Model-Theoretic Semantics as Structural Semantics*

Avery D Andrews, ANU

draft, July 21, 2003

In spite of its impressive empirical track record over more than three decades, Model-Theoretic Semantics[2] (MTS) of natural language has been persistently under attack from two opposed directions: from one direction, various kinds of 'Cognitivists' such as Jackendoff (1983, 2002) and Lakoff (1987) argue that its commitment to Realism render MTS irrelevant or even antithetical to the goals of natural language semantics, while from the other, 'Neo-Davidsonian Realists' such as LePore (1983), Larson and Segal (1995) and Ludlow (1999) argue that MTS doesn't succeed in connecting language to the world, and is therefore equivalent to the 'Structural Semantic' theories advocated by the Cognitivists, which don't try to connect language to the world, but merely try to give an account of the 'Logico-Semantic Properties and Relations'[3] (LSPRs), such as entailment.

If both of these critiques held up, it would look bad for MTS, but I will argue here that they don't. In the first section I will argue that MTS is a perfectly good generative theory of the LSPRs, with a modicum of psychological reality (at Marr's (1982) 'level 1', and even a bit higher), and should therefore not offend Cognitivists. In the second I will consider the fate of the deductive theories of the LSPRs that have typically been favored by Cognitivists, and will suggest some reasons why they have failed to thrive, the most important being that they contain excessive amounts of arbitrary information (basically unsupported speculation at Marr's level 2), relative to model-theoretic ones. The conclusion is that not only is MTS a good approach to the LSPRs, but that it is also the most suitable, at early stages of inquiry.

In the third section I will look at some of the work flowing from LePore's critique of MTS, the 'Absolute Semantics' approach of Larson and Segal (1995) and Ludlow (1999). A distinctive feature of this approach is that it abandons model-theory and the concomitant account of the LSPRs, and instead uses a 'T Theory' to connect sentences in the language being studied to a metalanguage. I will argue that from an empirical point of view, this approach is equivalent to a structure-transducing translation device, such as the 'semantic intepretation rules' proposed by Katz and Jackendoff, or 'glue logic' in LFG (Dalrymple 2000), and so does not constitutes progress in understanding linguistic phenomena.

The first three sections might be taken as indicating that the idea of connecting language to reality doesn't work, but this would be a paradoxical conclusion, because the rather successful MTS program is substantially motivated and informed by ideas about how language relates to reality. In the final section I will suggest that the intuitions that inform MTS derive from our built-in metalinguistic facilities for commenting on, maintaining and repairing linguistic communication. Since these facilities are based

*I am indebted to Barbara Partee and Ray Jackendoff for discussion of a predecessor of this paper; all errors are of course my own.

[2]Including not only 'standard' formulations based on Montague, but also Situation Semantics, Data Semantics, etc. I haven't yet thought about Game Theoretical Semantics, but suspect that it is equivalent to MTS w.r.t. the issues discussed here, although rather different in appearance.

[3]The term is from Larson and Segal (1995:3); Katz (1972:4-7) provided an extremely cogent and influential statement of this goal of semantic theory.

on our conceptions of language, the world, and the relations between them (a kind of part-whole relation), there is no conflict with Cognitivism, equally, the success of these facilities, to the extent that they work, suggests that the relationships as they appear to be are at least approximately true, hence there is no conflict either with Realism, although a Realistic attitude isn't essential for current linguistics.

1 MTS and the LSPRs

Although foundational discussions of MTS tend to focus on the idea of relating language to the world, it's a very important although sometimes neglected fact that it also provides a mathematical account of at least some of the LSPRs, especially entailment.[4] MTS analyses define entailment and other LSPRs by means of a bottom-up (compositional) assigment of semantic values to grammatical constituents, which ultimately determines possible patterns of truth and falsity holding over the sentences of the language. A set of sentences Σ is said to entail a sentence S if every assignment of values that makes all the sentences of Σ true also makes S true.

This style of account has the potential to seem problematic for cognitively-oriented linguists, because the semantic values can be extremely large (infinite sets of infinite-sized possible worlds), and there may be an infinite number of significantly different ones to consider (in the definition of entailment, every assignment of values that satisfies Σ must satisfy S). Since cognitivists want to explain how people perform various tasks, in this case compute/detect the LSPRs, they would clearly want to end up with a recursive enumeration, indeed one that is decideable for the LSPRs about which people have quick and reliable intuitions.[5] MTS doesn't constitute or clearly suggest the structure of a recursive enumeration, so does not in and of itself constitute full satisfaction of the cognitivist goals. But it still might be the case that it is a useful or even necessary stage in developing an account that does.

An instructive example is Optimality Theory (OT, McCarthy 2002), one of the most rapidly-developing areas of generative grammar in the last decade. In this approach, the well-formed structures are defined as the best, according scheme involving a ranked system of constraints, in a potentially infinite equivalence class of structures.[6] Since the equivalence classes are in the general case infinite in size (due to the possibility of deleting contentful elements or inserting pleonastic/epenthetic ones), there is no completely general way of enumerating the optimal structures. Nonetheless, it is also the case that instances of the theory can be rendered decideable by imposing additional constraints, and the strategy being pursued is to first try to find clean analyses that

[4]See for example the walkthrough in Chierchia and McConnell-Ginet (1990:74-75, 2000:84-86). There are substantial questions as to how to account for some of the 'non-declative' LSPRs, such as for example commands and the descriptions of the actions that would constitute compliance with them (I'm indebted to Ray Jackendoff (p.c.) for bringing this issue up).

