Logic and Natural Language.

1. Natural language and logical analysis
   1.1 Compositionality and Frege's puzzles
   1.2 Montague Grammar
   1.3 The nature of meaning-postulates
2. Towards a new theory of meaning and interpretation
   2.1 Anaphora in Montague grammar
   2.2 Discourse
   2.3 The fallacy of misplaced information
3. Theories of dynamic interpretation
   3.1 Discourse representation theory
   3.2 Situation semantics
   3.3 Quantification and anaphora
   3.4 Dynamic Montague Grammar
4. The state of the art
   4.1 Open problems
   4.2 Cognitive science
5. Selected References

1. Natural language and logical analysis.

Logicians have always found inspiration for new research in ordinary language that we use on a daily basis and acquire naturally in childhood. Whereas the logical issues in the foundations of mathematics motivated the development of mathematical logic with its emphasis on notions of proof, validity, axiomatization, decidability, consistency and completeness, the logical analysis of natural language motivated the development of philosophical logic with its emphasis on semantic notions of presupposition, entailment, modality, conditionals and intensionality. The relation between research programs in both mathematical and philosophical logic and natural language syntax and semantics as branches of theoretical linguistics has increased in importance throughout the second half of the twentieth century. This chapter reviews the development of a particularly interesting and lively area of interaction - the semantics of natural language. This emergent field is paradigmatic of the fruitful interaction between logic and linguistic research for the development of
empirically, cognitively adequate models of reasoning with partial information, sharing or exchanging information, dynamic interpretation in context, belief revision and other cognitive processes. A brief historical note on how the attitudes of logicians towards natural language changed through the twentieth century provides our starting point.

In the early developments of modern logic, roughly in the first half of the twentieth century, logicians tended to be rather suspicious of natural languages. Alfred Tarski, for example, contended that natural languages were hopelessly confused, because one can simultaneously assert something and claim of it that it is true, as in “I am telling you the truth”. He argued that a sharp distinction should be made between the object language to be truth-functionally interpreted, and its meta language in which interpretations and truth conditions are specified. Failing to make this distinction in natural languages gives rise to self-referential truth-functional paradoxes like the famous liar-paradox "I am lying", which expresses false information, if true, and if false, true information.(Tarski, Alfred (1956)). Tarski’s important contribution to logic defined the notion of truth in a model in terms of assignments to variables and satisfaction of formulas, where set theory is chosen as meta language to interpret the language of first order predicate logic. In a similar vein, Bertrand Russell had advocated the logical reform of ordinary language in order to clarify which expressions established reference and which did not (Russell, Bertrand. (1912)). A definite description like “the butler” needed such a reform, according to Russell, to reveal its essential quantificational nature, despite its seemingly referential character. Russell’s account of a statement “The butler killed Jack” reanalyzed it as (a) claiming existence of a butler, (b) claiming that there was no other butler, and (c) asserting that he killed Jack. Accordingly, any model in which there was no butler or more than one would falsify this
statement. Opposed to such logical reform of natural language and its inherent limitation to descriptive, truth-functional use of language was the Oxford school of ordinary language philosophers, led by John Austin, Peter Strawson, Paul Grice and John Searle. Although there are important differences between their views, they shared an interest in the way people use language to change the world and share information. Strawson argued that “The butler killed Jack” could simply not be evaluated in a model that had no butler or more than one, pleading for a proper account of presuppositions as conditions that made truth functional evaluation of statements possible. Grice gained fame for his view that some interesting forms of inference are not logical in nature, but may be cancelled upon change of context or addition of information. His maxims, rules of cancellable inference, captured certain aspects of proper conduct in sharing information, such as requiring that you should not state the obvious and provide sufficient relevant information. Austin and Searle developed the theory of speech acts with its core notion of felicity conditions, instead of truth conditions, for expressions, like commands or questions, that have no primary descriptive content and whose semantic behavior escapes truth-functional analysis.

Once mathematical logic had established a firm footing as an academic discipline after World War II and was branching out into recursion theory, model theory and set theory, some logicians and philosophers of language gradually developed an interest in a more systematic logical exploration of aspects of natural language. A host of intensional logics were developed, analyzing forms of reasoning with modalities or tenses in models with possible worlds. The development of formal grammars, automata theory and related mathematical models of natural languages in the then nascent academic field of linguistics merged some of the traditional concerns of logic and philosophical
theories of meaning with empirical linguistic inquiry into the syntax, semantics and pragmatics of natural languages. Though debates between philosophers, linguists and logicians were often acrimonious at the time, the changing climate at the turn of the century now favors cooperation among logicians, computer scientists, linguists and philosophers, fusing these diverse strands of traditional semantic inquiry into a common research agenda of how natural language is used to communicate information effectively.

1.1 Compositionality and Frege’s puzzles.

Gottlob Frege, the founder of modern logic, provided a core foundational principle for contemporary semantic theory by requiring that the meaning of an expression should be a function from the meaning of its parts and the way in which they are put together. This Principle of Compositionality serves as a major methodological constraint on the interface between the syntax, which generates a (fragment of) a natural language, and its semantics, which takes the form of either a recursive procedure to translate its expressions to formulas of a formal language or a set of rules that specify how its expressions are given meaning in a model. For a system of interpretation to be compositional, the syntax of the object language needs to be formulated in such a way that the semantic properties of each syntactic category are completely determined by it. Given this methodological requirement of Compositionality on a theory of meaning and interpretation, the two semantic puzzles which preoccupied Frege still constitute major foundational problems that direct much of contemporary linguistic and philosophical theories of meaning and interpretation. The first puzzle concerns the information expressed in identity statements with
coreferential noun-phrases (NP). The Fregean discussion is based on the question why

(1) Hesperus is Phosphorus

once was an informative identity statement to the Babylonian astronomers, who learned from empirical observations of the stars that (1) was true, whereas

(2) Hesperus is Hesperus

was never informative, eventhough (1) and (2) are both true statements and the NPs, all proper names, corefer to the one and the same object, the planet Venus. If coreferential expressions have the same semantic value, they must be substitutable for each other in any context without affecting its semantic value. But how can (1) then be informative, while (2), in which a coreferential expression is substituted, is completely uninformative? If a semantic theory is to account for such facts, it must allow for coreferential expressions to differ in semantic value. For this purpose Frege introduced the fundamental distinction between the reference (Bedeutung ) of an expression and its sense (Sinn ). Different proper names and other referential NPs may refer to the same object, but they could still differ in their sense, as sense determines reference and not vice versa. Identity statements are informative when they contain expressions with different senses, and they are true when their NPs are coreferential. Conditions of 'informativeness' hence cannot be identified with or reduced to truth-conditions. Perhaps there is more to the semantic value of an expression beyond its sense and reference, like its psychological associations, connotation or 'color', but that part of its meaning is considered subjective and should be disregarded in semantics, according to Frege, for it cannot be the source of publically communicable information.

