About "Natural Language versus Formal Language"

I presented "Natural Language versus Formal Language" as an invited speaker (together with Frederic Fitch, Bas van Fraassen, and *Richard Montague) in the joint symposium by that title of the Association for Symbolic Logic and the American Philosophical Association at their joint meeting, New York, December, 1969. I have appended a scan of a typed manuscript prepared just prior to that talk. It builds on material from my dissertation The Algebra of Intensional Logics, Univ. of Pittsburgh, 1966 (Director: Nuel D. Belnap), which also does "An Intuitive Semantics for First Degree Relevant Implications," contributed paper, meeting of the Association for Symbolic Logic, Chicago, May, 1967, (Abstract) "An Intuitive Semantics for First Degree Relevant Implications," Journal of Symbolic Logic, 36, 1971, pp. 362-363. And it is a precursor to my "Intuitive Semantics for First-Degree Entailments and Coupled Trees," Philosophical Studies, 29, pp. 149-168.


J. M. Dunn, January 19, 2011
Natural Language versus Formal Language

J. Michael Dunn
Indiana University

The comparison of natural languages and formal languages has become quite popular of late. The topic was on the program of the last International Congress for Logic, Methodology and Philosophy of Science in Amsterdam, and also on the program of the 1968 New York University Institute of Philosophy. I have read the published results of both meetings \([7,4]\), and I must confess that I am not quite sure what all the fuss is about.

On both occasions it was pointed out that a natural language typically differs from a formal one in that a natural language is ambiguous, vague, context dependent, and generally untidy. I agree that one can typically point to these differences, but frankly my reaction is, so what? We all know that untidiness has both its good points and its bad.

The very title of this symposium, "Natural Language versus Formal Language," suggests a certain opposition that I think is inappropriate. It sounds rather like the opposition, "Ford cars versus General Motors cars," but it seems to me that the opposition is more like that of "Ford cars versus John Deere tractors." It could be that both are useful, for different purposes.

Thus if one is writing poetry, it seems desirable to have a language that is ambiguous; but not necessarily if one is writing mathematics. And if one is talking to one's wife, vagueness might be a convenience; but maybe not if one is programming a computer.

Now surely the purpose which is of most interest to the people here today is that of "doing philosophy," as we say. With
this purpose in mind, one might think that there is a definite choice, at least in principle, between natural languages and formal languages. But I doubt that there would be any universal agreement on this choice, simply because I am sure that there would be no universal agreement about what doing philosophy involves. Indeed, if doing philosophy is an activity having something to do with the gaining of insights, these insights might just as well be produced by a Zen master's stick as by the use of either natural or formal languages.

But if one believes that part of doing philosophy involves attempts at constructing valid arguments, then it seems to me that one should be concerned with making explicit why they are valid, at least in problematic cases, and the best way of doing that is by means of a formal language. I am not suggesting that arguments should actually be written completely in a formal language, nor that the steps in arguments be numbered and labeled according to, say, Copi's rules. Mathematicians do not do this either, but the formal structure of their arguments is usually (though not always) clear enough that it could be reconstructed in some appropriate formal language if one so desired.

When the classical logic of Principia Mathematica was the only brand of formal language on the market, it was perhaps understandable for some philosopher to feel that his argument lost something in translation (often its validity) when it was formalized. But nowadays, what with modal logics, free logics, tense logics, deontic logics, epistemic logics, entailment logics, et al, this feeling deserves less sympathy. Of course it is always possible that a philosopher with such a feeling has some genuine insight
about what follows from what, and that this insight is not captured
by any extant formal logic. But insights being rather rare and
logical errors being all too common, a little shopping around among
at least the more well known formal logics would not hurt. And if
none of these fit, one is always free to try to knit one's own.

I have pointed out that logicians seem to be getting away
from the bugaboo, to paraphrase Ramsey out of context, "What we
can't say in PM we can't say, and we can't whistle it either."
Formal languages are becoming increasingly natural, and I believe
that this undercuts part of the supposition behind the topic of
this symposium.

