Informal Pragmatics and Linguistic Creativity

John Collier
Philosophy, University of KwaZulu-Natal, Durban, South Africa 4041
collier@ncf.ca

Abstract

Examples of successful linguistic communication give rise to two important insights: 1) it should be understood most fundamentally in terms of the pragmatic success of each individual utterance, and 2) linguistic conventions need to be understood as on a par with the non-linguistic regularities that competent language users rely upon to refer. Syntax and semantics are part of what Barwise and Perry (1983) call the context of the utterance, contributing to the pragmatics of the utterance. This full and distributed multichannel context determines meaning if anything does. On the standard account of context, context disambiguates the meaning of language, but it is at least as apt in many situations to say that language disambiguates context. In practice, the two work together, sometimes with more emphasis on one than the other. Reference should be understood in pragmatic terms (it is an act) and, since success is often achieved in non-standard, creative ways, any formalization of pragmatics can only be partial. The need for such an inventive approach to referring traces back to the need for language to be highly efficient, with expressions underdetermining their interpretation. Next, the shared semantic and syntactic regularities, which might seem to be independent of the context of an utterance, should be understood as also being part of that context. Past usage underdetermines how terms can be used since it allows for multiple projections. Successful reference with novel uses that are disambiguated by context can become the ground for new conventions.

1. Traditional pragmatics

For the wider reader of this paper, I will start by describing the traditional formal view of pragmatics as envisioned by Montague (1970a; 1970b) and well represented by Kaplan (1989). Then I will go to discuss how meaningful information is transferred, and how this allows channels through the environment. After that I will introduce the more novel parts of this paper which is based on work by Barwise and Perry (1983) and Barwise and Seligman (1997), but is also influenced by the work of C.S. Peirce (e.g., 1878; 1905) according to whom our representations develop and become more clear and reliable, as well as changing in other ways based on their inherent potentials through the formation of new hypotheses of how they can be used (abduction). This permits intelligible creativity in language.

The traditional formalist view of pragmatics takes the primary function of language to be representation that can be mutually understood. It takes an external view of language and tries to

*I am grateful for comments from two referees that helped me be clearer. This work was in part sponsored by a Rated Scholars Incentive Grant from the National Research Foundation of South Africa.
explain how it works, rather than a view of language from the perspectives of the participants. Language encodes thoughts of the transmitter into words that are decoded by the receiver, giving a shared representation. Utterances are signals that ‘carry’ the relevant meaning; the meaning is ‘in’ the signal as determined by shared meanings of the participants. Most terms are generals and are not context dependent, so pragmatics is restricted to indexicals and ambiguous terms and usages. On this account contexts disambiguate utterances.

This can be seen fairly clearly in Kaplan’s (1989) account of pragmatics in terms of character and content. Content is a function from propositions (or sentences), predicates and names to extensions, typically across possible worlds. These extensions are fixed for terms that have a fixed meaning, and then character = content. We can (we hope) determine if a sentence is true, or a term applies within a possible world, including the actual one, by examining the relation between the content and the world. Character, on the other hand, is a function from terms or sentences together with their context to contents. Characters, then, determine contents according to the usage of the term and the context. The most direct use of the theory is to explain indexicals like ‘I’. If I say ‘I am a philosopher’ this is true in the actual world, but not true in a world in which I stayed in my original profession of planetary scientist. Normally, usages of ‘I’ refer to the person uttering the sentence, and the context is provided by the circumstances of the world in which it is uttered. (I say ‘normally’ because there are strained examples that violate normal usage of ‘I’ that still have an extension.) The theory can be extended to demonstratives and to ambiguous and even vague terms, though with increasing complexity and potential complications.

