THE RELATION OF SYNTAX AND SEMANTICS AT THE LEXICAL LEVEL

K. Ashton

University of Auckland

O. This paper[1] is intended as a non-technical summary of a forthcoming mathematical treatise and consequently will begin by describing, informally, a minimal number of mathematical concepts without which any exposition of the theory would be impossible. This appeal to mathematics is intended not to provide an apparently erudite but unreadable description of known phenomena, but rather to make use of the deductive power of mathematical systems in order to explore certain basic laws governing the nature of language-like codes.

In this paper we shall examine the relationship between two views of semantics: on the one hand that meaning can be considered as an extra-linguistic phenomenon, for example in mentalistic or behaviouralistic terms; and on the other hand that meaning should be considered linguistically, roughly speaking that the meaning of a word lies in its usage[2]. We shall restrict our attention to the basic, meaningful units of language, to the morpheme, word and phrase[3], without analysing their structure. Such an analysis, together with a treatment of sentence structure will be reserved for a later paper and it will be at this later stage that we shall have to discuss such phenomena as transformation descriptions of sentences.

0.1 Our primary viewpoint will be to take language as a phenomenon to be analysed according to certain mathematical techniques[4] making no *a priori* assumptions as to the nature of models most apt to describe the results.

1. Preliminary mathematical notions.

Although we shall avoid the performance of mathematical deductions in this paper, it will be necessary for the reader to have some intuitive grasp of certain technical constructs. Where a purely mathematical result is required in the subsequent argument it will be quoted without proof.

1.1 A set is a collection of objects. Thus we may talk of the set of phonemes of a given language or of the set of environments in which a given word may occur. Note that we can form sets whose objects are themselves sets.

1.2 If every object belonging to a set A belongs also to a set B we say that A is a subset of B, or that A is contained in B, and write

$$A \subset B.$$

This relation between sets, and relations having similar properties are called orderings.

---

1. This paper is a development of ideas presented in Ashton 1968.

2. Wittgenstein, 1922.

3. Bloomfield, 1926.

4. An analysis of environments was given in Kunze 1967, and a representation as a lattice in Ashton 1969.

Consider the set whose only members are the letters 'a' and 'b'. This set, written as {a,b} has subsets {a} and {b}. We agree that any set **A** is a subset of itself (see the definition) and we further agree that the empty set, that is the set having no members and denoted by Ø, shall be a subset of all sets. Thus {a,b} has the subsets {a,b}, {a}, {b} and Ø, where Ø ⊂ {a} ⊂ {a,b} and Ø ⊂ {b} ⊂ {a,b}.

These latter facts we can represent clearly by means of the following diagram:

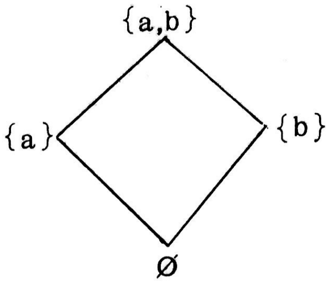


Figure 1.

As an exercise the reader may care to verify that the set {a.b.c.) has eight subsets forming the following diagram:

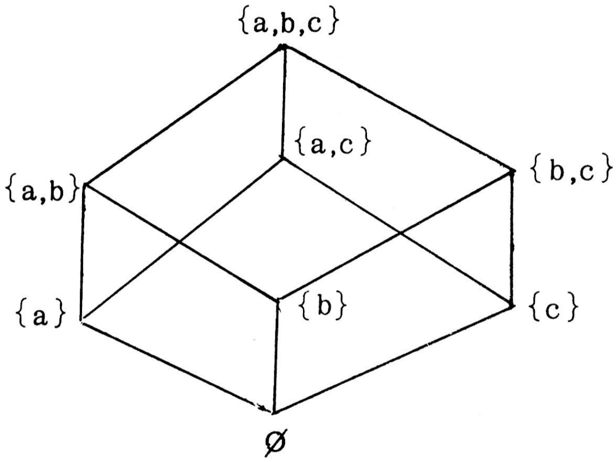


Figure 2.