The sense of an expression determines its reference in different situations, but even when the reference of an expression in every situation is determined, this
does not fix its sense uniquely. If we assume that the reference of a sentence is
its truth value, two sentences that necessarily have the same truth value in all
situations, e.g.

(3) Robin won the race
(4) Everyone who did not compete or lost in the race has done something Robin did not do

still differ in their Fregean sense. Similarly two distinct logical tautologies
which are both necessarily true may have different senses, and convey different
information depending on the situation in which they are used. The way in
which the sense is expressed also determines its meaning, which Frege called its
way of being given (Art des Gegebenseins), a notion still requiring clarification.
If the semantics of natural language is to account for coreference, inference,
efficient information sharing and reasoning, it should give a satisfactory
analysis of the Fregean notions of sense, art des gegebenseins and reference in
compositional semantics.

The second problem Frege presented as a central question to a theory of
meaning is related to the first one about informative identity statements. If such
statements or any other two statements with the same truth value are embedded
as sentential complements of verbs, the resulting statements may differ in truth
value. For instance,

(5) Robin believes that Hesperus is Phosphorus
(6) Robin believes that Hesperus is Hesperus

(5) may be false, whereas (6) must be true, even if Robin knows nothing of
Babylonian astronomy, or if he is not even aware of what the name 'Hesperus'
refers to. For Frege this meant that sentences embedded in that-clauses do not
refer, as they ordinarily do, to their truth value, but refer indirectly, i.e. they
refer to their customary senses and constitute opaque contexts. Substitution of
coreferential or extensionally equivalent expressions in such opaque clauses does not necessarily preserve the truth value of the entire sentence. Only if it is assumed that the believer knows that two NPs are coreferential can they be substituted in her belief reports. Ordinary predicate logic cannot account adequately for these puzzles, as two of its basic laws fail in such opaque contexts:

(i) substitution of logical equivalents may not preserve truth value

(ii) existential generalization may not be valid

The first problem was explained above, and the second one means that one cannot infer from a referential expression used in an opaque context that it actually has a referent. If John believes that the spy is watching him, even if someone is indeed watching him, he need not be a spy. It may even be that no one is watching him, but John erroneously believes someone, whom he believes to be a spy is watching him. If two sentences differ in their sense, they express different thoughts, as Frege would say. But to provide a full fledged compositional semantic analysis of these sense differences in terms of their information content still constitutes a major driving force of current research. It requires a satisfactory account of equivalence of 'information content', sufficiently fine-grained to explain when a statement expresses new information to someone in a particular context, which depends in part on the information that is already available to him.

As a first attempt at formalization of Fregean senses, Rudolf Carnap defined the intension of an expression as a function from a set of indices to the extension of the expression (Carnap, Rudolf (1947)). The indices could be given various kinds of interpretations; Carnap thought of them as indexing a set of states of affairs. But the later conception of them was based on Saul Kripke's semantics of modal logic, i.e. as possible worlds (Kripke, Saul (1963)). Carnap introduced
the notion of an *individual concept* as intension of proper names and referring expressions, as a solution to the Fregean informative identity puzzle. Coreferring NPs would differ in their meaning, if they are interpreted as individual concepts by different functions from indices to individuals. Different occurrences of the same referring expression would, however, have to be interpreted as having the same, constant reference. The problem is now transferred to the problem of telling functions apart, which has a clear set-theoretic criterion: functions are identical just in case they assign the same values to each argument. One consequence of this set-theoretic criterion of function identity is that all mathematical and logical truths together with all analytical ones are interpreted on a par by the same constant function, hence they have the same intension and are still substitutable in all contexts, even opaque ones. This leads to the *problem of logical omniscience* for accounts of belief reports that rely on possible worlds: assuming you are rational, if you believe any contingent truth, you must believe any of its logical consequences, including all tautologies or necessary truths (cf. Stalnaker, Robert (1984)).

The main difficulty appears to be to make precise according to which constraints our imagination creates alternatives which differ from our actual world in the relevant ways, but still resemble it sufficiently in other important ways. The desire to keep the meaning of the descriptive vocabulary of our language constant plays an important role here. It is significantly harder to imagine what the world would be like, if, for instance, a bachelor may be married or conjunctions would not require both conjuncts to be true, than it is to imagine what the world would be like, if it happened to be raining. The underlying issue is how to account for the difference between a change in the world versus a change in the meaning of an expression, while doing justice to the fact that both the world and the language are constantly changing.
1.2 Montague Grammar.

Montague Grammar, developed in the late sixties by the logician Richard Montague, provided a major step towards a compositional theory of interpretation of ordinary language, since it specified in all required detail how the semantic value of a sentence could be computed from its syntactic derivation (cf. Partee, Barbara (1997)). The form Compositionality took in the fragments of Montague Grammar is sometimes referred to as *rule-by-rule compositionality*, as each syntactic rule is mirrored by a semantic one, which specifies how the meaning of the input to the syntactic operation determines the meaning of its output. A central issue for the compositional theory of interpretation was to provide a uniform function/argument structure to sentences with quantified NPs and sentences with simple referential ones. Montague's insight was to interpret all NPs, independently of syntactic form, as denoting a set of properties of individuals, doing justice to Compositionality, and reducing them if possible to the familiar Fregean interpretation using standard first order representations of existential and universal quantifiers. Predicates deriving from the interpretation of verb-phrases (VPs) are then simply interpreted as properties of individuals which either are or are not in the set of properties interpreting the NP in subject position (Westerstahl, Dag, (this volume)). Forms of quantification in natural language which are known to not be expressible or definable in terms of the first order logical quantifiers can be similarly interpreted by this higher order notion of generalized quantifiers. E.g. *most students* is not first order definable, since its interpretation requires a one-to-one mapping between two sets dependent on a well-ordering by cardinality
Richard Montague initiated this higher order analysis of quantification for the logical quantifiers in his famous paper 'The Proper Treatment of Quantification in Ordinary English' (Montague (1974), p. 247-270), and recent research in natural language semantics has developed his insights considerably for linguistic purposes.

