Chapter 8

Towards a Formal Characterization of Semantic Topics

Howard Gregory

1 Preliminaries

In this paper I propose a model-theoretic approach intended to formally capture the notion of Sentence Topic (henceforth simply "Topic"). Such a formal characterization is still lacking in the linguistic literature to date (cf. the complaints to this effect by Chafe [7] [1]), and may be considered a necessary preliminary to further research on the behaviour of Topic as a Grammatical Relation.[2]

Basic model-theoretic approaches to sentence meaning (such as those based on first order logic) aim to capture the predicate-argument structure of a sentence (plus quantification, which will not be discussed here), but do not in general address the topic-focus structure of sentences. Clearly, the success-

[1] I am referring to the provision of a principled denotational semantics, not the specification of algorithms for particular computational processes. A popular example of the latter is Centering Theory [2], which formalizes the idea of (transitions between) attentional states but without elaborating on the semantic interpretation of the entities used.

[2] This is an abbreviated version of a longer paper presented at a Linguistics Departmental Seminar in April 1998 [18]. In the present version only *sentential* Topics are considered. The extended version, which is available from the author (*http://semantics.soas.ac.uk/ellip/howard3/homepage.html*), extends the theory to give an account of *intersentential* Topic relations using a modal logic based on the conclusions of the last section. The ideas contained in this paper were influnced by a number of courses both in the SOAS Linguistics Department and at the ESSLLI summer school in Aix-en-Provence, 1997. I am particularly grateful to Shalom Lappin, Ruth Kempson and Wilfried Meyer-Viol for valuable discussion; the comments of one of the anonymous reviewers were also very helpful. The theory presented here remains, of course, my own responsibility.

Proceedings of the Third ESSLLI Student Session
Ivana Kruijff-Korbayová (editor).

ful treatment of topic-focus structure requires a dynamic characterization of sentence meaning, such as that offered by update semantics [19] (which however has not been concerned with topic and focus as such) or the process-oriented accounts of LDS or Information Packaging [40, 41] [3]. However, such approaches do not come equipped with a denotational semantics against which their "intended meaning" can be checked. In order to provide such a semantics I turn to certain mathematical methods which have been widely used in computer science to provide a denotational semantics for programming languages. What is useful here is that computer languages, which have an inherently procedural semantics (in terms of functions between information states), can nonetheless also be given a denotational characterization [37], in which even the dynamic aspects of their meaning are captured. The intention here is to take a similar approach to dynamic aspects of natural language, abstracting away from the actual nature of the processing involved.

The proposal is based on the application of topology to predication, as presented by [33, 42, 35, 1]. There are a number of reasons for choosing this rather unusual approach.

First, the notion of Topic has been regarded in several Linguistic traditions as corresponding to the traditional definition of a Subject, i.e. that of which something is predicated. While Topic and Subject are not co-extensive in the syntax, Topics will here be treated as an underlying semantic or notional Subject in this sense [4] [24, 32, 22, p.7ff]. I distinguish between affirmative sentences which have such a notional Subject and those which do not (*categorical* vs. *thetic* judgements in the terminology of the Prague School, also introduced into generative linguistics by Kuroda [24]). To take Kuroda's examples in (8.1), 1 is a thetic judgement, comprising an act of recognition of an event type, while 2 is a categorical judgement, recognizing the relation

[3]These frameworks both present topic-focus structure as a function of information-processing and deny that it affects the denotational meaning of a sentence [46, 45, 41]). However, sentences differing only in topic-focus structure are known to have different truth conditions in the presence of certain focus-sensitive elements such as *only*. For example the following contrast based on focus intonation (from [31]) -

1. John only introduced LENA to Mary. - *false* if introduce(j,k,m) for some distinct *k*

2. John only introduced Lena to MARY. - *false* if introduce(j,l,k)

- is also forced when one of the non-subject NPs is topicalized:

1. Mary John only introduced Lena to. - *false* if introduce(j,k,m)

2. Lena John only introduced to Mary. - *false* if introduce(j,l,k)

I take this as confirming the necessity for some denotational semantics for information structure related phenomena (though I will not be discussing focus as such here).

[4]Not, of course, the same sense in which many generative theories speak of a "logical subject".

of an event type to a previously fixed individual, which is marked with *wa*:

(8.1)

1. *Inu* *ga* *neko* *o* *oikatete* *iru.*
   dog NOM cat ACC chasing is
   The dog is chasing the cat. (Look, the dog's chasing the cat).

2. *Inu* *wa* *neko* *o* *oikatete* *iru.*
   dog TOP cat ACC chasing is
   The dog is chasing the cat. (The dog, he's chasing the cat).

In this paper I concentrate entirely on the categorical judgement [5]. Predication (of an event or state description) of a fixed subject will be taken as central to the idea of topichood, and the approach selected can be seen as simply transferring to Topics a treatment which is commonly applied in computer science to the (computational) subject-predicate relation.