1.3 If A and B are sets, A ∪ B represents the union of A and B, that is the set of objects which are in A or in B or in both. Similarly A∩B represents the intersection of A and B, that is the set of objects in both A and B, that is, common to both sets. Thus

$$\{a,b\} \cup \{a,c\} = \{a,b,c\}$$

$$\text{and } \{a,b\} \cap \{a,c\} = \{a\}.$$

1.4 In terms of the above diagrams we may note that A ∪ B is the 'lowest' set lying 'above' both A and B in the diagram whereas A ∩ B is the 'highest' set lying 'below' both A and B. Such a structure of objects (sets in the above case) where, to each pair of objects there corresponds both a 'least greater' and a 'greatest smaller' object, also in the set of objects, is called a lattice and will be our basic algebraic device. It will suffice for present purposes if the reader accept that some lattices have a certain algebraic property not possessed by all lattices. These special lattices we shall call D-lattices[5].

5. This property is in fact that of distributivity and defines a more general type of lattice than Boolean lattices.

1.5 A function is a particular type of correspondence, mapping or relation between objects in one set and objects in another. If an object *a* in the first set corresponds with an object *b* in the second, *b* is said to be the image of *a*. Such a correspondence is then a function if and only if to each object in the first set there corresponds at the most one image in the second. The inverse relation of a function is that relation which reverses the direction of the original correspondence, that is to say that images become objects and vice versa. Such an inverse can be itself a function if and only if every member of the second set is an image of some member of the first set and if the original function corresponds one object to one image in all cases.

1.6 Suppose that a function f maps a set A into a set B and that the sets A and B are, in fact, lattices (e.g. the set $\{\{a,b\}, \{a\}, \{b\}, \emptyset\}$ forms a lattice in fig.1). The function f is called a homomorphism if it preserves unions and intersections. Thus if f(a) $= \alpha$ (i.e. α is the image of a under the mapping f) and f(b) $= \beta$ then

$$\alpha \cup \beta = f(a \cup b) \quad \text{and} \quad \alpha \cap \beta = f(a \cap b).$$

It will suffice if the reader understands that a homomorphism preserves the algebraic structure. In particular, $\alpha \subset \beta$ if and only if $a \subset b$. If f and its inverse are both functions, then if f is a homomorphism its inverse is also a homomorphism and, in this case we say that f is an isomorphism.

1.7 Let A be a set and let B $\subset$ A. If A is also a lattice and, under the same operations of union and intersection, B is a lattice in its own right, then B is called a sublattice of A. Not every subset of a lattice is a sublattice.

In a lattice the set of objects lying below any one given object is called an ideal[6] and always forms a sublattice. Thus in fig. 2 the set $\{\{a,b\}, \{a\}, \{b\}, \emptyset\}$ is the ideal corresponding to the object $\{a,b\}$ and $\{\{a\}, \emptyset\}$ is the ideal corresponding to $\{a\}$.

We shall need the following two mathematical results:

A: Every sublattice of a D-lattice is a D-lattice

B: Every homomorphic image of a D-lattice is a D-lattice. With these technical preliminaries we may proceed to our subject matter.

2. The semantic component

We shall make the initial, simplifying assumption that we have to analyse a corpus of simple, declarative sentences, in particular a corpus containing no negations, interrogatives or other transformational derivatives. This restriction is inessential but will serve to ease the subsequent arguments, it will in any case be further considered at the end of section four. Furthermore we can begin by analysing any sub-corpus, for there is a uniform method available by which the model may be extended as new data becomes available.

2.1 Since we wish to explore the relation between two differing views of semantics we must be able to model both aspects and we shall begin by positing a model for 'concepts' which can be named in the given language. These concepts may be imagined as existing in the mind of the informant or in the collective mind of a society and must not be confused with the symbols naming them. This set of concepts may be considered as having an ordering imposed on them by the speaker's notion of logical fact, the 'meaning' of a given concept being essentially its position within this ordering (or lattice as it transpires). Thus the meaning of a concept will be determined by the ideal that the concept determined in the semantic lattice.