[5]However there is always the possibility, occasionally mentioned by Chomsky, that explaining brain function might involve new physics, which would open the possibility that even routinely detectable LSPRs were produced by a non-computable system. At the moment there appears to be no reason at all to suspect that anything like this is the case.

[6]Most often, the structure consists of an 'input' and and 'output', linked by correspondence relations, and the equivalence classes are of structures with the same input, but there are other possibilities.

explain generalizations, and then work out the constraints that might enable people to compute these analyses (for example, many phonological analyses can be expressed with finite state constraints, for which production and generation of optimal structures is decideable, and not even very hard to implement for simple cases).

That MTS might be a good preliminary stage of analyses is suggested by Marr's (1982) system of levels of psychological reality for cognitive theories. Marr proposed the following three levels for such theories:[7]

(1) Level 1: an account of the function-in-extension computed by a proposed module, without reference to how this might be computed.
Level 2: an account of the algorithm used by the mind/brain to compute the function.
Level 3: an account of the physiological implementation of the algorithm.

The model-theoretic definition of entailment satisfies the requirements for a level 1 theory of this LSPR, although obvious not those for level 2.

The similarities of MTS to other generative theories goes further. Developing work by Evans (1985), Peacocke (1986) and Davies (1987) propose a 'level 1.5', intermediate between levels 1 and 2, which basically amounts to depicting the causal structure inside an Agent that is responsible for certain generalizations one might make about its behavior. For example the fact that in all languages, one gets almost identical structural possibilities for expressions designating entities (Noun Phrases) in the positions where these appear, is evidence that there is a common mental structure responsible for their format,[8] represented by the NP rule, or system of constraints applying to all NPs, typically found in a grammar.

Such inferences about the causal structures internal to an Agent are intermediate between mere descriptions of functions-in-extensions, and full descriptions of algorithms. And it is clear that MTS analyses can support such inferences, indeed, Evans' original example, and also the one discussed by Peacocke and Davies, is a truth-conditional semantics for a small language, where the issue is whether the truth-conditions of a sentence such as *Mary walks* are determined compositionally from semantic values associated with *Mary* and *walk*, for from a list specifying a semantic value for each subject-predicate pair. Larson and Segal (1995:56-62) make the same basic point, citing this literature.

So MTS analyses are not only like ordinary generative analyses in mathematically stating properties and relations of utterances, they also capture (or fail to capture) generalizations about them. Nonetheless, the fact that they don't provide much guidance about performance is certainly a problem, and if there were an alternative approach that did, and was roughly equivalent in performance in other respects, we would certain prefer it.

Once upon a time there appeared to be one, the proposal to account for the LSPRs in terms of structural rules formulated over 'semantic representations', as cogently

[7] For reasons which I don't understand at all, Chomsky (2000:188) claims that Marr's levels don't apply to the study of the language faculty. Since no argument for or substantial explanation of this claim is provided, I here ignore it. A similar but slightly different system of levels is presented in Larson and Segal (1995:546-548).

[8] *pace* Suppes (1976,1979, 1982).

described in (Katz 1972),[9] and proposed in different forms by Jackendoff (1973, 1983), Lakoff (1970), and others. If one of the deductive accounts had developed well, it would have been inherently preferable to MTS, for the purposes of linguistics, since by characterizing the LSPRs in terms of structural and therefore presumaby recursively enumerable concepts stated over finite representations, it would take us closer to having an account of how people manage to perceive the LSPRs (and perform the everyday tasks that they seem to be involved with, such as accomodating presuppostions). But they did not; in the next section I will make some suggestions as to why.

But before doing this I want to point out that the failure of the Deductive Approach is a significantly different issue from another which has a close historical relation to it, the idea of Representational Semantics. Representational Semantics, as practiced by Jackendoff, Fodor (1985), and others, is the decision to focus first on working out a good format for meaning-representation, deferring the problem of providing an explicit characterization of either the LSPRs or the relation between language and the world. Although widely derided by practitioners of MTS (most notably Lewis 1972) as mere 'translation into Markerese', this is arguably a viable strategy, vindicated by for example Zwarts and Verkuyl's (1994) success in providing a MTS for Jackendoff's semantic representations, and has produced extensive results in lexical semantics, where the acheivements of MTS and other kinds of putative 'real world' semantics have been rather modest. Representational Semantics is therefore a claim about the best order to do things in (representations first, details of logic second), and is independent of the issues being considered here.

I suggest that the reason this works is that on the one hand, there are a great many different syntactic options for expressing essentially the same semantics (as demonstrated for example by Cooper and Parsons's (1976) work on alternative representations of quantivfier scope), and on the other, people are quite good at judging whether a given syntactic proposal is capable of supporting an adequate account of entailment. So one can develop a format for semantic representation by focussing mainly on its adequacy for connecting to the syntax and expressing intuitively accessible facts about semantic patterns (such as causativization and detransitivization relations), without a great deal of explicit concern for defining the LSPRs, since, with a bit of common sense, this is likely to come out all right in the end.