Traditional pragmatics, following Montague, takes as its ideal formal languages. Montague is precise about this:

There is in my opinion no important theoretical difference between natural languages and the artificial languages of logicians; indeed, I consider it possible to comprehend the syntax and semantics of both kinds of language within a single natural and mathematically precise theory. On this point I differ from a number of philosophers, but agree, I believe, with Chomsky and his associates. (1970a)

In (well-constructed) formal languages each symbol has a unique interpretation, and can be applied generally. Well-formed formulas each have a unique propositional correlate with an unambiguous reference, or meaning. Formal languages are maximally precise idealizations of imprecise natural language. The problem with indexicals and similar terms that Kaplan’s approach was designed to accommodate is that they show that language itself is not maximally precise. The solution to this problem, within the formalist context, was to introduce a way to re-establish this precision.

2. Elements of an alternative non-formalist view

Before turning to our account of pragmatics I will present the situation semantics of Barwise and Perry (1983) that inspired it. Their account, like ours, combines information transmitted through language with information transmitted through circumstances to determine content, which in

their case is a situation.¹ This information perspective is similar to the traditional view in that it views linguistic communication externally in terms of its overall functioning, rather than from the subjective perspectives of the participants, but it makes generality an abstraction, grounded fully in actual utterances. An utterance is part of an overall situation and it is the overall situation that conveys the situation (possibly absent) that is the content of the utterance. The information conveyed by the utterance, assuming it is successful, restricts the possible meanings to a smaller set (though the utterance may also draw attention to the idea that there is a situation to be attended to – see the third example in section 3 below). In order to understand how information is conveyed by utterances in a situation it is useful to look at work by Barwise and Seligman on information flow (1997). Their theory is rather technical and not presented in a very transparent way, but fortunately only some aspects need to be explained for my purposes here. Nonetheless, the ideas are grounded in a highly formal theory of information transfer that has the properties required by the idea of information flow in situations in Barwise and Perry (1983).

**Situation semantics**

In situation semantics individual utterances are primary and generals like sentence forms and terms are abstractions from particular instances.² Note that this completely inverts the usual idealization of natural languages as a sort of formal language. Not surprisingly, this inversion has far-reaching consequences: 1. Utterances are (parts of) situations that carry information about situations, not instances of generals. 2. All situations are particulars in the world; meaning is in the world. 3. Interpretation is situation dependent (almost always, rather than just with some recalcitrants like indexicals and demonstratives): so generals associated with instances by regular use (conventions), the conventions also sometimes (often?) being situation dependent. 4. The context (part of the situation that contains the utterance) disambiguates indexicals and other ambiguous terms, as on the traditional account, but the utterance can also be better said to disambiguate the situation in many cases. 5. The logic of situations is nonstandard. This is because, as Benson Mates pointed out, only the traditional content can be expected to follow standard logic. When we bring in extralogical elements like situations adding them together (conjunction) can greatly change the meaning.

Collier and Talmont-Kaminski (2005, in review) adopt this basic framework, but argue that when all is said and done we can determine what propositions are invoked and recover standard logic. But this depends on situations and propositions being able to be made commensurate. I will not address in detail in this paper. The significance of point 5 above is that the logic of information flow (Barwise and Seligman 1997) has the same nonstandard logical character as situation semantics. The reception of situation semantics concentrated mostly on the semantics, and left most of the rest in the classical form (a gross misinterpretation of its radical nature, which to judge by some comments of referees of this paper is still rampant) partly by trying to

---

¹ This contrasts to some extent with Rączaszek-Leonardi, and Scott Kelso (2008) who take a more dynamical view of the interpersonal process. I believe our approach is not incompatible with theirs, not the least from direct discussion with Rączaszek-Leonardi. Bach (2002) also give primacy to pragmatics (or at least parity), but he makes no advances not already implicit in Barwise and Perry (1983). Needless to say, we are sympathetic with his move.

² One of the referees failed to see the significance of this. It means that all of the mutual abstract knowledge of generalities of syntax and semantics are part of the context of utterance. I assume this throughout.
assimilate it to other approaches to semantics. However, we are not so interested in the semantics per se as in the primacy of utterances as situated individuals over generals and its consequences.