In the PTQ-article Montague defined the interpretation procedure for any expression of the natural language fragment in two steps: first the translation of the syntactically disambiguated natural language expression to an expression of the formal language of the intensional logic, and second, the interpretation of that formal expression in the modeltheoretic semantics. The formal language hence intervenes in this PTQ-framework between the natural language and the specification of meaning in terms of truth conditions in a possible worlds model. Montague himself emphasized that there is no real need for such an intermediate level of 'logical form' representing the meaning of an expression of natural language. In another article, 'English as a formal language' (Montague (1974), p.188-221) he defined the semantic interpretation directly as a mapping from the syntactic structure of the English expressions to appropriate modeltheoretic objects and functions. The reason why the PTQ-approach with indirect interpretation via a formal language has prevailed over this direct approach is largely practical. It is easier to see, for instance, the quantifier scope dependencies as differences in the linear order of quantifiers in a formula than to follow the notation of complex modeltheoretic functions assigning values to variables and all of the alternatives of such assignments. It should, however, not be lost out of sight that formulas merely encode their semantic interpretation in modeltheoretic terms. We can in principle characterize any number of languages which can be used to intervene between the natural language and the
It is well-known that the intensional logic used as the language of 'logical form' in PTQ could be replaced without loss of expressive power by a language in which quantifiers could range over reference-points (worlds at a certain time) (cf. Gallin, Daniel(1975)). This seems to have initiated a research trend of fleshing out the formal language with different parameters which initially belonged to the modeltheoretic realm. In the newest dynamic theories of interpretation elements of the non-linguistic context, as well as variable-assignment functions themselves may be quantified over (cf. section 3 below).

The general theoretical question remains at this point whether a many-sorted first order predicate logic would ultimately suffice for the semantics of natural language or whether it still needs to be enriched with tools or techniques of a higher order logic. The desiderata of axiomatizability, decidability and completeness for the logic of natural language play still an important background role in such new developments.

1.3 The nature of meaning-postulates

In PTQ meaning-postulates are formulas of the intensional logic which are true, not only in all possible worlds, but in all possible models. In other words, by defining a set of meaning-postulates one characterizes which among all logically possible models provide a plausible interpretation for the natural language interpreted. They were, for instance, required to capture the necessary truth of analytic statements, i.e. statements which are always true due to the meaning of their descriptive vocabulary, such as *Bachelors are unmarried*. But
meaning-postulates were designed to do a quite diverse number of jobs, e.g. the reduction of intensional formulas to extensional ones in contexts where existential generalization and truth-preserving substitutivity of coreferential expressions held. They also guaranteed that proper names are Kripkean rigid designators, i.e. referred to the same individual even in opaque contexts. As meaning-postulates were only required to be wellformed formulas of the intensional logic, there was no constraint on the kinds of tasks they could be designed for. The Montegovian strategy was to hardwire all semantic structure one might ever wish to use into the modeltheoretic rules and then weed it out by meaning-postulates whenever appropriate. For instance, even a simple extensional sentence like

(7) Mary walks

required higher-order and intensional types of logical expressions, computing its truth value at each possible world by determining of all properties which ones applied to a function from all worlds to Mary at each world. By making all functions in PTQ total (i.e. defined for every argument of the appropriate type), one would have to determine of anyone who could walk, whether he did walk at each possible world, before one could determine whether (7) was true. The inefficiency and computational intractability of admitting only total functions made this strategy rather unattractive for any computational implementation of the inferences that PTQ could otherwise account for so beautifully.
2. Towards a new theory of meaning and interpretation

2.1 Anaphora in Montague grammar

In PTQ pronouns are bound by using rules which syntactically replace the first occurrence of an indexed pronoun in a sentence, in a common-noun phrase with a relative clause or in a VP by the antecedent NP which semantically binds any subsequent pronoun bearing the same index in that sentence. These 'quantifying in' rules apply to any kind of NP - proper name, existential or universal - , and any complex sentence, nested relative clause or complex VP. If the NP introduced by such a rule is quantificational, its translation has scope over any scope-bearing expression already present in the sentence, common-noun phrase (CN) or VP quantified into. This technique provides a universal compositional method of accounting for scope-ambiguity in natural language by syntactic disambiguation. The traditional ambiguities of universal and existential NPs are thus accounted for on a par with the ambiguities in intensional contexts - the *de re* and *de dicto* readings of NPs. The example discussed in 1.2 *John believes that a spy is watching him* would be generated by quantifying in both NPs *John* and *a spy*, if there actually was a spy watching John. But if the spy were a mere figment of John’s imagination, the interpretation would proceed directly, composing *a spy* with the property *watch him* and not allow for existential generalization, inferring existence of an actual spy, outside John’s belief world. From a linguistic point of view these rules of quantifying in overgenerate spurious ambiguities and fail to account for some essential differences in anaphoric potential of the three different kinds of NPs. Since the rules apply to
any NP, for each NP in a sentence there are always at least two syntactic derivations of that sentence, one direct and one indirect derivation using a quantifying in rule. Even for the extremely simple extensional (7) PTQ would generate at least two non-trivially different interpretations, one directly composing the subject with the VP and the other quantifying in the subject into a proposition with a free variable. Semantically these distinct derivations would be logically equivalent, so these syntactically driven differences have no semantic effect. If this is the price one has to pay for the compositionality of the semantic interpretation, it would be relatively harmless. But despite its universality, there are very natural interpretations of NPs in intensional contexts which cannot be accounted for by quantifying in rules. Examples of such sentences started emerging in the philosophy of language as early as 1962 (Geach, Peter, (1962)) and are based on the fact that it is not possible to evaluate an NP at a possible, non-actual world, retain its value while accessing another world. For instance, in (8)

(8) John tries to catch a fish and wants to eat it

one would like to interpret 'a fish' *de dicto*, with the fish John is trying to catch and then use that fish as referent of the subsequent pronoun. For such a coreferential *de dicto* reading the quantifying in rule would have to be applied after the two intensional VPs are conjoined, which would produce only a *de re* reading. This would be counterintuitive as it allows for the possibility that John does not know that what he wants to eat is a fish, in a situation where (8) is true. Further examples which demonstrate essential limitations on the technique of quantifying in are called e-type pronouns, after Gareth Evans.