6. Actually a principal ideal.

As an example of this ordering consider the concepts represented by the words *human, mammal* and *animal.*

Then "human" $\subset$ "mammal" $\subset$ "animal"

in that every human is a mammal and every mammal is an animal. It should be noted that we are not imposing any *a priori* logical structure on the set of concepts, but only representing the beliefs of the speaker; thus if he believes that all linguists are mad, then the relation

"linguist" $\subset$ "mad person"

will appear in his semantic lattice. Thus this lattice will reflect the cultural use of the concepts rather than their logical use.

2.2 In order to discover more of the nature of this lattice we shall represent concepts more specifically by identifying each concept with the set of all objects, concrete or abstract, which possess the attribute determined by that concept. Thus the concept named as *cow* will be represented by the set of all bovines, the concept 'fast' by the set of all objects possessing the attribute of speed. In certain cases especially where a relation is denoted, a concept may be representable not by a set of objects but rather by a set of pairs of objects or, in general, a set of n-tuples of objects. Thus when concepts such as "thought" and "act" are coupled, as in a verb, the result may be represented by a set of pairs where the first object is an object capable of thinking and the second object is a "thought". (This pair forming is a feature of the concept of a "transitive Act".) The finer details of this representation of concepts by sets is a matter for consideration when we explore the methods by which language forms larger units from smaller; for the moment we need only posit the assumption that concepts can be represented as sets in some such manner.

We now make an assumption concerning the mental abilities of our informant, namely that given two concepts he can conceive the 'union' and 'intersection' of these concepts and decide whether these derived concepts may be of interest to him or not. Thus given the concepts "male" and "bovine" he can conceive the intersection, viz. "bull" and the union, viz. "those objects which are either male or bovine".

We also make a linguistic assumption, to be used later, that he can name all such conceptions, possibly with the use of phrases.

By the first of these assumptions we ensure that the lattice of concepts is a D-lattice; that is to say, an assumption that our informant's conceptual universe is ideo-logical suffices to guarantee that it forms a D-lattice, however peculiar his beliefs as to what constitutes logic may be.

2.3 We may further note that ideals in this lattice constitute what we may term semantic categories. Thus all concepts which carry the attribute of animality lie precisely in that ideal headed by the concept "animal". Such ideals also constitute D-lattices.

3. The syntactic component.

3.1 We next consider the set of meaningful units of the language, be they morphemes, words or phrases[7]. To each such a unit corresponds a set of environments, namely those strings of words with a single gap in them which become phrases or sentences of the corpus when the given unit is placed in the gap.

---

7. In the Bloomfieldian sense.

We form a lattice [8] by identifying each such unit with the set of its environments and by ordering these sets by the subset relation. An immediate difference between our two lattices now becomes apparent if we consider any two units. Given their respective sets of environments it by no means follows that there will be any word or phrase whose set of environments coincides with either the union or the intersection of the given sets. Thus the set-theoretic union and intersection need not be in the syntactic lattice: instead the 'union' in this lattice will be that unit whose set of environments is the smallest which contains the given sets and the 'intersection' is defined similarly. This immediately implies that the syntactic lattice need not be a D-lattice.

3.2 It is moreover a feature of all languages that there are words whose sets of environments have no elements in common and which do not both lie in a single set of environments corresponding to some other word or to some phrase. This feature is in fact a necessary consequence of the existence of grammatical categories without which the syntactic code could not exist. This feature also suffices to guarantee that the syntactic lattice cannot be a D-lattice. We have therefore formalized an essential distinction between the structures of the syntactic and the semantic lattices.

3.3 We may note that an ideal in the syntactic lattice represents a distribution class of words and phrases, but not a grammatical category. All units in a given ideal can be linked by a chain of common environments or, perhaps more to the point, the usage of any given unit is given by its position in the syntactic lattice and, equivalently, by the ideal which it heads.

