References


--- (in progress) *Syntax and semantics of adjunct predicates*. Ph.D.: UMass/Amherst


---

Determiners and the Logical Expressive Power of Natural Language

Edward L. Keenan†

Lawrence S. Moss

0. Notional and Notational Preliminaries

We may consider that a major function of natural language is to enable us to raise questions and make assertions about the world. It is natural then to consider the precise semantic contribution to these assertions made by the various sorts of expressions a natural language provides. Here we are concerned with the expressive contribution of English Determiners (det) and with comparing the expressive nature of det with that of other categories.


On most of these approaches an extensional model for English is given by a non-empty set E of entities. And an extensional property is just a set of entities. The set P of all properties of the model is the denotation set for *common noun phrases* (CNPs). In general, given a model E, we use \( \mathcal{P} \) to denote the set of logically possible denotations (relative to E) of expressions of category C.

Proper nouns (e.g., *John*) denote sets of properties called *individuals*, where a set I of properties is an individual iff for some \( e \in E \), \( I = \{ q : e \in Q \} \). To say that an individual I “has” some property \( q \) is just to say that \( q \subseteq I \). We commonly use the symbol \( n \) to refer to the number of individuals of a model, and we use \( n \) for the number of individuals that have \( q \).

We use \( I \) for that property which all individuals have. And where \( p \) and \( q \) are properties, we use \( p \land q \) for the property of being both \( p \) and \( q \), that is, the property an individual has iff he has \( p \) and he also has \( q \).

We use \( P \) to denote the set of all sets of properties. So each individual \( I \) is an element of \( P \), and in fact, \( P \) is the denotation set for *full noun phrases* (FPs). Finally we use \( FP \) for the set of functions from \( P \) into \( P \). So this set provides the set of logically possible denotations for (one place) det.

(We refer the reader to Keenan (1982) for a semantics equivalent to the above in which there is no \( P \), \( P \) is a primitive, and individuals are defined directly as subsets of \( P \) satisfying certain conditions. On all views \( P \) is isomorphic to the power set of the set of individuals.)

1. Determiners and the Conservativity Constraint

We treat one place det (det's) as expressions which combine with one CNP to form an NP. They include expressions such as *every*, *no student's*, *more of John's than of Mary's*, *at least two but not more than ten*, etc. We refer the reader to Keenan and Stavi (K&S) for a copious list of det's in English.
Semantically $\text{Det}_1$'s are interpreted by functions from $P$ into $P^*$. Earlier work, most exhaustively in K&S, established that in general the functions which can interpret elements of $\text{Det}_1$ are conservative, where:

$\text{Def 1: A function } f \text{ from } P \text{ into } P^* \text{ is conservative iff for all } p, q \in P, p \in f(q) \text{ iff } (p \wedge \neg q) \in f(q)$$

So to say that $f$ is conservative is to say that $f(q)$ are $p$'s iff $q$'s are both $q$'s and $p$'s. E.g. every student is a vegetarian iff every student is both a student and a vegetarian. We note that the functions which can interpret "non-logical" det's such as John's and more male than female (as in more male than female students failed the exam) as well as "logical" det's are conservative.

2. Generalizing Determiners

The initial questions of expressive power we consider concern the relative expressive power of distinct subclasses of det's. We shall first extend the class of det's considered and then define the subclasses of interest.

We may think of $k$-place det's, $\text{Det}_k$'s, as expressions which combine directly with $k$ CNP's to form an NP. They will be interpreted by functions from the set $\mathcal{P}^k$ of $k$-tuples of properties into $P^*$. Keenan and Moss (K&M) give several reasons for treating the italicized expressions in (1) below as $\text{Det}_k$'s:

1. a. more students than teachers
   b. every author and critic
   c. John's brown cats and dogs
   d. fewer French workers than students
   e. the 27 students and teachers

For example, we want the two CNP's on the same level of structure so that they may both be under the scope of the adjectives in (1c) and (1d); equally they behave in similar ways with regard to number marking and the role they play in selectional restrictions. Further we cannot represent the ambiguity in (1b) by treating author and critic as a conjunction of CNP's which combine with the Det every, on the one hand, and by treating every...and... as a Det which combines directly with (author, Critic) on the other hand. Additional sorts of reasons are given for treating more...than... (and other numerical comparatives such as not more than twice as many...as...) as a Det (see also Napli 1983). One such reason is given in Theorem 1 below, which uses the interpretation of more...than... given in Def 2:

$\text{Def 2: more than } t = \{q : [x \wedge q] > [r \wedge q]\}$

So $q$ is a property which more students than teachers have iff the number of students ($s$) with $q$ is greater than the number of teachers ($t$) with $q$.