Information flow

I think it can be taken for granted that language involves information transfer in all instances in which it is successful (Collier and Talmont-Kaminski in review). It is helpful, then, to get a clear idea of what information transfer involves. The basic components are transmitter, receiver, channel and message (Shannon and Weaver 1949). Often left out, but in Shannon’s diagrams, are coding and decoding processes by the transmitter and receiver, respectively. These do not contribute to his communication theory directly. A message is something that the receiver can decode to reduce the possibilities to a smaller number. Logically, this is a more restricted proposition. Note that this assumes that there are specific possibilities at both the transmitter and receiver end. Presumably this involves the coding and decoding processes, and I will return to this later. In any case, as it is, the theory has nothing directly to do with meaning. Interestingly, the most informative messages are the ones with the least redundancy, since these restrict the possibilities the most. Noise entering the channel can introduce ambiguity, but this can be reduced through redundancy in the message; the minimum required was calculated by Shannon. Reliable channels tend to be highly redundant because the amount of noise is difficult to predict in many cases, such as in human interactions. Eliminating noise, however, is not always desirable: noise can be promoted to information with novel possibilities (e.g., mutations, creative misunderstandings). Most of the time noise just leads to confusion, but in some cases it can be creative. What makes this creativity possible? I will return to this in section 3 below.

In Shannon’s theory something carries the information, but it is not specified beyond what is required to make up combinations that are distinct from each other. Typical examples would be Morse code and the bits that flow in a computer, but information can also be carried by the modulation of sound, or of radio waves, and many other carriers. I proposed (1986) a notion of a physical information system with elements that could physically combine in various ways in order to carry distinct messages. Although the combinations could be strings, as in DNA (at least to a first order), higher dimension combinations are possible, which I called arrays. The important requirement on the physical system other than the possibility of multiple distinct combinations is that the elements are fairly stable during the information carrying process, as well as across different instances of similar processes. The arrays need to be stable only enough to complete the carrying of a specific message. It should be clear, however, that information is carried by particular tokens, not by general kinds. However, if the information is to convey anything meaningful then it must convey something about the transmitter being in some kind of state, in other words that the source falls under some predicate. So both types and tokens are required for meaningful information flow. Barwise and Seligman (1997) take this into account in their basic concept of an infomorphism, which is a non-accidental relation between two classifications, in which tokens fall under specific types. The conditions on an infomorphism require that a received token falls under a type that is correlated with the type that the correlated sent token falls under. They construct channels from this concept, and then show how channels can be combined to describe information flow in a distributed system. We (Collier and Talmont-Kaminski 2005, in review), as well as, implicitly, Barwise and Perry (1983) use the distributed
nature of information flow through channels within contexts rather than a single channel from
the producer to the receiver, as is presumed on traditional models.

Some examples might be helpful here to the uninitiated reader. A simple example given by
Barwise and Seligman (1997) is a signal by a flashlight. The components might be a bulb,
battery, switch and case. The channel is basically a connected series of infomorphisms from
switch to bulb through the mediation of battery and case. A channel does not need to be
sequential as a whole, but various parts do, e.g., the switch sends information to allow the flow
of electricity through the case to the bulb, so that the bulb has information about the state of the
switch. This can be used to signal from a distance. The receiver, knowing the light comes from a
flashlight, receives the signal produced by the person controlling the switch. Note that
infomorphisms may be directed (hold between classifications in one direction, but not the other).
In the flashlight case, if the channel is functioning the switch also has information about the bulb,
but this sort of converse relationship need not hold. This makes the logic of information flow
nonstandard.

A more sophisticated example comes from phylogenetic taxonomy, the classification of
organisms. Traits can be coordinated according to presence or absence in different species.
Assuming descent with modification, we can use parsimony calculations to determine probable
relatedness. This can give us information about probably descent, giving information about
evolutionary pattern. Biogeography can also provide information about descent: proximity of
related species and historical information about migration vectors can be used to determine
probable relatedness. These channels can be combined to give more robust information about
probable descent. Other channels in the distributed system that makes up evolutionary history
may give further information. Note that in this case the direction of information flow is one way.
Changing a classification will not change the pattern of descent.

