DOCUMENT RESUME

ED 076 083 FL 004 114

AUTHOR	Zatorski, R. J.
TITLE	Recovering Scenes from Linguistic Representations.
INSTITUTION	Monash Univ., Clayton, Victoria (Australia).
PUB DATE	70
NOTE	21p.; In Linguistic Communications, 2, 1970. Paper read at the Conference on Theoretical Linguistics University of Sydney, Australia, August 1970
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Deep Structure; *Descriptive Linguistics; English; Graphemes; Language Research; *Linguistic Theory; *Models; Morphemes; Phonology; *Semantics; *Sentence Structure; Surface Structure; Transformation Generative Grammar; Transformations (Language); Transformation Theory (Language)

ABSTRACT

This paper cites the inadequacy of transformational generative grammar theories in their attempts to describe the meaning of a given sentence. The author sees the specification of meaning involving the recovery of the particular section or sections of the world model communicated or represented by the sentence. As a corollary, the author argues that sentences of English are essentially representations of "scenes" in the world model and not representations of objects. The author proposes a model for sentence analysis which seeks to recover meaning by rejecting the notion that sentences are "nothing but" objects and that their "disambiguations" are the representations of these objects. Recovery of scenes from sentential representations is considered equivalent to meaning-specification for language. (VM)

In: Linguistic Communications, 2, 1970.

Paper read at the Conference cn Theoretical Linguistics, August, 1970, University of Sydney. An abbreviated version was presented at the 13th Congress of A.U.L.L.A. August, 1970, Monash University.

RECOVERING SCENES FROM LINGUISTIC REPRESENTATIONS

R. J. Zatorski

University of Melbourne

A. The "standard" theory of transformational grammar has been described in Chomsky (1969) as

$$1. \qquad \Sigma = (P_1, \ldots, \underset{\substack{\downarrow \\ S}}{P_{\underline{i}}} \ldots, \underset{\substack{\downarrow \\ P}}{P_{\underline{n}}}),$$

where P_1 is the initial phrase marker, $P_{\underline{i}}$ is the deep structure and $P_{\underline{n}}$ the surface structure; and where P is a phonetic and S is a semantic representation. In this account, S is derived from $P_{\underline{i}}$ by rules of semantic interpretation (Katz and Postal, 1964) and P - from $P_{\underline{n}}$ by phonological rules. P_1 is generated by the base lexical insertion does not occur beyond $P_{\underline{i}}$. The statement above that S and P are representations should not escape attention. It is supported by a number of other remarks, including one (p.1 of mimeographed version) which designates phonetics and semantics as "two universal language-independent systems of representation", one specifying sound and the other - meaning. The former is said to have been given a specification in Chomsky and Halle (1968); the latter lacks a "reasonably concrete or well-defined 'theory of semantic representation' to which one can refer".

Chomsky then examines in considerable detail a series of challenges to the "standard theory" and concludes that the "semantically based" grammar proposed by McCawley (1968b) is merely a notational variant of the "standard theory". Furthermore, he accepts the conclusions of Akmajian (1968), Dougherty (1968a) and Jackendoff (1967) that the properties of surface structure P_n also contribute to the derivation of semantic representation, and goes on to propose that 1 should be altered to

2. $\Sigma = (P_1,\ldots,\underbrace{P_i}_{S},\ldots,\underbrace{P_n}_{P})$

Currently, Chomsky's theory has the following reconstructed form:

3. (i) Base: $(P_1,\ldots,P_i)$;

(ii) Transformations: $(P_i,\ldots,P_n)$;

(iii) Phonology: $P_n \longrightarrow$ phonetic representations;

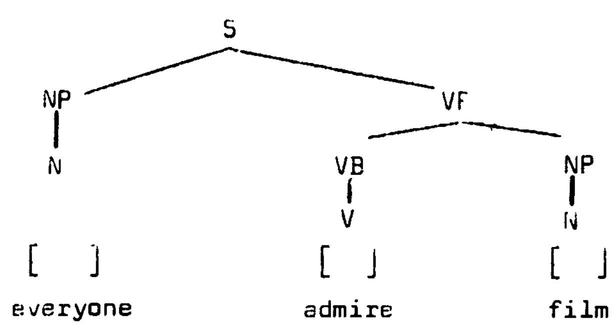
(iv) Semantics: $(P_i,\ P_n) \longrightarrow$ semantic representation (the grammatical relations involved being those of P_i, i.e. those represented in P_1).

On this view, the phonetic equivalent of, e.g.,

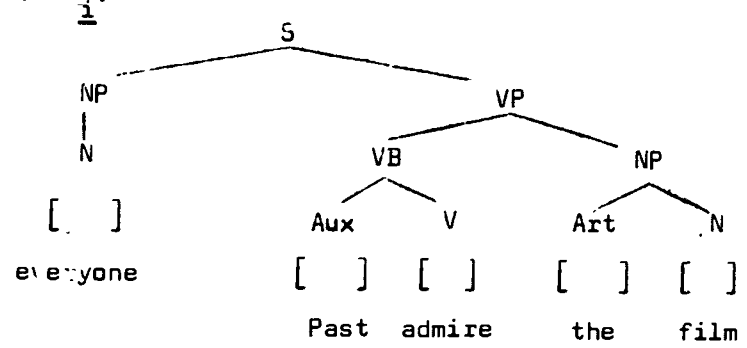
4. The film was admired by everyone,

is assumed to be one of the two representations of the complex linguistic object:

5. (i) P_1



(ii) $P_{\underline{i}}$:



(iii) $P_{\underline{n}}$: The film was admired by everyone.