Theorem 1: In any model with at least two individuals the range of more...than... has larger cardinality than $P$ itself. Hence, it cannot also be the range of any one place function (conservative or not) on $P$.

Theorem 2: A function $f$ from $P$ into $P^*$ is conservative iff for all $p_1, p_2, q \in P$, if $p_1 \wedge q = p_2 \wedge q$ then $p_1 \in f(q)$, iff $p_2 \in f(q)$.

Theorem 2 says that the one place conservative functions are just those which cannot discriminate between properties $p_1, p_2$ which have the same meet with their argument $q$.

Generalizing below, we use $q$ as a varying range over $k$-tuples of properties, the $i^{th}$ element of which is denoted $q_i$.

$\text{Def 3: A function } f \text{ from } P^k \text{ into } P^* \text{ is conservative iff for all } p_1, p_2 \in P \text{ and all } q \in P^k, \text{ if } p_1 \wedge q_i = p_2 \wedge q_i \text{ for } i = 1, 2, \ldots, k \text{ then } p_1 \in f(q) \iff p_2 \in f(q)$

So the $k$-place conservative functions are just those which cannot distinguish between properties $p_1, p_2$ having the same meet with each argument $q_i$.

We use $\text{CONS}_k$ for the set of $k$-place conservative functions and $\text{CONS}$ for those functions in $\text{CONS}_k$ for some $k$.

3. Subcategories of Determiners

We shall here semantically subcategorize det's according as the functions which interpret them satisfy one or another condition. Specifically we shall distinguish "logical" det's such as every, all but two, more than ten, more...than..., etc. From "non-logical" or "real world" det's such as John's, more French, than British, all...but John, etc. And secondly, among the logical det's we shall distinguish cardinal det's from non-cardinal ones.

Informally, a function from $P$ into $P^*$ will be called a cardinal function just in case:

- whether it puts a property $t$ in the set it associates with $q$ is determined by how many individual $q$'s are $t$'s. For example the functions which interpret more than ten, between five and ten, and infinitely many are obviously cardinal functions, since if we know how many $t$'s are $q$'s we know infinitely many are, whether between five and ten are, etc.

Generalizing to $k$-place functions, we say that $f$ is a cardinal function just in case:

- whether it puts $t$ in the set it associates with a $k$-tuple $q$ of properties is determined by how many individual $q_i$'s are $t$'s, for each $i$ between 1 and $k$. For example, if we know how many $q_1$'s are $t$'s and how many $q_2$'s are $t$'s then we know whether $t$ is a property that more $q_1$'s than $q_2$'s have, so the function which interprets more...than... is a two place cardinal function.

For the record:

$\text{Def 4: A function } f \text{ from } P^k \text{ into } P^* \text{ is a cardinal function iff for all } s \in P \text{ and all } k\text{-tuples } p, q \text{ in } P^k, \text{ if } [s \wedge p_i] = [t \wedge q_i] \text{ for } i = 1, 2, \ldots, k \text{ then } s \in f(p) \iff t \in f(q)$

$\text{CARD}_k$ will denote the set of $k$-place cardinal functions and $\text{CARD}$ the set of functions which are in $\text{CARD}_k$ for some $k$. We note that for all $k$, $\text{CARD}_k$ is a proper subset of $\text{CONS}_k$ and that
CARD is a proper subset of CONS. So all (k-place) cardinal functions are (k-place) conservative functions. Finally, of course we define a (k-place) denotet to be a cardinal denotet just in case its denotation always lies in CARD_k.

Turning now to the larger class of logical denotets, we shall find a property common to the denotations of those denotets intuitively judged to be "logical" and then for each k, we define LOG_k to be the set of functions from P into P^* which have this property.

The intuition behind the definition of LOG_k is exemplified by the difference in truth conditions between (2a,b) on the one hand and (2c) on the other:

(2) a. Exactly one house is white.
   b. Every house is white.
   c. John's house is white.