(9) If Mary dates a guy her parents disapprove of, they will make his visit miserable.

(10) Every woman who loves a man kisses him.
In (9) the pronoun *his* refers to any guy Mary dates and her parents disapprove of. In (10) *him* refers to any man loved by a woman. The readings PTQ will generate with such bound pronouns necessarily give widest scope to the existential antecedent NPs, contrary to our intuitions which tell us that neither (9) nor (10) must be interpreted as being about a specific existing individual. Another objection already mentioned to the universal treatment of NPs by the quantifying in rules is based in the fact that universal NPs in relative clauses cannot bind pronouns in the VP, but existential ones and proper names do, cf. (11)-(14).

(11) * A woman who kissed every man left him
(12) A woman who kissed a man left him
(13) A woman who kissed Jim, left him
(14) No woman who kissed a man left him

Quantifying in universal NPs hence needs to be restricted in a principled way to prevent such bindings as in (11) to arise, but the PTQ rules are entirely unrestricted. In the generative linguistic literature a host of facts concerning the difference in anaphoric potential of the three kinds of NPs has been reported, which any proper semantic theory of anaphora should take into account. Just to mention a few of the most interesting facts, consider the anaphoric dependencies in (15)-(17).

(15) His mother loves John
(16) His mother loves a/every man
(17) A woman loves her

Proper names allow for backwards anaphora, whereas existential or universal NPs generally do not, as we see in (15) and (16), although the pronoun can still be bound by another antecedent or be interpreted deictically.. In (17) PTQ
would allow for a coreferential interpretation of the subject and the pronoun, if the subject were quantified in, contrary to our intuitions. In (18) inverse scope

(18) A flag was hanging from every window

of the NPs shows that the syntactic linear order of the NPs may be inverse to their preferred semantic interpretation. Such clear linguistic facts concerning the different anaphoric potential between NPs should be explained in a satisfactory and universal account of anaphora, which departs more radically from some of the fundamental assumptions of variable binding in formal languages which PTQ inherited from its logical tradition. What must be revised is the intrinsic connection between scope and variable binding.

The problem of spurious ambiguities was tackled first in (Cooper, (1983)) by weakening compositionality in a precise and constrained way. His grammar was allowed to generate ambiguous sentences, and it did not include any quantifying in rules. Hence meaning was not completely determined by syntactic form and the grammar did not embody the rule-by-rule compositionality of PTQ. Instead, the semantic interpretation must choose for any NP it evaluates whether to determine its semantic value immediately or to put it 'in storage', placing it on hold in a stack of NPs whose interpretation is deferred to a later point. When a stored NP is retrieved for evaluation, it receives scope over everything that is already interpreted at that stage of the semantic interpretation. This NP-storage technique circumvented one linguistic objection to quantifying in rules, but it still required overgenerating in the syntax, where gaps or empty NPs are generated to be bound by *wh*-quantifiers, but may be filtered out in the semantics, if the quantificational structure is deviant. No appeal is made in Cooper's framework any syntactic notion of illformedness in such semantically uninterpretable strings. The problems with cross-world quantification and the *e*-type pronouns remain, however, since this framework has no means to keep
track of information already obtained about the referent of a pronoun, lacking any notion of context and dynamic binding. The problem of logical omniscience outlined in section 1.1 demonstrated that the logical mechanics from predicate logic would require substantial revision, if they were to simulate how context, prior information and external situation of use might be used to draw inferences from given information. The characteristic topic-neutrality of logic is seen as one of the sources of the problem of logical omniscience in possible worlds semantics. A beginner in logic often experiences how difficult it is to rid oneself of the natural topicality and context dependent aspects of reasoning. For instance, learning the disjunctive law, disjoining a proven formula with any arbitrary formula, one must learn to consider formulas that may be completely irrelevant to the first one. To make more precise what it means for two sentences to be about the same topic, or to be relevant to each other, a more sensitive notion of the informative content of a sentence in a context is required. The requirement that all semantic functions be total should also be relinquished, for definite referential NPs could fail to pick out a referent at some worlds or presuppositions of other expressions could fail, and sentences with uninterpreted constituents should be neither true nor false (pace Russell and vivat Strawson). Yet another problem of a more metaphysical nature faced possible world semantics. What are they, if not mere formal entities that serve to distinguish contingent from necessary truths? If an individual at two different worlds may have no properties in common, in what sense is it still one and the same individual? Kripke, along with other possible world semanticists reviving an Aristotelian essentialism, argued that some properties were essential to an individual, most notably the properties concerning its origin. Personal identity would only break down when such essential properties would be lost (cf.
Kripke, Saul (1972/1980)). Other possible world semanticists, especially David Lewis (cf. Lewis, David (1983)), took the extreme opposite view and argued that individuals can never be the same across possible worlds, but are rather related by a much weaker counter-part relation and need not have any common properties. The philosophical debate continues to be lively and provides a plethora of philosophical options on choices of primitives, views on identity of individuals, properties and propositions (cf. Almog, Joseph et al. (eds.) (1989), and Stalnaker, Robert (1984)). But the need for possible worlds in the semantics of natural language is now also disputed by the theories of dynamic interpretation. Although intensional contexts and opacity phenomena obviously require tools beyond mere extensional first order predicate logic, a reinterpretation of possible worlds is called for as possible states of information an interpreter can be in. Modality should accordingly range over possible updates of the given information, not over metaphysically possible states of the world. This epistemic turn of natural language semantics opens up the way for dynamic interpretation, where the core concept is updating a given information state, by using a sentence in discourse to be interpreted as an instruction to add new information or constrain the given information in a particular way. By shifting away from the Fregean focus on truth functional meaning to a theory of informative content of sentences in discourse, but still characterizing various inferences as truth preserving operations on given information, the semantics of natural language is now merging the traditional issues of truth functional meaning with pragmatic issues of context-dependent interpretation, interpretation as action between people and sharing of information.

