classifying Verbal Deception***

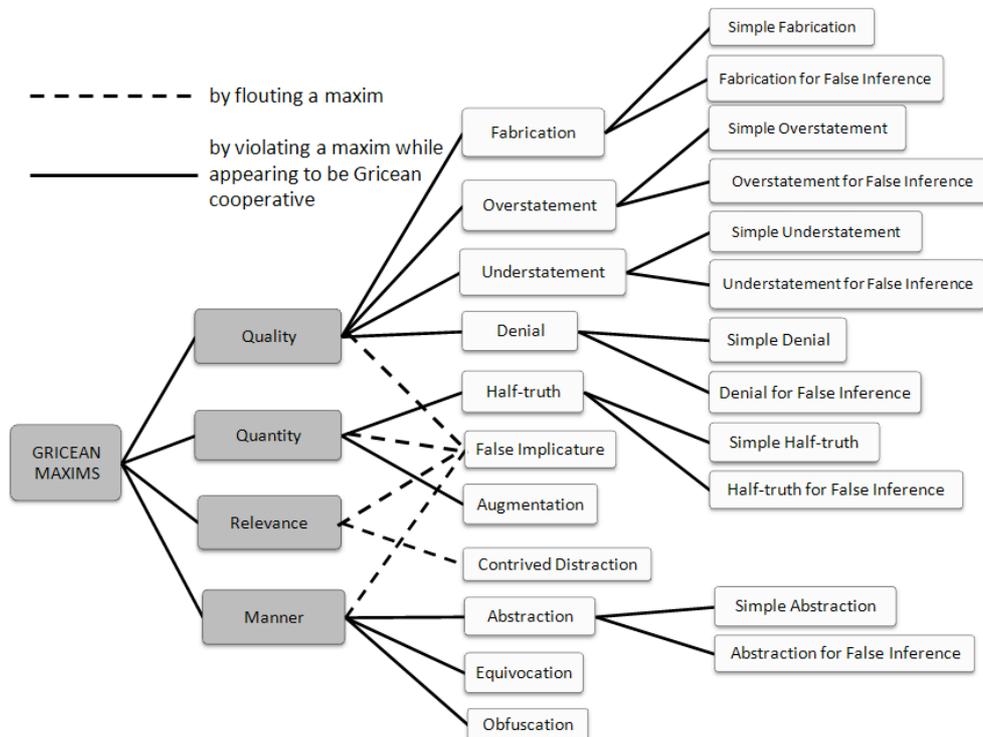
Earlier we discussed several different approaches to classifying verbal deception types, and in many respects these different approaches complement each other. Accordingly, in this section we draw upon what we consider to be the most promising aspects of them, and present two classification systems, each corresponding to one of the two major distinctions pertaining to verbal deception that we introduced earlier. Thus, we will first present a classification of deception types based on their relation to the Cooperative Principle, and then present a classification rooted in the notion of Speaker Intentions. Roughly speaking, one might say that the first of these classification systems focuses on the language of deception, while the second focuses on what the speaker is trying to do when he or she engages in an act of verbal deception. We should note, however, that what we are proposing still constitutes a work in progress and should therefore be viewed as only giving a flavor of our general approach.

### **4.2.1 A Classification based on the Cooperative Principle**

As we have seen, one of the standard ways of talking about verbal deception is in terms of Grice's Cooperative Principle, but we are unaware of any attempts to systematically organize the different types of verbal deception in terms of their relation to the four maxims that are so central to it. In this section, we try to remedy this deficiency.

When a speaker attempts to deceive another, the Cooperative Principle is of necessity implicated. One of the maxims of Quality, Quantity, Relevance and Manner is involved because speakers deceive either by (1) stealthily *violating* one of them while appearing to be Gricean cooperative, or by (2) obviously *flouting* one of them so that the hearer is aware of the violation. With the exception of three of the deception types we have identified, False Implicature, Contrived Distraction and Pretending to Lie, our deception types all fall into the first category, that is, they deceive by the non-obvious *violation* of maxims. Figure 1 shows our classification of verbal deception types based on these maxims. Dotted lines show the types that deceive by flouting a maxim and solid lines show the types that deceive by violating a maxim.