In situation semantics information is carried by situations, part of which is the utterance. The
situation is itself a particular (a token), but the information is useless unless it indicates the
transmitter is in some kind of state, that is it falls under some predicate, so that a proposition is
conveyed. It should be kept in mind, though, that the rest of the situation (the context) can also
convey information, and there may be multiple channels that carry relevant information, and
these can vary considerably in duration and retention of the information. Channels are formed by
constraints on the flow of the tokens that carry information. As Shannon (Shannon and Weaver
1949) pointed out, a constraint is a form of information because it restricts possibilities. This is
reflected in the duality of types and tokens in Barwise and Seligman’s (1997) approach: tokens
and types can be exchanged so it is the nominal types that carry the information and the tokens
provide the classification. This may be surprising at first, but consider the following: a detector
can detect whether an electron is in an up or down spin state and allow only those in an up state
to pass. It gives information about the electrons. However, we can use the spins of the electrons
to get information about what kind of detector they have been through (or what state it is in, if it
can switch its polarity). Given that both types and tokens, properly understood, can convey
information, in situation semantics almost anything can carry contextual information:

• general semantic conventions
• local semantic conventions
• ongoing perceptions and possible perceptions of the environment
Meaning

For two classifications A and C related by an infomorphism, it is natural to talk of tokens of C being about tokens of A when classification C is artificial, and perhaps when C is selected. In the artificial case the relation is usually intentional (accidental infomorphisms are ruled out in Barwise and Seligman’s approach). Intentionality does not apply easily to cases in which the infomorphism has been selected, but I think that Terry Deacon’s idea of ententionality (2012) is useful here. It refers to processes that are induced by things that might not be present. Given this strong analogy to intentionality we might want to extend the notion of meaning to cover ententionality. However there are also natural infomorphisms that can be used as signs intentionally or entionally, but are purely mechanical. These will also involve dual classifications (the detector example I used above might be a good case if we see the electrons as carrying information about the state of the detector whether or not they are observed). I have argued that all causal processes carry information, and that this is what distinguishes them from other correlations (Collier 1999, 2012). Causation in itself is surely not meaningful. I think it is best to conclude that not all infomorphisms carry meaning, but do carry information. I leave the boundary between such cases and ones that are ententional open here, as it has no bearing on what I say below.

Despite an infomorphism allowing meaningful aboutness, the A and C classifications need not be the same, or even commensurate in general. However, satisfying the channel conditions in a given case is necessary and sufficient for information flow. Thus the channel (including context) may go from a speaker’s mind to her words to the receiver’s ears and then his mind and the receiver may thus have information about what the speaker intended to convey. This communication need not be complete or perfect, and probably seldom is. Moreover, the receiver may add aspects either consciously or unconsciously that alters the intended communication, and the speaker may have communicated things she never intended. In any case, none of the speaker’s thoughts, words or the receiver’s detection and interpretation of them need have the same classifications for some information to be transferred successfully. The notion of a common meaning is much more slippery than the traditional view would have it.

Note once more that the overall channel may contain both symbolic (mental and verbal, among others) and nonsymbolic constraints that are distributed. It is the infomorphisms that tie things together and carry whatever meaning is communicated.

Selection theory has been suggested as a ground for natural content or meaning in the guise of teleosemantics (Millikan 1984, Matthen 1988). In selection there is a tuning of genes and traits with the environment so that the classifications correspond in general. However this does not directly extend to a theory of communication. One problem with this model as one of
communication is what I call the symmetry problem (Collier 2008). Note that the correlation of genes and traits (and also with environmental features) goes both ways, so the information relations are symmetrical. This is not how we usually think about ‘aboutness’. Whether or not this is a problem in biology I set aside here, but it applies to communication if it is thought of just as a correlation. Relative time can solve the problem if we assume that the information is conveyed causally and this determines the direction of communication. This has obvious merits. But furthermore, infomorphisms, like Barwise and Perry's logic of situations, do not obey standard logic, and C having information about A does not in general imply that A has information about C. Changing what I think might change the information I get from what you say, but it does not change what you think. This is a more general condition than temporality alone. So the symmetry problem has a solution in the formal theory of information flow, and allows for real symmetries that can occur in mutual understanding. Nonetheless, we still need something similar to evolutionary tuning for linguistic communication that can establish appropriate infomorphisms. This is to be found in mutual habitual usages, what might be called conventions. I will describe how these might be created in the next section.