Second, the approach seems to have a natural relation to the idea, going back to Strawson [36] and taken up within Generalized Quantifier Theory (GQT) [3, 43, 44, 28], that the topical role of NPs is related to the process of verification of propositions (cf. also [30]). In this tradition NPs have a natural function as topics in narrowing down the space to be searched, by fixing a restriction set; and in cases where a non-empty restriction set cannot be found the processing is blocked. While I argue below that this is inadequate in itself as a characterization of Topics, the basic insight that the restriction set plays an essential role in the processing of a sentence (see [28]) is retained. I develop this essentially along the lines of Strawson [36], invoking his theory of reference failure and its sensitivity to topic-focus structure, though without commitment to particular algorithms for verification, or to Strawson's particular treatment of definite descriptions. Under the system being proposed here, the apparently process-based idea of verification is incorporated directly into the semantics, via the less committal notion of a "finite affirmation" which will form the basis of the logic [1, 42]. Effectively it invokes the (hardly controversial) idea that a proposition is not affirmable unless it has a model, which is both a constraint on processing and a semantic statement.

Finally the mathematical structure used (a topological system in the sense of [42]) is known to form a framework congruent with a generalized form of situation semantics ([4, p.75-8], 2 below) - to whose ontology it gives an interpretation which is perhaps more intuitive than other forms of the theory. Situation semantics itself is adopted as a framework for wider research of which this forms a part, especially inasmuch as it forms the semantic

[5] Thetic judgements may possibly be treatable as having a dummy or "stage topic" [12], but I will not consider this here. In [18] I introduce categorical and thetic judgements as two subtypes partitioning the CONTENT type of affirmative sentences.

underpinning of HPSG [6]. However, it is hoped that the approach outlined here will enable a more promising dynamic treatment of discourse than has yet been obtained within standard situation theory (cf. [13, 18]).

In this paper the property of "topicalizability" for an NP [7] will be given a formal characterization related to the notion of logicality in Generalized Quantifier Theory (GQT) [3, 44, 20, 27], namely as the class of non-logical NPs, or those which fix the individuals in their denotation. This notion of logicality is then related topologically to the ability of the NP denotation to fix a model for the interpretation of the sentence. Actual topichood is distinguished from intrinsic topicalizability by the NP actually fixing such a model, whereas other topicalizable NPs (like all non-topicalizable ones) may be absorbed into the psoa [8]. The structure resulting from this analysis is then used to characterize the dynamic process by which non-topics in one sentence may become topics in the next [12].

Most (intensional) model-theoretic approaches assume at least the following apparatus: on the one hand a set W of worlds, times or situations, on which a structure such as an accessibility relation is defined and against which a proposition is evaluated; and on the other hand a domain of discourse D to which expressions in the object language are assigned by the interpretation function I. A dichotomy thus appears to exist between the objects inhabiting W and D. However there are good reasons for supposing that objects in W, at least times and situations [5, 10], may equally occur as arguments of predicates, in such a way as to have an extension in D. Part of the motivation for my proposal is the converse idea that the referents of NPs show similar behaviour in potentially inhabiting W as well as D, and that this ambivalence can be used to capture the distinction between NPs as arguments and NPs as topics [9].

[6] Its interaction with GQT has recently been given considerable attention and shown to be amenable to formal treatment [10].

[7] In this paper generics are excluded from consideration.

[8] The abbreviation *(p)soa* is used for the situation-theoretic term "(parametrized) state of affairs".

[9] cf. the work of Tsiplakou on clitic doubling [38, 39], in which Topic NPs, like Tense, are treated as database labels in a Labelled Deductive Systems framework.

2 Introduction to the formal proposal

2.1 Topological notions

Topology can be thought of as an abstract set-theoretical generalization of geometry, in which a space X is thought of as a set of points with certain structures defined on it. In topology proper most of the properties normally associated with geometry are abstracted away from, while all the "relevant" properties to be retained are captured by defining a class of "open sets" from the subsets of X. This class must be closed under finite intersections and arbitrary unions; it is then "a topology on X" [34, p.92].

These ideas have a natural interpretation in model-theoretic semantics. First of all an "open set" (a set without its boundary points) includes those instances which can be affirmed to be within the set [42]. Furthermore, if open sets denote finitely affirmable information, then any disjunction and any *finite* conjunction of such pieces of information will also be affirmable, and hence open [42]. Such a system, interpreted set-theoretically, yields a topology on some underlying space X. This space (or set of points) will in fact turn out to be the "model" for this logic (each point "supports" a piece of information iff it is an element of the corresponding open set; if S is the open set corresponding to the piece of information σ, then $x \models \sigma$ iff $x \in S$).

To make this more explicit, I assume (i) an algebra $\langle P, \leq \ \rangle$ of affirmable pieces of information constructed from the corresponding propositional [10] geometric logic so that it is partially ordered by entailment ($p \leq p'$ in P iff $p \Rightarrow p'$ in the logic [11]; also (ii) a family $\mathcal{E}$ of *partial* truth-assignment functions e from P to $\{0\ 1\}$. P is closed under disjunctions and finite conjunctions (corresponding to a "geometric" or "affirmative" logic as explained above), and $\langle P, \leq \ \rangle$ constitutes an algebra with the operations $\vee$ and $\wedge$ defined in the normal way (I ignore negation), and with a top element *true* or $\bigwedge \emptyset$ and a bottom element *false* or $\bigvee \emptyset$ [12]. The functions e of $\mathcal{E}$ must be homomorphisms preserving this structure [13]; i.e. any homomorphism which assigns a truth value to a subset Q of P must assign the same truth value to $\bigvee Q$ and (if finite) $\bigwedge Q$ (also to any superset of Q, as every subset of P must be "upper closed").