The same object has a representation on the semantic level via a reading (or set of readings) assembled with the aid of appropriate projection rules. The "meaning" of the sentence 4 is identified in this approach with the recovery of the structure 5 (i)-(iii) from its phonetic and its semantic representations.

Whatever the "message content" communicated by the sentence, therefore, the current version of the transformational grammar accounts only for those aspects of it which have overt linguistic manifestations via subcategories ($\underline{+}$ animate, $\underline{+}$ feminine, etc.), features ($\underline{+}$near*, $\underline{+}$penetrable**, etc.) and semantic markers (Physical Object, Adult, Male, etc.). On this view, the "reality" of the action represented in

---

* Jacobs and Rosenbaum (1968)

** Langendoen (1969)

6. John reads books

is exclusively that recoverable from a structural-linguistic account which exhibits the identity of the lexical features on <u>John</u> and <u>books</u> with the features specified for the subjects and the objects of <u>read</u> respectively, and showing that the semantic selection restrictions have been satisfied for each set of homonym distinctions, i.e., for each derivable reading.

The basic theory of transformational grammar (Lakoff, 1969) likewise defines a "semantic representation (SR) of a sentence" (p.3), though not in terms of projection - amalgamation on the Katz and Fodor (1963) model but in terms of the assembly of logical predicates whose argument places include presupposition, topic, focus, etc., and which are subject to the operation of global and local derivational constraints (rules). The theory is rather obscure concerning the exact level at which lexical insertion occurs and no illustrative model for English appears to have been designed.

McCawley (1968) [illegible] that in the generative semantics model only indices and logical predicates will constitute terminal elements of semantic representations but makes no positive proposals about mapping from such representations to morphemes. Insofar as the generative semantics theory may possibly be addressing itself to the specification of meaning at some "pre-morphemic" level, there may be some affinity between its conception of language and that advocated in the present paper; to the extent, however, that semantic representations appear to constitute disambiguations in the sense of Katz (1966) - in this case via the logical operations - there is no significant difference between the two theories in the conception of "sentential meaning" and we may follow Chomsky in treating the "basic" theory as a variant of the "standard" theory.

With respect to the latter, however, it will now be argued that since it is descriptive of only the structure of the object: sentence alone, it is completely adequate only for a restricted class of sentences, of which

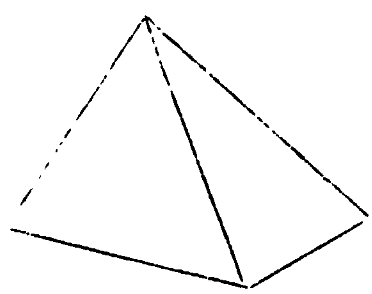
7. (i) <u>John</u> is an element of a sentence;

(ii) colourless green ideas sleep furiously, etc.

are typical and that for all other sentences of English the specification of meaning involves the recovery of the particular section or sections of the world model communicated or represented by the sentence. As a corollary, it will be argued that sentences of English are essentially representations of "scenes" in the world model and not representations of the object 2. and furthermore, that they cease to be such and become representations of the object 2. when no scenes can be recovered from them or when the recoverable scene is "disregarded" in favour of the sentential structure itself, i.e. when the sentences themselves become objects or "linguistic scenes" in the world model.

B. It may be useful to pose the question simultaneously in two communicational domains, to broaden the perspective. In graphics, the picture:

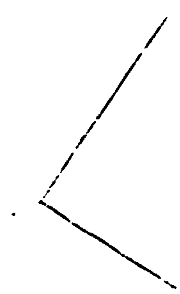
8.



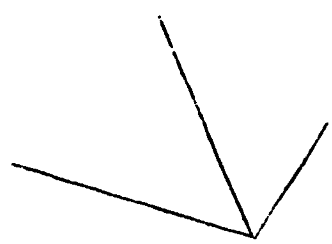
has the status equivalent to that of a sentence and, like any sentence, can be "looked at" in two distinct ways: as an object and as a representation.