2 The Failure of the Deductive Approach

By contrast with model theory, the Deductive Approach is much more in the style of conventional linguistic theory. Sentences are supposed to get finite-sized 'semantic representations', over which inference and perhaps other kinds of rules are supposed to apply to define the LSPRs (Katz 1972). Since it employs methods similar to those used in other areas of linguistics, one would expect linguists to have pursued it, but this didn't happen. Katz's framework[10] failed attract followers or develop empirically, and

[9] Developing earlier proposals such as Katz and Fodor (1964), Katz (1966).

[10] One question a reader might ask is why Katz wasn't included amongst the 'Cognitivists' discussed above. The reason is that although his early work was consistent with Cognitivist assumptions, he never showed much concern for issues of psychological reality, and in later work (1981), moved to an explicitly

while Jackendoff's and spme other Representational Semantics proposals have been much more successful, the proposed inference rules have never been delivered, Jackendoff's last attempt appearing in Jackendoff (1983).

I suggest that the proximal reason for this is that there is a much greater burden of arbitrary formal decision-making that has to be tolerated by people trying to set up a deductive system rather than a model-theoretic one. People with some exposure to logic might be aware of Prawitz-style natural deduction systems, various kinds of Gentzen systems, tableaux logic and Hilbert-style axiomatizations, and there are doubtless many more less well-known types, together with the possibility of cooking up something completely different on one's own. Furthermore each general type has innumerable variants, with different choices of axioms and inference rules. And we have very little insight as to whether there are any empirical effects of choosing one rather than another. These choices, with unknown (and possibly zero) empirical significance are highly entangled with choices that might make a difference, constituting a substantial obstacle to development and acceptance of a specific system. The effect of this is perhaps exacerbated by the fact that linguists are trained to justify formal choices in terms of rather local empirical effects in accounting for data items, rather than for example their contribution to large-scale mathematical properties.

By contrast, although MTS accounts are not free from problems of arbibrary decisions, these seem to in general be considerably less severe. In the case of standard model theory for the connectives *and*, *or* and *not*, for example, the alternatives that one can think of are either obviously equivalent (such as using 1 and 0 as truth-values rather than **T** and **F**), or clearly different (such as using three or four truth-values instead of just two). There are of course cases that are more difficult, such as those involving the propositional attitudes, but, by and large, where there are substantively different proposals, people have a sense that there is a reasonable prospect of finding facts to motivate a choice between them (Barbara Partee, p.c.). I should perhaps point out that what I take to be relevant here is not the objective existence of fewer hard-to-assess alternative analyses in MTS, but rather the smaller number of subjectively salient possibilities that will distract the attention and dissipate the focus of a normal investigator.

If the proximal cause is indeed an excessive requirement for arbitrary decision-making, is there a deeper reason for that? I suspect that there is, and suggest that it is a matter of over-specification of level 2-relevant information, which has no empirical warrant in the available data, and so must be made up on an arbitrary basis. An ordinary deductive system is not of course exactly what we need to explain the LSPR data of natural-language semantics, because the LSPRs are recognized so quickly and routinely that they presumably have decision-procedures, and fast ones.[11] But still, formulating a recursive enumeration of a property or relation is a good first step towards writing a recognizer—if you're writing in Prolog, and are lucky or clever, the enumerator might even be able to run as a recognizer without modification. This means that it embodies a lot of specification about how the computations are being done, even if there's not enough to fully specify an algorithm. But the linguistic semanticist trying to invent a semantic theory probably doesn't have any empirical evidence about these

non-psychological position about the nature of language.

[11] Although semi-algorithms that terminate in the cases normally encountered in practice, but not necessarily always, would also be a possibility.

computations, and so just has to make something up.

An MTS account on the other hand doesn't provide any kind of computational specification, so fewer arbitrary decisions have to be made. This makes things easier, but carries a price: ultimately we want to explain how people detect the LSPRs, which we can assume implies that we need a deductive theory (indeed, one with a fast though possibibly incomplete decision method that performs as well and as badly as people do). As pointed out by Hornstein (1988), a purely referential, externalist semantic theory can't do this, and neither an an MTS analysis using things that are too big to fit into heads.

So far, the choice might look balanced, since we do ultimately want a recursive enumeration (and more), and the Deductive Approach clearly has the potential to provide this, in spite of the problem of arbitrariness, but there are two further considerations. The first is Marr's general point is that it is sensible to first characterize the properties and relations being computed (level 1), and then procede to work on how this is done (levels 2 and 3). With an MTS analysis, we can first give a relatively clean account of the facts, and then procede to working on the rather messier issue of how people manage to be aware of them. The second is a more specific technical point, which is that a model theory can be a very useful tool for working on a deductive theory.