2.2 Discourse
The PTQ technique of quantifying in to obtain wide scope readings with bound pronouns was already shown to have some inherent shortcomings from a natural language point of view. If we would include anaphoric dependencies that arise between sentences into our consideration, we see that no simple generalization of the quantifying in rules can ever account for such forms of binding in discourse. Binding of pronouns is not postponed until we reach the end of a sequence of sentences. It is rather a more dynamic process, where the interpretation of a sentence in a sequence is constrained by what information is gained from preceding sentences and whatever common background is supposed, and in turn constrains the ways in which the information exchange may be continued. In discourse, universal NPs are again more limited in their anaphoric potential than either existential ones or proper names, as in (19)-(21).

(19) *Every woman kissed a man. She left.
In (19) the pronoun she cannot be referentially dependent upon the universal NP every woman in the preceding sentence. Existential NPs, definite descriptions or proper names may corefer with pronouns across sentences, as in (20) and (21).

(20) A/The woman kissed a man. She left.
(21) Jane kissed a man. She left.
Since the PTQ quantifying in rules do not apply to sequences of sentences, this characteristic difference in anaphoric potential of universal versus referential NPs cannot really constitute a principled objection to quantifying in. But if these rules were generalized to apply to sequences of sentences, quantifying in would not always give the desired result. For instance, to generate a bound reading of

(22) Only one student is reading. He is sitting at the table.
the NP *only one student* should be quantified into the sequence 'he₁ is reading. he₁ is sitting at the table'. But then the interpretation is weaker than intuitively needed, requiring only that there be precisely one student who has both properties of reading and sitting at the table. It does not rule out other reading students, as (22) does seem to require. Hence a generalization of the quantifying in rules will not make the correct predictions for bound pronouns in discourse. (cf. Gamut (1990), chapter 8). But the similarity between binding pronouns across sentences as in (19)-(22) and within sentences as in (11)-(14) is striking. Any semantic theory of binding should not only account for the difference in anaphoric potential of the three kinds of NPs, but also admit generalization to the interpretation of pronouns in discourse. This insight has been a driving force behind the development of the theories of dynamic interpretation discussed in section 3.

2.3 The fallacy of misplaced information

The information one may get from interpreting an expression depends on a host of different parameters. Consider an utterance of a simple sentence in (23).

(23) My husband and I invited her for dinner today.

The direct situation of use determines for instance the reference of indexical expressions like 'I' and 'today', but common sense knowledge may be necessary to understand what a dinner-invitation means, and linguistic knowledge will help determine who 'her' could refer to. Informative content arises as a relation between these parameters, the syntactic form of the expression used, its meaning, and the external world. In Barwise and Perry (1983) meaning is considered such a dependency of many parameters. They stress that a sentence
may be used in different situations to convey different information, which is why communication in natural language is so efficient. The sentence used in (23) could be used in a different situation to express completely different information. Sentences may be used to describe parts of the world, situations, and the reference of a sentence should be the set of such described situations, more like Austin’s view, and not merely a truth value as Frege would have it. The meaning of the sentence partially determines which situations it can be used to describe. But other contextual parameters come into play when we interpret the use of the sentence in a particular situation as giving us information about a particular topic. To assume that the entire informative content of the use of an expression is determined solely by its interpretation is what Barwise and Perry call the fallacy of misplaced information.

The performative hypothesis, popularized in pragmatic theories of meaning (e.g. Searle, John (1970)), proposed to analyze any sentence as subordinate to a performative first person verb - e.g.

(24) It is raining.

(25) I inform you that it is raining.

This is a clear example of the fallacy of misplaced information, as it attempts to put information about the situation of use overtly into the described situation. Similarly, the Russellian analysis of definite descriptions, which analyzes any definite description as referring to any unique individual who satisfies the describing properties is prone to the fallacy of misplaced information, since definite descriptions can be used to give us information about the situation of use, which is distinct from the described situation. Denying that proper names can be used in contexts to contribute to its interpretation, as many direct reference theorists have claimed along with Kripke, is another instance of the fallacy. If I introduce myself, this is a meaningful communicative act, because
the addressee gets the information how to call me, and knows henceforth how to refer to me. If names had no meaning beyond referring directly to their bearers, it would not be possible to explain how we do extract useful information from such an introduction. Informative identities are informative because the two coreferring expressions each contribute a different property of being so named. If I report to you Jane's belief that her husband is happy with the sentence

(26) Jane believes Jim is happy

I invite you to draw the inference, which is an implicature, revokable upon further information, that Jane herself would report this belief using the proper name Jim. If I had used instead of (26),

(27) Jane believes her husband is happy

a different implicature would be invited. But both (26) and (27) may be true even when Jane denies that Jim is her husband. A Fregean theory of reference could not ever account this, as it avoids context-dependent parameters of language use. In a dynamic theory of interpretation the shifting reference of indexicals and demonstratives should be accounted for in constructing a context from which conclusions may be drawn which could also contain indexical expressions. The central assumption of classical logic, that context dependence should be avoided or eliminated, is no longer viable, once reasoning of human interpretors in natural language has become the target of investigation.
3. Theories of dynamic interpretation

3.1 Discourse Representation Theory

Discourse Representation Theory (DRT) was developed by Hans Kamp (Kamp, 1981), van Eijck, Jan and Hans Kamp (1997)) partly in response to the anaphora problems Montague Grammar was facing. But it was also motivated by a more philosophical concern with the nature of reference, meaning, inference and interpretation as traditional inquiry of philosophers of language. A closely related theory of dynamic interpretation File Change Semantics was developed independently by Irene Heim (Heim, Irene (1982)). The main ideas and concepts of DRT are presented here in the context of the presentation of natural language semantics. For a more comprehensive introduction the selected references contain some excellent expositions of the theory.