(a) As an object in two dimensions (henceforth: 2D), the graph in 8 has a description which may be assumed to consist of a transformational elaboration of a one-to-one mapping from a specification of its parts and attributes (lines, junctions), and relations ($\pm$ coincidence, $\pm$ congruence, $\pm$2D orientation) into natural language morphemes or equivalent graphic units, say, graphemes. Such an object has no representation in two dimensions; its graphic or linguistic description is directly mapped from its abstract specification i.e. <u>its description is its only possible disambiguation</u>. As an example, consider two of the parts into which the object in 8 can be articulated:

9. (i)

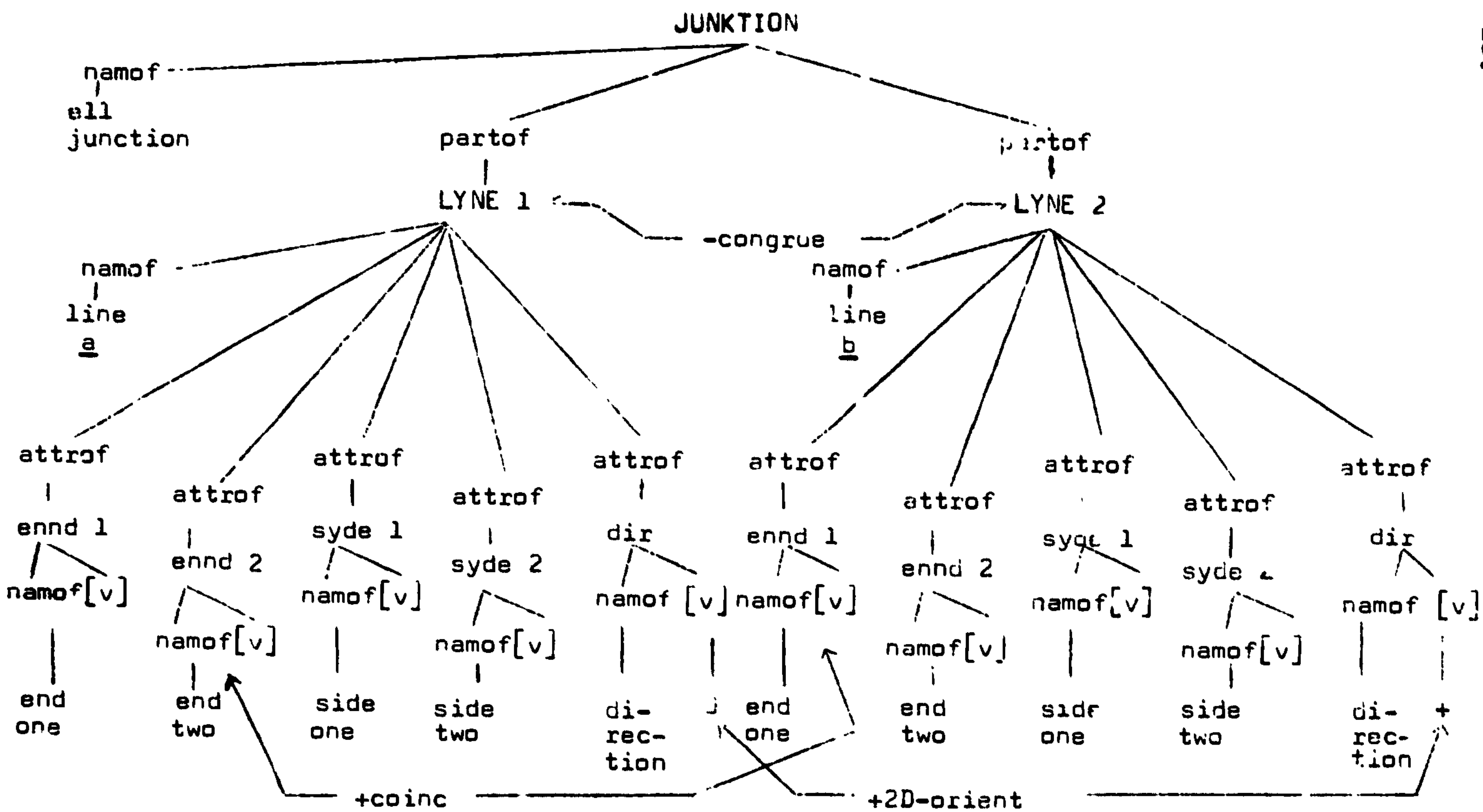


(ii)



We can assign to 9 (i) a structural specification which enumerates its parts and attributes and makes explicit their mutual relations, e.g.

10.


JUNKTION
namof
all
junction
partof
LYNE 1
partof
LYNE 2
-congrue
namof
line
a
namof
line
b
attrof
ennd 1
namof[v]
end
one
attrof
ennd 2
namof[v]
end
two
attrof
syde 1
namof[v]
side
one
attrof
syde 2
namof[v]
side
two
attrof
dir
namof [v]
di-
rec-
tion
attrof
ennd 1
namof[v]
end
one
attrof
ennd 2
namof[v]
end
two
attrof
syde 1
namof[v]
side
one
attrof
syde 2
namof[v]
side
two
attrof
dir
namof [v]
di-
rec-
tion
+
+coinc
+2D-orient

The relations governing the assembly of 10. are:

+coincidence ⇛ ⟨<u>ennd 2</u> (LYNE 1) = <u>enna 1</u> (LYNE 2)⟩

+2D-orientation ⇛ ⟨v (<u>direktion</u> (LYNE 1)) = 0

v (<u>direktion</u> (LYNE 2)) = + ⟩

-congruence ⇛ <LYNE 1 ≠ LYNE 2>

where the upper case entries designate objects; lower case entries: word equivalents; lower case entries, underlined: attributes; ---of : relations; {± coincidence, ± 2D-orientation, ± congruence} : meta-relations.