One of the problems with deductive theories is making sure that they are consistent, and, more generally, identifying some non-entailments of sets of sentences as well as some entailments. This is likely to be hard, since one has to show that a chain of deduction steps of arbitrary length can never derive certain conclusions. But model-theories can be kept consistent by having negation change truth-values, and being careful about how truth is assigned to sentences involving self-referrence (Kripke 1975, (Etchemendy and Barwise 1987)). It is also often easy to produce examples of non-entailment, by presenting a (possibly finite, small) interpretation which satisifes all the sentences of Σ but not S. And finally, given a model-theory, it is usually easy to make up some deductive rules which are provably sound, so one can focus on getting them to deliver the desired entailments without worrying about gross misbehavior. It will then probably be possible, although not necessarily easy, to find a provably complete axiomatization, or a proof that the model-theoretical account is non-axiomatizable.[12]

A closely related advantage is that it is also very easy to add meaning-postulates (constraints on models) expressing axioms of well-known algebraic systems such as various kinds of lattices, so that mathematical results can be applied immediately. This is a major feature of Link's work, for example.

The lower arbitrariness burden of MTS, and its usefulness as a tool for later development of deductive theories, clearly makes it desireable to produce a model-theoretic account first, or at least to have one alongside the deductive account that one might by trying to produce. One can see this strategy pursued to good effect by for example Chierchia (1984:45-53) and Link (1998:133-161). More generally, doing at least some MTS first is to follow Marr's (1982) recommendation to work on the lower levels before the higher ones.

Granting this, linguists might still be puzzled at the apparently rather low level of

[12]In the final section I will briefly discuss the possible significance of non-axiomatizable MTS analyses of natural languages.

interest in developing deductive accounts of various areas in semantics. Optimality theory affords an interesting contrast. Although, as noted above, OT is also nonconstructive, there has from the beginning been a lot of attention given to the problem of how to find the optimal candidates. By contrast, the interest in such issues in formal semantics would seem to be rather muted.

I think this difference can be explained by differences in the circumstances of MTS and OT. In OT, most of the phenemona under investigation have been well-known for decades or even longer, and have well-known generative analyses. The mathematical aspects of OT on the other hand are novel and in many areas not very well understood: for example it is a genuinely interesting question whether it is possible to calculate optimal correspondences for reduplicative identity is a reasonably efficient manner. By contrast, formal semanticists work with theories which have been under mathematical investigation for decades, and whose mathematical properties are mostly familiar and well-understood. So, I propose, the reason that there's not so much interest in axiomatizing theories in MTS is that in the early stages of inquiry, where the outline of the facts isn't clear, it's actually not very interesting.

Of course, later on, when the MTS analyses have acheived good empirical coverage, it becomes more interesting to know whether they are axiomatizable or not, and also to procede to more psychologically oriented investigations of how the computations work, perhaps along the lines of Laird and Byrne (1991, 1993). But the first thing to do is to nail down the facts, and MTS would appear to be the easiest way to to that.

A question on might ask is how it is that syntax has managed to progress, in spite of having the same problem of excess level 2 specification that I have claimed exists in the Deductive Approach. One thing to note is that syntax is indeed quite balkanized into different formal approaches, to a greater extend than formal semantics is, it seems to me. Most of these emerged in some form or other in the seventies, and have persisted until the present. In syntax, there doesn't seem to be any substantial alternative to the level 2-overcommitted theories, so we simply have to live with the situation.

3 Re-stating Truth-Conditions

So far I have been discussing the virtues of MTS as a technique for mathematically characterizing the LSPRs, treating it as an apparently technically superior alternative to the Deductive Approach found in traditional Representational Semantics. This is in agreement with the position of LePore (1983), who argued that, in spite of standard claims to the contrary, MTS was incapable of relating language to the world, and therefore failed in the same way as classic RS. The claimed difficulty is that MTS defines truth relative to an interpretation (the assignment of semantic values to the expressions of the language), but doesn't specify how the correct interpretation (the actual world) is to be identified, so that we don't get an account of what it is for a sentence to be true 'absolutely', but only relative to an interpretation. I am here concerned with neither the validity nor the significance of this claim, but just with the nature and effects of the proposed solution.

LePore's proposed fix was to supplement the clauses of an MTS treatment with additional clauses such as those of (2), which would pin down the intended interpreta-

tion:

(2) a. The extension of “Barbara” = Barbara

b. (x)(x satisfies “sekoilee” iff x is confused)

Since these appeared as supplements to the truth-definition, they constituted an additional facility, not altering the account of logical consequence, but adding to it an account of the intended interpretation.

However, in further developments of LePore’s aproach, Larson and Segal (1995) and Ludlow (1999) appear to have dropped the model-theory, thereby loosing its ability to describe the LSPRs. Since I’m not sure to what extent this move is endorsed by other proponents of an ‘Absolute Semantics’ position, I’ll call it ‘Neo-AS’.

In this approach, the semantics for a natural language is to be specified as a ‘T theory’, which is a theory about an object language L, framed in a metalanguage M. The T-theory is an axiomatizable theory (working via a bottom-up definition of a relation beween semantic values and parts of sentences) delivering a set of theorems of the form (3), where S is a name or description of an L-sentence, and p is an M-sentence with the same truth-value as the sentence referred to by S (adapted from Larson and Segal 1995:25, Larson and Segal 1997:179):

(3) S is true if and only if p.

It often happens that L and M are the same language, say English, in which case one gets T-theorems such as (Larson and Segal 1995:27, Larson and Segal 1997:181):

(4) *Jill knows Kate* is true if and only if Jill knows Kate.

