Models of Intentions in Language

William C. Mann
SIL International

Intention is one of the significant topics in linguistics, partly because of the important roles that it is given outside of linguistics. All of the classic domains of text interpretation -- law, religion, literary studies, philosophy and the communication sciences, give prominent roles to intention and the one who intends. In common culture, the so called "folk linguistics," the role of intentions turns out to be significant as well, providing influential starting points for technical work. The importance of the concepts represented by the word "intention" is not reduced by the fact that it is controversial, nor that some schools of thought would deny that it is relevant or tractable for study.

The key questions for this paper are: What models of intention are there? Where do they come from? How do they differ? It is a survey with evaluative comments. It has a restricted focus. In particular, it focuses only on intention accompanying language, and only on adult human language. For convenience, the terms "intention" and "goal" are used in this paper as synonyms. The term "linguistics" represents a very inclusive notion, broader than any academic department. Although the scope of the paper is open to all of the literatures of all of the disciplines that deal scientifically with human language, its coverage is necessarily partial and irregular, so that important contributions may have been missed. It is descriptive, seeking to identify models rather than to define them.

What is intention? i.e. what do various documented notions of intention have in common? The verb "to intend," is generally used to describe situations in which, minimally, some individual has in mind some situation, not yet actualized, along with a disposition to prefer that the situation be actualized. Related terms include desire, want, wish, hope and prefer. Comparable phrases include to carry out an intention, to implement an intention, to act on an intention, to commit to an intention, and many more. The term intention sometimes, but not always, includes the notion that the person who has that intention in mind also has in mind a commitment to try to actualize that intention. Context usually makes it clear whether commitment is present; the paper uses both senses.

A good starting point for considering notions of intention is Intention in the Experience of Meaning, (Gibbs 1999). Gibbs has an excellent and extremely broad description of the uses of intention concepts in various literatures, including not only General Linguistics, Computational Linguistics, Cognitive Linguistics, Sociolinguistics, Psycholinguistics and various similar work but also art and literary studies.

As Gibbs amply demonstrates, the topic of intentions and language has both defensive and offensive (that is, constructive) aspects. His general purpose is to defend the use of the relevant terms and concepts in many fields, especially showing their legitimacy and promise in Cognitive Linguistics and other cognitive studies. The early chapters lay out the controversies well, including the "death of the author" movement in literary studies. It turns out that, even recognizing that significant uncertainties arise, the legitimacy of developing and using such concepts is extremely defensible.

Intention has been a controversial idea in various schools of psychology for decades, especially since the period when behaviorism was dominant. The effect is still with us, but Gibbs shows how these concerns can be used as precautions rather than prohibitions.

**Particular Models of Intention from Gibbs:**

The first purpose of this paper is to bring together a diversity of mutually relevant concepts, partially surveying models of intention in linguistics and philosophy. A second purpose is to comment on particular ideas.

Gibbs’ book is also a convenient place to start exploring various models of intention. In the process of concentrating on the legitimacy and benefits of using the concept of intention, the book identifies four different models, discussed below. These models represent a synthesis and presentational simplification of prior work, represented in (Krauss and Fussell 1996), where each of them is represented by extensive discussion of multiple papers.

The first model is the well known “code model” of communication, which has no role for intention at all. It is given to provide a starting point, and Gibbs quickly rejects it. I reject it as well. It fails to capture the notion that

---

1 The code model, sometimes called the transmission model, basically consists of the following: <<Communication using language is in essence exchange of ideas, which are propositional. Language is a code, analogous to Morse code. Speakers encode their ideas by encoding the corresponding propositions in language, which
people choose to communicate, and also that understanding and misunderstanding differ only if there is intent. We see below that its failings are beyond counting.

The code model of communication is in deep disrepute. It is widely rejected outside of linguistics (see Craig 1999), and also often within linguistics. Craig sees most of the various schools of communication theory as defining themselves in contrast to, and rejection of, the code model (p. 125).

The second model is called the Simple Intentionalist model. In this model, each language user has personal intentions which are used as the basis of expression. Each also has a personal world view which is used as a conceptual basis. Even in dialogue, the views and intentions of interacting persons are unrelated. One of many available objections to the Simple Intentionalist model is that it has no possibility for accounting for the collaborative nature of dialogue. A student interacting with a teacher would simply be pursuing personal goals, not responding.