4. Mappings between the lattices

A dictionary presents a relation between the lexicon and the set of denotations or concepts. The relation from lexicon to concept fails to be a mapping (or function) insofar as homonyms exist and the inverse relation fails to be a mapping insofar as synonyms exist, but both these problems are, as will be seen, easily solved.

We may indeed agree that a word has a certain denotation but it may be argued that the word points towards the speaker's meaning for that denotation and can thus be interpreted only from his usage of that word which should, ideally, reflect the position of the denotation in his semantic universe. We wish to consider how true this reflection is able to be.

4.1 If the meaning of a word is defined by its usage, and the same argument will of course apply to the other syntactic units, then two words of closely related meaning should have similar distribution. In particular if for two concepts a and b, a $\subset$ b, then whatever can be said positively of b can also be said of a and so the environments of *b* should form a subset of those of *a* (where *a* and *b* now denote the names of a and of b.) In terms of our mathematical model this means that if we derive a new semantic lattice [9] by reversing the ordering of the structure described in section 2, then a mapping between the syntactic lattice and this derived semantic lattice will be a homomorphism. Thus the hypothesis that meaning is usage will be justifiable exactly to the extent that such homomorphisms are possible. We shall refer to the syntactic lattice henceforth as ν and to this derived semantic lattice as $\mathcal{C}$.

That at least some degree of homomorphism should be possible is required by the condition that a language can represent the world as the speaker conceives it.

Note, in particular, that if usage defines meaning then there should be a homomorphism from ν into $\mathcal{C}$, whereas if usage is determined by meaning then there should be a homomorphism from $\mathcal{C}$ into ν.

8. See footnote 4.
9. Since $\mathcal{C}$ is isomorphic to the original lattice its introduction causes no complication.

4.2 Let us first consider mappings from $\mathcal{L}$ to ν. If we admit the existence of true synonyms we must place all such synonyms having one given meaning into a set (an equivalence class in algebraic terms) and treat each such class as a single unit. This process will not change the fact that the resulting version of ν cannot be a D-lattice. By this means we ensure that there is a mapping, in the technical sense, from $\mathcal{L}$ to ν. We further repeat our earlier assumption that all concepts in $\mathcal{L}$ have a referent in ν and further that no unit in ν is vacuous in that it refers to no concept, thus ensuring that we have a mapping f from $\mathcal{L}$ onto the whole of ν.

If now f were a homomorphism it would follow from *B* (section 1.7) that ν would be a D-lattice, which we know it not to be. Hence no mapping from the whole of $\mathcal{L}$ onto the whole of ν can be a homomorphism.

4.3 Let us now consider relations from ν to $\mathcal{L}$. To ensure that any such relation is a mapping we must ensure that there are no unresolved homonyms present. To do this we assume that homonyms are recognizable in ν (see section 5.2) and use the lexicographer's traditional device of indexing such terms. In this we need only to add the above assumption that all units in ν have some denotation and we ensure that a relation g, from ν to $\mathcal{L}$ is a mapping.

Now ν is a lattice and so, if g is a homomorphism, then the image $g(\nu)$ of ν will be a sublattice of $\mathcal{L}$ and will therefore, by *A* (section 1.7), be a D-lattice. But by removing homonyms and synonyms as described we ensure that the inverse of g is also a homomorphism from a D-lattice $g(\nu)$ to ν and so, by *B*, ν would appear to be a D-lattice; again a contradiction. We therefore conclude that there can be no homomorphic mapping of the whole of ν onto $\mathcal{L}$.

4.4 We may clarify this situation by making use of the mathematical fact that there can be a homomorphism between two lattices if and only if there is a mapping from the ideals of the one lattice to those of the other, this mapping preserving order (in terms of set-theoretic inclusion of the ideals.) Thus there is a homomorphism from ν to $\mathcal{L}$ only if, to every distribution class of units of ν, there corresponds a semantic category, that is a meaning; conversely there is a homomorphism from $\mathcal{L}$ to ν only if, to each "meaning", there corresponds a word with a given usage. Neither of these latter conditions therefore can hold and thus the assumption of the equivalence of usage and meaning cannot hold uniformly throughout the lexicon.