Larson and Segal note that theorems of this sort give an appearance of triviality, but show that this is false because of the possibilty that L and M might be different. For example if L was French and M was English, the T theory could be usefully used to interpret French sentences, and they give an argument that it can be interpreted as a theory of what the French sentences mean. However one also might note at this point that there are plenty other kinds of devices that can implement a translation relation, such as Jackendovian rules of semantic intepretation, which advocates of AS don’t want to recognize as constituting semantic theories.

Technically, the proposed T theory works by setting up a relation Val between utterances and parts of utterances, and ‘semantic values’ such as T (for true) and various entities in the world, such as people. If for example the person Kate is named *Kate*, the theory will contain an axiom:

(5) Val(x, *Kate*) iff $x =$ Kate

And for an intransitive verb such as *ponders*, there will be an axiom such as:

(6) Val(x, *ponders*) iff x ponders

Then there will be a semantic composition rule for the subject-predicate relation:

(7) Val(T, [$_{\mathrm{S}}$ NP VP] iff, for some x, Val(x, NP) and Val(x, VP).

The mathematical flavor of these formulations suggests that they might deliver an account of the LSPRs, but further developments indicate that this is not intended, and won't be provided.

So we find that many of the T rules proposed in Ludlow (1999) make quite unconstrained use of the facilities of ordinary English. Here for example is the rule for the interpretation of *before*-conjunction from Ludlow (1999:120) (σ is a sequence of individuals used for evaluating quantifiers):

(8) Val(T, [$_{\rm S}$ TP1 before TP2], σ) iff Val(T, TP1, σ) before Val(T, TP2, σ)

In a conventional MTS account of tense, there would be a collection of time moments or intervals with various kinds of ordering imposed on them, enforcing entailments such as that if A happens before B and B happens before C, then A happens before C.[13] But there doesn't seem to be anything to immediately enforce such an inference in Ludlow's account. The mere absence of mathematicalese doesn't necessarily mean that a rigorous account of the LSPRs can't eventually be provided; the relevant virtue of mathematicalese is that it is an idiom crafted over millenia to support sound reasoning about complicated situations, so it is merely an aid rather than a necessity for rigor, but it does mean that the rigorous account of the LSPRs hasn't yet been provided, since the logic of the concepts in the metalanguage has not yet been nailed down solidly enough to support any kind of reasoning.

That this is policy rather than oversight is clearly indicated by the discussion of inference in Larson and Segal (1995:74-75), where it is proposed that the LSPRs are to be accounted for by rules of inference formulated in the metalanguage, constituting a 'logic module' that is part of the speaker's overall linguistic competence. Although this looks like a proposal to abandon the demonstrated results of the MTS account of the LSPRs in favor of the stalled Deductive Approach, this appearance could be avoided by doing model theory over the metalanguage. A more troubling issue, I think, is the question of exactly what the T Theory is now acheiving, stripped of its capacity to specify the LSPRs.

The problem is that all it seems to be doing is translating between formalized languages, a function that can be performed by many devices that are not generally regarded as semantic theories. Indeed, the overall structure of theory doesn't look any different from that of classic Representational Semantics: the syntactic representations produced by a grammatical theory are paired with expressions in a metalanguage, which is supposed to in some manner state their meaning and provide a basis for the LSPRs. Perhaps some claim could be made to the effect that the way in which the translation and statement of meaning are acheived in Neo-AS are in some way 'more correct', than in classic RS, but even this claim is undercut by a rather odd move that Larson and Segal make (which I don't find endorsed in Ludlow 1999).

They propose that the T Theory is internally implemented by speakers as a device that derives metalanguage sentences from object-language sentences. The first step towards this conclusion is the 'T hypothesis', which says that a native speaker's knowledge of the semantics of their language should be expressed as a T theory (Larson and Segal 1995:33, 1997:187):

[13] For an introduction to the mathematical aspects of such analyses, see Landman (1991).

(9) *The T hypothesis* A speaker's knowledge of meaning for a language L is knowledge of a deductive system (i.e. a system of axioms and production rules) proving theorems of the form T that are interpretive [ADA: that is, specify meanings] for sentences of L.

An uncertainty factor in the interpretation of (9) is exactly what 'knowledge of a deductive system' is suppposed to consist of. Larson and Segal's suggestion is that the brain computes T-sentences in a 'Language of Thought', and that the results of these computations are what is passed to the brain modules responsible for inference, etc. (Larson and Segal 1995:40, 1997:194).

On this construal, it is entirely clear that the T theory is functioning merely as a translation device from the object language L into the metalanguage M, which has become a form of internally represented Markerese in the general manner of Jackendoff, but without the detailed linguistic motivation for its syntactic properties. As a result, the account is open to the standard criticism of Markerese, that the Markerese sentences themselves require a statement of their truth-conditions (Lewis 1972:169). Given where Larson and Segal are starting, I find this a truly bizarre conclusion for them to come to, but I don't see any other way to interpret what they are writing here, and to accomodate their proposals about the 'logic module'.

Perhaps a partial reason for it is the observation that T-Theories can have psychological reality at something like level 1.5; from this one might derive the conclusion that a T Theory had to be entirely psychological. But the conclusion doesn't follow, due to the fact that a theory of language-world connections might be partly pscychological and partly not (Partee 1988). In particular, as Partee suggests, the composition rules could be psychological, but at least some of the reference-determining rules non-psychological (determined by causal links, social relations, etc.). The psychological basis for part of the theory is enough to explain the level 1.5 effects, it is not necessary to try to stuff the entire thing into the head.