**Figure 1:** Gricean maxim violations by verbal deception strategies



We now briefly discuss how the maxims are violated or flouted by the various deception types. We should first point out that False Implicature can be generated by flouting any of the four maxims. As already discussed, conversational implicatures are generated when the hearer knows that the speaker only appears to have flouted a maxim, and so fills in the suggested meaning to make sense of the utterance; and false implicatures are simply conversational implicatures in which the literal meaning is true, but the suggested meaning is false. Any of the four maxims can be flouted to generate a discrepancy between the literal meaning and the implied or suggested meaning.

**Quality:** Of the twelve deception types that we have identified, four violate the maxim of Quality: Fabrication, Overstatement, Understatement, and Denial, including all of their sub-types. Fabrication violates Quality because it involves making up a false story, so that the surface form of the utterance itself is false. Overstatement and Understatement violate Quality because they modulate an aspect of something mentioned in the proposition, thus

affecting the truth value of what is said. Denial is a direct violation, because the speaker knowingly contradicts something believed to be true.

**Quantity:** Two of our types violate the Quantity maxim: Half-truth, and Augmentation. Half-truth violates Quantity because it conceals part of the truth, hence saying less than is required. Augmentation violates Quantity (and also Relevance, as discussed below) because it adds trivial and distracting information to the truth, hence saying more than is required.

**Relevance:** As just mentioned, Augmentation also violates Relevance, because it adds something extra to the truth that is not exactly relevant to the central issue. Meanwhile, Contrived Distraction often flouts (rather than merely violating) the Relevance maxim. One might think that a generalization of Contrived Distraction—Evasion—would make more sense in this context, and that Evasion would also have to be regarded as flouting Relevance. However, we have not included Evasion as one of our deception types because it is not clear that being evasive in general is a case of intentional verbal deceit. On the other hand Contrived Distraction, we believe, does involve intentional deception on the speaker's part, precisely because the distraction is contrived. Furthermore, it is likely to be clear to the hearer that Relevance has been flouted, and that the speaker is trying to distract the hearer to attend to some other, apparently more important, matter. Most other kinds of evasion are transparent refusals to discuss the matter at hand and therefore are not (intended to be) deceptive.

**Manner:** The strategies that violate the Manner maxim are Equivocation, Obfuscation, and Abstraction. Equivocation and Obfuscation violate the Manner maxim because they use ambiguous or confusing language. Abstraction violates Manner because it involves using generalized language to avoid stating the issue as it is. Because Abstraction involves omission by generalization, it follows that Abstraction also violates the Quantity maxim. One might think about it in terms of 'primary' and 'secondary' violations. By definition, Abstraction primarily violates manner, but this manner violation is framed in such a way as to also violate Quantity, thereby deceiving H. So Manner is primary (and therefore shown in Figure 1) while Quantity is the secondary violation (and is therefore not shown in Figure 1).

Finally, there is the interesting case of Vincent and Castelfranchi's Pretending to Lie deception type, which the reader might have noticed, does not appear in Figure 1. With re-

spect to violating or flouting maxims, the principle governing this deception type is the converse of the principle governing all of the others. This is because in case of Pretending to Lie, the speaker hopes that the hearer will wrongly believe that in fact he or she is *not* being cooperative and is attempting to violate the Quality maxim (i.e., to not tell the truth). This is in contrast to the other deception types, in which the speaker, while violating or flouting a maxim, assumes that the hearer thinks that he or she *is* being Gricean cooperative.