The table below specifies word equivalents for each of the relations in 10.

Table 1

Complex relation	Simple relation	Word-equivalent
← partof →	= partof →	have as part comprise be assembly of
	← partof	belong to be part of
← attrof →	= attrof →	have as attribute characterise + Pass
	← attrof	be attribute of characterise
← namof →	= namof →	be called designate + Pass name + Pass have as name
	← namof	be name of be designation of designee
← valof →	= valof →	have as value be valued
	← valof	be value of
→ "down", ← "up" the tree structure		

Among the meta-level conventions imposed are the following:

11. (i) on every level of articulation, successive integers are assigned to unbroken sequences of identical elements and relations;

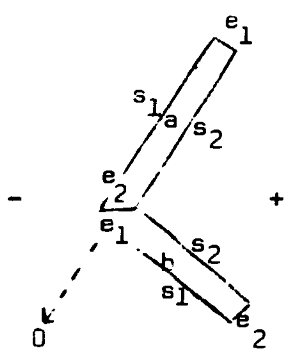
(ii) [v] on ennd and syde are specifications of length via a set of points on a point array;

[v] on direktion can be 0, + or −; LYNE 1 is arbitrarily assigned direktion 0. The areas + and − are computed in the way schematically illustrated in 12 below.

(iii) every object and attribute has a name in language; all relations, including meta-relations, have mappings to language on the lines of table 1.

(iv) = maps to: be equal; $\neq$: be not equal.

12.

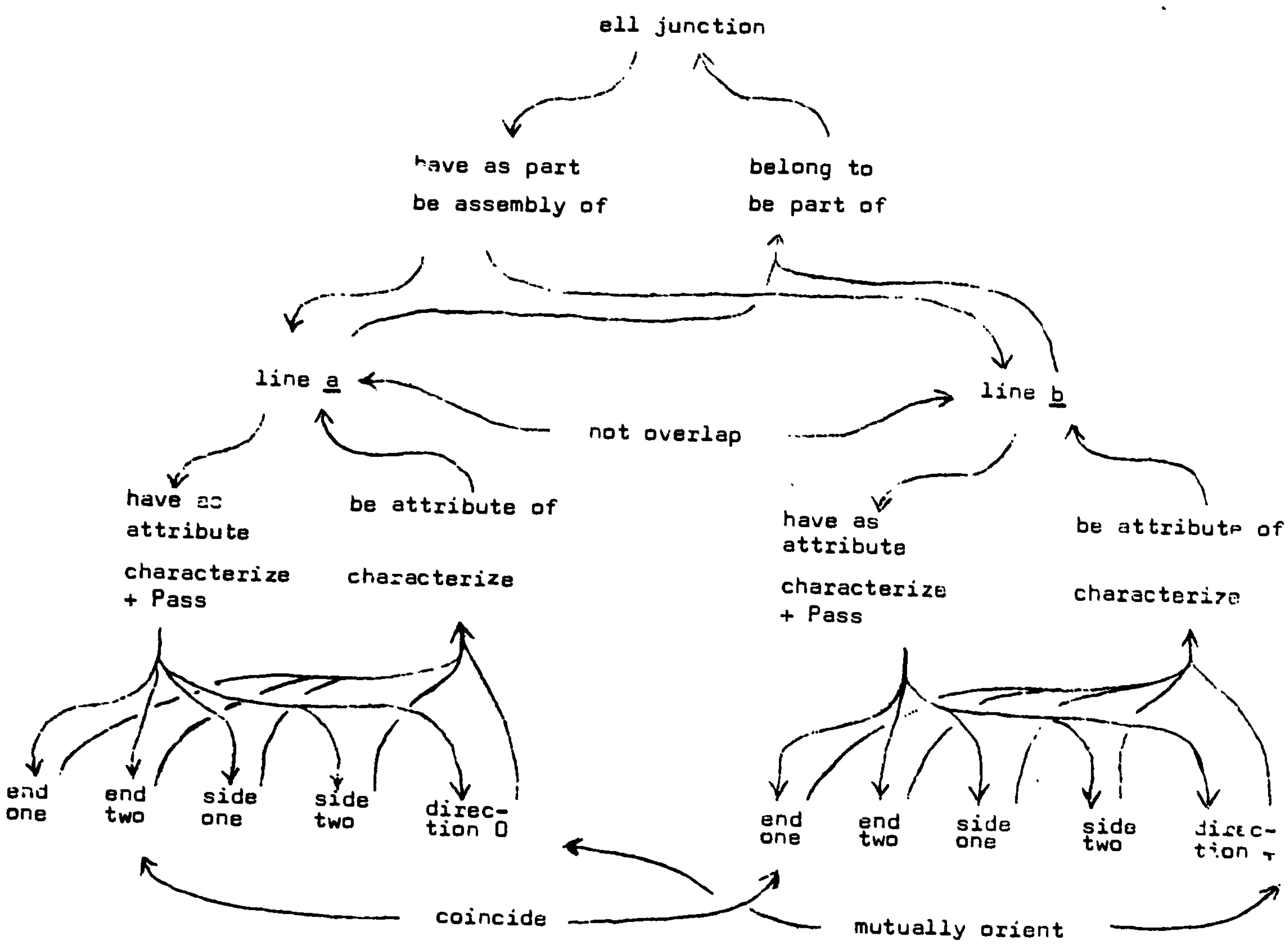


The specification of the picture/object 9 (ii) will add to 10 a new LYNE 3 with direktion: plus, and a new set of meta-relations:

13. Rel: + coinc $\Rightarrow$ ennd 2 (LYNE 1) = ennd 2 (LYNE 2) = ennd 2 (LYNE 3) $\rangle$

+2D-orient $\Rightarrow$ $\langle$ v(dir(LYNE 1)) = 0; v(dir(LYNE 2)) = +; v(dir(LYNE 3)) = + $\rangle$

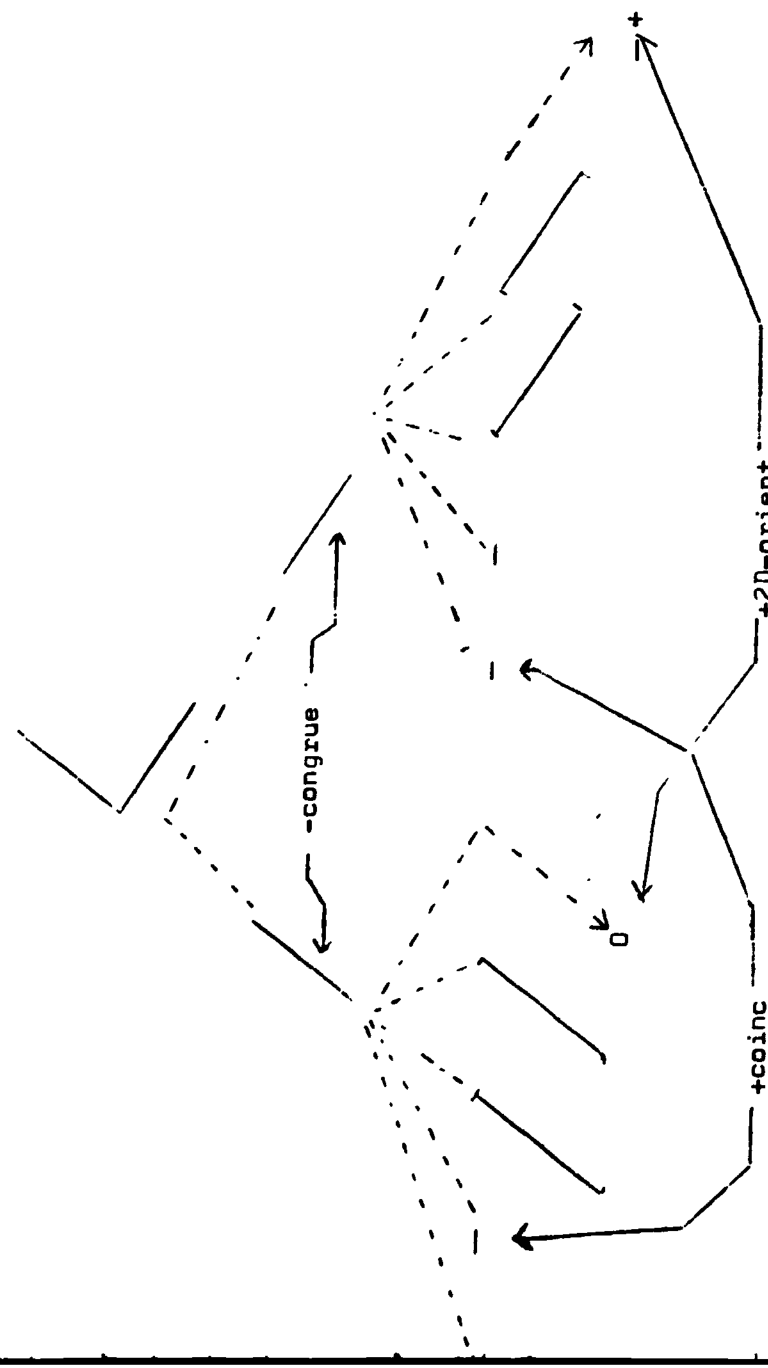
−congrue $\Rightarrow$ $\langle$ LYNE 1 $\neq$ LYNE 2 $\neq$ LYNE 3 $\rangle$


ell junction
have as part
be assembly of
belong to
be part of
line a
not overlap
line b
have as attribute
be attribute of
characterize + Pass
characterize
have as attribute
be attribute of
characterize + Pass
characterize
end one
end two
side one
side two
direc-tion 0
end one
end two
side one
side two
coincide
mutually orient

By abstracting the name nodes in 10 and mapping all relations into their word equivalents, we obtain a one-to-one specification of the structure 9 (i) in 14.