The third model is called Perspective Taking. Language users are aware of the intentions and world view of their addressees, and they always adopt the intentions and world view of the addressee in directing language to that addressee. This has the advantage of giving a kind of recognition to the interactive nature of dialogue, and a kind of democratic or peer respect to the addressee’s views and biases.

However, the Perspective Taking model has some problems as well. If I always adopt (my view of) others’ intentions and world view while addressing them, then I do not expose my own intentions and world view, and others may have no access to them. So if I always use Perspective Taking, my addressees cannot. Similarly I may have no access to their intentions based on their use of language, making true perspective taking impossible. Also, there is no obvious way to extend the Perspective Taking model to the task of addressing an audience with mixed world views and intentions.

The fourth model is called the Dialogic model. Here the intentions that shape the language of interaction are not imposed by one or both of the parties to a dialogue, but rather are arrived at jointly, in effect negotiated. The results of the negotiation are jointly determined constructs, which may not be quite the same as intentions held by either party before the dialogue began. (Participants may in addition hold individual personal goals.)

This model of intention, like Perspective Taking, has the advantage of giving a kind of recognition to the interactive nature of dialogue, and a kind of democratic or peer respect to the addressee’s views and biases. It goes further by negotiating agreement, creating a communicative regime in which the other party’s intentions and view can have effects as well as respect. In this model, language understanding is based on an essential way on a shared set of intentions, not pre-specified as in the previous models but created by some sort of tacit collective bargaining.

Unfortunately the Dialogic model has problems as well. One problem concerns the negotiation of goals to be jointly held. In what language is this negotiation conducted, and, in that language, how is understanding achieved, and how can the success of the negotiation (even between adversaries) be guaranteed or recognized? The language of the negotiation cannot be the language that is used after the negotiation, since there is not yet a set of jointly held intentions. But if that pre-agreement language turns out to be adequate for the negotiation, it might be adequate for general communication as well. Possibly, not all such negotiations succeed. Empirically, participants do not appear to switch languages when they come to rapport. So the process of coming to negotiated goals may not be necessary, or may not be possible, or may not be reliably discoverable.

Even if there are weaknesses in these models, Gibbs should not be faulted for sketching them in the way that he has. His purpose is to present an attractive diversity of useful looking ideas, without working out details. The ideas support the notion that interesting research can be done based on models of intention. (He does not choose between these models, and rejects only the code model.) Gibbs presents a very strong defense for using intentions as a key element in modeling language. These models are only a small part of his presentation.

A Single Adequate Model of Intention in Language

Gibbs’ models allow us to reexamine the whole scientific effort toward seeking one or more satisfactory models of intention. We must seek to clarify the nature of the search for models. Typically in the sources that I have drawn on, implicitly or explicitly, the search has focused on finding one adequate model. I relied on this basis in rejecting Gibbs’ models. But this basis is not appropriate.

is then transmitted to recipients. Recipients decode the propositions, thus recovering the ideas. No intentions are referenced in encoding or decoding. For intention, it is the null model.

2 It is interesting that (according to Craig) many fields outside of linguistics regard the code model as what linguists believe about communication.
Given the complexity of languages, cultures and interacting people, and the very diverse needs represented in different communication situations, it is not reasonable a priori to expect that one model fits all situations. Arafat and Bush may indeed be using a different intentional model than the travel agent with the customer or the mother with her child. If indeed there is a single model of intention for all purposes, then the burden of demonstration should be on whoever makes such a claim.

It is far more reasonable that, in describing the phenomena at hand, we expect to describe them with a structured collection of models, selected and applied in a way that depends on situations. It is not necessary to have a full taxonomy of models before recognizing that multiple models prevail.

Kinds of intention that are not plausible as the sole form of intention, and thus required in all cases, can be very plausible as partial accounts. This includes all of Gibbs’ models. Surely Perspective Taking applies to some participants’ behavior in some interactions, and the negotiation of goals as in the Dialogic model occurs in some interactions. The Simple Intentionalist model applies sometimes (and traces of it are found in theories of development of language abilities.) Even the Code Model may apply sometimes. It is more reasonable for us to expect a combination of intention types and intention-dependent processes to be found, rather than a single, presently unimaginable scheme that applies to all human interaction in all situations. The further description below of models from the literature reinforce this expectation.