#### 4.2.2 A Classification based on Speaker Intentions

When a person decides to deceive another, he or she always has a purpose or goal. From the perspective of the communicative act, the speaker is always trying to do something with respect to the hearer. For example, and perhaps most typically, the deceptive speaker has the goal of getting the hearer to believe something that is not true. We think that the most informative and psychologically descriptive way in which to systematically organize the different types of deception is to view them as strategies that speakers use to attain goals of this kind with respect to their hearers. As already indicated, we refer to the goals that a speaker has with respect to the hearer's beliefs as the speaker's belief-manipulation goals for the hearer. Through such goals, speakers seek to establish or maintain in their hearers a particular mental state relative to the propositions with respect to which they intend to deceive their hearers. Partly inspired by the work of Chisholm and Feehan (1977), we propose four belief-manipulation goals of verbal deception that speakers use, either independently or in combination, to deceive their hearers. We refer to these four belief-manipulation goals as the goals to Acquire, to Continue, to Cease, and to Prevent, and their basic characterizations are as follows:

- Acquire: S wants H to come to believe something is true when S believes it is false,
- Continue: S wants H to continue believing something is true when S believes it is false,
- Cease: S wants H to stop believing something that S believes is true,
- Prevent: S wants to stop H from coming to believe something that S believes is true.

All of these goals assume that the speaker has some initial belief about the hearer's belief state (i.e., about what the hearer does or does not know or believe), and that the purpose of the deception is for the speaker to somehow influence the hearer's belief state. Thus, we

can think of deception strategies as “actions” that attempt to convert the speaker’s initial belief state about the hearer’s beliefs into a target belief state, as shown in Table 2. In the representations that follow, “bel (A, true (P))” and “bel (A, false (P))” should be taken to mean that individual A believes (and this can include believing correctly or incorrectly) that the proposition, P, is true or false, respectively. Meanwhile, we use “suggest (Qc, true (P))” to indicate that S, the speaker, believes that the proposition Qc might suggest or even warrant the inference that the proposition, P, is true or (in the case of “suggest (Qc, false (P))” false. Finally, it is important to emphasize that an understanding of speakers’ verbal deception strategies does not require a determination of whether or not those strategies are successful. To be sure, rational speakers will expect, or at least, hope, that their deception strategies will succeed, but that does not mean that we need to know whether their hearer’s actually are deceived.

Table 2: Initial Belief State and the Target Belief State for each of four belief manipulation goals.

P represents the proposition whose truth or falsity is the nexus of the deception, Q represents the proposition that S expresses in order to influence H's belief about P (i.e., what S actually says), and Qc represents the Gricean-cooperative proposition that S would express instead of Q if S were not deceiving H (what S "ought" to say). Note that all of the conditions shown pertain to S’s beliefs.

	Initial Belief State (of S)	Action	Target Belief State (of S vis à vis H)
Acquire	(1) false (P) (2) $\neg(\text{bel (H, true (P))}) \wedge \neg(\text{bel (H, false (P))})$	assert (S, Q)	bel (H, true (P))
Continue	(1) false (P) (2) bel (H, true (P)) (3) suggest (Qc, false (P))		bel (H, true (P))
Cease	(1) true (P) (2) bel (H, true (P))		$\neg\text{bel (H, true (P))}$
Prevent	(1) true (P) (2) $\neg(\text{bel (H, true (P))}) \wedge \neg(\text{bel (H, false (P))})$ (3) suggest (Qc, true (P))		$\neg(\text{bel (H, true (P))}) \wedge \neg(\text{bel (H, false (P))})$

We will now discuss the four belief manipulation goals in more detail, first in a somewhat formal manner, and then with the help of examples. As can be seen from Table 2, in both the Acquire and Continue goals, S believes that P is false but also believes that by as-

serting Q, H will come to, or continue to believe that P is true. The difference between the Acquire and Continue goals is that in case of the Acquire goal, S believes that H has no prior belief pertaining to P, that is, that H neither believes that P is true, nor that P is false, and S intends H to acquire the new (false) belief that P. In contrast, for the Continue goal, S believes that H already believes P (which S believes is false), and wants H to continue believing P. In this case, H's initial belief state is already the desired belief state, but the situation is such that S needs to say something to maintain the status quo with respect to H's belief. In particular, we assume that the situation precludes H from asserting  $Q_c$  (a non-deceptive and Gricean cooperative utterance) because doing so would either imply or be equivalent to acknowledging that P is indeed false.