Considering $\mathcal{E}$ as a set E, a partial order $\sqsubseteq$ may now be defined on it, so

[10] The term "propositional" may be slightly confusing here, as the variables in the logic do not represent propositions but psoas. They are propositional in the sense that they have no arguments.

[11] The entailment may either be logical or in more interesting cases imposed by constraints [4].

[12] So that $true \wedge p = p = false \vee p$.

[13] Or "frame homomorphisms".

that e $\sqsubseteq$ e' iff $\forall$p$\in$P e(p)=1 $\Rightarrow$ e'(p)=1. Thus a poset $\langle$E $\sqsubseteq\rangle$ is obtained, on which more structure will be defined as necessary. $\langle$E $\sqsubseteq\rangle$ corresponds to a parametrized model for P, and will henceforth be referred to as "the model", with e $\models$ p iff e(p) = 1. However E also corresponds to the set of points X over which P was defined to be a topology. Because of this equivalence the e $\in$ E will be referred to interchangeably as "points" of the set underlying the topology or as homomorphisms on the propositional algebra, depending which point of view is convenient at any one time.

This construction results in a Topological System [42] (pt X, ΩX) in which *pt X* is the set of points and ΩX the frame for some underlying set X. Thus in the notation used so far, E = pt X, and P = ΩX for a Topological System. This brings out the point that both E and P are structures defined on a single underlying set. This contrasts with more familiar forms of intensional logic, where the set of worlds W and the domain of individuals are disjoint and the relation between them unclear. In the system used here a topological system will be defined on D as its underlying set (2.2), enabling individuals to serve as "worlds" in a modal logic ([18]).

Note that E will have a bottom element "$\perp$"; this is the point that supports only *true*, i.e. combinations of affirmations which are either logically true or stipulated to be true by constraints on the infon algebra (cf. [4]). In situation-theoretic terms it supports the "logical" domain of propositions and properties as opposed to the "informational" domain of situated information [14].

The important point about the resulting set of points is that considered as a space, it has the following characteristic. Unlike more familiar spaces, in which any two distinct points occupy separate positions, in the space being discussed here one point may be a refinement or specialization of another, in that its position is more precisely pinpointed but not separate (i.e. it is a non-Hausdorff space, not satisfying the Hausdorff separation axioms). Thus the points in the model are partial objects, with the ordering $\sqsubseteq$ representing partiality (relative to some maximal element).

The partial order on the points can be used to express the updating of the model during discourse. If a proposition is interpreted using a point e as model then e represents the existing state of the discourse before p, while the information contained in p enables e to be refined to e$'$. Thus if S is the set of contextual psoas (including any presuppositions of p), then e $\models$ $\bigwedge$S, while e$'$ $\models$ ($\bigwedge$S $\wedge$ p). This will be used to formalize what is meant by a statement being *about* a topic - it is information about an individual if it enables the position of that individual in the logical space to be refined.

[14] cf. [8] for a clear discussion of these notions within situation semantics.

2.2 Topologizing the Universe

In the previous section a Topological System was constructed out of a finite affirmative logic, following [42, 33]. One established property of such systems is that the ordered set of models (pt X) is isomorphic to the underlying set of points X. In this section a topology will be defined on the domain of discourse D of Generalized Quantifier Theory, and it will be argued that under certain conditions GQT expressions form such a topology on D, thus creating a topological system. The discussion will try to substantiate the following claims:

(8.2)
1. GQT expressions which when combined with a predicate set P are reducible to the form $x \in P$, for some fixed individual x, form a topology on D.
2. The same set of expressions corresponds to the set of categorical judgements in natural language, i.e. to sentences with a Topic. Thus categorical judgements are those expressions which denote an open set in $\mathcal{P}$D.
3. The NP combining with the predicate set to give the required expression must have a non-logical GQ as its denotation, because that denotation fixes x with respect to a specific partial model. (This non-logicality gives the correct empirical characterization of a sentence Topic, as will be further argued in subsequent sections).

The results will be used as the basis of the logic of Topics set out in subsequent sections. It is also intended to contribute further non-trivial constraints to the GQT characterization of natural language.

At a basic level the powerset of D, $\mathcal{P}$D, already yields a topology on D if all sets are taken as open. However this "discrete topology" is not the one of interest here. The aim is to construct a "frame" in the sense of the previous section. The results summarized there guarantee that the set of models for such a frame will be isomorphic to D, and the propositional formulas in the frame will be equivalent to open sets holding at the points of D.

It should be noted that in general the set of sentential formulas comprising an NP denotation combined with a predicate will not yield a topology, even allowing for the standard restrictions placed in GQT on NP denotations in natural language. Such formulas are not closed under the required operations. This is clear, for example, in (8.3) where the NP is a non-specific indefinite (a cardinal GQ) and the formulas are not closed under conjunction of predicates.

(8.3)
1. A million Greeks support Panathinaikos.
2. A million Greeks support Olympiakos.
3. A million Greeks support both Panathinaikos and Olympiakos.