The first part of this paper, along with many of the cited references, commits an error which could be called the Single Comprehensive Model Fallacy. Conclusions are drawn, and models rejected, on the basis that a particular model could not possibly always be in use. Such reasoning may have unfortunate consequences beyond the error of rejection. Models rejected as the sole model may subsequently be overlooked after multiplicity of models has been accepted. Desire for simplicity has its place, but when we are faced with a poorly understood collection of phenomena, partially and loosely described, the expectation of simplicity is often simple foolishness.

It is more appropriate to seek a consistent covering set of models of intention, and to expect new models or new variations to come in with new kinds of data. If this orientation is correct, then work toward creation of a model of intention should somehow anticipate that a comparative phase will follow, and that the model need not account for all human language use.

A Set of Attributes of Intentions

Moving on, we find that many models of intention visible in the literature are not on Gibbs’ list or in (Krauss and Fussell 1996). There is extensive work on grounding and mutual knowledge, some involving intended changes accomplished using language, in (Resnick, Levine et al. 1991). Among close neighbors of linguistics, philosophy deals extensively with language and intention, e.g. (Tuomela 2000), (Bratman 1987), (Bratman 1990). Two lines of work by Grice are well known: one on the involvement of intention in meaning and the other on Conversational Implicature, (Grice 1975). Both have prompted an extensive constructive response; see for example (Levinson 1999). In Artificial Intelligence, work on Computational Linguistics deals with language and intention, but so does AI work on computational agents, represented in part by (Wooldridge and Rao 1999). The abundant work on computer authorship (text generation) almost universally includes intentions in an activity called text planning. See (Rubinoff 2000) as a representative token. Work on computational semantics (or model-theoretic semantics) sometimes includes intention or interactants’ goals as part of the “context,” i.e. the knowledge of each other held by interacting persons, see (Bunt and Muskens 1999). However, those intentions are not generally in an organizing position in the interpretation process. Also, such models can be capable of using knowledge of intention but not capable of producing it. Work on understanding of dialogue also makes extensive use of intention; see (Poesio, Cooper et al. 1999). Cognitive sciences, including Cognitive Linguistics and Cognitive Psychology, also develop models of intention, as seen in Gibbs work above and (Clark 1996) below. Intentions are widely acknowledged in Social Psychology as well; see (Higgins and Kruglanski 1996). There are also models of intention in linguistics, e.g. (Sperber and Wilson 1995), most prominently in many kinds of work in Linguistic Pragmatics.3

Faced with such an abundance of sources, it is not possible to cover enough ideas by continuing one model at a time. We must leave the list-of-models approach and deal with distinctions and details at a slightly finer grain. The descriptions below are drawn from linguistics and its close neighbors, including philosophy.

3 Beyond all of these, there are patterns in language that represent the satisfaction of highly recurrent intentions. Patterns used to satisfy recurrent intentions become conventionalized to such a degree that their connection to intentions is no longer recognized. Thank you in English is an example. "Purpose clauses" are another. Such constructs are often studied without direct use of concepts of intention.
For a diversity of models that are in use, we will focus on gross distinguishing characteristics, calling them attributes. These attributes are perhaps not as neat as one might expect. Some of the attributes represent distinctions that do not always exhaust the alternatives, and some represent concepts usually thought of as degree concepts. Not all are sharply defined, and even if sharp are not always discernible in accounting for natural cases. There are also a few dependencies between them, so that an intention with one of the attributes may also need to have a certain value of one of the others. Despite these complicating factors, it is useful to identify them.

These attributes are discussed:

1. Activeness
2. Partialness
3. Priority
4. Tacitness
5. Immediacy
6. Interaction-configuring
7. Intended to be recognized
8. Jointness
9. Sharedness
10. Structuredness
11. Complementarity
12. Conventionality

**Attributes of Individual Intentions**

**Activeness:** What kind of thing is intended? Perhaps the most important distinction for modeling the intentions that accompany language use is a contrast between intended actions and intended effects. Intended effects typically are states of affairs that the intender desires or prefers, while intended actions typically involve some identifiable process within the capacities of the actor(s).