4.5 All ideals in $\mathcal{L}$ are necessarily D-lattices and the lattice ν contains many sublattices in the form of ideals which may be locally D-lattices. The possibility therefore remains open, in any given language/culture situation, that certain semantic classes may be represented in a 'regular' manner within certain distribution classes of words and phrases. It is also possible that complex assemblages of distribution classes[10]; such as which define grammatical categories, may map 'regularly' into the concept lattice but, in this case, they cannot coincide exactly with any similar assemblage of semantic categories.

4.6 Summarizing this result, we see that language betrays reality on two levels; firstly insofar as the formation of concepts is an arbitrary decomposition of a continuum into a finite, structured set (that is, a lattice in our model) and secondly that, in order to operate as a code, the symbols by which these concepts are represented can no longer faithfully represent the structure on the set of concepts, which structure is the essence of meaning. Only locally, that is to say over restricted semantic fields, can this coding process possibly remain faithful to the conceived structure.

10. These assemblages are Boolean functions of distribution classes.

4.7 A corollary of this result concerns the use of semantic criteria for translation between languages and for language learning. Let one language have the lattices ν_1 and $\mathcal{L}_1$ and the other the lattices ν_2 and $\mathcal{L}_2$. In order to translate an expression composed of items from ν_1, into an expression composed or items from ν_2, taking account of the semantic structure of the two languages, it is necessary to find three mappings; f from ν_1 to $\mathcal{L}_1$, g from $\mathcal{L}_1$ to $\mathcal{L}_2$ and h from $\mathcal{L}_2$ to ν_2. The divergence of structure between $\mathcal{L}_1$ and $\mathcal{L}_2$ that is between the conceptual universe of two speakers or speech communities, is a well known problem; a considerable problem in the case that the two languages involved are mutually exotic but still a barrier to communication between two individuals speaking ideolects of the same one language. The lesser problem, namely that different languages use their respective lexicons to label different concepts, is easily solved by forming definitions where needed. This problem however concerns the whole structure of $\mathcal{L}_1$ and $\mathcal{L}_2$ and limits the extent to which g can be a homomorphism or rather to which g can be represented by a set of local homomorphisms and this in turn limits the degree of accuracy of any uniform process of translation.

As we have shown, the mapping f cannot be a homomorphism over the whole of ν_1 but may possibly be so for fairly large subsets of ν_1. The mapping h, on the other hand, is much more restricted (as stated in section 4.5), homomorphism being restricted to sets which map into distribution classes and so, apart from restrictions imposed by g, h will ultimately determine the effectiveness of a translation process at a lexical level. This function h is, however, the function involved in any attempt to give semantic criteria for syntactic usage. Hence the problem of translation, at this level, can be solved with just that same order of crudity as is provided by a semantically orientated grammar. It follows that a criterion which we have demonstrated to be inadequate for the linguist may well be justifiable in certain pragmatic situations.

4.8 It will be recalled that, at the outset, we posited a corpus consisting of declarative sentences only. Let us consider, for a moment, say, the notion of negation. It is a feature, deriving solely from the fact that $\mathcal{L}$ is a D-lattice, that the subset of negated concepts will form a sublattice of $\mathcal{L}$ homomorphic to $\mathcal{L}$ itself. Thus these types of lexical transformations correspond to endomorphisms of $\mathcal{L}$, that is to homomorphisms of $\mathcal{L}$ into itself. In that ν is not a D-lattice a corresponding property cannot obtain over the whole of the syntactic lattice but only within, at the best, certain distribution classes.

5. APPENDIX: EXTENSION OF THE MODEL TO PHRASE AND SENTENCE LEVELS

Pursuing our aim of avoiding technical discussion we cannot readily provide any comprehensible description of such extensions but instead will simply describe the process to be followed.