The structure in 14 specifies an ell junction exclusively in terms of its natural language name nodes; an analogous account in terms of graphic symbols would yield (in a simplified account):

15.



The structural specification in 10 enables us to generate a linguistic description of an ell junction in terms of statements. A statement is derived by abstracting the names of two adjacent nodes and the verbal equivalent of the directed relation between them. Examples of statements are:

16. (i) end two characterize line a;

(ii) line b be part of ell junction; etc.

Paragraphs are derived by abstracting the names and the relations connecting more than two adjacent nodes, e.g.;

17. (i) end one be attribute of line b. Line b be part of ell junction.

(ii) ell junction have as part line a. Line a characterize + Pass direction zero; etc.

In general, statement and paragraph derivation involves abstracting from 10 the following:

18. $S = \text{Node}_i \leftarrow \text{of-relation} \rightarrow \text{node}_j.$

$$P = \left[\begin{array}{l} \text{Node}_i \leftarrow \text{of-relation} \rightarrow \text{node}_j \\ \text{Node}_j \leftarrow \text{of-relation} \rightarrow \text{node}_k \end{array}\right]$$

Meta-relations linking nodes constitute the conditions governing the validity of the element specified by the nearest common node dominating the two linked nodes. Thus congrue in 10 which relates LYNE 1 and LYNE 2 gives rise to the statement: ell junction iff line a not overlap line b.

A linguistic description of the structure in 10 is the set of all statements derivable from it by following each internodal "path", once only in either direction. We may thus have :

19. (i) Ell junction have as part line a, line b.

(ii) Line a characterize + Pass end one, end two, side one, side two, direction zero.

(iii) Line <u>b</u> characterize + Pass end one, end two, side one, side two, direction plus.

(iv) Ell junction iff line <u>a</u>, line <u>b</u> not overlap.

(v) Ell junction iff mutual orientation be direction zero characterize line <u>a</u>, direction plus characterize line <u>b</u>.

(vi) Ell junction iff end two characterize line <u>a</u> coincide end one characterize line <u>b</u>.

A large number of descriptions of the structure 10 equivalent to that in 19 can be derived by abstracting statements "up" and "down" its tree-structural specification. Assuming that the statements in 19 constitute kernel-predicates subject to interpretive logical operations and linguistic transformations, we can gain an idea of the scope of such operations and transformations by comparing 19 with some "surface" version thereof e.g.

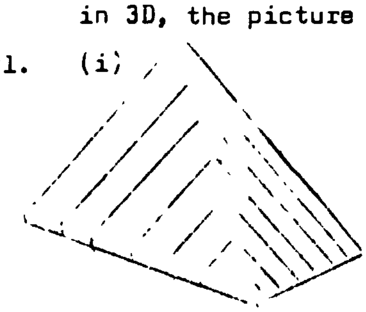
20. (i) The ell junction consists of lines <u>a</u> and <u>b</u>.

(ii) Lines <u>a</u> and <u>b</u> are each characterized by (have as attributes) ends one and two, and sides one and two but whereas line <u>a</u> is characterized by direction zero, the direction of line <u>b</u> is plus.

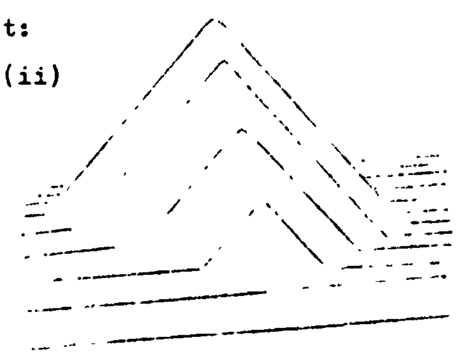
(iii) A junction is a valid ell junction when lines <u>a</u> and <u>b</u> which are its parts do not overlap, when the mutual orientation of lines <u>a</u> and <u>b</u> is such that <u>a</u> has direction zero and <u>b</u> has direction plus, and when end two of line <u>a</u> coincides with end one of line <u>b</u>.

(b) As a representation in 2D of a solid polyhedron, existing in 3D, the picture 8 may depict:

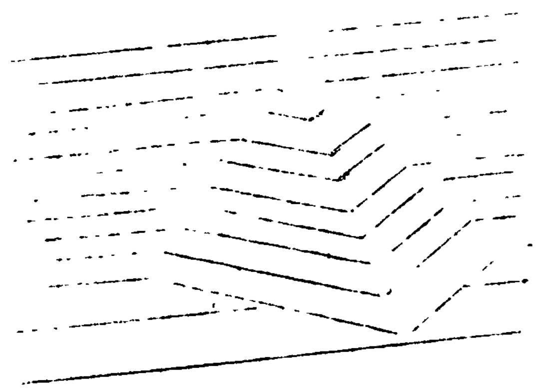
21. (i)



(ii)



(iii)



in which the elements of 8. are not individual parts of an object but contribute graphic representations in 2D of the elements of solid polyhedra specifiable (i.e. for which descriptions exist) only in 3D, such as edges, surfaces and corners, mutually related by convexity, concavity, adjacency, 3D-orientation, etc. These elements and relations must be mapped from each of the three abstract specifications underlying the three different objects in 21. Simultaneously, of course, each of the graphs in 21 (i) - (iii) constitutes a "graphic" object in 2D in the sense of section (a) above.