Whatever conclusion we come to about Larson and Segal's total psychologization of the T theory, I think this discussion raises substantial questions as to what the goal of 'stating truth conditions' really amounts to. In a sense, a sentence in a language is already a perfectly good statement of its truth-conditions; all one can hope to do is re-state them, in a manner that does something that the original statement does not. In the MTS approach, what is acheived is an account of the LSPRs, and I think that many times, in concrete discussions of the semantics of some expression or constructions, people use the phrase 'the truth-conditions' as an equivalent for 'the entailments' (they mention some entailment, and then characterize it as one of the truth-conditions, which an adequate semantic theory must capture). The format of restatement can also be used to express hypotheses about language typology (Jackendoff's program, and, in a very different vein, the 'Natural Semantic Metalanguage' program of Wierzbicka and her colleagues (Wierzbicka 1996, Goddard 1998, Wierzbicka and Goddard 2002, and many other works), but there is a widespread conviction that simply reformatting the content of the sentence does not constitute doing semantics. Semantics, we are told, requires connecting language to the world, but how can we actually do that? Connecting sentences to metalanguage sentences that re-state the truth-conditions is the proposed program, but the practical consequences of doing this appear to be no different from

those of translation. Therefore the utility of T Theories for this purpose is unclear.

An additional confusion factor is that status of 'disquotational' knowledge in semantics. It is often claimed that an important part of people's semantic knowledge is their knowledge of T sentences such as (5) (e.g. Higginbotham (1988:30-31). Although people clearly do know such things, I question whether it is sensible to see them as central to semantics, because one could perfectly well imagine a species of clever but somewhat unreflective beings who had a complex language, but showed no signs of having any disquotational knowledge about it. For example, you could tell them that the beer was in the refrigerator, behind the milk, and they would succeed in going to find the beer, but wouldn't have words for *means* or *true*, and might for example be unable to translate sentences between languages, even though appearing to be in many respects able to use the languages. Their linguistic knowledge wouldn't be as rich as ours, at least in some respects, but it would seem odd to say that it had no semantics, or that it had semantics but that they didn't know it.

It is therefore not so clear that T Theories actually do anything useful, and, if they don't, LePore's critique of MTS loses much of its force, since the idea that MTS is incomplete is not very convincing if the proposed supplementation doesn't acheive anything that can't be done by a translation device. On the other hand, the job done by MTS, characterizing entailments, is clearly relevant to characterizing meaning, since one can think of the function of a (declarative) sentence as being to assign the world to a classificatory category, and the entailments of a sentence are a full statement of the position of that category in the classification system. I think this is a defensible concept of 'meaning', with obvious relevance to such issues as the adaptive utility of language, which are part of the motivation for connecting language to the world.

4 A 3rd Person view of Language and the World

If one accepts what has been said so far, one might be inclined to write off the idea of trying to connect language to the world, but this would be to ignore the striking fact that the successful program, MTS, seems to have been almost entirely motivated by it. Indeed, I suggest that one of the reasons why MTS apears to be less afflicted with arbitrariness than deductive approaches is that it benefits from a lot of implicit constraint from our intuitions about the language-world relationship, such as that sentences have truth-values, names refer to people and other sorts of entities, etc. In other words, our ideas about *how* language relates to reality have played a major and perhaps essential role in helping us build theories of the LSPRs that work (imperfectly, etc., in the usual manner of empirical inquiry), in spite of the difficulties adduced above in the attempt to 'officially' connect sentences to real world truth-conditions.

My suggested answer to this paradox is that the intuitions behind MTS derive from our innate metalinguistic facilities for monitoring, maintaining, and repairing communication (this is closely related to the points made at the conclusion of Partee 1988). Consider for example the predicate *true*. Work by Wierzbicka and her colleagues (Wierzbicka and Goddard 2002) indicates that this is a semantic universal, found in all languages, and not straightforwardly definable in terms of other words found in all languages, and therefore a semantic primitive. In English, one of the functions of

true is to acknowledge previous contributions to the discourse in order to contradict or qualify them. So this dialog seems rather incoherent, presumably because it violates the expectation that assertions are expected by default to constitute additions to the common ground:

(10) A: It's raining outside

B: # It's raining outside, but not in the catchment area

It's much better using *true* in the first clause:

(11) A: It's raining outside

B: That's true, but it isn't in the catchment area

In a more pedantic style, one could repeat the first clause as a *that*-complement of *true* without creating incoherence: 'It is true that it is raining outside, but . . .)'.

Another metalinguistic Wierzbickian primitive is *say* [*something*] (*about* [*somebody/something*]). This can be used to get information needed to repair a communication that isn't working. For example if somebody is going on about *Jane*, and what they're saying doesn't seem to make sense, you can ask 'who are you talking about (="saying something about")', and they might answer 'I'm talking about the Jane character from the Orson Scott Card novel *Xenocide*'. For English, when an *about* argument is included, there appears to be a constraint that what is said be intelligible, whereas no such constraint is imposed when the *about* argument is missing:

(12) a. The demon said (to Dante) 'Pape Satan aleppe'

b. # The demon said about Dante 'Pape Satan aleppe'

c. The demon said about Dante 'This guy needs a permit to get in here'

So the frame *say about* appears to involve some kind of notion of 'proposition', while the one without it involves a more inclusive notion of 'utterance', including ones that are inarticulate gibberish.