There has been much philosophical work on intended actions, sometimes considered without simultaneously considering intended effects (desired states, directing action toward particular results, etc.). Such work has strongly influenced models of language use. For example, Tuomela's book *Cooperation: a philosophical study*, is a detailed development of notions of joint actions, actions characteristically performed by a group of two or more persons (Tuomela 2000). The actions studied are predominately intentional, and in this view intentions of joint actions may be accompanied by subintentions of individual participants, intending actions which serve to carry out their individual roles.

Tuomela's work is not particularly focused on language, nor on groups of only two participants. But Tuomela's notions of joint action are applied in Clark's *Using Language* (Clark 1996). Clark describes a wide range of situations in which the concept of joint actions is both necessary and powerful in accounting for interaction using language.

Past work in both linguistics and philosophy has made extensive use of notions of intention, but often what was seen as intended was some sort of effect or outcome rather than an action. So, for example, Austin, in the course of defining *perlocutionary* and eventually defining *perlocutionary act*, says "Saying something will often, or even normally, produce certain consequential effects upon the feelings, thoughts, or actions of the audience... and it may be done with the design, intention, or purpose of producing them..." (Austin 1975 (1955) p 101. He is working with both intended effects and intended acts, and gives the general impression that the acts are selected in part in order to produce particular effects. In another place he says "In considering responsibility, few things are considered more important than to establish whether a man intended to do A..." and he goes on at length to distinguish the English forms: *intentionally, deliberately* and *on purpose*.

---

4 Although this might seem to be a clear and exhaustive dichotomy, it is not. Many cases and categories are not easily classified. Still, it is helpful to consider cases where the contrast is clear.
In the light of the widespread use of notions of intended effect, it is curious to find recent work that focuses on actions to the exclusion of effects. It is understandable in philosophy, where action theory is a well established focus. However some philosophers put focial attention on both intended action and intended effect, along with the interactions of the two. One of the influential philosophers with this orientation is Michael Bratman. In (Bratman 1987), he rejects "the methodological priority of intention-in-action," (which represents a strong philosophical tradition) making a fundamental assumption that humans are in an essential way planning agents (p. 5-9).

So, for models of intention, there is a design issue of whether to recognize intended effects, whether to have an action effect contrast, whether to have intentions representing an overlap of intention and action and whether to have a residue that is neither. Bratman notes that in such issues philosophy draws significantly on artificial intelligence (Bratman 1990). For examples see (Cohen, Morgan et al. 1990), in the same volume.

**Partialness:** It may be possible, in studies of intentional actions by others, to restrict one’s focus to fully specified intentions. This would yield a certain sort of simplicity. However, if we are to think effectively about our own actions, or write computer programs which function as agents, or otherwise model the intentions we encounter, we must represent partially specified intentions such as we find in plans and in the processes which develop plans. Current research must balance short term simplicity with the risk that work that does not represent partially specified intentions will not generalize.

**Prioriness:** As Bratman notes, there is a tradition in philosophy of studying the intentions that immediately accompany actions, often called intention-in-action. (See (Searle 1983),(Goldman 1970),(Anscombe 1963),(Davidson 1980).) There may be qualitative differences between an intention-in-action which is inseparably attached to an action that is in progress and an intention that is not so attached. There is thus a design decision in modeling intentions about how each is represented, and how the two are related. For prior intentions, but perhaps not for intention-in-action, there are also issues of how commitment to an intention is made or abandoned.

**Tacitness:** The intentions involved in language use are generally not consciously experienced. Some sort of contrast may be needed in models, and it may be consequential. For example, if conscious and tacit reasoning differ, the difference may affect what sorts of reasoning or even emotional influences might apply to particular intentions.

**Immediacy:** Varieties of intentions that inherently must be acted upon immediately are, by that attribute, subject to distinctive methods of recognition, understanding and pursuit. For example, a question may express an intention that the parties immediately jointly pursue the intention of the speaker knowing the addressee’s response to the question. Immediacy is part of the understanding of questions, and although postponing a response may be negotiated, it constitutes a rejection of answering the question immediately.

**Interaction-configuring:** Some actions in interaction are intended to alter the configuration of the interaction. For example, intentions may be added or removed from the set of mutually accepted joint intentions that are regulating a dialogue. These actions may form a very small class for which the recognition, acceptance and rejection methods may be specific to that class. If these methods are distinct, then it is worthwhile to treat these intentions in a distinct manner.