The core claim of DRT is that interpretation should be considered a dynamic process, in which discourse representation structures are constructed representing the information and the anaphoric dependencies expressed in a sequence of sentences. Such information is true in a model just in case there is an structure-preserving embedding of the reference markers and the descriptive conditions relating them, which constitute the representation, into that model. The conditions in the representation arise incrementally from the interpretation of the sequence of sentences by application of the construction rules for DRSs. A condition is a property or relation with an appropriate number of reference-markers as arguments, which function in certain respects like context dependent referring variables. DRSs consist of different levels of conditions, where a reference marker is accessible from a lower level only if it is declared at a
higher level. Negation, modality, conditionals and quantifiers create deeper levels of embedded structure in the DRS.

The DRS-construction rules require that a proper name introduces a reference marker in the top level of the representation, which remains accessible to any lower level, thus capturing the semantic property of names that they always take widest scope or refer rigidly to one and the same referent no matter what context they occur in. An indefinite NP introduces a new reference marker into the given level, which may be a lower one, and the predicate in its CN is attached as property of the reference-marker. Definite NPs are treated differently; they must be identified with an accessible reference-marker present in the given or any higher level of the representation. Since pronouns are definite NPs too, their reference-marker is unified with the accessible reference-marker of their antecedent NP. Clauses with universal NPs force a split of the DRS into two levels, where the information in the universal NP is represented as a property of a new reference-marker in the first deeper level, and the information expressed in the remainder of the sentence by conditions in a new, deeper, subordinate level. The embedding conditions of such a split of levels requires that every verifying embedding of the conditions in the first deeper level can be extended to a verifying embedding of the conditions in the next deeper level. Some illustration of these DRS-construction rules are given below, and its analysis of (12) which constituted a problem for PTQ's account of anaphora is presented.

(28) A man came in. He sat down.

An indefinite NP is represented by introducing a new reference marker \(x\) and attaching the CN as a property of \(x\), and representing the remainder of the sentence as a property of \(x\) too. The pronoun which is anaphoric to the
indefinite NP is represented by its own reference marker $y$ and identified with $x$, the marker for its antecedent.

(29) $\begin{array}{c}
x \\
\text{man (x)} \\
\text{come in (x)} \\
\text{sit down (y)} \\
y = x
\end{array}$

The DRS in (29) is true in a model $M = < D, I >$ (where $D$ is a domain of individuals and $I$ the usual set-theoretic interpretation-function assigning sets of $n$-tuples to $n$-place predicates) iff. there is a verifying embedding $f$ mapping $x$ and $y$ into the same individual in $D$, and $f(x) \in I(\text{man}), f(x) \in I(\text{come in}), f(y) \in I(\text{sit down})$.

If the antecedent in (28) would be a singular definite determiner, as in (30),

(30) The man came in. He sat down.

the construction of the DRS would essentially be the same, with the sole difference that the reference marker for the definite description should already be available either in the DRS representing preceding sentences, or as part of the assumed common ground of the discourse.

A universal NP gives rise to a split of the DRS into a top level containing reference markers for proper names and all referential NPs, if any were so far represented, a first deeper level which represents the CN in the universal NP, and a second deeper level which represents the VP of the sentence. Sentences to be represented after the split are processed at the top level above the split structure.

(31) Every man came in. * He sat down

The first sentence is (31) is represented as:
The second sentence should be represented as a condition of $x$, but due to the structure of the levels $x$ is inaccessible from the top level, where the second sentence is to be processed. So the bound variable reading of the pronoun in (31) is excluded, although a deictic reading is still available.

Indefinite NPs in the restrictor of universal NPs which bind pronouns in the VP formed a major problem for a PTQ style quantifying in account of binding. In DRT such anaphora are accounted for by the accessibility conditions between levels of the DRS. Sentence (10) is represented according to the construction rules for the DRS as in (10a).

(10) Every woman who loves a man kisses him

(10a)

With the embedding conditions for a subordinating construction (10a) is interpreted as true in any model where any man loved by any woman is kissed by her, no matter how many men each woman loves.

Deictically used referential NPs are directly referential to an individual in the immediate situation of use. In DRT such a directly referential link is represented by an external anchor, which is an ordered pair consisting of the reference marker for that NP and some object in the immediate situation of use. Such external anchors are themselves not parts of the DRS but rather constrain
the set of verifying embeddings of the DRS into the model. The semantic content of a deictically used referential NP is completely determined by the associated external anchor, but the information someone may get from the use of such an expression is partly dependent upon the form of the NP itself.

3.2 Situation Semantics

Situation Semantics (SS, cf. Barwise, Jon and John Perry (1983) and Seligman, Jerry and Lawrence Moss (1997)) is a theory of dynamic interpretation which does not rely on a syntactic level of representation for anaphoric dependencies. Instead, information structure is constructed from semantic objects, which may or may not be parts of the actual world. Meaning arises as a relation between linguistic expressions, the context of use (including time of utterance, speaker, audience, location), linguistic and logical constraints, and the external world. Despite this important difference with DRT, the two theories are significantly similar in the insights and logical tools they offer to linguistic analyses. The primitive objects in SS are $n$-place relations, individuals, locations and polarities. They constitute events or situations, e.g. $< l,<< \text{walk, Mary }>, 1 >>$ represents a situation of Mary walking at $l$ and, $< l,<< \text{kiss, Mary, John }>, 0 >>$ a situation at $l$ in which Mary does not kiss John. Indeterminates or parameters act like reference-markers for locations, relations, individuals and polarities, and are equally constituents of situations. They are assigned appropriate values by partial assignment functions, or by context-dependent speaker connections to parts of the external world.

For example, a definite description can be interpreted as referring to an individual, determined by the speaker-connection, which is customarily called a referentially used definite description. A definite description which is so used to
refer to an individual does not require its descriptive properties to be true of the referent. This is commonly recognized to be possible, when speaker reference is at stake. In SS this usage is called the value-laden use of definite descriptions. But a definite description can also constrain a situation, picking up an individual to contribute to another situation. This is called the value-free, or attributive use of a definite description.

