Speech Acts and Recognition of Insincerity

William C. Mann
SIL
bill_mann@sil.org

Jörn Kreutel
SemanticEdge GmbH
and Universität Potsdam
joern.kreutel@semanticedge.com

"Oh what a tangled web we weave when first we practice to deceive" – Scott[1]

Abstract

From the earliest years of speech act theory, sincerity, or the absence of it, has been one of the defining aspects of speech acts and their uses. It remains prominent today, but models of communication often give it little function. How could a model of dialogue be designed so that the sincerity of speech acts could be defined and examined? How could natural language understanding and generation programs recognize or use insincerity? Is sincerity part of a collection of speech phenomena that could share implementation methods? The issues are complex, but approachable.

What are appropriate recognition criteria for sincerity? Are the sincerity-conditions described by Austin or Searle adequate guides to recognition of insincerity? No.

Other ways of using assertions have a formal resemblance to insincere assertions. Several of these ways involve statements by a speaker who does not believe those statements. Not all of these ways involve deception. Examination of a collection of similar ways to use language leads to a much more accurate, possibly adequate, guide to recognizing the absence of sincerity.

This paper examines relationships between (in)sincerity and other language phenomena. Focusing on irony, exaggeration and understatement, it also identifies several others that share characteristics with sincerity, and thus might benefit from joint work on definitions, formalization and computational model building.

[1]Walter Scott's Marmion Canto 6,Stanza 17

Overview

Imagine a future computer system that has a strong capacity to understand, and perhaps participate in, many different kinds of natural language interaction. We would expect that part of the understanding process would focus on speech acts and rely on speech act interpretation processes. To do this, a theoretical basis would be needed, including all of the common aspects of each distinct kind of speech act. In addition to act identification and propositional content identification, the system would have to judge whether the act was sincere. This judgment is necessary because the consequences, the grounded understandings from particular speech acts are very different for acts judged to be insincere than for sincere acts.

Speech acts are defined in a way that includes a *sincerity-condition*. An act is judged insincere or sincere according to its conformity to its sincerity-condition. Correct formulation of sincerity-conditions is essential to sincerity recognition.

This paper examines sincerity-conditions as they are identified or defined in foundational work of Austin and Searle, and finds those formulations inadequate. The inadequacy has to do with improperly labeling some sincere ways of using language as insincere.

Finding those definitions inadequate, the paper makes a number of observations that appear to provide a basis for creating more adequate fresh definitions.

Background of Sincerity in Pragmatics

To some, sincerity might seem to be just a topic in psychology or sociology, but it has a long history in linguistics as well. Since the inception of speech act theory in Ordinary Language Philosophy (see (Austin, 1975)), the sincerity aspect of speech acts has been recognized. Searle reformulated Austin's conceptual scheme, again making sincerity one of the prominent concepts (Searle, 1969). The continuing development and use of speech act theory is a major theme of linguistic pragmatics, and the topic is still being developed in philosophy as well. For example, in Habermas' *Theory of Communicative Action* (Habermas, 1984), it is one of the three aspects of validity of speech acts.

Speech Act Type	Sincerity-Condition
Request	S wants H to do A
Assert, state (that), affirm	S believes P
(ask a) Question	S wants this information
Thank (for)	S feels grateful or appreciative for A
Advise	S believes A will benefit H
Warn	S believes E is not in H's best interest
Greet (on encounter)	none
Congratulate	S is pleased at E

Figure 1: Eight Speech Acts and Sincerity-conditions by Searle

More than with any other variety of speech act, people think about sincerity associated with assertions. The range of potentially insincere acts is much broader than just assertions, certainly including commissive acts (promises), and congratulations. Requests, questions, acknowledgments and various other acts also raise sincerity issues.

Austin said that for certain acts (including assertional acts) to be performed sincerely, the speaker must have "the right thoughts and feelings." Searle said that for certain insincere speech acts the speaker "purports to have (beliefs, intentions...) that he does not have." Both of these statements presume that speech acts are based on certain mental states of the speaker, and if a particular utterance is to be sincere, it will correspond to the speaker's mental state in a certain way.[2]

Part of the interest in sincerity surely comes from its involvement with deceptions (a larger category) and with lying. Another part surely arises from an episodic effort in philosophy to relate language to certitude.