There doesn't appear to be a *refer* primitive, but a notion of reference can be expressed using *say about*:

(13) When the demon said 'this guy', he was saying something about Dante Alighieri

It would be interesting to investigate to explore various philosophy of language issues concerning truth, propositions and reference in terms of the exponents of the metalinguistic Wierzbickian primitives in different languages. Suppose we read on the website of a philosophy department that this week's seminar is 'The Ethics of Cannibalism', delivered by Fred Stokes, and it really is the case that somone called Fred Stokes is giving a seminar on that topic. Is the site saying something about Fred if, in spite of this superficial conformity with the facts, the information was actually planted by malicious hackers who aren't aware of Fred and his interests and merely chose the name and topic at random, in an effort to discredit the department? If the intuition held across cultures, it might reflect a universal connection between the *say about* and *because (of)*

primitives; along the lines that the appearance of 'Fred Stokes' in the utterance doesn't render the utterance *about* philosopher Fred unless its appearance is in some sense *because of* ihim. Investigating such issues with the Wierzbickian primitives might offer the prospect of getting information from many diverse cultures, including ones where well-trained philosophers of language are thin on the ground.

And the fact that such expressions are defensible as universal primitives indicates that, unlike the previously introduced hypothetical creatures without disquotational knowledge, humans do have cognitive facilities for reflecting upon, and linguistic facilities for commenting on, certain kinds of relationships between language and the world. One could regard MTS as a project of taking some of the intuitions grounded in these facilities, and adding some math to make a theory. It's interesting that this theory is quite successful in some areas, such as quantifiers, and arguably rather problematic in others, such as belief contexts; it is possible that that the innate facilities provide some support for a science of semantics, but not enough to do the whole job. Or perhaps the situation is just rather complex, and we haven't managed to unravel it all yet. There is also an interesting contrast between ideas which seem intuitively accessible, such as reference and the extensions of 1 place predicates, and those where more substantial mathematical understanding is required, such as two-place predicates and quantification.

Accepting the idea that our intuitions about truth and reference are based on an aspect of our psychology, a facility for perceiving relationships between utterances and the wider world, one might draw the conclusion that some sort of Anti-Realist position was justified. But this would be wrong. The fact that we have evolved so as to perceive these relationships is *prima facie* evidence that they exist (although not that everything we immediately perceive about them is true—the bug detectors of a frog do detect bugs, but also airplanes in the distance), and the fact that we can use them in a reasonably successful theory of the LSPRs corroborates their existence.

One of the more intriguing potential consequences of a Realist position is that there can be non-axiomatizable entailment relations, such as those of second-order logic, or branching quantifiers.[14] If, by reflecting on various features of the use of a language L, we were persuaded that some such theory were correct for L, and also assumed that the routinely detectable LSPRs (the ones that we take to count as data for linguistics) must be at least recursively enumerable, it would follow that there would be a gap between the LSPR data of linguistic semantics for L, and the complete facts of logical consequence for L, as implied by our intuitions of how the sentences of L (as conceptualized by us) relate to reality (as conceptualized by us).

As a result, there would be a divergence between what a hard-boiled empirical linguist might take as the best account of linguistic data (the axiomatizable theory that delivers the LSPRs), and what someone with Realist philosophical inclinations might take as the best account of the relationship between language and the world. And the Realist view might prove the more useful one in the long run, since rigorous linguistic theory might someday be able to embrace questions of what one might call an 'ecological' nature, such as those raised in Partee (1988) and Ludlow (in press). But equally, it seems to me that current day-to-day linguistic semantics can and therefore probably

[14] Hintikka (1974), empirically challenged by Fauconnier (1975).

should take an agnostic position with respect to issues such as Realism: linguistic semantics will of necessity primarily deal with how we conceptualize the world we're in, and the question of the status of the elements of the conceptualization (individuals, stages, properties, scenes, facts, amounts of stuff, flaws, average men, etc.) doesn't have to be decided in order to get on with the work.

Bibliography

Chierchia, G. 1984. *Topics in the Syntax and Semantics of Infinitives and Gerunds*. PhD thesis, University of Massachussetts, Amherst MA. distributed by the Umass Amherst GSLA.

Chierchia, G., and S. McConnell-Ginet. 1990. *Meaning and Grammar: An Introduction to Semantics*. MIT Press.

Chierchia, G., and S. McConnell-Ginet. 2000. *Meaning and Grammar: An Introduction to Semantics (second edition)*. MIT Press.

Chomsky, N. 2000. *New Horizons in the Study of Language and Mind*. Cambridge University Press.

Cooper, R., and T. Parsons. 1976. Montague grammar, generative semantics, and interpretive semantics. In B. Partee (Ed.), *Montague Grammar*, 311–362. Academic Press.

Dalrymple, M. 2000. *Lexical Functional Grammar*. Academic Press.