For example, one can use the NP in (32)

(32) The woman in the red skirt is tall.

to refer to Mary wearing a red skirt in situation $s$. For such a value-laden interpretation we fix the resource situation $s$, the speaker connections $c$ and represent it as $d, c \upharpoonright [\text{the woman in the red skirt}] (s) = \text{Mary}$. Third party reports of what was said may of course use other NPs to continue the reference to Mary, e.g. if Mary is also reading in $s$ coreference is established with (33).

(33) She said that this reader is tall.

In the attributive use of this definite description the interpretation is a relation between situations and individuals, whoever fits the descriptive properties. The condition of being tall is still a constituent of the interpretation of (33), but none of the individuals are. In order to get the attributive use the describing properties are not constituents of the interpretation, and it picks out an individual, if the resource situation contains an individual who satisfies the properties.

Other uses of definite descriptions are still possible, like in appositive clauses, where their reference is already determined by the context, as in John, the neighbor I play tennis with, is a nice guy. and the description contributes new properties to it, or functional uses, where reference is made to the role itself, not to whoever plays the role in any given situation, as in The next president must be elected. In evaluating any sentence containing a descriptive referential NP,
we have to be particularly careful in determining which situation is described, and cannot in general conclude that its truth value remains the same, if we consider a larger situation of which the situation described is part or another situation which does not contain the individual referred to as constituent.

In SS anaphora and other dependent NPs are interpreted dynamically by incrementally extending partial assignment functions. The core idea here is that the interpretation of an NP in a given context is an action which may affect the context in a systematic way. Current research is focussing on the details of an inductive definition of such dynamic interpretations of expressions of all categories as context-changing actions and its relation to the standard static satisfaction conditions of ordinary predicate logic. SS relies on indexing rules which operate on parsed sentences before their interpretation, such that every NP bears a unique referential index and every dependent NP is coindexed in subscript with its antecedent superscript. The interpretation of these indices form a crucial part in the dynamics of the procedural interpretations.

An important SS construction of semantic objects needed in the interpretation of universal NPs are the parameterized sets. They are semantic objects in which certain constituents are still undetermined, i.e. the parameters which need to be determined by extensions of a given assignment function. For example, in interpreting our old example of E-type anaphora

\[(10) [\text{Every woman who loves [a man}]^j \text{ kisses [him}]^k]

we need to form a parameterized set \(X\) which contains all pairs consisting of an individual \(a\) and an assignment function \(g\) such that kissing holds between \(a\) and the object \(g(j)\). All such \(g\) are supposed to be defined on the same indices. We verify (10) on this parameterized set, given an initial assignment function \(f\), if for each \(g\) every \(a\) in its set \(X_g\) is a woman who loving the man \(g(j)\) (cf. Barwise, Jon (1987) for more discussion and details).
3.3 Quantification and anaphora

Further discussion of anaphoric binding for theories of dynamic interpretation leads to the most recent developments and open research problems. One central issue is the interpretation of plural anaphora. They can be bound by singular universal antecedents, as in (34).

(34) Every woman kissed John. They left him.

The semantic operations with which the group consisting of all the women who kissed John is constructed as an appropriate referent for the plural pronoun in (34) is a central question of research. The converse of this issue is illustrated in (35), where the plural pronoun has a numerically appropriate plural antecedent.

(35) All women gathered in the room. They were wearing a badge.

The antecedent is an argument of a collective verbal predicate, denoting a property which can only be attributed to groups, not to the individuals constituting the group. The pronoun is, however, an argument of a distributive predicate denoting a property true of each member in the group of women. Some semantic operation is required to divide the group of all women as a single unit into the set of individual women (cf. Lønning, Jan-Tore (1997)).

Binding of plural anaphora by an antecedent in the scope of a universal NP cannot cross sentential boundaries, as (36) shows.

(36) Every father of two children sends them to Montessori school.

#They love it.

And furthermore two occurrences of plural anaphora bound by the same antecedent can be interpreted as referring collectively and distributively within one sentence as (37) illustrates.

(37) Mary and John invited their parents to their place.
The interpretation of (37) which is intended here makes the first anaphoric reference to the parents of each of Mary and John, but the second anaphoric reference to the place where they live together. Such issues of collective and distributive reference and predication provide a wealth of new puzzles for natural language semantics, which seem to lend themselves very well for analysis in these dynamic theories of interpretation which allow for a specific part-whole structure on their domains of reference-markers (cf. esp. Roberts, Craige (1987), Kadmon, Nirit (1987), Landman, Fred (1996)).

A third important problem for theories of anaphoric reference is called the 'proportion problem', illustrated by (38).

(38) Most women who love a man kiss him

The DRT analysis seems to predict (38) is true in a situation in which Jane loves Jim but does not kiss him, Paula who loves Peter does not kiss him, but Edith who loves Eric, Eduard and Evert kisses the three of them. The quantification merely counts the cases of a woman and a man loved by her, and counts Edith three times in verifying instances, whereas Jane and Paula are counted each only once in two falsifying instances. Solutions to this proportion problem have been proposed using the SS notion of parameterized sets, which suggest clearly that the dynamic interpretation should be constructed from the interpretation of expressions and constituents in all syntactic categories.

3.4 Dynamic Montague Grammar

The question remains whether the determinism of the syntax-semantics interface, as was required by the Fregean Principle of Compositionality, should be adhered to. It is clear that Montague's rule-by-rule Compositionality is not
adhered to in DRT, for there is only one syntactic rule putting determiners and common nouns together into noun phrases, but there are at least four different rules of DRS construction for NPs, depending on whether its is a proper name, a pronoun, an indefinite, existential or a universal NP. Critical of DRT for abandoning the Compositionality Principle, but overall motivated by much the same evidence, Groenendijk and Stokhof (1991) developed Dynamic Montague Grammar (DMG), as a compositional theory of interpretation of discourse. An update of a given information state relates partial assignments to variables, constituting the current information state, to their continuations, preserving the prior assignments and adding values to new variables. Besides the customary context-independent variables, the formal language is enriched with discourse variables, comparable to the reference markers of DRT, functioning like context dependent names that create dynamic bindings across sentential boundaries. Logical constants may be interpreted either dynamically or statically, depending on their desired degree of stability across updates of the information states. Dynamic conjunction is for instance sensitive to the order of presentation, and hence not commutative, as the ordinary static conjunction is in first order predicate logic.