There is a converse development which also tends to undercut
the distinction between natural and formal languages. Roughly put,
natural languages are lately appearing to be more formal. What
with the work that Chomsky and others have done on generative
grammar, it is no longer clear that the so-called natural languages
such as English are not formal languages after all. It is now
frequently conjectured that the grammatical sentences of, say, English
are recursive. It is true that the transformational rules most often
suggested for generating them are context-dependent, whereas the
rules for generating the well-formed sentences of a formal logic are
typically context-independent. This, however, seems to be but a
mere difference in detail (which is not to say that detail cannot be
extremely interesting).

Now it might be thought that the difference between a natural
language and a formal language arises not at the level of syntax,
but instead at the level of semantics. The most extreme view that
might be taken here is that of the formalist: a formal language,
by definition, is regarded as uninterpreted. I learned very early to avoid quarreling with definitions. But I cannot help pointing out that many of the languages created by logicians are regarded as interpreted. Indeed, Gödel’s famous completeness and incompleteness theorems take full advantage of such interpretations.

Now it is quite true that the semantical theories proposed by logicians for their formal languages have typically differed from the semantical theories of natural languages proposed by linguists. Thus the semantics of formal languages has typically centered around a Tarski-type recursive definition of truth. This definition can be quite complex in the case of some of the more sophisticated formal languages whose semantics is some variant of the Kripke semantics for modal logic. (Professor Montague has a general semantical theory for such languages, calling them all pragmatic languages.) But still the basic objective is to define under what conditions a sentence is true (relativized to a model, a world, a history, a speaker, or what have you).

On the other hand, the semantical theories of natural languages proposed by linguists have typically avoided the notion of truth altogether. Instead they have tried to provide “readings” of sentences by some ideal representation of their semantic structures. Simple semantic components, for example Katz’s semantic markers which are supposed to stand for simple ideas, are strung together in such a way so as to provide an unambiguous reading. These semantic theories have concentrated on the ways that readings are generated from sentences, and of central interest here has been the disambiguation of sentences into their different readings.
At the risk of oversimplifying, the quickest way to characterize the difference between the logicians' and the linguists' semantic theories is to mobilize Quine's division of semantics into the theory of reference and the theory of meaning. Logicians have talked as if they have been concerned with the former, and linguists as if they have been concerned with the latter.

Lately, however, there have been developments on the side of the logicians which challenge this easy dualism. Thus it has been argued that an account of meaning for both formal and natural languages alike should proceed via a Tarski-type truth definition. In defense of this, Donald Davidson says [1]:

There is no need to suppress, of course, the obvious connection between a definition of truth of the kind that Tarski has shown how to construct, and the concept of meaning. It is this: the definition works by giving necessary and sufficient conditions for the truth of every sentence, and to give truth conditions is a way of giving the meaning of a sentence. To know the semantic concept of truth for a language is to know what it is for a sentence--any sentence--to be true, and this amounts, in one good sense that we can give to the phrase, to understanding the language.

Professor Montague has also argued on several occasions for handling the semantics of natural languages in the same general manner as the semantics of formal languages, and his "English as a Formal Language I" is an ingenious and penetrating application of this point of view.

Now let us leave aside the obvious problems caused for this approach by the fact that there are many sentences in natural languages that appear to be neither true nor false (questions, imperatives, etc.). These cases need not necessarily vitiate the Montague-Davidson approach. First, because there may be
ways of handling the troublesome cases by some extension of the concept of truth (perhaps along the lines suggested in Michael Dummett’s paper “Truth,” and as actually carried out in the Belnap approach to the logic of questions and the the Castaneda approach to the logic of imperatives). Second, because even though the Montague-Davidson approach might not be appropriate to all of English, it might still be applicable to that large fragment of English consisting of ordinary declarative sentences.