Even for those unfamiliar with the habits of Greek football supporters, the above inference obviously does not go through. The set of formulas of GQT therefore cannot constitute a topology, even when the GQs concerned are confined to the denotations of natural language NPs. The problem with (8.3) is clearly to do with the fact that the Subject NP can denote non-identical sets, even keeping the model constant. If such cases were eliminated (if the same Greeks were involved in each of the two conjuncts) then the inference would be valid. Such a condition might be expressed as follows:

(8.4) **Fixed individual condition (FIC)** (provisional form)
A predicate set P is open iff for every model $\langle D, I\rangle$ there is a fixed possibly composite individual [15] x such that $x \in P$ is affirmable.

This condition will now be related to the corresponding NP denotations.

$\mathcal{P}$D itself is a poset ($\leq$ corresponding to $\subseteq$) and also a frame, though not a very interesting one. The relevant subframe of $\mathcal{P}$D will be formed as follows. A GQ $\{X\} \subseteq \mathcal{P}D$ is a function from predicate sets to {*true*, *false*}. The question here is under what conditions such a function constitutes a frame homomorphism. This will be the case if {X} is closed under union and finite intersection, i.e. disjunctions and finite conjunctions of {X} will be *true* under that function if any X is. (8.3) above illustrated that this is not true for cardinal NPs. In fact in general it never can be for cardinal NPs for the following reason: all X are required to have an intersection with a set N satisfying given cardinality condition. Moreover, all unions and finite intersections of X's must meet the same condition. However, $X_1 \cup X_2$ can meet this condition only if the NP is monotone increasing, whereas $X_1 \cap X_2$ only can if the NP is monotone decreasing; and if the cardinality condition is of the type "exactly n", then in general neither condition is met. This does not apply, however, to cases such as intersective NPs, where the requirement is that each X must include a fixed subset of N. In this case $(N \cap X_1) = (N \cap X_2) = (N \cap (X_1 \cup X_2)) = (N \cap (X_1 \cap X_2))$. Assuming that infinite conjunctions are unaffirmable as explained above, this means that {X} is a frame, or a topology on D.[16].

[15] I assume that D is closed under summation [29], and by "individuals" I include i-sums.

[16] A cardinal NP can only satisfy this requirement if there is an i-sum such that the intersection $N \cap X$ for all $X \in NP$ is the singleton set containing that i-sum and that the i-sum is of the required cardinality. Note that this has the effect of transforming a GQ into a name, for which the required closure under Boolean operations follows immediately.

This leads to the following form of the FIC (8.5).

(8.5) The **fixed individual condition** (revised form)

1. An NP satisfies the FIC iff N ∩ X comprises a fixed, possibly composite individual (henceforth simply "individual") for every X in its GQ denotation.

The following are an immediate consequence of (8.5).

(8.6)

1. An NP satisfying the FIC constitutes a frame homomorphism on its predicate sets.
2. An NP satisfying the FIC induces a topology on D in which its predicate sets are the open sets. Note the correspondence with the spatial definition of an open set X, which is that for some fixed individual x, x is inside X.

A predicate set P thus can be thought of as slotting in to a ready-made topology established by its Topic NP, which affirms it as *true* by providing an x such that $x \in P$ (as in the preliminary form of the FIC above (8.4)).

It is possible now to work from the other direction, considering formulas corresponding to GQT predicates as propositional [17] variables in a logic. Let Σ be the set of such formulas, algebraicized as in the previous section. These are to be assigned truth values by partial models each of which is a frame homomorphism. In GQT it is more normal to think in terms of total models. However, I take the class of models assigning N ∩ P to the individual x as being a partial model giving the same assignment to N, P and N ∩ P and undefined elsewhere. This model e will be a frame homomorphism and the ordered set of models $\langle E, \sqsubseteq \rangle$ together with $\langle \Sigma, \leq \rangle$ will form a topological system as outlined [18]. However E will now be isomorphic to the original domain D, under the following isomorphism *i* (8.7):

(8.7) $\langle E, \models, \Sigma, \wedge, \vee \rangle$
$\langle D, \in, \mathcal{P}D, \cap, \cup \rangle$

1. (a) $e \models \sigma \rightarrow i(e) \in i(\sigma)$

[17] The justapposition of "predicate" and "proposition" here may seem paradoxical. However it should be remembered that "propositional" here means only that they do not have variables as arguments; they are not propositions in the sense of denoting a truth value. They may be thought of as "sentence contents", like *psoas* in situation theory (hence notated as $\sigma \in \Sigma$). I assume that in categorical judgements the content of the Topic NP does not form part of the *psoa* but plays a role in fixing the model.

[18] Any formulae to which no $e \in E$ assigns *true* will be outside the algebra.