The truth (or falsity) of (2a) and (2b) is determined once we have specified which objects have the property and which objects have the property of being white. But to know whether (2c) is true we must in addition know which individual John is and which object he "has". Thus, John's carries real world information in a way in which neither exactly one nor every does. More specifically, to say in a given state of affairs just which function from P into P^* interprets John's we must be able to discriminate one individual from another. But to say which functions interpret logical denotets as exactly one and every we do not have to discriminate one individual from another in the sense of knowing which properties this individual has which that one doesn't. We may of course have to know the number of individuals with a certain property (as in the case of cardinal denotets) but it doesn't matter which individuals they are. And we may have to assess the proportion of individuals with p that have q, as with denotets such as all, most, two-thirds of the, etc., but again it doesn't matter which of the individuals with p have q as long as the proportion is satisfied.

More formally then we may define a function from P into P^* to be logical just in case it remains invariant under permutations of individuals. The formal statement of this definition is somewhat more "remote" than in the case of cardinal functions and space will not permit the length of discussion necessary to justify that the definition captures the intuition sketched above. The reader not concerned with the formal details may omit the formal definition (at least on first pass) noting only that we have found a property which discriminates the possible denotations of (pretheoretically judged) logical denotets from non-logical ones. We may note further here that all cardinal functions are logical on the proposed definition, but the converse fails. E.g. denotets like every, most, all but two denote in LOG_1 but not in CARD_1. Finally, we note that the definition of logical is not specific to denotets but applies to the denotation set associated with any category. So it makes sense to ask what are the logical elements among any denotation set? among two-place predicate denotets? etc. And in the last section of this paper we compare, with interesting results, the nature of logical elements among denotets with those of many other categories.

Defining "logical": A permutation of the individuals of a model is just a one-to-one function from the set of individuals onto itself which preserves that boolean structure. The extension proceeds along the following lines: first, if a has been extended to an automorphism on a set K then for each S element of K, set a(S) = {a(s): s element of S}. This extends a to an automorphism on the power set of K. Similarly if a has been extended to an automorphism on sets B and D then we extend it to an automorphism on the sets of functions from B into D as follows: for each such function f, a(f) is that function from B into D which sends each a(b) to a(f(b)). And we may now define:

Def 5: An element d of any denotation set is automorphism invariant iff for all basic automorphisms a, a(d) = d.

We now define LOG_k to be the set of those functions in CONS_k which are automorphism invariant. LOG_k denotes the set of those functions in LOG_k for some k. And an element of Det_k is logical iff it always denotes in LOG_k.

We note further that in all models, CARD_k is a proper subset of LOG_k, and in all but the model with just one individual, LOG_k is a proper subset of CONS_k.

4. Questions of Expressive Power

(3) a. Suppose that English could take their denotations freely in the set of functions from P into P^*. Is there any significant sense in which we could say more then we currently can?
   b. Suppose that English had no non-logical denotets. Would we suffer a significant loss of expressive power? E.g. can we always mimic the effect of possessive denotets (John's, every student's, etc.), however ad hocly, with logical denotets?
   c. Can anything we can say using the full class of logical denotets be paraphrased (however ad hocly) using only cardinal denotets?
   d. Does allowing k>1 place denotets increase expressive power? E.g. can we in any interesting sense say more with LOG_2 than with LOG_1?
   e. Can we account for why there seem to be so many logical denotets and so few logical expressions among most other categories, e.g. the place predicates (NP's)?

(QUERY 3e) arises in a natural way from the answers to (3a-d).

To answer these questions we need a "significant" measure of the expressive power of the different classes of denotets in question. Trivially for example if we had no non-logical denotets in English we would lose the ability to denote conservative functions which were not automorphism invariant. But this loss hardly seems significant. We are not primarily interested in which functions from P into P^* we can name. Rather we think of the role of denotets as being inherently ancillary—their "purpose" is to allow us to refer to sets of properties in terms of properties. A reduction in the class of denotets of English then may be deemed significant if it results in a loss of sets nameable by NP's of the form Det + CNP. To make this notion precise:
Def 6: For QCP and H any set of functions whose range is included in $P^*$, we say that $Q$ is $H$-expresible iff for some $h \in H$ and some $x$ in the domain of $h$, $Q = h(x)$. We let $F_H$ denote the family of $H$-expressible sets.

Note: the elements of $F_H$ vary with the choice of $P$.