Indeed, both Montague [in 7, p. 276] and Davidson [in 1] have claimed that the Tarski truth definition can be straightforwardly applied so as to provide a satisfactory semantics for that fragment of English that consists of the literal translations of the formulas of the classical predicate calculus.

Should we identify the meaning of a sentence with its truth conditions? I do not ask this as a metaphysical question like the question are numbers really classes of equinumerous classes? If classes of equinumerous classes behave enough like numbers, then at least we have some sort of isomorphism, and that is all we need for certain purposes dear to logicians. But the present question is whether truth conditions do behave enough like meanings. Clearly meanings determine truth conditions, but I do not find the converse so obvious.

Now there is an obvious truism that we can take to justify the claim that truth conditions determine meaning. Thus consider the sentence 'Snow is white'. A truth condition for this sentence is: 'Snow is white' is true iff snow is white. I readily agree that if one knew and understood this truth condition, then one would know the meaning of the sentence "Snow is white" (for the
unexciting reason that one must understand this sentence in order to understand the truth condition, which is formulated by using this very sentence itself). If truisms such as this were the only thing in the wind, I would not bother to turn my head.

But typically when people claim that truth conditions determine meaning, they go on to say some profound but ultimately silly things, such as that any two logically equivalent sentences have the same meaning since they have the same truth conditions. This leads quickly to the view that any two logically false sentences (or any two logically true sentences) are synonymous. We get the most striking application of this line of thought in Wittgenstein's *Tractatus*, where he says [9]:

4.461 Propositions show what they say: tautologies and contradictions show that they say nothing.
A tautology has no truth-conditions, since it is unconditionally true: and a contradiction is true on no condition.
Tautologies and contradictions lack sense.

I simply do not believe that the sentence

\[(A) \ (\exists x)(y)(y \in x \equiv y \neq y)\]

has the same meaning as

\[(B) \ 1 \neq 1,\]

nor do I believe that they are both meaningless, even though I grant that they are both logically false.

This Tractarian view survives today in the best logic texts. Jeffrey, in his text *Formal Logic* [5] says a little more than most authors to justify that the truth table rules of valuation give meaning to the connectives. Thus he says [p. 15]:

The rules of valuation make no mention of the meanings of sentences; they are couched entirely in terms of truth-values. Nevertheless, the rules of valuation determine the
meanings of compound sentences in terms of their ingredient sentence letters, for we know the meaning of a sentence (we know what statement the sentence makes) if we know what facts would make it true and what facts would make it false. Now if we have this information about the letters that occur in a sentence, the truth conditions supply the corresponding information about the whole sentence.

A little later [pp. 30-31] in discussing contradictions, Jeffrey says:

The sentence
\[ \text{It is and is not raining} \]
is only apparently about the weather, just as the sentence
\[ 2+2 = 4 \text{ and } 2+2 \neq 4 \]
is only apparently about numbers. In fact the two sentences have exactly the same truth conditions: in all possible cases, both are false.

I think that we can avoid the necessity of Jeffrey's conclusion while yet agreeing, in a trivial sense, that the meaning of a sentence is determined by its truth conditions. Thus let \( p \) be the sentence 'It is raining' and let \( q \) be the sentence '2+2 = 4'. By standard truth table considerations it follows that \( p \& \neg p \) is true iff \( p \) is true and \( \neg p \) is true, that is, iff \( p \) is true and \( p \) is false. Similarly, \( q \& \neg q \) is true iff \( q \) is true and \( q \) is false. The question bluntly then is whether the condition that \( p \) is true and \( p \) is false is the same condition as that \( q \) is true and \( q \) is false. I think it is not.

Notice that it is no argument against me to reply that the first is a contradiction meaning \( p \) is true and \( p \) is not true, while the second is also a contradiction meaning \( q \) is true and \( q \) is not true, and that of course any two contradictions have the same meaning. This only pushes the question with which we began up into the metalanguage.