In the Cease goal, S believes that P is true and believes that H also believes that P is true. By asserting Q, S wants H to cease believing that P is true, the minimal condition required to achieve this goal. Finally, in the Prevent goal, S again believes that P is true, but S also believes that H has no belief pertaining to P. Here, S wants to prevent H from coming to believe P and thus asserts Q which S believes is incompatible with and will not suggest P. As in the case of Continue, S's need to assert Q arises from the particular circumstances of the situation.

Since Acquire and Cease goals have to do with changing the initial belief state of the hearer, and Continue and Prevent on the other hand involve maintaining it, we can characterize Acquire and Cease as belief-changing goals, and Continue and Prevent as belief-maintenance goals.

Now, to see how verbal deception types can be used as strategies for converting S's initial belief state vis à vis H's beliefs into S's desired target state for H, we present an example for each of the four goals.

### **The Acquire Goal**

To illustrate the Acquire goal, we will again use the example of the mean-spirited student, S, seeking to mislead another student, H, about the capital of Turkey. The mean student knows that the capital of Turkey is Ankara, not Istanbul, but wanting his victim to acquire a false belief, says, "Istanbul is the capital of Turkey". This is an example of the deception type Simple Fabrication being used as a strategy for the Acquire goal, a strategy that, in this example, has the following features:

- (a) P, the false proposition that Istanbul is the capital of Turkey, is the proposition that S wants H to *acquire*
- (b) Q expresses this same false proposition and is what S, assuming that H does not know that Q is false, actually communicates to H
- (c) Thus, Q is equivalent to P

So, S attempts to achieve the Acquire goal of inducing a false belief, P, in H, by asserting Q, which in this example, is a Simple Fabrication intended to convert the initial state of H having no belief pertaining to P, to S's target state for H of having H believe that P is true. The table below shows Simple Fabrication as a strategic action for converting the initial belief state of the Acquire goal to its target belief state:

Initial Belief State (of S)	Action: assert (S, Q) Q = Simple Fabrication	Target Belief State (of S vis à vis H)
(1) false (P) (2) $\neg(\text{bel}(H, \text{true}(P))) \wedge$ $\neg(\text{bel}(H, \text{false}(P)))$	(1) false (Q) (2) Q=P	bel (H, true (P))

### The Continue goal

A father, H, thinks that his daughter Susan, S, is a well-behaved girl and never does anything wrong. Susan knows that her father holds this belief about her, but in fact she often behaves badly when he is not around. One day, Susan breaks a priceless antique vase. Her father asks her what happened to it. Susan, wanting her father to continue thinking that she is a good girl, says, "The cat broke it." This is an example of the Continue goal being achieved using the Fabrication for False Inference strategy because Susan intends to falsely imply that (because she was not responsible for breaking the vase) she is indeed a well-behaved girl. So, we have the following features for the Fabrication for False Inference strategy for the Continue goal:

- (a) P is the proposition that S considers false but believes that H considers true, namely, that S is a well-behaved girl. S wants H to continue believing P,
- (b) Q expresses another proposition that S also knows to be false, namely, that the cat broke the vase, but which S nevertheless communicates to H,

(c) S believes that H has no belief pertaining to Q and that it is therefore new information for H,

(d) S believes that Q does not suggest the falsification of P to H (i.e., will not cause H to surmise or conclude that P is false)

Note that in this example, if S were not deceiving H, the Gricean-cooperative proposition (Qc) that she would express would be that it was she, S, who broke the vase. However, S knows that expressing this might prevent her from achieving the Continue goal because H would no longer believe that she is a trouble-free girl. So S attempts to achieve the goal of causing H to continue believing P (which S knows is false) by asserting Q, a Fabrication for False Inference in this case, through which S can maintain the initial state of H believing P is true. The table below shows Fabrication for False Inference as a strategy for attaining the Continue goal whereby S can maintain her initial belief state as the target belief state:

Initial Belief State (of S)	Action: assert (S,Q) Q = Fabrication for False Inference	Target Belief State (of S vis à vis H)
(1) false (P)	(1) false (Q)	bel (H, true (P))
(2) bel (H, true (P))	(2) $\neg(\text{bel (H, true (Q))}) \wedge$	
(3) suggest( Qc, false (P))	$\neg(\text{bel (H, false (Q))})$	
	(3) $\neg\text{suggest (Q, false (P))}$	

### The Cease goal

In an example discussed earlier, a mother, H, knows that it was her six year-old son, S, who emptied the cookie jar of all the cookies. Knowing this, she says to her son “You took the cookies, and I told you not to.” The boy, wanting his mother to cease believing that it was he who took the cookies, contradicts her, saying “I didn’t take them. It wasn’t me.” Here we have a case of the Cease goal being attained using the strategy of Simple Denial. Notice, importantly, that several other strategies were available to the boy. For instance, as in the example of the Continue goal discussed above, he could have used Fabrication for False Inference by asserting that someone else took the cookies (thereby inviting the inference that it wasn’t he who took them). The Simple Denial strategy for the Cease goal has the following features:

- (a) P is the proposition that S considers to be true and believes that H also considers true. S wants H to cease believing P, that is, to stop believing that it was he, the boy, who took the cookies,
- (b) Q expresses the proposition that S considers false and is what S actually says to H,
- (c) Q is the negation of P

In this example, S attempts to achieve the Cease goal of causing H to stop believing P by asserting Q, a Simple Denial intended to convert the initial state of H believing P is true, to the target state of H believing P is false. The table below shows Simple Denial as an action to convert the initial belief state of the cease goal to its target belief state:

Initial Belief State(of S)	Action: assert (S,Q) Q = Simple Denial	Target Belief State (of S vis à vis H)
(1) true (P)	(1) false (Q)	$\neg(\text{bel} (H, \text{true} (P)))$
(2) bel (H, true (P))	(2) $Q = \neg P$	

### The Prevent goal

Here we again use an example, presented earlier, this time the example of the man, S, who gets home late because he had a long meeting running into the early evening, and then went to see his girlfriend. When, on getting home, his wife queries him, he tells her that he had a long meeting into the evening with the goal of preventing her from coming to believe the truth. This is an example of Simple Half-truth being used as a strategy for the Prevent goal, and it has the following features:

- (a) P is the true proposition that S went to see his girlfriend, and about which S believes H has no belief. S wants to prevent H from coming to believe P,
- (b) Qa is the allowed-to-be revealed part of a larger proposition that is a conjunction of simple propositions, a proper subset of which, Qh, S seeks to keep hidden from H,
- (c) Q is Qa

To elaborate the second of these features, let Qc represent the Gricean-cooperative proposition that S would express instead of Q if he was not attempting to deceive H, namely the

complex proposition that he had a long meeting in the late afternoon, *and* left work later than usual, *and* then went to see his girlfriend. We can then view  $Q_c$  as consisting of two subsets of propositions,  $Q_a$ , the allowed set of propositions that S believes that he can safely divulge, and  $Q_h$ , the set of hidden propositions that S wishes to conceal (in the present case, that he went to see his girlfriend) because it either implies or is equivalent to P. So, S attempts to prevent H from coming to believe P by asserting Q, in this case a half-truth, which allows him to maintain the initial state of H having no belief pertaining to P. The table below shows Simple Half-truth as an action to convert the initial belief state of the prevent goal to its target belief state:

Initial Belief State (of S)	ACTION: assert (S,Q) Q = Half-truth	Target Belief State (of S vis à vis H)
(1) true (P)	(1) $Q_c = \{Q_a, Q_h\}$	
(2) $\neg(\text{bel}(H, \text{true}(P))) \wedge$ $\neg(\text{bel}(H, \text{false}(P)))$	(2) suggest (Qh, true (P)) (3) $\neg$ suggest (Qa, true (P))	$\neg(\text{bel}(H, \text{true}(P))) \wedge$ $\neg(\text{bel}(H, \text{false}(P)))$
(3) suggest (Qc, true (P))	(4) $Q = Q_a$	

The examples we have just presented demonstrate how different types of verbal deception can be construed as strategies or actions to achieve different belief manipulation goals, and up to this point we have presented a different deception type for each of our four belief manipulation goals. In fact, however, several different deception types can be used as strategies to attain each of the four goals. Furthermore, most deception types can be used to achieve more than one of the goals, although not all strategies are applicable to all goals.

In elaborating the Acquire goal, we used the example of Simple Fabrication. But Simple Fabrication can be used as a strategy to achieve other goals as well. For example, it can be used to achieve the Prevent goal. To show this, we can consider a scenario similar to the Half-truth example given above for the Prevent goal, in which a man, S, gets home very late one night because after work he went to visit his girlfriend. When he reaches home, his wife, H, asks him, "Why are you so late?" In order to prevent his wife from coming to believe the truth, he says, "I had to go to the library to do some research." So, we have the following features for Simple Fabrication as a strategy for achieving the Prevent goal:

- (a) P is the proposition that S considers true, namely that he went to visit his girlfriend. S believes that H has no belief pertaining to P, and wants to prevent her from coming to believe P, which is the nexus of deception,
- (b) Q expresses the proposition that S considers false and wants H to come to believe, that is, the proposition that S went to the library to do some research,
- (c) S believes that Q will not suggest P to H (i.e., will not cause H to surmise or conclude P).

Qc is the Gricean-cooperative proposition that S would express instead of Q if S were not deceiving H, that is, the proposition that he went to see his girlfriend after work. S attempts to achieve the goal of preventing H from coming to believe P by asserting Q, a simple fabrication in this case, with the intention of maintaining the initial state of H as having no belief pertaining to P. The table below shows Simple Fabrication as an action to convert the initial belief states of the Prevent goal to its target belief state:

Initial Belief State (of S)	ACTION: assert (S,Q) Q = Simple Fabrication	Target Belief State (of S vis à vis H)
(1) true (P)	(1) false (Q)	
(2) $\neg(\text{bel}(H, \text{true}(P))) \wedge$ $\neg(\text{bel}(H, \text{false}(P)))$	(2) $\neg(\text{bel}(H, \text{true}(Q))) \wedge$ $\neg(\text{bel}(H, \text{false}(Q)))$	$\neg(\text{bel}(H, \text{true}(P))) \wedge \neg(\text{bel}(H, \text{false}(P)))$
(3) suggest (Qc, true (P))	(3) $\neg$ suggest (Q, true (P))	