Now the classes $H$ of interest here are of course $\text{CARD}$, $\text{LOG}$, $\text{CONS}$, and for every $k$, $\text{CARD}_k$, $\text{LOG}_k$, and $\text{CONS}_k$. And for $H$ and $K$ any of these classes we may compare their expressive power by comparing $F_H$ and $F_K$. Unfortunately the results of the comparison sometimes depend on the choice of $P$. For example in a model with only one individual the $\text{LOG}_1$-expressible sets are the same as the $\text{CONS}_1$-expressible sets, whereas this identity fails in all models with more than one individual. To eliminate this dependency we define:

Def 7: For $H$ and $K$ any of our classes of interest, we write $F_H \subset F_K$ iff (i) and (ii) below hold:

(i) for all $P$, $F_H$ is a subset of $F_K$

(ii) for $P$ sufficiently large, $F_H$ is a proper subset of $F_K$. (To say that $P$ is sufficiently large is to say that for that $P$ and all larger $P$, $F_H$ is a proper subset of $F_K$.)

We now answer our questions of interest:

Theorem 3: (a) $F_{\text{CARD}} \subset F_{\text{LOG}} \subset F_{\text{CONS}}$

(b) for all $P$, $F_{\text{CONS}} = P^*$

Theorem 3b answers query (3a) in the negative. For any $P$ and any $Q$ of $P$ there is a one place conservative function $f$ and a property $s$ such that $Q = f(s)$. (We may in fact choose $s$ to be the trivial property 1, the property which all individuals have. So restricting $s$ to conservative functions does not in principle restrict the sets we can refer to with full $\mathcal{L}_P$.)

Theorem 3a answers queries (3b) and (3c) in the affirmative for the case $k = 1$. Theorem 4 below shows that the result generalizes in a nice level by level way:

Theorem 4: (a) for all $k$, $F_{\text{CARD}} \subset F_{\text{LOG}} \subset F_{\text{CONS}}$

(b) $F_{\text{CARD}} \subset F_{\text{LOG}} \subset F_{\text{CONS}}$

We should note here that the examples of property sets which show that the inclusions above are proper are in practice expressible. For example, let $Q$ be an infinite property (one had by an infinite number of individuals, such as e.g. the property of being an even number). Then the set every $(q)$ is not in $F_{\text{CARD}}$ and hence not in $F_{\text{CARD}}$ for any $k$. Since every denotes in $\text{LOG}_1$, that set is in $F_{\text{LOG}_1}$. Thus the expressive contribution of every cannot be mimicked by any cardinal dets, however ad hoc, since there are sets such as every$(q)$ above which are not equal to $(q)$ for any $k$-tuple of properties $f$ and any $k$-place cardinal function $f$.

Similarly we can show that among the acceptable interpretations of some person's one or more things are sets not in $F_{\text{LOG}}$ and hence not in $F_{\text{LOG}_1}$ for any $k$. things can be interpreted as the trivial property 1. The algebra $P$ however must be infinite and the possessor property is infinite in our example and probably necessarily. We note that for any $P$, every finite set of properties is in $F_{\text{LOG}}$, whence if $P$ is finite then every set of properties is in $F_{\text{LOG}}$. Query (3d) is mostly answered by Theorem 5 below:

Theorem 5: (a) $F_{\text{CARD}} \subset F_{\text{CARD}_k+1}$ for $k = 1, 2$ and probably all $k$

(b) $F_{\text{LOG}} \subset F_{\text{LOG}_k+1}$ all $k$

(c) $F_{\text{CONS}} = F_{\text{CONS}_k} = P^*$ all $k$, all $P$

For example, we can show that for certain pairs of properties (p,q) the value of $\text{more} \ldots \text{than}$... at that pair $\neq (\mathbb{R}, q)$, any property $s$, any $\mathbb{R} \in \text{CARD}_k$. So with respect to cardinal and more generally logical dets the addition of $k \geq 1$ place dets does increase expressive power. But this is not the case for the full class of conservative functions. All sets are $\text{CONS}_k$ expressible so in principle we could not express more sets with $\text{CONS}_k$ functions. In practice however things might be otherwise. For while it is true that for any set $Q$ of properties there is a $\mathbb{R} \in \text{CONS}_k$ such that $f(1) = Q$ we have no general way to construct a determiner expression (however complex) in English which could be interpreted as this $f$. Moreover there appear to be serious constraints on the freedom with which we may interpret dets (of any degree). K&S prove an indefiniteness result concerning DET_1 which suggests that the size of the denotation set for DET_1 grows faster than the freedom with which we may interpret complex dets. More important for our purposes here, there are severe constraints on the interpretation of syntactically simple sets (mostly they denote increasing, automorphism invariant functions) and hence of the sets we can express with them. As a category then, DET provides many logical expressions and constrains the denotations of its lexically simple ones. Below we compare DET with other categories in these (and other) respects:

<table>
<thead>
<tr>
<th>Category</th>
<th>Size of D(f)</th>
<th>No. of AI elements</th>
<th>Lex Free?</th>
<th>Free?</th>
</tr>
</thead>
<tbody>
<tr>
<td>large</td>
<td>DET_1</td>
<td>$2^n$</td>
<td>$2(n+1)^2 + 2$</td>
<td>no</td>
</tr>
<tr>
<td>cats</td>
<td>N^P</td>
<td>$2^{n+1}$</td>
<td>$2(n+1)$</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>AP</td>
<td>$2^{n+1}$</td>
<td>$2^n$</td>
<td>no</td>
</tr>
<tr>
<td>small</td>
<td>DET_1</td>
<td>$2^n$</td>
<td>$2^n$</td>
<td>yes</td>
</tr>
<tr>
<td>cats</td>
<td>DET_2</td>
<td>$2^n$</td>
<td>$2^n$</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td>DET_1</td>
<td>$2^n$</td>
<td>$2^n$</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td>N^Pabs</td>
<td>$2^n$</td>
<td>$2^n$</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td>S = F_0</td>
<td>$2^n$</td>
<td>$2^n$</td>
<td>yes</td>
</tr>
</tbody>
</table>

n.a.
Space does not permit a serious discussion of the entries in the table. By way of minimal explanation of the entries we note: $n$ is the number of individuals of the model and is assumed finite. $P_3$ is the category of 3 place predicate (transitive verb phrases); $P_2$ and $P_1$ are understood analogously. $AP$ is adjective phrase, and $AP_{ab}$ is absolute adjective phrase, a subcategory of $AP$ which includes e.g. male but not tall—so those $AP$'s which actually determine properties. $NP_{prop}$ is the subcategory of proper $NP$'s.

A category $C$ is called free if for every model, any element $d \in 1_C$ is denotable, that is, there is an expression $e$ in category $C$ such that for some interpretation $m$ of the language, $m(e) = d$. $C$ is called lexically free if the element in question can always be chosen from among the lexical expressions of category $C$.

Large categories are ones the size of whose denotation sets is a hyperexponential function of the size of the universe of individuals. That is, the size of the denotation set increases as an exponential function of $n$. For small categories on the other hand the size of the denotation set increases merely as an exponential of a polynomial in $n$, or less.

We may now summarize the major generalizations from the table:

(4) Small categories are lexically free, large categories are not.

(5) For large categories the number of automorphism-invariant elements in $1_C$ (so the number of possible denotations of logical constants of category $C$) increases as an exponential function of the size of the universe (and increases much faster for $1_{C_1}$ than any other large category). For small categories the number of $AU$ elements is fixed and independent of the size of the universe.

So from (4) we note that categories with large denotation sets constrain how lexical items in the category may be interpreted. E.g. lexical $NP$'s in general denote individuals (and perhaps a handful of other sets, expressed by $NP$'s such as everyone, someone, etc.)

And from (5) we have a partial answer to query (3b): No matter how big the universe there are only two distinct possible denotations for logical expressions of category $P_1$ (verbs), They are basically the denotations of exist and not exist (extensionally). So English can but in principle provide more than two extensionally distinct logical expressions of category $P_1$. It can however provide infinitely many expressions of category $D_{C_1}$ and in an infinite universe they can all denote differently.

Footnotes

* The authors would like to thank the Max-Planck-Institut für Psycholinguistik for support during the writing up of this paper.

1. We may then show that $f : P^k \rightarrow P^k$ is conservative iff $\bigwedge_{i=1}^{k} \bigvee_{q_i \in R(Q)} p_i \wedge r_i \in R(Q)$. This characterization of conservativity is comparable to the one originally given for one-place functions.

2. We use the same symbol $a$ for the basic automorphism and its various extensions.

References


Zwarts, F. (1982) "Determiners: a Relational Perspective," in