Intuitively, \( p \& \neg p \) and \( q \& \neg q \) describe different situations, granted that neither situation is realizable. What we need is a
a semantics that is sensitive to this intuition.

I may as well let any who do not know me in on a little secret at this point. I was a student of Belnap and Anderson's at the University of Pittsburgh, and I am one of those crazy people who think that there is something in their system E of Entailment (and in the other similar relevant logics that have been developed). I believe that there is a sense of 'entails' (or 'implies') in which it simply is not true that a contradiction entails (or implies) any old sentence whatsoever. It thus becomes extremely critical that not just any two contradictions are synonymous. For if p&¬p were synonymous with q&¬q, then since it is true that q&¬q entails q, then by substitution of synonyms salva veritatae, it would be true that p&¬p entails q. Having made a clean breast of my motivations, I hope that they will not be held against me as I continue.

I mention at this point that both Professor van Fraassen and myself have developed semantical ways of ruling out p&¬p's entailing q. I, in my dissertation [2], in terms of q's possibly being about some topic that p&¬p is not about, van Fraassen [8], in terms of some fact forcing p&¬p which does not force q. Both of these semantics lead to completeness proofs for a very narrow fragment of the system E, namely those sentences of the form A entails B, where A and B are purely truth functional (the so-called first degree entailments), and it is very difficult to see how these semantics might be generalized so as to take care of all of E (with entailments entailing entailments, etc.). Furthermore, both these semantics suffer from the defect that they are formulated in terms of concepts that are out of fashion in logic (topics that sentences are about, facts that force sentences to be true).
Let us then recall which concepts are in fashion, and I shall try my best to talk that language in trying to communicate the notion that not just any two contradictions are synonymous. The standard realization of a proposition as found in Montague, Kaplan, Scott and others is a mapping from possible worlds (or reference points, situations, call them what you will) to truth values. This corresponds to the principle that different meanings can be distinguished by different situations with different truth values, i.e., by different truth conditions. But it has the unfortunate consequence that (relative to a given set of situations) there is only one contradictory proposition, simply because there is only one constant false mapping.

However, we need modify this picture only slightly to provide a kind of extensional apparatus that allows us to distinguish contradictory propositions from one another. Starting from the intuition expressed above that a contradiction can be true in some situations (of course, unrealizable) in which some other contradiction is not true, we can identify a proposition with a relation from a set of situations into the set \( \{T, F\} \), where every situation is related to at least one of \( T \) and \( F \). A contradictory proposition is then such a relation where \( F \) if in the image of every situation. There can then be many different contradictory propositions. These can be distinguished by a situation such that one proposition has \( T \) in its image while the other does not.

What this means as far as the modeling of truth functional logic is concerned is that a valuation is a relation from sentences into the set \( \{T, F\} \), rather than a mapping, and of course it is required that every sentence be related to at least one of \( T \) and \( F \) (we shall eventually speculate upon what happens if we drop this
last requirement). This relation is determined inductively in
just the classical truth table way. Thus

1) \(-A\) is T iff \(A\) is F,
   \(-A\) is F iff \(A\) is T;

ii) \(A \& B\) is T iff \(A\) is T and \(B\) is T,
    \(A \& B\) is F iff \(A\) is F or \(B\) is F;

iii) \(A \lor B\) is T iff \(A\) is T or \(B\) is T,
    \(A \lor B\) is F iff \(A\) is F and \(B\) is F.

Note that in each of i)-iii), we need two clauses, one giving
truth conditions and the other giving falsity conditions. We cannot
rely upon the standard intuition that a sentence which has been given
the value T is not F.

We can already give a semantical explication of one of the
principal features of entailment, namely, that \(p \& \neg p\) need not entail
\(q\). For there is a valuation in which \(p \& \neg p\) receives the value T and
yet \(q\) does not. This is a valuation in which \(p\) receives both the
values T and F, while \(q\) receives the single value F.