5.1 The algebraic process which describes the ordering of 'meaningful' units into the lattice ν may readily be applied to strings of units from ν; in this case we derive an analysis of sentence types in terms of the substitutability into differing environments of differing distribution classes.[11] Each such distribution class is a sublattice of ν and "sentence space" can then be modelled as the lattice product of sublattices derived in this manner[12]. In an ideal situation each such sublattice would correspond to a semantic category, that is an ideal in $\mathcal{L}_1$ on which basis sentences could be interpreted[13] in terms of ordered relationships of concepts. It follows immediately however that the non-existence of homomorphisms at the lexical level will carry over to this extended model, since a product of lattices is a D-lattice if and only if all its components are D-lattices. Our general results concerning the 'ineptitude' of language will thus be duplicated at this more realistic level.

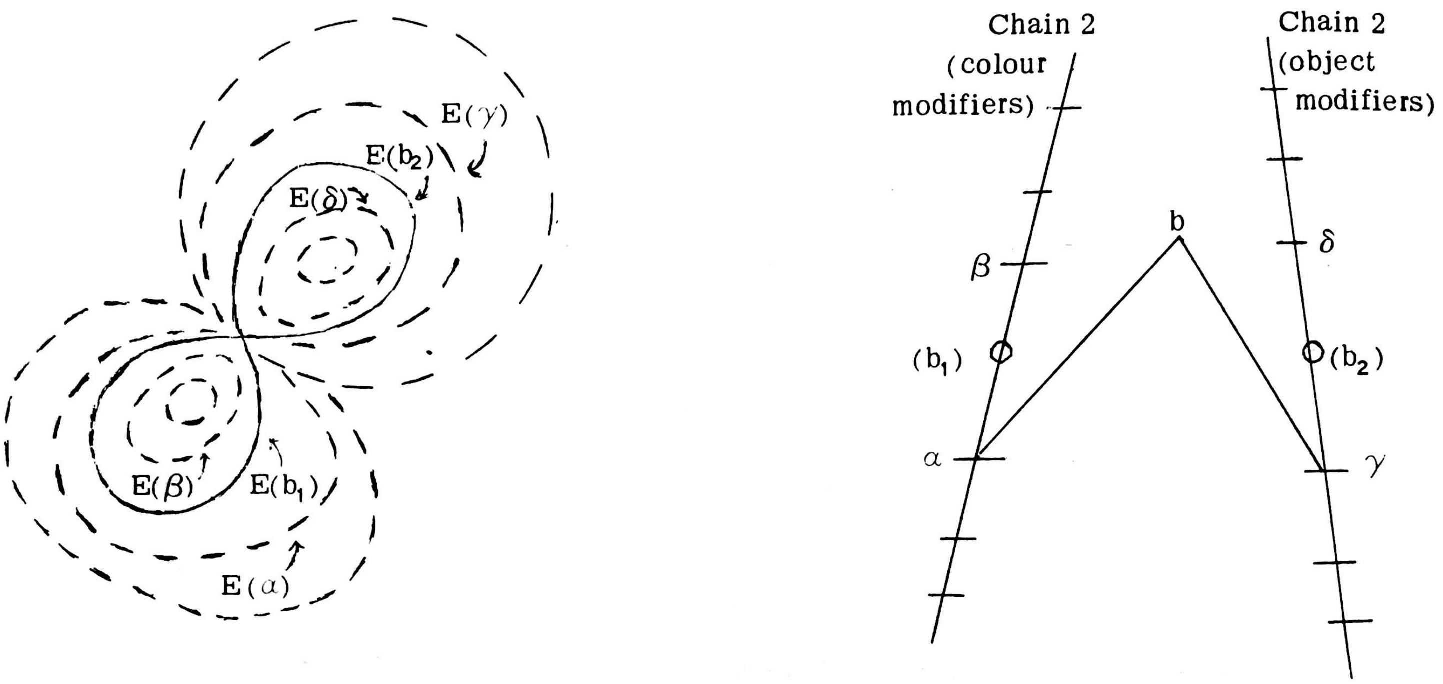
It may also be noted that endomorphisms will again play the rôle of certain types of transformations. (Cf. section 4.8).