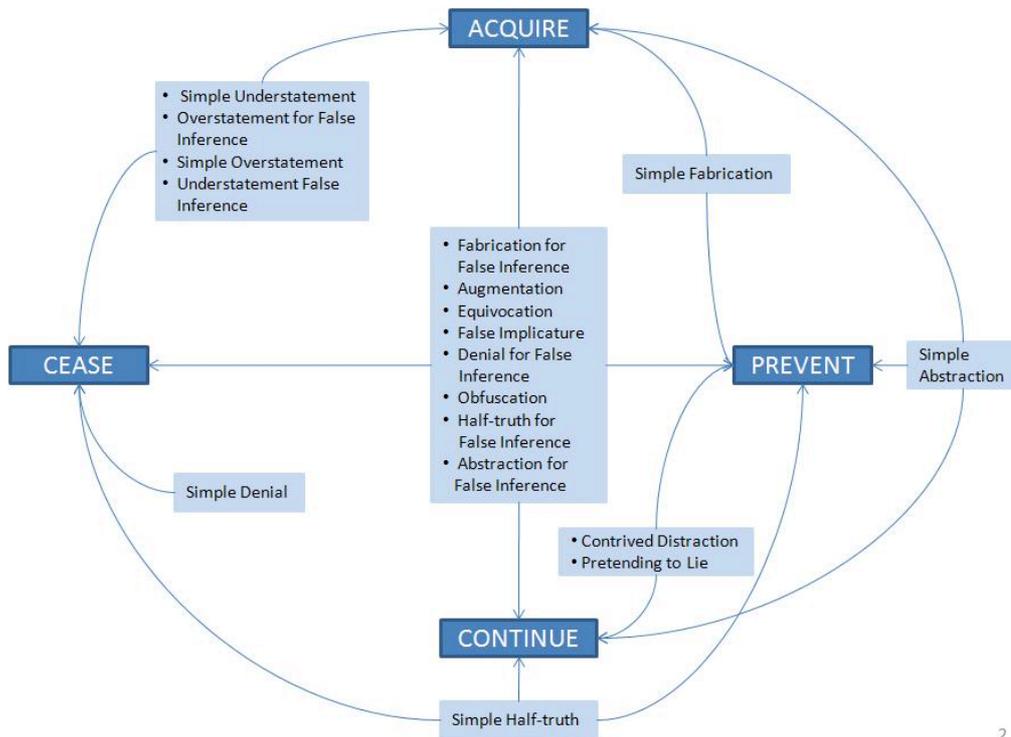
By presenting an example of how Simple Half-truth and Simple Fabrication can both be used to achieve the Prevent goal, we have demonstrated that in principle, different verbal deception types can sometimes be used as strategies for the same goal. Furthermore, we have shown how the same type (Simple Fabrication) can be used as a strategy for different goals (Achieve and Prevent). However, not all types can be used as strategies for all four goals. Consider, for example, Simple Denial. Simple Denial can be used to achieve the Cease goal, but it cannot be used for the Continue and Prevent goals. In the case of the Continue goal, Simple Denial would violate the initial condition  $\text{bel}(H, \text{true}(P))$  when in fact S believes P is false. Since H already believes that P is true when S believes that it is false, it would make no sense for S to contradict P. In the case of Prevent, Simple Denial would violate the maintenance of belief state that H has no belief pertaining to P. P is

something that S wants to hide from H. Again, it would make no sense to deny P when S wants to keep H from coming anywhere near P!

So, our general conclusions are that (a) in principle, most types of verbal deception can be used to attain more than one belief manipulation goal, (b) in principle, several different verbal deception types can be used to attain any one belief manipulation goal, and (c) not all verbal deception types can be used to attain all belief manipulation goals. To conclude, therefore, we present in Figure 2 a proposal as to which verbal deception types can be used as strategies for which belief manipulation goals.

As we have mentioned before, but feel we need to emphasize, all of this is work in progress and the connections indicated in the figure still need to be verified. We thus present them here only to illustrate our basic idea. There remains much work to be done but our hope is that what we have presented in this chapter will stimulate thought and discussion.

Figure 2: Verbal Deception Types as strategies used to achieve four belief manipulation goals