(b) $e \models \sigma \wedge \sigma' \rightarrow i(e) \in i(\sigma) \cap i(\sigma')$
(c) $e \models \sigma \vee \sigma' \rightarrow i(e) \in i(\sigma) \cup i(\sigma')$

2. (a) $x \in P \rightarrow i^{-1}(x) \models i^{-1}(P)$
   (b) $x \in P \cap P' \rightarrow i^{-1}(x) \models i^{-1}(P) \wedge i^{-1}(P')$
   (c) $x \in P \cup P' \rightarrow i^{-1}(x) \models i^{-1}(P) \vee i^{-1}(P')$

In contrast to standard (intensional) approaches which use a set of worlds disjoint from the universe of discourse and in which the relationship between the two remains unclear, in the system outlined here the frame (ΩD) and points (pt D) are constructions on the single underlying set D. Because of the isomorphism in (8.7), D is serving as a set of partial models, or worlds. Perhaps less intuitively, but crucially for the claims of this paper, the individuals of D are thus partial objects with a subsumption ordering corresponding to $\sqsubseteq$. In fact the central claim is that discourse is a process of continually refining the individuals of D by adding information in the form of psoas. In [18], a relation on D corresponding to $\sqsubseteq$ is in fact given in the form of an accessibility relation, and made the basis of a modal logic of topics. [19]

2.3 Topics and models

The model supporting a proposition is constrained by linguistic information. This is obvious for temporal logic, where the presence of tense or temporal adverbials will normally constrain the temporal parametrizations of the model which render the proposition "true". In the case of a thetic judgement this may perhaps be the only constraint on the model. However, I claim that for categorical judgements the same holds of certain NPs - namely, those which function as Topics. A psoa as a piece of information is not self-sufficent, but must be supported by a situation, or partial model. The latter must anchor (assign reference to in the model) any outstanding parameters of the psoa which have not already been absorbed. Thus the function which anchors the outstanding parameters of a psoa will (together with the location function) assign the psoa a truth value. Here I ignore the location parameter and questions of temporal logic, and concentrate on the relevant "argument" parameter. I argue that a Topic is precisely the parameter of a psoa which remains unabsorbed.

[19] The title of this section was chosen to contrast with that of "Eliminating the Universe" [21, p.52f]. The ontological dilemma posed by individuals through their unknowability, which there serves as a reason for their elimination, is here turned into the motivation for a theory of discourse in which the topological principle "no access to points except via open sets" is used model a process of refining the location of individuals in a given conceptual space.

To take a well-known example, it is often maintained [36] that a sentence containing "The King of France" (and evaluated at a time later than 1848) cannot be assigned "true" or "false" if this NP is Topic ("*the King of France* is bald", where the Subject acts as default Topic). According to Strawson this sentence is neither true nor false precisely because the NP lacks a referent, though the same will not necessarily hold for a corresponding non-Topic NP.

(8.8)	*The King of France* is bald.	undefined
	Mary saw the *King of France* at the station.	false
	The King of France Mary saw at the station.	undefined
	Mary, *the King of France* spoke to her at the station.	false

In (8.8) the sentences in which *the King of France* is not Topic can be asigned false because the NP is absorbed into the predicate, and its failure to provide a model does not matter. The precise notion of reference for definite descriptions invoked by Strawson, and the corresponding idea of a truth gap, are controversial (cf. [11] inter alia). From the present point of view what is important is not exactly the idea of reference but the ability of the NP to provide a model for the sentence, i.e. whether it can be considered as describing a point in the model, a function from P to $\{true, false\}$. In the next section I will discuss what it means in topological terms for an NP (Generalized Quantifier) to be able to describe a point. Here I simply note (i) that the empty set cannot describe a point, and (ii) that the universal set, which is the denotation of certain NPs whose N′ set is empty [28], cannot describe any point other than $\perp$, in other words it cannot support any information other than what is logically true or true by definition.

In 3, I will characterize the ability of an NP to provide a model for a sentence in terms of the conditions imposed on its GQ. The semantics proposed will also allow a dynamic characterization of the ability of a non-Topic NP to become the topic of a subsequent sentence. This in turn is used in [18] to support a more general account of the development of topics in discourse.

3 Sentential Topics

3.1 Generalized Quantifiers and Topichood

Turning to the GQT perspective, an GQ is a subset of $\mathcal{P}(D)$, but as shown by [3, 44, 21], GQ denotations of natural language NPs are confined to a highly restricted class of such subsets, the conditions for which themselves reflect the "logical topicality" of language in a broader sense [20]. However, even these conditions do not capture the notion of Topic in the grammatical sense (they allow all NPs, including non-specific indefinites or cardinal NPs). In fact, it was argued above that the crucial characteristic of such non-topicalizable NPs is precisely their independence of particular models for their interpretation, i.e. their *logicality*. Topicalizable NPs, by contrast, have a semantics in which the identity of elements in the relevant set intersections (henceforth "pivot sets") is important. It is suggested in [15] that topicalizable NPs are precisely the class of non-logical NPs [21], and furthermore that the idea of a situation plays an important role in fixing their meaning. That account however was not formalized. Here a somewhat more formal account will be offered incorporating some of those ideas.

Given the denotation of an NP as a set of sets whose membership is determined by particular set-theoretic relationships with a restriction set, I first distinguish between those NPs which fix the identity of the individuals in their denotation and those which do not [20, 25]. Prototypical examples of the former class include NPs with definite or possessive Dets, as well as proper names, while prototypical examples of the latter are NPs with cardinal Dets; in this paper I will confine my attention to these types.