11. See Kunze 1967, Revzin 1966 and Marcus 1967.
12. See Ashton 1968.
13. The process having points in common with Katz and Postal 1964.

5.2 In conclusion we shall attempt to illustrate the notion of semantic and syntactic lattices in a consideration of a resolution of a case of homonymy. To do this we must appeal to a notion of phrase analysis not explicitly dealt with in this paper; we consider only the simplest type of linkage[14] interpreting the result in terms of a simple intersection.

Let us consider the lexical item 'bright' both as a modifier of colour words and as a modifier of object words and consider phrases constructed from this item and the items 'green' and 'lake'. Designate these three items by 'b', 'g' and 'l' respectively. Let α and β be items having environments contained in and containing the environments of b as a colour modifier respectively.

Let γ and δ have a similar relationship to b used as an object modifier. (Our *a priori* knowledge of b justifies the existence of such items.) These set of environments and the corresponding section of the lattice ν are illustrated in figure 3 below.



$E(b) = E(b_1) \cup E(b_2)$

$E(b_1) = E(b) \cap E(\alpha)$

$E(b_2) = E(b) \cap E(\gamma)$

Figure 3

Suppose that we had failed to recognize the homonymity of b and had attempted a semantic analysis of the possible phrases derivable from these lexical items, that is to discover their semantic relationship.

We should derive the phrases *bright green* (bg), *bright lake* (bl), *green lake* (gl), *green, bright lake* (g.bl) = *bright, green lake, bright green, bright lake* (bg,bl) and *bright green lake* (bgl) = *green, bright green lake;* the contractions and synonyms being derived from the semantic facts that anything that is green and bright is also bright and green, and that *bright green* $\subset$ *green.*

14. See Weinreich, 1961, section 3.

We first form the only lower half of a sublattice consisting of all intersections derivable from 'b', 'g' and 'l'.

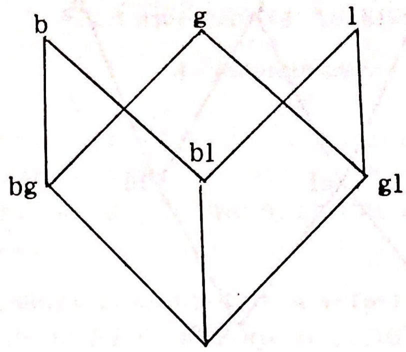


Figure 4.

We find here that bg, bl is indistinguishable from bgl, that is that an ambiguity results.

In the same way let us now consider the only partial sublattice of the intersections from four items b_1, b_2, g and l, the two usages of b having been distinguished by indexing. It should be noted that this distinction is definable in the set of environments in terms of intersections as indicated in figure 3. In figure 5 we represent this partial sublattice, with the symbol * representing culturally or logically inadmissible concepts and with broken lines representing orderings which will consequently become vacuous.