Austin and Searle's defining characteristics for sincerity seem to be appropriate, but closer examination indicates that there are systematic exceptions.

According to Austin and Searle, for a statement by S, "Today is Tuesday," to be made sincerely, S must have a certain thoughts. In this case, S must believe that the day of saying the statement is Tuesday. The sincerity-condition of an assertional speech act such as this requires that the speaker believes the asserted proposition. So one of the effects of performance of the act is to communicate that the speaker believes the asserted proposition. Similarly, the sincerity-condition for commissive acts is for the speaker to intend to do what has been promised.

To bring more of the range of sincerity-conditions into view, the table in figure 1 is an extract from (Searle, 1969), p. 66-67, a table by Searle in which he defines 8 types of speech acts, with their sincerity-conditions. Clearly this is only an open, representative list. Based on the same book, we could add this:

Promise	S intends to do A

All but one of these can, according to Searle, be performed insincerely. All of the sincerity-conditions are different, but to a degree they share predicates: *believes, wants, feels grateful or appreciative, is pleased, intends to do.*

So, to recognize (in)sincerity in an interaction or written text, they present eight closely related tasks, each of which involves an assessment of the mental state of S. Below we focus on asserting.

Belief and Deception

Beside insincerity, there are other ways of using language that also involve the speaker making statements without any associated belief that those statements are true in the speaker's world of daily life. We will examine three other ways of using language that together challenge the adequacy of Searle's sincerity-condition for assertions. The three are *irony, exaggeration and understatement.* We call them *ploys.* Along with a possibly deceptive assertion, here is an example of each:

1) "I will send you the money after I get my first paycheck." - possibly deceptive assertion

2) "All of Bill Gates' influence is due to his good looks." – irony

3) "Every time the Beatles had a concert, ten million fans showed up." – exaggeration

4) "The Beatles had a few fans." - understatement

Using irony, as in example 2), involves saying something which is completely opposite to the intended meaning. The speaker expects that hearers will quickly recognize that the statement is not believed by the speaker. No deception is involved.

Use of irony violates Austin and Searle's sincerity-conditions on speech acts. Those conditions label ironic speech as insincere. Yet ironic language is usually understood for what it is. It is not insincere. Rather it draws a certain kind of attention to what is meant, and it requires the hearer to construct what is meant.

[2]It is not always the speaker whose mental state is at issue. Rather, using existing participant framework notions, especially of Levinson (Levinson, 1988) and McCawley (McCawley, 1989), we can often identify another participant in the act whose mental state is the one actually at issue. Space limitations prevent discussing this further here.

One of the important conclusions that can be drawn from comparing irony to insincerity is this:

The traditional defining conditions for insincerity are inadequate as recognition criteria for insincerity.

Recognition of insincerity cannot depend only on judging speaker's belief.

Consider exaggeration, as in 3). It is not credible that the speaker believes an exaggeration. Exaggeration is another ploy for manipulating the attention of the hearer, in this case to the scale on which the assertion depends. So it is like using a superlative such as "huge," but often with more marked effect.

Like ironies, exaggerations would be labeled as insincere by Searle's sincerity-condition. Yet they are not understood as insincere. The ploy is such that the hearer is not led to believe that the speaker believes an absurd statement. There is no deception and no appearance of insincerity. So again in this case, recognition of insincerity cannot depend only on judging speaker's belief.

Now consider understatement, as in 4). The effects of exaggeration and understatement are very similar. Both draw attention to the scale on which a degree-related assertion has its force. Unlike exaggerations, for understatements generally the speaker does believe the understatement, and more.[3] Again, there is no deception and no appearance of insincerity. These two ploys are quite similar, but their belief conditions are opposites.

Each of these ploys succeeds only if the hearer can determine, easily and with confidence, that the speaker believes the opposite (for irony), or significantly more (for understatement), or less (for exaggeration), than what is said. Ease of recognition (of the incompatible character of the assertion) is essential. None of them involves the speaker hiding his or her beliefs. When what is said is compared with speaker's belief, the ploys differ, but they all in similar ways draw attention to what the speaker is saying.

These examples together suggest that for sincerity and insincerity, attempted deception is a vital unstated element of the sincerity-condition of assertions, thus part of constituting the speech act. Similar arguments are expected to apply to commissives (promises) and perhaps to other speech acts as well.