In 2.2 above a formal distinction was presented between "topicalizable" NPs which are able to provide a model for the sentence and "non-topicalizable" NPs which are not. However, not all topicalizable NPs in a sentence need be actual Topics. In (8.9), in sentence 1 the Object is non-specific and cannot be a topic. In 2 (based on [7]) the Object is definite and is presumably known information, but nonetheless on the most natural reading of the sentence is not Topic.

(8.9)
1. John is buying *a Greek island*.
2. I saw *your wife* at the party. [7]

The characterization of actual Topichood, as opposed to Topicalizability, therefore requires an account of the process by which situations are built up in the discourse.

---

[20] "The combined effect of CONS(-ervativity) and EXT(-ension), namely Domain Restriction, is a kind of *logical topicality* condition." [20, p.56].

1. A function D .... is *conservative* iff .... if $A \cap B = A \cap C$ then $D(A)(B) = D(A)(C)$
2. D satisfies *Extension* iff .... with $A, B \subseteq E$ and $A, B \subseteq E'$, $D_E(A)(B) = D_{E'}(A)(B)$.

[21] As a corollary of this, it was claimed there that topicalizable NPs are the complement of those which can appear postverbally in existential constructions (assuming Lappin's [25] analysis of the latter as cardinal, and hence logical, NPs). While this seems to be largely true, this question requires further research, as there appear to be counterexamples.

However in both cases the non-topic NP can naturally become Topic in a subsequent sentence.

(8.10)
1. It is almost unspoilt.
2. She seemed to be having a good time.

Erteschik-Shir (op.cit.) describes the changing informational status of the NP in this kind of discourse (where it acts as presentational focus and then topic) by the metaphor of first adding a file-card to the top of the stack (focus) and then having it available as the most salient background information (topic). Although I do not make any formal use of the popular image of file-cards here, it is essentially this process that has to be captured. In terms of the account given so far, this would seem to require the following elements:

(8.11)
1. In the first sentence, the open set representing the predicate should be representable as a relation, one term of which is a Generalized Quantifier. This Generalized Quantifier will not in general be anchored to the individual which forms the Topic of the sentence, nor necessarily to any individual.
2. For the second sentence, the topic situation includes this second Generalized Quantifier, and should also bear a specifiable relationship to the Topic of the first conjunct.

I assume the standard GQT account of the combination of (extensional) Direct Object NPs with a predicate as given by Cooper [9] [22]. As usual the NP denotation for the Direct Object, notated GQ_2, is the set of all sets Y satisfying a given relation with an N' set N_2 or individual n_2. The relation R of the transitive verb denotes a set of ordered pairs $\subseteq A \times B$ where A and B are the sets of (descriptively speaking) potential "actor" and "goal" arguments respectively. The denotation V of a transitive verb is then the set of ordered pairs $<a, GQ_x>$ such that $B \in GQ_x$:

(8.12) Transitive verbs in GQT [9]
R = denotation of verb relation in D × D
A = potential "agents"
B = potential "patients"
GQ_x = denotation of an NP X

1. $V = \{\langle a, \mathbf{GQ}_x \rangle : \{b : \langle a, b \rangle \in R\} \in \mathbf{GQ}_x\}$

[22] cf. also [26, p.273-4] for a convenient summary.

2. GQ_{NP2} = {X: exp(N, N∩X)} as above
3. VP = {a: ⟨a, $\mathbf{GQ}_{NP2}$⟩ ∈ V}

The transitive VP thus denotes the set X (⊆ A) of all x such that ⟨x, $\mathbf{GQ}_2$⟩ is an element of V, and this then combines with the Subject NP denotation in the normal way.

It is clear from the semantics just given that an Object NP cannot by itself anchor the psoa, since its combination with the transitive verb denotation does not fix a unique pivot set in the sense required (its set denotation still contains the unanchored variable x). This is true whether the Object is definite or indefinite. The role of anchoring the psoa, i.e. that of Topic, falls to the last NP to be combined; since this is by default the Subject (following the syntactic structure), the Subject will be the default Topic. However I assume that the order of combination may be reversed, either through overt Topicalization of the Object or by some other mechanism with the Object remaining *in situ* [23].

As for the information contained in the Object (non-topic) NP, it is "absorbed" into the predicate, providing an additional constraint on the open set characterizing the final proposition. In the GQT chracterization just given, this is seen in the incorporation of its GQ denotation into the description of the predicate set (the VP denotation). However, the crucial point here is that this combination will also have the effect of restricting the denotation of the Object NP to a single set (the intersection of its N′ set with the set {b: ⟨a, b⟩ ∈ R} specified by the verb) - the latter being *one* of the elements of its GQ set). Unlike the NP denotation itself (for a non-topicalizable NP), this set S partially fixes a particular (possibly composite) individual (call it *y*) of a given cardinality. This is then available as a model for subsequent sentences.