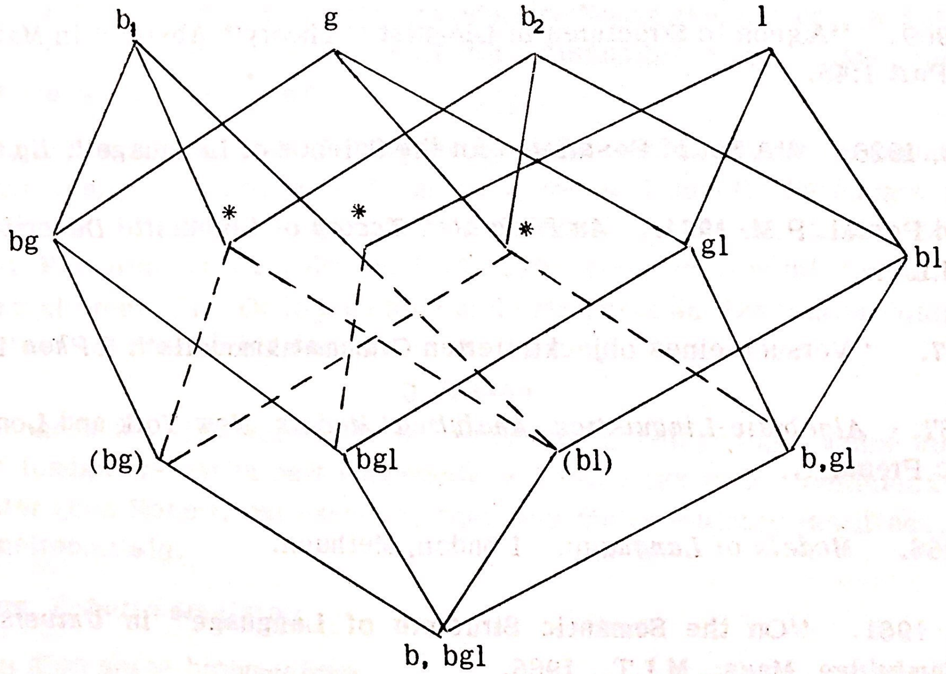


Figure 5

Finally we represent the modified partial sublattice from which all vacuous items have been removed, leaving a section of a D-lattice (but not a Boolean lattice) with no ambiguities remaining.

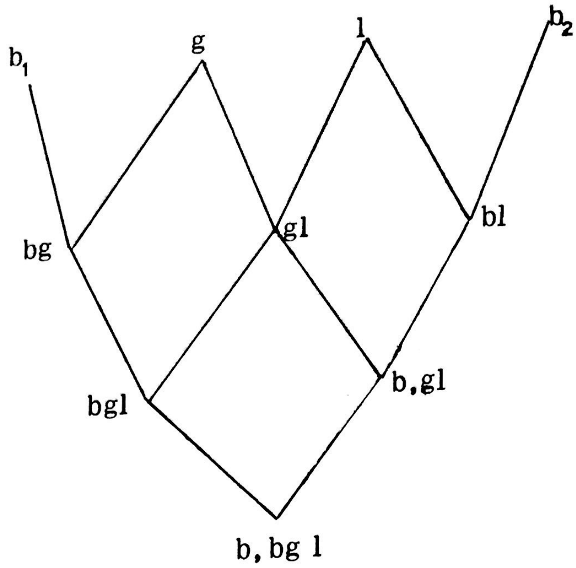


Figure 6.

5.3 The results presented in this paper refer to a certain mathematical structure and are true of natural languages insofar as this structure is a model for certain aspects of language. That certain deductions from this model seem to explain actual linguistic phenomena is evidence towards the validity of such a model.

BIBLIOGRAPHY

Ashton, K., 1968. "Towards a Mathematical Model for Semantics" *Te Reo* 10-11: 1-8.

Ashton, K., 1969. "Algebraic Structures in Linguistic Theory" Abstract in *Mathematical Chronicle* 6, Part 1:45.

Bloomfield, L., 1926. "A Set of Postulates for the Science of Language". *Lg* 7.

Katz, J.J. and Postal, P.M. 1964. *An Integrated Theory of Linguistic Descriptions*. Cambridge, Mass., M.I.T.

Kunze, J., 1967. "Versuch eines objecktivierten Grammatikmodells". *ZPhon* 20.

Marcus, S., 1967. *Algebraic Linguistics: Analytical Models*. New York and London, Academic Press.

Revzin, I.I., 1966. *Models of Language*. London, Methuen.

Weinreich, U., 1961. "On the Semantic Structure of Language" in *Universals of Language* Ch. 7. Cambridge, Mass., M.I.T., 1966.

Wittgenstein, L., 1922. *Tractatus Logico-Philosophicus*. London.