**Intended to be recognized:** Ever since Grice wrote about intentions that are intended to be recognized (Grice 1975), there have been active attempts to incorporate his insight into larger frameworks. Use of such intentions appears to have far broader scope than Grice discussed. Just as for the interaction-configuring intentions, these have distinctive recognition, acceptance and rejection methods and so can benefit from special treatment.

**Jointness:** There are actions that are inherently those of multiple actors. They may be physical actions, such as moving a large table, or language use actions, such as negotiating a price. Clark has argued extensively for recognizing joint actions in language use. (Clark 1996) His underlying assumptions generally follow Tuomela.

Joint actions are typically undertaken on the basis of joint intentions. If we see intentions as states of mind, and minds as personal and individual, (and if we recognize interaction on a sensory basis but deny any sort of telepathy,) then joint intentions cannot be directly analogous to joint actions. Joint actions are observable in a way that joint intentions are not. Joint actions typically have individual actions that are in some sense subordinated to the joint action, and such individual actions are also intentional. But see Sharedness below.

**Attributes of Collections of Intentions**

The attributes below are typically involved with multiple intentions of a single intender, although the first one is not necessarily multiple, but has multiple intenders.
Sharedness: Multiple parties can share particular intentions. For example, two people may each intend to wash dishes until the pail of dirty dishes is exhausted. Sharing of an intention does not make it a joint intention, and having two parties committed to a shared intention does not produce cooperation. (For example, the dishes may be washed by the one who arrives first.) Even so, identifying intentions as shared can be consequential in modeling interactions.

Structuredness: Many discussions treat intentions only in isolation. However in practice, intentions are often formed, negotiated and carried out in compatible sets, being restricted so that commitment to the various intentions is a single act. Plans may include a structure of subplans, and choices in planning may involve commitment or abandonment of large intentional combinations of this sort, combinations that may include more than one actor. Providing for this (or not) is a design decision in modeling. There is already an extensive literature that implicitly relies on structured intentions in language use. It includes the text planning literature, of course, but also various kinds of "dialogue acts" and various constructs called "dialogue games." (For example, see (Stolcke, Reis et al. 2000) and (Mann 1979), (Mann 1988).)

Complementarity: Given a joint intention or comparable construct, individual actions to fulfill the intention may simply be additive, with each actor making some contribution to the total, and the intentional structure may reflect this.

Alternatively, there may be distinctive roles for the parties, with the intentions interlocking (complementary) so that the fulfillment of the joint intention depends on the parties fulfilling their distinct individual intentions. Such intentions are complementary.

Tutoring exemplifies this. In tutoring there is a shared joint intention that the student will come to master some body of ideas or skills, and accompanying this intention are intentions about what the teacher will do and what the student will do. In one phase, teachers will periodically test students' knowledge, and on those occasions students will attempt to exhibit their knowledge. The testing and exhibition of knowledge are intentional, and their interaction makes the intentions complementary.

Convenationality: For recurrent kinds of situations, and the corresponding recurrent kinds of intentions, the intentions function as problems to be solved. As a consequence, the actions which are taken function in part as solutions to problems. Patterns arise and become acculturated or conventionalized, becoming conventions of interaction and language use.

Tutoring also exemplifies this. Tutoring is not reinvented every day on demand. Rather it is developed as a collection of plans, skills, strategies, habits and choices that are created incrementally over periods of years. These patterns of intentions affect the language used in ways that are every bit as conventional as the lexicon.

Other intentions, such as the intention to express appreciation for a favor, or the intention to acknowledge an expression of appreciation, also fall into patterns and become conventionalized. For recurrent situations that are culturally specific, such as giving and responding to praise, conventions of different cultures also differ. The Western practice of acceptance of praise and expression of gratitude, or the Asian practice of denial and humility, are more easily treated as conventional patterns of intentions rather than as responses generated immediately on the basis of need.

The list of attributes above could surely be made longer and refined in detail, based on the extensive literature and many ongoing investigations. The listed attributes represent an informal core of notions that are necessary for representing any major share of human linguistic experience. There are some dependencies between attributes, so that not all conceivable combinations represent genuine possibilities. Even so, the projected complexity is formidable.