Thus, in general, when a graph is comprehended as <u>a representation</u> of a solid, its elements are, in turn, representations of the elements of that solid; when it is comprehended as <u>an object</u> in 2D, each of its elements is itself an object.

If the analogy holds for language, then the sentence:

22. John saw a colourful ball

must be simultaneously:

(a) an object in 2D, e.g.

23. $[_S [_{NP} \text{word}_{\underline{\text{John}}}] [_{VP} \text{word}_{\underline{\text{saw}}}, \text{word}_{\underline{\text{colourful}}}, \text{word}_{\underline{\text{ball}}}]]$

with individual morphemes having the status of "strings-of-letters",

lexical items, etc., i. e. parts of the object: sentence 22. linked by such relations as: follow, etc.; simultaneously, it must be:

(b) a representation of one of a set of events in 3D, underlying the sentences in 24 below (dimensions are used here in a deliberately loose sense; verbal utterances, as objects, are assumed to reside in 2D; for simplicity, time is not considered at all):

24. (i) a male person called John perceived with his eyes a gay and exciting social event called ball and devoted to dancing;

(ii) a male person called John perceived with his eyes a multicoloured inflated object called ball and used in sport;

(iii) a male person called John perceived with his eyes a social event attended by people wearing multi-coloured dresses, called ball and devoted to dancing.

which depict the elements of 22 not as objects: word$_{\underline{John}}$, word$_{\underline{ball}}$, etc., associated by the relation +follow, but as representations in 2D of structures specifiable only in 3D, namely:

25. (i) <u>a male person</u>

whose structural description includes the label <u>John</u>; or a mapping to the phrase: <u>a colourful ball</u>.

(iii) <u>a multicoloured inflated object, etc</u>.,

whose description includes a label or a mapping to the phrase: <u>a colourful ball</u>.

(iv) <u>a social event attended by people wearing multi-coloured dresses, etc</u>.

whose description also maps to: <u>a colourful ball</u>, where (i) is associated with one of (ii)-(iv) by the relation:

(v) perceive with eyes

which maps to: <u>see</u>.

Thus, if we designate a representation as R , a surface manifestation (in words, in lines of a graph, etc.) as SM and a structural specification of an object, event, etc., which I subsume under the term "scene", as SS, the graphic object 8 can be defined as:

26.

$$O_8^{2D} = SM_{O_8^{2D}} + SS_{O_8^{2D}}$$

Likewise, the object 21 (i) can be defined as:

27.

$$O_{21(i)}^{3D} = SM_{21(i)}^{3D} + SS_{21(i)}^{3D}$$

When, however, 8 is viewed as a representation of 21 (i) we have

28.

$$R_{O_8^{3D}} = SM_{O_8^{2D}} + SS_{O_{21(i)}^{3D}}$$

Note, that 28 takes no account either of the SS of the object 8, or of the SM of the object 21 (i), both being irrelevant to representation.

By analogy, the situation for the sentence 22 is:

29. (i) $$S_{22}^{2D} = SM_{S_{22}^{2D}} + SS_{S_{22}^{2D}}$$

(ii) $$E_{24(i)}^{3D} = SM_{E_{24(i)}^{3D}} + SS_{E_{24(i)}^{3D}}$$

(iii) $$R_{E_{24(i)}^{3D}} = SM_{S_{22}^{2D}} + SS_{E_{24(i)}^{3D}}$$

The definitions in 29 illustrate a number of important points.

To being with, in direct contrast with Chomsky's view

that the sentence e.g. 24 (i) is a representation constructed on {22,23} , where 22 = $P_{\underline{n}}$ and 23 = $P_{\underline{i}}$ in the sense of Chomsky (1969), 24 (i) is in fact a representation of the event E^{3D} and not a "disambiguation" of any sentence. In general, representation involves SMs in a lower dimension and SSs in higher dimensions. Sentences can, of course, be translations of one another in a single dimension and this is precisely what the lexicalist semantic model specifies.

Secondly, again in direct contrast to the current lexicalist theory, representation of the eventunderlying 24 (i) by means of 22 in no way involves the structure 23 of the object 22. This conclusion directly contradicts the notion of "reading-amalgamation" by means of projection rules.

Most importantly, however, no existing version of the theory of transformational grammar takes account of the structure of the event E^{3D}. Insofar as the recovery of the meaning of a sentence may be correctly identified with the recovery of whatever the given sentence is a representation of, in the sense of this paper, neither the lexicalist nor the generative semantics model of language succeed in recovering it, although both attempt to derive meaning so as to avoid the stygma of Bloomfieldian structuralism. The roots of this problem may lie in the belief commonly held among linguists that sentences are "nothing but" objects and that their "disambiguations" are the representations of these objects. However, to maintain this is to deny language its communicational function.