A major bone of contention is the need for a pre-semantic representational level at which anaphoric dependencies are captured. DRT claims that such a syntactic representational level is essential, whereas SS and DMG claim they do better without. Compositional reformalizations of DRT have been presented, and alternative dynamic systems are being explored (cf. Muskens, Reinout (1995) and Muskens, Reinout et al. (1997)). The arguments are far from conclusive and an ultimate assessment of these issue must depend on the development of much more substantive and detailed semantic analyses of various linguistic phenomena.
4.1 Open problems

The great deal of attention which has been devoted to anaphora in natural language semantics has spurred generalizations of such informational dependencies in other categories than NPs. In Partee (1973) it was pointed out that tenses function very much like pronouns, in that their temporal reference can be determined deictically by the non-linguistic context, or depend on a referential, existential or universal antecedent. Sometimes these antecedents are adverbial, but they can also be verbs themselves.

A second analogy between NPs and VPs is commonly recognized. Mass NPs as *some gold, more peace, all furniture* are seen to be analogous to certain kinds of descriptions of events, since both may contain parts of the same kind, e.g. part of some gold is gold and part of an event of John walking is also an event of his walking. Count NPs are on par then with event descriptions which include some inherent endpoint, like the NP *part*, whose denotation does not contain the same man as part, and *John walking a mile* whose denoted event does not contain another walking of a mile by John. These two analogies play a very important role in developing a compositional semantic theory of tense and aspect, of temporal reference and quantification. Such a theory has obvious consequences for philosophical views on the nature of events and their identity- and individuation-conditions. Hinrichs (1986) and Partee (1984) provide accounts of nominal and temporal anaphora using tools of DRT theory, but the topic has grown into a fruitful field for interdisciplinary research in logic and linguistics.

Another important area of current research is the semantics of generic expressions. There is an important distinction between generic statements which refer to a kind as an abstract object and statements which are essentially
of quantificational form binding cases by a default operator. The two kinds of generic statements are illustrated in (39) and (40).

(39) Elephants are rare

(40) Elephants have valuable teeth

The main semantic difference between reference to kinds and default quantification resides in the fact that only the default quantification allows for exceptions, i.e. in the case of (40) an elephant whose teeth have been cut off. A host of linguistic evidence supports the distinction, and it is especially interesting to study the interaction with anaphora. We see that generic statements with universal NPs seem to allow binding of pronouns across sentential boundaries more easily, as in (41) which is a statement of a rule.

(41) Every player chooses a pawn. He puts it on square one.

Further observations which form explananda for natural language semantics are bindings which change the referential type as in (42).

(42) There is a beaver in the creek. They build dams

In (42) there is first reference to an individual beaver, but this serves as antecedent of a pronoun which refers to the entire species. The converse dependency is possible too, although it appears to be more restricted as in (43).

(43) Beavers build dams. I saw one/*him in the creek

A systematic account of such type-changing bindings is a topic of much current research. (cf. Carlson and Pelletier (1995)).

4.2 Cognitive science
To conclude this assessment of the developments in natural language semantics some questions about the entire research program should be addressed from a more general perspective. The renewed contact between logical theory and linguistic analysis prompts the question what kind of theory of inference
semantics is after. Should it be a theory about the inferential abilities of idealized, competent users of a natural language, or should it be a theory of actual inferences exhibited in human linguistic behavior? Frege's abhorrence of psychological interpretations of logical laws had promoted a stark separation of logical and psychological research on inferential processes. Most psychologists nowadays are still apt at pointing out that abstract mathematical laws do not explain their actual data, because 'people are not rational', 'human beings are no machines' or 'error is only human'. Yet the program of modelling inferential processes in natural language understanding by abstract logical representations has certain explanatory claims in cognitive science as a substantial contribution to a general theory of human cognitive capacities.

Here the classical Chomskyan distinction between a theory of competence and a theory of performance can clarify this apparent conflict. As a theory of inference, natural language semantics disregards the parameters of individual variation, cases of inferential failure, and normalizes its concepts by abstracting from actual practice and performance. Its empirical base is essentially the intuitive judgements of its users, not measured in a quantitative manner. Psychological theories of cognitive capacities are rooted in experimentally gained evidence from actual, quantitatively measurable inferential behavior. As in any science, they too make fundamental assumptions about their subject matter, excluding certain parameters in the experimentation as irrelevant to their explanatory goals, and stabilizing the context of their experimentation by a host of 'ceteris paribus' clauses which rarely receive any independent justification. Both forms of theorizing are empirical in nature, essentially falsifiable, and have genuine predictive power. But they contribute to our understanding of human cognition at quite distinct levels. A theory of error in linguistic processing is immediately relevant and perhaps even part of a psychological
theory of inferential processes, but it would not be of immediate interest to natural language semantics. But like aphasia studies can provide us with arguments concerning the modularity of the brain and its cognitive functions, such a theory of inferential failure may be able to provide evidence concerning the modularity of the brain for inferential processing and the interference with other cognitive functions.

If we see the two kinds of cognitive theory contribute explanatory insights at different levels, they can be considered respectively as characterizing the algorithms of inferential processes and characterizing the actual implementations of such algorithms in the human wetware.

It is however important to emphasize again that both areas of research regard inferential processes as central theme in a theory of human cognitive capacities.
Suggested Further Readings
The field of Natural Language Semantics is truly booming in academic and industrial research institutes.
(to be completed)
Handbook of Logic and Language
Handbook of Semantics
Handbook Haegeman
Invitation to Cogsci
Meaning and Grammar
GAMUT

5. Selected References


Heim, Irene (1983), 'File change semantics and the familiarity theory of


Partee, Barbara Hall (1997), Montague Grammar. In J. van Benthem and A. ter