## References

- Adler, J. E. (1997). Lying, Deceiving, or Falsely Implicating. *The Journal of Philosophy*, 94(9), 435-452.
- Bavelas, J. B., Black, A., Chovil, N., & Mullett, J. (1990). *Equivocal Communication*: Sage Publications, Inc.
- Bolinger, D. (1973). Truth is a Linguistic Question. *Language*, 49(3), 539-550.
- Bowers, J. W., & Bradac, J. J. (1982). Issues in Communication Theory: A Metatheoretical Analysis. In M. Burgoon (Ed.), *Communication Yearbook 5*. New Brunswick, NJ: Transaction Books.
- Bowers, J. W., Elliott, N. D., & Desmond, R. J. (1977). Exploiting Pragmatic Rules: Devious Messages. *Human Communication Research*, 3, 235-242.
- Bradac, J. J. (1983). The Language of Lovers, Flowers, and Friends: Communicating in Social and Personal Relationships. *Journal of Language and Social Psychology*, 2, 141-162.
- Burgoon, J. K., Buller, D. B., Guerrero, L. K., Afifi, W. A., & Feldman, C. M. (1996). Interpersonal Deception: XII. Information Management Dimensions Underlying Deceptive and Truthful Messages. *Communication Monographs*, 63(1), 50-69.
- Castelfranchi, C., & Poggi, I. (1993). Lying as Pretending to Give Information. In H. Parret (Ed.), *Pretending to Communicate* (pp. 276-291): De Gruyter, Walter, Inc.
- Chisholm, R. M., & Feehan, T. D. (1977). The Intent to Deceive. *The Journal of Philosophy*, 74(3), 143-159.
- Clark, H. H., & Brennan, S. E. (1991). Grounding in Communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on Socially Shared Cognition* (pp. 127-149). Washington: APA Books.
- Grant, C. K. (1958). Pragmatic Implication. *Philosophy*, 33(127), 303-324.
- Grice, H. P. (1975). Logic and Conversation. In P. Cole, & J. L. Morgan (Eds.), *Syntax and Semantics volume 3: Speech Acts* (pp. 41-58). New York: Academic Press.
- Hopper, R., & Bell, R. A. (1984). Broadening the Deception Construct. *Quarterly Journal of Speech*, 70, 288-302.
- Levine, T., Lapinski, M., Banas, J., Wong, N., Hu, A., Endo, K., et al. (2002). Self-Construal, Self and Other Benefit, and the Generation of Deceptive Messages. *Journal of Intercultural Communication Research*, 31, 29-45.
- McCornack, S. A. (1992). Information Manipulation Theory. *Communication Monographs*, 59, 1-16.
- Meibauer, J. (2005). Lying and Falsely Implicating. *Journal of Pragmatics*, 37, 1373-1399.
- Metts, S. (1989). An Exploratory Investigation of Deception in Close Relationships. *Journal of Social and Personal Relationships*, 6, 159-179.

- Scott, R. L. (1977). Communication as an Intentional System. *Human Communication Research*, 3, 258-268.
- Turner, R. E., Edgley, C., & Olmstead, G. (1975). Information Control in Conversations: Honesty is Not Always the Best Policy. *Kansas Journal of Sociology*, 11(1), 69-89.
- Vincent, J. M., & Castelfranchi, C. On the Art of Deception: How to Lie While Saying the Truth. In H. Parret, M. Sbisà, & J. Verschueren (Eds.), *Conference on Pragmatics, Urbino, 1979* (pp. 749-778)

---

<sup>1</sup> We are grateful to Ilya Farber, Sanford Goldberg, Tei Laine, Maria Miceli, and David Pautler for helpful comments and suggestions at various stages of this work.

<sup>2</sup>A good example of this are the false denunciations in advertisements paid for by the US Swift Boat Veterans for Truth, a politically conservative group of veterans of the Vietnam War. The purpose of their advertisements was to undermine the credibility of John Kerry's candidacy in the race for the 2004 US presidency.

<sup>3</sup> As will be discussed later, there are circumstances under which the speaker might believe that such a response could suffice. For example, he might decide to say something so obviously zany as to lead his wife to believe that he had lost his mind. This, he might conjecture, would lead to an immediate changing of the subject, an avoidant strategy that we refer to as *Contrived Distraction*, and incidentally, a strategy that might even be successful using an assertion the speaker believes to be true.

<sup>4</sup> We call it a "belief-manipulation" goal because the goal pertains to H's belief or knowledge state.

<sup>5</sup>We find many of the names given to these strategies somewhat confusing because they deviate from the ordinary language meanings of those terms. For instance, the ordinary language meaning of "reticence" is simply a reluctance to disclose something, regardless of its presumed truth value, whereas in Vincent and Castelfranchi's account, the reluctance concerns material that the speaker believes to be false.

<sup>6</sup> In our own analysis, we would consider this to be a case of Understatement.