Illocutionary Force

The same point about the inadequacy of the classical definitions of sincerity can be made by considering illocutionary force. Reconsider examples 1) through 4) above. The illocutionary force commonly assigned to assertion 1) agrees with the statement itself. The illocutionary force commonly assigned to 2) through 4) is an altered force, representing the speaker's obvious intent. It is opposite to what was said for 2) less than what was said for 3) and more than what was said for 4). In that sense they behave like kinds of indirect speech acts, and it would probably be useful to classify them as such.

In all four examples, what is said is obviously not compatible with the speaker's thought. The list of possible bases for incompatibility is open and quite diverse. It may be logical, emotional, motivational, a cultural taboo or a host of others.

In 1) that incompatible character is not obvious; in the other three, it is. This again suggests that Austin, Searle and successors had in mind images of deception when they were describing insincerity. The major role given to sincerity by Habermas (Habermas, 1984) also seems implicitly to have this character.

For the future, it might be best to make this aspect explicit, and continue to work with sincerity of speech acts as involving attempted deception.

Other Communicative Techniques with a Family Resemblance to Insincerity

One of the aims of this paper is to facilitate work on sincerity. This includes work in Ordinary Language Philosophy, in formalization of phenomena for models of language function, and development of computer-based models and agents capable of using human languages.

If the prevailing definitions are inadequate, as we claim above, then redefinition and reconception are called for. That rework is not here. When that work is done, the work might benefit from having a broad view of the interaction of language use and speakers' beliefs. In that hope, we now examine a loose collection of such uses of language.

In addition to attempted deception, irony, exaggeration and understatement, we now consider assertions that arise in pretending, play acting, written fiction, quotation, teasing, mistaken speech, forced speech, "confidence games,"[4] impersonation, deliberate misunderstanding, covert deliberate obscurity, legal representation of a client, overconfidence, politeness, outward respectful manner, and feigned ignorance. Each of these has a literature and most have some theoretical development. Many of them do not involve deception.

Figure 2 presents this arbitrarily chosen list, together with an indication of whether deception is commonly involved, and also the manner in which the speaker departs from believing the statement. The right hand column, labeled "Requires Assessing PM?" is about whether judging the sincerity of the item requires some estimation of the thoughts of the speaker (private memory, PM).

People regularly participate in these language uses, or

[3] This can depend on details of the assertion. Consider "The Beatles had only 100 fans." It could be an understatement which the speaker does not believe.

[4] One of these was the subject of the movie *The Sting*.

Language Use (ploy)	Dimension of Speaker's not Believing	Typically Deceptive?	Requires Assessing PM?
Irony	opposite belief	No	Yes
exaggeration	degree scale	No	Some cases
understatement	degree scale	No	Some cases
pretending	imagined world	No	No
play acting	imagined world	No	No
written fiction	imagined world	No	No
quotation	representing another speaker	No	No
mistaken speech	speaker commitment	No	No
legal representation of a client	role fulfillment toward a set of beliefs	No	No
forced speech	speaker commitment	not by speaker	No
"confidence games"	intended deception	Yes	Yes
impersonation	identity of speaker	Yes	Yes
covert deliberate obscurity	intention to communicate	Yes	Yes
politeness	apparent beliefs from convention	No	No
overconfidence	degree of confidence	Yes	Yes

Figure 2: Uses of Language in which the Speaker does not Believe What is Said

read about them, mostly with no extreme difficulty in deciding what is going on. (Of course, where deception is intended, we may be deceived.) In formal and computational models of natural language, there is not yet much to say.[5]

Certainly, for understanding ordinary natural language communication, the traditional characterizations of sincerity of speech acts must be supplemented. The table above may help in considering how more effective definitions might be constructed.

Use of sincerity and insincerity form the basis for some of the more complex phenomena. Because sincerity is at the base of some, modeling of sincerity can be expected to facilitate modeling more complex varieties of language.

Statements, beliefs and networks of beliefs

As all of us know by experience, effective lying does not proceed only on a statement by statement basis. We must present a view of ourselves that appears to be appropriately consistent, motivated and based on our immediate factual world. It requires a system of beliefs, commitments, intentions, interpretations of events and more. We call this a *self-presentation* (SP). In the literature of sincerity that is not oriented to speech acts, the focus is often on personality and habitual aspects of personal life. Often particular acts are taken as sincere because they come from people who have been judged trustworthy. For an interesting philosophical discussion of these aspects, see (Williams, 2002). The idea of a self-presentation (SP) is not about these aspects of sincerity.