C. In a recent paper, Harwood (Harwood, 1970) argued that the current problem for the theory of generative grammar is one of deciding how to deal with situations in which the various subsets of a paraphrastic set of surface structures are "derived from different initial strings $Y_1 \ldots Y_i$, without any representations of the paraphrase relation between them". An example is provided by the set:

30. S_1 an apple is in the box;
    S_2 it is an apple that is in the box;
    S_3 is an apple in the box?
    S_4 a box contains an apple;
    S_5 it is an apple that the box contains;
    S_6 does a box contain an apple?

where the subsets S_{1-3} and S_{4-6} are derived from the deep structures 31 (i) and (ii) respectively:

31. (i) $Y_1 = [_S [_{NP} [_N \text{ apple }]][_{VP} [_{Cop} \text{ be}] [_{PP} \text{ in box }]]]$

(ii) $Y_2 = [_S [_{NP} [_N \text{ box }]][_{VP} [_V \text{ contain}] [_{NP} [_N \text{ apple }]]]]$

In general, Harwood argues, linguists are unanimous in acceptingthe structures S_{1-6} as members of a single paraphrastic set, despite the fact that no means currently exist by which the paraphrastic relation between Y_1 and Y_2 in 31 could be exhibited. Harwood further points out that paraphrastic sets must also represent entailment, i.e. that the set in 30 must be supplemented with surface structures such as:

32. S_7 an apple and nothing else is in the box;
    S_8 a box contains only one apple;
    S_9 a box contains at this moment only one apple;

and that this poses the need to postulate either one or the

other of the following:

(a) a single "deeper" deep structure X underlying all S_{1-9}, viz.:

33. $X \Rightarrow \{S_{1-9}\}$ where $X \neq Y_i$

with the accompanying relaxation on the meaning - preservation constraint on grammatical rules; or

(b) an abstract expression A, specifying a set of disjunctions $\{A_{a-i}\}$ mapping to the set $\{Y_{1-i}\}$:

34. $\{A_{a-i}\}$ = many-to-many mapping $\Rightarrow$ $Y_1 \Rightarrow \{S_{1-3,7}\}$

$Y_2 \Rightarrow \{S_{4-6,8,9}\}$

in which case it might prove possible to retain meening preserving rules, at least between Y and S. The introduction of the expression A in turn requires the specification of the level at which the "content" morphemes of natural languages (designated by Harwood as M1 morphemes) enter the derivation.

We can now apply the results of our investigation to the problems posed by Harwood.

It will be seen firstly that the set S_{1-6} in 30 does address a single event; the alternation of the two morphemes apple and box in the subject position and the alternation of in and contain constitute a shift in focus, i.e. in the structure of representation, and almost certainly reflects the order in which the structural specification of the underlying event is "read-off" or abstracted for sentence-generation. S_{7-9}, on the other hand, either:

(a) address three different event specifications;

(b) represent mappings from different "zones" of a single structural specification;

(c) derive from a single specification which also underlies S_{1-6}, but via a set of intermediate "logical" operations;

(d) derive from the combination of (b) and (c).

Whatever the eventual answer to this question, S_{7-9} will have different scene specifications from S_{1-6} and Harwood's options 33 and 34 must be rejected.

We can now also say something about the level at which morphemes are introduced. If a sentence is a representation of a scene then the morphemes used in its generation are in turn representations (or SMs) of elements in the abstract structure of the scene and constitute arbitrary symbolic labels which map into morphemes, graphemes, etc. The crucial issue is that of the status of S_{7-9} with respect to scene specification; its resolution clearly involves the acceptance of the notion that the recovery of scenes from sentential representations is equivalent to meaning-specification for language.

REFERENCES

1. Akmajian, A. 1968

On the analysis of cleft sentences, mimeographed, M.I.T.

2. Chomsky, N. 1969

Deep structure, surface structure and semantic interpretation, mimeographed, Indiana Univ. Linguistics Club

3. Chomsky, N. and Halle, M. 1968

The Sound Pattern of English, Harper and Row

4. Clowes, M. B. 1970

On the interpretation of line drawings as simple three-dimensional scenes, mimeographed, Laboratory of Experimental Psychology, University of Sussex

5. Dougherty, R. 1968a

A Transformational Grammar of Coordinate Conjoined Structures. PhD Dissertation, M.I.T.

6. Harwood, F. W. 1970

On the introduction of morphemes in a generative grammar, mimeographed, University of Tasmania

7. Jackendoff, R. 1967

An interpretive theory of pronouns and reflexes, mimeographed, M.I.T.

8. Katz, J. J. and Fodor, J. A. 1963

The structure of a semantic theory, Language 39:170-210

9. Katz, J. J. 1966

The Philosophy of Language, Harper and Row

10. Lakoff, G. 1969

On generative semantics, mimeographed, University of Michigan

11. McCawley, J. D. 1968

Lexical insertion in a transformational grammar without deep structure, mimeographed, 4th Regional Meeting of the C.L.S., April,1968

12. Zatorski, R. J. 1970

Picture-language interaction in artificial intelligence, Proceedings of the Computer Communications and Display Conference, Sydney

R. J. Zatorski