3. Pragmatist pragmatics and linguistic creativity

Now, having laid the foundations, we can start with the novel part of this paper. Collier and Talmont-Kaminski (2005, in review) presuppose the basic structure of the previous section. Unlike the standard view, in which contexts disambiguate utterances, on our view utterances often, even typically, disambiguate contexts. Because of the open-ended nature of contextual factors that might be shared by the speaker and the listener, there is no way to fully define character. Pragmatics therefore extends to all interpretation of natural language, and context includes all factors involved in a speaker to listener channel but the transmitted utterance. From this it follows that conventional usage, which on the situation semantics account is a part of context, is malleable in the sense that we can exploit shared aspects of context in order to convey meanings in nonconventional ways. This malleability allows channels in context to interact with each other (and themselves) in order to allow new meanings to emerge, based on the rearrangement of classifications. This permits information to be communicated that might not be able to be communicated otherwise.

There are some restrictions on this sort of creativity, however, and they are important for the potential meaningfulness of novel usages. First, the new classifications are a result of shared context, so they are already present, at least implicitly. Not just any alteration or replacement of conventions would be meaningful. Furthermore, overlapping contexts between the communicators are required for common propositions. A particular utterance might facilitate this in novel ways, but the potential has to be immanent for both the speaker and listener.

Once a novel usage has been made, it can become habitual within a smaller or larger group, and may eventually come close to the literal sort of meaning embraced by the traditional view. Adopting our view allows the explanation of meaningful linguistic innovation, but at the expense of giving up standard logic as the logic of communication. We do not see this as a special problem, however, since our view also allows the recovery of traditional logic once the meaning in a specific context has been fixed. Once representational meaning is determined, it can be fit

more or less well into the idealized structure of traditional semantics. This allows us to use traditional logic to analyze meanings and relations of meanings. There will be limits to formalization because communication often needs only sufficient overlap in meaning for the circumstances, and need not be fully precise, unlike formal languages. At least part of the appeal of traditional formal methods, we think, comes from the possibility of this sort of post facto formalization. It was a mistake, however, to try to analyze the pragmatics of natural language on the model of formal languages.

I now turn to some examples to illustrate some of the points I have made.

The first example is due to John Perry (1994) who demonstrated the point with some coloured blocks borrowed from his child. The point is originally Wittgenstein’s.

As John Perry (1994) observes, the amount of information required to disambiguate the referent depends upon both the intended referent and the context we wish to pick it out of. The relevant considerations can be seen using Wittgenstein’s (1953, §2) example of the language employed by builders to identify the building elements that they require. Given a language made up of words such as ‘block’, ‘pillar’, ‘slab’ and ‘beam’ a builder can inform an assistant what element is now required by him, so long as any of the available elements of the particular general type would be appropriate. This may be because all of the elements of one type are the same or, equally, because the differences between them are not significant in relation to what the builder requires: in the first case the reason is tied to the complexity of the context; in the second it is tied to the pragmatic needs of the utterer. So, even if we specifically need a pillar which has been painted in barber shop stripes, all we need to say is ‘pillar’, so long as all of the available pillars have been painted in barber shop stripes. (Collier and Talmont-Kaminski in review).

This example shows that it is sometimes more appropriate to think of the words as disambiguating the context than vice versa. The shared world of the builders creates a common context that serves as a channel for the information in the words. The builders know what they need to do in general, but they need the specifics for that action. The word provides that disambiguation.