Finally, the new discourse referent *y* thus made available, and which may itself become a Topic subsequently, depends on the Topic *x* of the sentence S in which it is introduced; in fact any model *f* provided by *y* must hold constant all truth assignments of the model *e* provided by *x* (e ⊑ f). This can be seen as an extension of the subsumption relation (cf. the isomorphism between individuals and partial models presented in (8.7) above). The analogue of ⊑ on D will be a relation such that for every open set X, x ∈ X → y ∈ X. Clearly this relation does not hold in such a simple form (otherwise *y* would have to be a member of the N′ set used to fix *x*). However for the open set P which is the predicate of S, both x ∈ P and y ∈ P. Since P is open, *y* is also a model for P, which is therefore a non-null intersection of the neighbourhoods of *x* and *y*. Let the neighbourhood of *y* be termed an

[23] These issues are discussed further in [16, 17].

"extended neighbourhood" of x, and P a "gate" from the neighbourhood of x to that of y. However the partial models e fixing x and f fixing y, as defined in 2.2 above, each represents a class of models. The class of models fixing y must not only be the class fixing the relevant N′ set and predicate set y but is also constrained to overlap x and thus to include all models for the latter. The model f for y will be an upper bound $\mathrm{e} \sqcup \mathrm{g}$ (such that $\mathrm{e} \sqsubseteq \mathrm{f}$ and $\mathrm{g} \sqsubseteq \mathrm{f}$, where g is any model in which the NP used to introduce y has a denotation). Transferring this reasoning to D, the relation of y to x can be considered an extension of the refinement relation whereby although y is not in the N′ set containing x, the intersection of its neighbourhood with that of x is fixed, and any further fixing of y by open sets will fix the neighbourhood of x.

(8.13) **Extended subsumption**

1. The extended subsumption relation $\sqsubseteq'$ holds between any $\mathrm{x}, \mathrm{y} \in$ D iff
   (a) $\mathrm{x} \sqsubseteq \mathrm{y}$
   (b) $\mathrm{y} \in \mathrm{S}$ ($\mathrm{S} \in$ neigbourhood(x))
   (recall $\mathrm{y} \in \mathrm{S}$ implies that y is partially fixed and S is open)
2. Any open set added to neighbourhood(y) overlaps neighbourhood(x) (it intersects with at least one set in neighbourhood(x)), and so fixes points in the neighbourhood of x.

Intuitively, any open set which further fixes y will also be information about x, because it gives information about the neighbourhood of x. In the next section this idea will be used as the basis of a modal logic.

4 Conclusion

In the preceding sections I have argued that the topological system outlined is an appropriate formal model of topichood. Combined with assumptions taken from Generalized Quantifier Theory it not only respects the standard account of NP denotations but interacts with it to predict the difference between topicalizable and non-topicalizable NPs, and to suggest a mechanism by which a non-topic NP in one sentence can serve as a topic for the next. In [18] this last aspect of the theory is applied to certain problems concerning the accessibility of topical information in more complicated discourse situations, and an attempt is made to extend it to intersentential aspects of topichood.

Many areas however have been touched on here in a rather programmatic way. Issues which need attention include the following:

(8.14)

1. Clarification of a number of formal issues including several mentioned in the text.
2. Generics, which have been excluded from consideration here [24].
3. Extending the theory to thetic judgements so as to include them in a unified treatment.
4. Negation. Classical negation is not supported by the affirmative logic invoked here, and while other treatments of negation are possible, it is necessary to ascertain that this will not have undesirable side-effects on the theory.

Bibliography

[1] S. Abramsky. *Domain Theory and the Logic of Observable Properties*. PhD thesis, QMC, University of London, 1987.

[2] B.Grosz A.Joshi and S.Weinstein. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2):203–225, 1995.

[3] J. Barwise and Robin Cooper. Generalized quantifiers and natural language. *Linguistics and Philosophy*, pages 159–219, 1981.

[4] J. Barwise and J.Etchemendy. Information, infons and inference. In *Robin Cooper K.Mukai and J.Perry eds. 1990*, pages 33–78.

[5] J. Barwise and J.Perry. *Situations and Attitudes*. MIT Press, Cambridge MA, 1983.

[6] G. Carlson. *Reference to Kinds in English*. Outstanding Dissertations in Linguistics. Garland, New York, 1980.

[7] W. Chafe. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In C. Li, editor, *Subject and Topic*, chapter 2, pages 25–55. 1976.

[8] Richard Cooper. Persistence and structural determination. In *J.Barwise J-M.Gawron G.Plotkin and S.Tutiya eds. 1991*, pages 295–309.

[9] Robin Cooper. *Quantification and Syntactic Theory*. Reidel, Dordrecht, 1983.

[10] Robin Cooper. The role of situations in generalized quantifiers. In S. Lappin, editor, *The Handbook of Contemporary Semantic Theory*, pages 65–86. Blackwell, Oxford, 1996.

[24] One reviewer suggested that this was because they provide a counter-example to the theory. In fact it is not clear that this is the case. In particular it may be noted that the GQ denotations of generic NPs satisfy the closure properties described in 2.2 above, and hence may be treated as equivalent to individuals, which of course is the semantics given for generics by Carlson [6]. This raises the possibility that Carlson's semantics may be derivable from the more general approach to Topics suggested here.

[11] K. Donnellan. Reference and definite descriptions. In D. Steinberg and L.Jakobovits, editors, *Semantics*, pages 100–114. 1971. first published in Philosophical Review lxxv (1966), 281-304.