We can also give a semantical explication of perhaps the
most controversial feature of entailment, namely, that \(-p \lor (p \& q)\)
need not entail \(q\) (the failure of the so-called rule of disjunctive
syllogism). Let me give this explication in the context of
examining the supposed proof of Lewis's that a contradiction entails
everything.

The proof starts out by supposing that \(p \& \neg p\) is true. We
then detach \(p\) by the rule of simplification, and from \(p\) we obtain
\(p \lor q\) by the rule of addition. Next we obtain \(\neg p\) from our supposition
of \(p \& \neg p\) by another use of the rule of simplification. So far, O.K.
But finally we claim that \(q\) follows from \(\neg p\) and \(p \lor q\) by disjunctive
syllogism. In producing this proof for a class, it used to be my
habit to motivate this last step by telling the following story. "So on our assumption that \( p \land \neg p \) is true, we have obtained that one of \( p \) or \( q \) is true. But we have also obtained \( \neg p \), which says that \( p \) is not the true one. So \( q \) must be the true one." When I was once telling this story, some wise guy yelled out, "But \( p \) was the true one--look again at your assumption."

That wise guy was right. If we assume that \( p \land \neg p \) is true, we are thereby assuming that \( p \) is both true and false, and hence it should not be surprising that \( p \land (\neg p \lor q) \) comes out true under that assumption, while \( q \) might still be false.

Do not get me wrong; I am not claiming that there are sentences which are in fact both true and false. I am merely pointing out that there are plenty of situations where we suppose, assert, believe, etc., contradictory sentences to be true, and we therefore need a semantics which expresses the truth conditions of contradictions in terms of the truth values that the ingredient sentences would have to take for the contradictions to be true.

I must unfortunately remark that these particular insights have not as yet given the degree of illumination regarding entailment that might be expected. In particular, they have not led to completeness proofs for the system \( E \) of Entailment. But I have obtained reasonably intuitive completeness results based upon this framework for an extension of \( E \) called \( R \)- mingle. These I now announce for the first time, and they should not be confused with earlier purely algebraic completeness results obtained by both my colleague Robert K. Meyer [6] and myself [3]. The intuitive results are rather like those of Kripke for intuitionist logic, and even more like those of Thomason for the system of Professor Fitch's Symbolic Logic, though with assignments allowed to give both
the values T and F. Furthermore, the extension to R-mingle with quantifiers seems not a bother. The basic idea is that instead of doing the classical thing of interpreting an n-ary predicate as a propositional function (a mapping from the n-tuples of objects in the domain into \{T, F\}), we rather interpret the predicate as a propositional relation (a relation from the n-tuples of objects of the domain into \{T, F\}, with the requirement that every n-tuple be related to at least one of T and F).

The reason why we do not obtain a semantics for the system E in this framework is that it is difficult to rule out p&q's entailing qv-q, since p&q is always false and qv-q is always true. Thus whenever the antecedent is true, the consequent is true (it being always true); and whenever the consequent is false, the antecedent is false (it being always false). Thus we are stuck no matter how we try to falsify the entailment, and yet the entailment is not a theorem of the system E. One way out that suggests itself is to let qv-q have no truth value, and we could naturally arrange this by allowing an assignment in which q was related to neither T nor F. Then p&q could be given the value T (by giving p both the values T and F), while qv-q is not given the value T (by giving q no value whatsoever). This works for first degree entailments, but there are vast problems both of an intuitive and a technical sort in generalizing this approach to entailments nested in entailments.

In closing, let me urge that even if the particular approach I have suggested for distinguishing contradictions semantically is not to your liking, still something should be done in this area. It may be good logic to say that any two contradictory sentences are logically equivalent, but I would think that it would be bad
linguistics to say that any two contradictory sentences have the same meaning.