Approaching sincerity from an interest in speech acts, we are in effect committed to enabling accounts of single acts, generally performed one at a time. Yet communication does not proceed as a set of independent acts. Acts are linked to context, and to other acts by other interacts or self. When a statement is received, if its sincerity is to be examined, no variety of knowledge is excluded. Various kinds of knowledge about the speaker, the subject matter, the occasion and the reasons for speaking contribute to judging sincerity, as well as to a number of other processes that apply to each statement. For example, ambiguity resolution, based on the plausibility of various readings, interacts with sincerity.

All such processes require a holistic use of available knowledge, interrelated knowledge that forms a network supporting interpretation. This means that judging sincerity, as part of overall language interpretation in communication, requires comparisons using a diversity of kinds of knowledge and techniques.

In ordinary interaction, when a speech act is performed and its sincerity is at issue, there is a prior history of knowledge of the speaker's thoughts, immediate purposes, cultural assumptions and more. There is a prior context of the interaction as well, so that there are already commitments in place, intentions being pursued and ideas under discussion. There is always to some degree a stable reconciliation of the parts, so that a new communication is judged for compatibility with a somewhat consistent network. A speech act is judged for sincerity relative to this larger body of knowledge. The SP is constituted of these sorts of knowledge, and when a speech act occurs, it becomes part of the self-presentation of the speaker.

Given this interrelated character of assertions in practical use, it is important to choose a representation of speech act effects that allows multiple collections of related information, networks that are similar in content but with differences, each network having its own kind of consistency.

In order to meet these requirements, we propose a memory organization for the information which must be examined to judge the sincerity of speech acts that can be sincere, and to make it possible for a speaker to be insincere. A hearer's memory, for ex-

[5]The commonalities between these ploys may justify exploring sharing parts in formalizations and implementations. Also, irony shares features with metaphor, so the possibilities of sharing are extended. See (Ortony, 1979) p. 108-111

ample, has four partitions. Two of them represent the private thoughts (PM_H) and public self-presentation (SP_H) of the hearer. The other two represent the private thoughts ($PM_{S/H}$) and public self-presentation ($SP_{S/H}$) of the speaker, in the hearer's view. Each participant's active memory is organized to reflect the possibility that what a speaker presents is not always consistent with what he or she thinks.

For all four of the modes of speaking above, deception, irony, exaggeration and understatement, what is said must be compared with beliefs attributed to the speaker.

Conclusions

Sincerity has been an important aspect of speech act theory ever since Austin and Searle introduced the theory. It was foundational; for certain speech acts to be performed sincerely, the speaker had to have certain thoughts: e.g. the speaker must believe what is asserted. Sincerity is part of the definitions of such acts.

However, if we explore recognition of insincerity, the definition is too imprecise to use.

There are sincere acts that are labeled insincere by the classic view. The classic definitions do not correspond to sincerity as we know it. When we compare deception, irony, exaggeration and understatement, we find that insincerity is involved with attempting to deceive rather than simply holding certain thoughts.

In order to recognize sincerity (or any of the other three ways of speaking), comparisons between networks of beliefs are required. They all, along with similar ways to use language, might be facilitated by using the four-partition model of active memory described above, the elaboration of which will be subject to further research.

References

J. L. Austin. 1975. *How to do Things with Words.* Harvard University Press.

J. Habermas. 1984. *The Theory of Communicative Action*, volume 1. Beacon Press.

S.J. Levinson. 1988. Putting linguistics on a proper footing: Explorations in goffman's concept of participation. In P. Drew and A. Wooton, editors, *Erving Goffman: Exploring the Interaction Order*, pages 161–227. Northeastern University Press.

J.D. McCawley. 1989. Participant roles, frames and speech acts. *Linguistics and Philosophy*, 22:595–619.

A. Ortony. 1979. *Metaphor and Thought.* Cambridge University Press.

J. Searle. 1969. *Speech Acts.* Cambridge University Press.

B. Williams. 2002. *Truth and Truthfulness.* Princeton University Press.