The second example is similar, but the shared world is much more restricted. As in the Wittgenstein example, the function is not representation but a request. It demonstrates how little can be needed for effective disambiguation. I go into a coffee shop in Beijing, and the only symbols I see are Chinese characters, which I do not understand. I say ‘latte’. The barista says something like ‘rattee’ and I receive a latte and hand over more than I think it will cost and receive change. Note that I need not speak the language; we need only a similar word and a relevantly overlapping context. ‘Latte’ disambiguates the context and establishes what should happen. The coffee shop and conventions about them serve as a channel. This would not work in a taxi. (But I found that other approaches for establishing common context would work, if I know where the destination is – now I could try google maps).

The third example we used to demonstrate the role of nonlinguistic aspects in defining context (Collier and Talmont-Kaminski in review). It also demonstrates how an utterance can lead to a more refined channel for communication. Two hunters A and B are hunting deer:
A: ‘Over there,’ pointing casually towards the woods
B: acts confused
A: nods upwards towards the sky over the woods
B: sees agitated birds over woods, says, ‘Oh.’
The perceptions and context of hunting serve as a channel. ‘Over there’ helps to disambiguate, allowing the location of the animal to be transmitted. There is also a use of nonverbal conventions (pointing and nodding), and a refinement of the channel through B’s making the connection between the utterance and what it means by connecting two facts he hadn’t connected before. Basically, B has created a novel category that opens the channel. The precise nature of the category B created does not matter; it might be complex or simple, general, or restricted to this instance.

The final example is a case of linguistic novelty and conventionalization in a specialized environment. I first encountered it in discussions by formal pragmatists as a difficult case. In a restaurant the manager says to a waiter, ‘The roast beef wants a glass of water.’ Normally this would be nonsense because roast beef can’t want anything. In the context, however, it could be a new innovation, referring to the diner who ordered roast beef by using a word for what they are ordered. But it could also be a convention established by repetition of similar usages, the convention having caught on because of convenience or some other reason. The original use might have been a perhaps playful innovation, and aspects of playfulness and belonging to an ‘in’ group may lead it to be retained. The restaurant context and its normal practice permit the communication, but in a novel case of the usage (in the history of this restaurant) common knowledge about restaurant practices and perceptions creates a channel. This involves a novel classification and potential generalization to similar cases. Furthermore, this could occur many times in many restaurants, given common situation kinds, permitting generalization, and fixation of a new linguistic usage. We could analyze the case as an abbreviated definite description, but this seems to me to be unnecessarily complex. To do that would also undermine the similarities between this case and the previous three.

Conclusions

There are many conclusions that can be drawn, but my main concern here is the creation of novelty. I first review the main points of our view: Contexts can be made up of a variety of factors including linguistic conventions (global and local), background knowledge, occurring perceptions, intentions, and many other things that can be exploited. Common reference requires coordination of classifications, though information transfer does not. Only enough clarity for the situation’s pragmatic purposes is required. Language does not stand on its own, but is located in the world, and is impossible without context. Traditional semantics and logic in terms of successful reference can be recovered (within limits) after meaning is established, explaining its appeal. Utterances can reduce ambiguity of contexts just as much as the other way around. Shared circumstances, which are often not verbal and need not be fully conscious, permit novel classifications.
Novel usages can spawn new conventions. These can become fixed and lead to new general terms. The novelty (creativity) is constrained by the context, which further ensures that the novelty is functional (has value). The novelty is a new class under which instances (tokens) can be classified. This is exactly Peirce’s abduction (method of hypothesis). Abduction is the reasoning process by which we invent a class under which to place our observations. It is often objected that there are no constraints on this creative aspect of scientific thought, so it should not be called reasoning. However at least in the sort of linguistic creativity I invoke here, the possibilities are highly constrained and the meaning is ready at hand to be used because it is implicit in the context. The approach ties together the classification based information model of communication with novel meanings. It explains both the possibility of creativity and why it is most likely to be functional.\(^3\)

References


Collier, John. 1986. ‘Entropy in evolution’ Biology and Philosophy 1, 5-24


\(^3\) Perhaps this is also true of scientific theory innovation, but examining that would take me too far afield.

Collier, John and Konrad Talmont-Kaminski. In review. ‘Reference as a pragmatic achievement’ submitted *Pragmatics and Cognition*.