[12] N. Erteschik-Shir. *The Dynamics of Focus Structure*. forthcoming.

[13] J-M. Gawron. Quantification, quantificational domains, and dynamic logic. In S. Lappin, editor, *The Handbook of Contemporary Semantic Theory*, pages 247–267. Blackwell, Oxford, 1996.

[14] J.Barwise J-M.Gawron G.Plotkin and S.Tutiya. *Situation Theory and its Applications*, volume 2. CSLI, 1991.

[15] H. Gregory. Topic noun phrases, specificity and resource situations. Term Paper, SOAS, 1996.

[16] H. Gregory. Clitic doubling, agreement and grammatical relations. Term paper, SOAS, 1997.

[17] H. Gregory. Predicate composition and the relational hierarchy. unpublished abstract, SOAS, 1998.

[18] H. Gregory. Towards a formal semantic characterization of topichood. Upgrading paper, SOAS, April 1998.

[19] J. Groenendijk and M.Stokhof. Dynamic predicate logic. *Linguistics and Philosophy*, 14:39–100, 1991.

[20] E. Keenan. The semantics of determiners. In S. Lappin, editor, *The Handbook of Contemporary Semantic Theory*, pages 41–63. Blackwell, Oxford, 1996.

[21] E. Keenan and L.Faltz. *Boolean Semantics for Natural Language*. Reidel, Dordrecht, 1985.

[22] K.E. Kiss. *Discourse Configurational Languages*. Oxford Studies in Comparative Syntax. OUP, Oxford, 1995.

[23] Robin Cooper K.Mukai and J.Perry. *Situation Theory and its Applications*, volume 1. CSLI, 1990.

[24] S-Y. Kuroda. The categorical and the thetic judgement. *Foundations of Language*, 9:153–85, 1972-3.

[25] S. Lappin. The semantics of 'many' as a weak determiner. *Linguistics*, 26:977–98, 1988.

[26] S. Lappin. Donkey pronouns unbound. *Theoretical Linguistics*, 15(3):263–285, 1988-9.

[27] S. Lappin. Semantic types for natural language. Inaugural Lecture, SOAS, February 1997.

[28] S. Lappin and T.Reinhart. Presuppositional effects of strong determiners: a processing account. *Linguistics*, 26:1021–1037, 1988.

[29] G. Link. The logical analysis of plurals and mass terms: a lattice-theoretical approach. In R.B editor, *Meaning, Use and Interpretation of Language*. Walter de Gruyter, 1983.

[30] T. Reinhart. Pragmatics and linguistics: an analysis of sentence topics. *Philosophica*, 27:53–94, 1981.

[31] M. Rooth. Focus. In S. Lappin, editor, *The Handbook of Contemporary Semantic Theory*, pages 271–297. Blackwell, Oxford, 1996.

[32] S. Rothstein. *The Syntactic Forms of Predication*. PhD thesis, MIT, Cambridge MA, 1983.

[33] W. Rounds. Domain theory and its applications in language, logic and computation. Course given at the European Summer School in Language, Logic and Computation, Aix-en-Provence, August 1997. (slides available from http://ai.eecs.umich.edu/people/rounds/).

[34] G. Simmons. *Introduction to Topology and Modern Analysis*. McGraw-Hill, 1963.

[35] M. Smyth. Topology. In S.Abramsky D.Gabbay and T.Maibaum, editors, *Handbook of Logic in Computer Science*, volume 1 of *Oxford Science Publications*, pages 641–761. OUP, 1992.

[36] P. Strawson. Identifying reference and truth-values. In *Logico-linguistic Papers*, pages 75–95. Methuen, London, 1971. first published in Theoria xxx (1964), 96-118; also in Steinberg and Jakobovits eds. (1971), 86-99.

[37] R. Tennent. *Semantics of Programming Languages*. Prentice Hall International Series in Computer Science. Prentice Hall International (UK) Ltd., 66 Wood Lane End, Hemel Hempstead, Herts., 1991.

[38] S. Tsiplakou. On the representation of focus in greek - an lds perspective. *SOAS Working Papers in Linguistics and Phonetics*, pages 233–250, 1995.

[39] S. Tsiplakou. *Focus in Greek: its structure and interpretation*. PhD thesis, SOAS, University of London, 1998.

[40] E. Vallduvi. *The Informational Component*. Garland, New York, 1992.

[41] E. Vallduvi. Structural properties of information packaging in catalan. In K.E. Kiss, editor, *Discourse Configurational Languages*, pages 122–152. OUP, Oxford, 1995.

[42] S. Vickers. *Topology via Logic*. CUP, 1989.

[43] D. Westerstahl. Logical constants in quantifier languages. *Linguistics and Philosophy*, 8:387–413, 1985.

[44] D. Westerstahl. Quantifiers in formal and natural languages. In D. Gabbay and F.Guenthner, editors, *Handbook of Philosophical Logic*, pages 1–131. 1989.

[45] R.Kempson W.Meyer-Viol and D.Gabbay. Topic and focus structures: The dynamics of tree growth. Abstract, MS SOAS, 1998.

[46] R.Kempson W.Meyer-Viol and D.Gabbay. *The Deductive Flow of Language*. Blackwell, Oxford, in preparation.