A Note About Persistent Intentions:

In some fields of study, a special category of "intentions," sometimes called fixed or persistent intentions, has been used as a variant of other intentions. For human language use these would include such personal desires as maintaining one's credibility, appearance of friendliness, stance of respect for addressees, appearance of self control and many other comparable desires. These are distinguished by being generally neither introduced explicitly nor dismissed upon successful satisfaction. They are seldom pursued explicitly, seldom pursued alone or directly, and often regarded as stable features of personality or culture. They do not seem to function much in means-ends reasoning. Rather, they are more like constraints or tendencies that influence choices in communication. They do not seem to be useful as subgoals in the ordinary way.

Thus in the family of intentions that impact human use of language, they are anomalous. They certainly represent genuine phenomena of some sort, but such phenomena deserve separate study rather than study integrated with other intentions. For those reasons I have not included them in this presentation.
**Intentions in Philosophy and Linguistics**

On the narrow subject of intentions, we have seen how linguistics and philosophy have been intertwined. There has been an impressively large amount of interaction on a broad range of subjects, represented in only a token way by this particular topic. All of the work cited above is a tiny portion of the total.

In this broad interaction, linguistics is functioning mainly in roles of observer and consumer. Philosophy is in more of a producer role. Some work may occasionally be motivated by a perception that it is needed in linguistics, but such cases seem rare. (Clearly the desire for accounts of personal responsibility, personal intention and intentional action, along with the need to extend these notions to social groups, have shaped the work in philosophy. But these concerns are not distinctively linguistic.)

Philosophy functions for linguistics as a source of ideas, a locus of invention. It also functions as a kind of quality control, a way of checking that conceptual schemes do not embody particular kinds of errors. Philosophy provides no guarantees, but it is a corrective force, genuinely beneficial.

It would be simplistic to try to say that one discipline leads and the other follows. Neither has enough unity for that. Sometimes linguistics has used complex concepts of intention that I have not found discussed in philosophy. An example would be joint intentions of effects, directly analogous to the joint intentions of actions that accompany joint actions discussed by Clark. Another example would be the intentions in dialogue game theory (Mann 1979; Mann 1988), which posited configurations of joint intentions and joint actions, structurally combined with complementary individual goals, involving intention of both effects and actions. They include both prior intentions and intentions-in-action, all represented in culturally shared conventions of interaction. These conventions were employed by means of a novel collection of speech act like devices, a special variety of actions.

We have seen that some conceptual frameworks used in linguistics are derived from chains of development in philosophy. What of the rest? There are certainly original inventions. There are also cases where accounting for data leads to novel insight. The remainder seems to come from commonly held views, the so-called "folk linguistics," of the home culture of the researchers.

Scientific views of intention arise to a significant degree from prevailing pre scientific views. The common ways of talking about language organize study. Metaphors of what people do, including the *container metaphor* and the *conduit metaphor*, point to transfer of ideas and change of belief as the most accessible foci for intentions in interactive language use. These are given primary roles in semantics, pragmatics, philosophy and general linguistics, often without examining other alternatives. So, for example, we note that the code model has a culturally supported credibility. The perspectives in the Perspective Taking model and the outcomes of communication are all taken to be intellectual in character. These easy assumptions also reflect folk linguistics, our western culture. Along with philosophy, it is a powerful influence on the study of intentions, and of course many other linguistic topics as well.

**Conclusion**

What has been accomplished in the informal survey above? How can it be consequential in future work?

Some models of intention in human language have been identified. In comments on them, problems that arise when they are used as comprehensive models have been identified. By gathering from the literature some attributes of models of intentions, it has become easier to generate many more models, and also to discriminate between models. It suggests from the literature how intentions are linked to concepts of action, consequence, plan, commitment and other concepts as well, and thus suggests ways to improve on the pattern of isolated study of intentions in human language.

Perhaps the most important outcome concerns partial models. This paper raises the issue of whether the search for valid models of intention should accept partial models, and concluded firmly that it should. This suggests setting aside the method of rejecting models for which there are counterexamples. It suggests that there is a Simple Comprehensive Model Fallacy that should be avoided in work on this topic, and in work on other suitably comparable topics as well.

**References:**