Davies, M. 1987. Tacit knowledge and semantic theory: Can a five percent difference matter? *Mind* 96:441–462.

Etchemendy, J., and J. Barwise. 1987. *The liar : an essay on truth and circularity*. Oxford University Press.

Evans, G. 1985. Semantic theory and tacit knowledge. In *Collected Papers*, 322–342. Oxford University Press.

Fauconnier, G. 1975. Do quantifiers branch. *Linguistic Inquiry* 555–567.

Fodor, J. D. 1985. Situations and representations. *Linguistics and Philosophy* 8:13–22.

Goddard, C. 1998. *Semantic Analysis, A Practical Introduction*. Oxford University Press.

Higginbotham, J. 1988. Contexts, models and meanings: A note on the data of semantics. In R. Kempson (Ed.), *Mental Representations: The Interface Between Language and Reality*, 29–48. Cambridge Universith Press.

Hintikka, J. 1974. Quantifiers vs. quantification theory. *Linguistic Inquiry* 5:153–177.

Hornstein, N. 1988. The heartbreak of semantics. *Mind and Language* 3:3–37.

Jackendoff, R. S. 1973. *The Role of Semantics in Generative Grammar*. MIT Press.

Jackendoff, R. S. 1983. *Semantics and Cognition*. MIT Press.

Jackendoff, R. S. 2002. *Foundations of Language*. Oxford University Press.

Johnson-Laird, P. N., and R. M. J. B. *et al*. 1993. *précis* of and peer commentary on *Deduction*. *Behavioral and Brain Sciences* 16:323–380.

Johnson-Laird, P. N., and R. M. Byrne. 1991. *Deduction*. Hillsdale NJ: Lawrence Erlbaum Associates.

Katz, J. J. 1966. *The Philosophy of Language*. New York: Harper and Row.

Katz, J. J. 1972. *Semantic Theory*. Harper and Row.

Katz, J. J. 1981. *Language and other abstract objects*. Oxford: Basil Blackwell.

Katz, J. J., and J. Fodor. 1964. The structure of a semantic theory. In *The Structure of Language: Readings in the Philosophy of Language*. Englewood Cliffs NJ: Prentice-Hall. first published in 1963, *Language* 39:170-210.

Kripke, S. A. 1975. Outline of a theory of truth. *Journal of Philosophy* 72:690–716. reprinted in R.L. Martin (ed) *Recent Essays on Truth and the Liar Paradox*, Oxford University Press.

Lakoff, G. 1970. Linguistics and natural logic. Technical report, Phonetics Laboratory, University of Michigan. also in Davidon and Harman (1972) *Semantics of Natural Language, Reidel*.

Lakoff, G. 1987. *Women, Fire and Dangerous Things*. University of Chicago Press.

Landman, F. 1991. *Structures for Semantics*. Kluwer.

Larson, R., and G. Segal. 1995. *Knowledge of Meaning*. MIT Press.

Larson, R., and G. Segal. 1997. Knowledge of meaning and theories of truth. In P. Ludlow (Ed.), *Readings in the Philosophy of Language*, 179–199. MIT Press. excerpted from Larson and Segal (1995).

LePore, E. 1983. What model theoretic semantics cannot do. *Synthese* 167–187.

Lewis, D. 1972. General semantics. In D. Davidson and G. Harman (Eds.), *Semantics of Natural Language*, 169–218. Dordrecht: D. Reidel.

Link, G. 1998. *Algebraic Semantics in Language and Philosophy*. Stanford CA: CSLI Publications.

Ludlow, P. 1999. *Semantics, Tense and Time*. MIT Press.

Ludlow, P. in press. Referential semantics for i-language. In . Blackwell Publishers. URL:
`http://www-personal.umich.edu/~ludlow/online_pubs.htm`.

Marr, D. 1982. *Vision: a Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: W.H. Freeman.

McCarthy, J. J. 2002. *A Thematic Guide to Optimality Theory*. Oxford University Press.

Partee, B. H. 1988. Linguistic facts and psychological facts. *Mind and Language* 43–52.

Peacocke, C. 1986. Explanation in computational psychology: Language, perception and level 1.5. *Mind and Language* 1:101–23.

Suppes, P. 1976. Elimination of quantifers in the semantics of natural language by the use of extended relation algebras. *Revue International de Philosopnie* 117-118:243,259. reprinted in Suppes (1991:191-205).

Suppes, P. 1979. Variable-free semantics for negations with prosodic variation. In *Essays in Honor of Jaakko Hintikka*. Dordrecht: D. Reidel. reprinted in Suppes (1991:207-216).

Suppes, P. 1982. Variable-free semantics with remarks on procedural extensions. In T. Simon and R. Scholes (Eds.), *Language, Mind and Brain*. Hillsdale, N.J.: Lawrence Erlbaum. reprinted in Suppes (1991:229-243).

Wierzbicka, A. 1996. *Semantics, Primes and Univerisals*. Oxford University Press.

Wierzbicka, A., and C. Goddard. 2002. *Meaning in Universal Grammar: Theory and Empirical Findings, 2 vols*. Benjamins.

Zwarts, J., and H. Verkuyl. 1994. An algebra of conceptual structure: an investigation into jackendoff's conceptual semantics. *Linguistics and Philosophy* 1–18.